Advances in Cognitive Sciences
Vol 26, No. 1, Spring 2024
Research Article
Pages 91-108

A review of emotional prosody research based on functional magnetic resonance imaging and event-related potential techniques and an introduction to the Persian emotional speech database (Persian ESD)

Nayereh Joodi[1*], Niloofar Keshtiari[2]

1. PhD Candidate of Linguistics, Allameh Tabataba'i University, Tehran, Iran
2. PhD in Psycholinguistics, Cluster of Languages of Emotion, Freie Universität Berlin, Berlin, Germany

Recieved: 9 Nov. 2023
Revised: 7 Apr. 2024
Accepted: 6 Jun. 2024

Keywords
Emotional prosody
Event-related potential (ERP)
Functional magnetic resonance imaging (fMRI)
Persian emotional speech database (Persian ESD)

Corresponding author
Nayereh Joodi, PhD Candidate of Linguistics, Allameh Tabataba'i University, Tehran, Iran

Email: Nayerehjoodi4@gmail.com

doi.org/10.30514/icss.26.1.91

Abstract

Introduction: Without prosody, effective verbal communication is unattainable. Prosody transmits both linguistic and emotional information. The present study aims to 1) determine brain regions active during the processing of emotional prosody and 2) discover the mechanism of understanding emotional prosody. This is to be achieved by reviewing a body of research performed on emotional prosody using neuroimaging techniques, functional magnetic resonance imaging (fMRI), and recording event-related brain signals (ERP).

Methods: This study was conducted from 2005 to 2021 using healthy individuals as subjects, and sentences as stimuli were systematically picked and analyzed. All these studies were available through PubMed and Google Scholar web search engines.

Results: A single hemisphere is not involved in processing prosodic information. Speech sounds are processed in three stages by the ear, brain stem, thalamus, and primary auditory cortex. The first step is to extract phonological characteristics from prosodic cues in the right auditory cortex. Second, recognition of emotions in speech, representing meaningful phonological sequences in the right hemisphere's posterior/posterior parts (STS). In the third step, this research evaluate and interpret expressed emotions cognitively. This process involves the simultaneous activity of two hemispheres of the frontal cortex.

Moreover, factors such as the difficulty of the test task, the quality of the stimuli, and how the test is designed affect the neural mechanism of the brain.

Conclusion: Studying emotional prosody requires access to standardized research tools for the language under study. In the second part, this study reports on the process of designing, producing, and validating the first comprehensive emotional speech database for Persian (Persian ESD) developed by Keshtari and colleagues (2015).

Citation: Joodi N, Keshtiari N. A review of emotional prosody research based on functional magnetic resonance imaging and event-related potential techniques and an introduction to the Persian emotional speech database (Persian ESD). Advances in Cognitive Sciences. 2024;26(1):91-108.

Extended Abstract

Introduction

Effective verbal communication is impossible without prosody, which transmits both linguistic and emotional information. This study aims to identify the brain regions involved in emotional prosody processing and uncover the mechanisms involved in understanding emotional prosody.

Methods

Using databases such as PubMed and Google Scholar,

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

we searched for keywords such as emotional prosody, speech processing, event-related potentials (ERPs), and functional magnetic resonance imaging (fMRI) published between 2005 and 2021 on emotional prosody processing in healthy adult subjects' brains. The search results for these keywords included more than 100 articles related to emotional prosody processing, an essential part of processing and understanding emotions. The subjects in all selected studies were healthy and right-handed regarding hearing, vision, and neuropsychology, and only studies using sentences as stimuli were included. The authors include only information regarding the present study questions in tables 1 and 2, and they avoid expressing technical details related to imaging and recording the brain signal, processing the images and signals, and avoiding issues related to statistical processing steps for the results of each study.

Results

Studying emotional prosody provided models for the stages of emotional tone processing. As one of these models that has been considered in subsequent research, the model of Schirmer and Kotz (11) was used to obtain more information about the brain areas involved in processing emotional speech sounds (2, 19). In the ear, brain stem, thalamus, and primary auditory cortex, speech sounds are processed and decoded in three stages. The first step is to extract phonological characteristics from prosodic cues in the right auditory cortex. In the second step, the right hemisphere's posterior/posterior parts (STS) represent meaningful sequences of phonological elements. Evaluation and cognitive interpretation of expressed emotions constitute the third step, which produces simultaneous activity in two hemispheres of the frontal cortex (figures 1 and 2 show these steps).

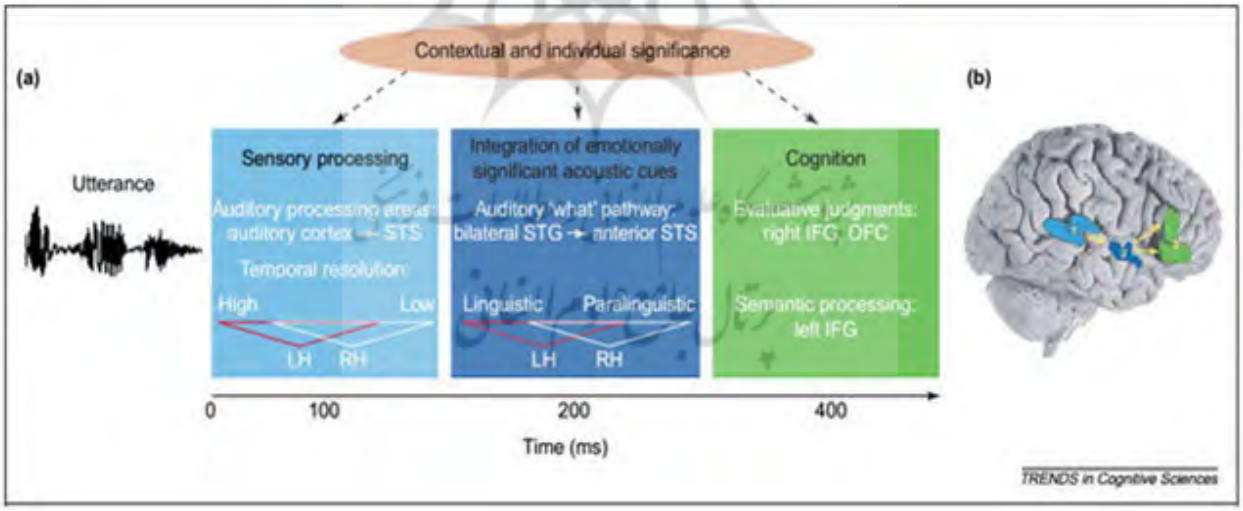


Figure 1. Model for the processing of emotional prosody (11, p. 25)

Notably, according to the socio-cultural characteristics of their native language, the speakers use these phonological characteristics (such as fundamental frequency, intensity or loudness of the voice, speed of speech, and voice quality (breathy, whispering, sharp, growling, and the like) to produce and use emotional and attitudinal understanding (5, 24, 25).

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

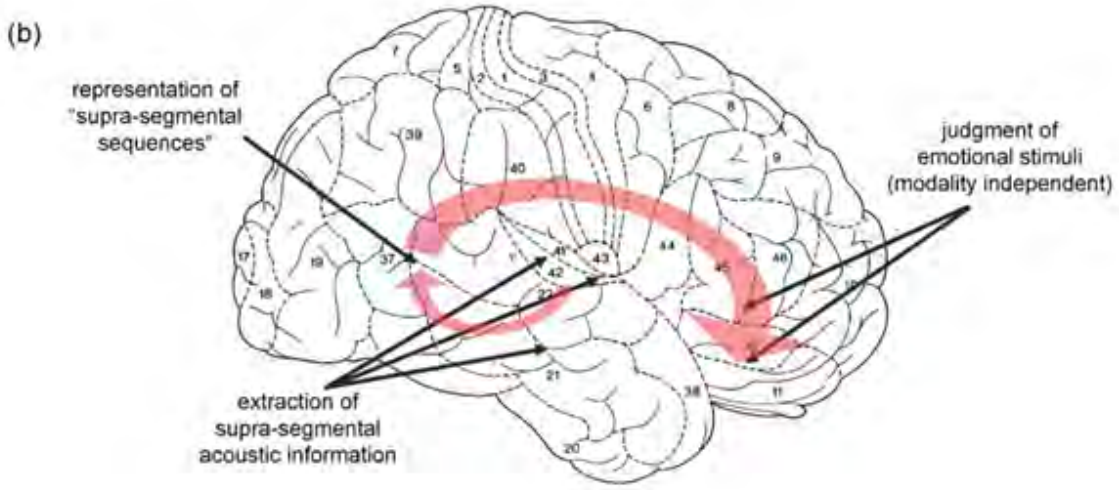


Figure 2. Three successive steps for the processing of emotional prosody (20, p.261)

Conclusion

Processing speech sounds is confined to just one hemisphere of the brain. It is also crucial to understand that various factors, such as the difficulty level of the test, the quality of the stimuli, and the design and execution of the test, influence neural mechanisms. Given the functional complexity of linguistic prosody, analyzing the neural structure of emotional prosody seems more straightforward. Clinical and neurological research has shown that these two types of prosody process differently in the brain. Therefore, speech prosody should be examined according to its multi-level division and application rather than viewed as a general concept (5).

Given the intricate and multifaceted nature of language, it appears essential to analyze the emotional prosody of speech independently, owing to its complexity and diverse characteristics. Although it provides detailed instructions on how to create verbal-emotional communication, attention must also be paid to how syntax, meaning, and pragmatics influence the tone of speech (both emotional and linguistic).

For linguistic studies, particularly neurological studies, linguistic data are examined outside their natural context due to technical limitations. In order to obtain accurate information about how language is processed and understood, it is essential to carefully design the appropriate task and select and formulate the appropriate stimuli. Therefore, a standard form of emotional prosody in any language is necessary to research emotional prosody (8). As manipulated stimuli are usually used in these experiments, it is also necessary to identify the effective phonological characteristics (fundamental frequency, duration, intensity, and the like) associated with emotional, attitudinal, and motivational prosody in each language using appropriate perceptual and behavioral research.

In order to support researchers conducting proper research on Persian language processing, the steps involved in creating and maintaining the “Persian emotional speech database” will be explained. These databases can help researchers design and conduct neurological experiments based on ERP or fMRI techniques to investigate how the emotional prosody of Persian speech is processed and perceived in the brains of Persian speakers. The database was previously designed and produced at Freie Universität Berlin in collaboration with a research group (26). However, this information was only available in English. This study aimed to build an authentic database of emo-

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

tional speech in Persian.

The database contains a set of 90 validated novel Persian sentences classified into five basic emotional categories (anger, disgust, fear, happiness, and sadness), as well as a neutral category. These sentences were validated in two experiments by a group of 1,126 native Persian speakers. The sentences were articulated by two native Persian speakers (one male, one female) in three conditions: 1) congruent (emotional lexical content articulated in a congruent emotional voice), 2) incongruent (neutral sentences articulated in an emotional voice), and 3) baseline (all emotional and neutral sentences articulated in neutral voice). The speech materials comprise about 470 sentences. The validity of the database was evaluated by a group of 34 native speakers in a perception test. Utterances recognized better than five times chance performance (71.4 %) were regarded as valid portrayals of the target emotions. Acoustic analysis of the valid emotional utterances revealed differences in pitch, intensity, and duration, attributes that may help listeners to correctly classify the intended emotion. The database is designed to be used as a reliable material source (for both text and speech) in future cross-cultural or cross-linguistic studies of emotional speech, and it is available for academic research purposes free of charge. This tool can benefit research in various fields, including neurology of language, psychology of language, clinical linguistics, speech therapy, and speech synthesis. To access the database, please contact the second author.

Ethical Considerations

Compliance with ethical guidelines

The present study has a review nature. Besides, there is no doubt that all the research introduced in this review had a valid code of ethics.

Authors' contributions

First author: drafting the article, revising it, and being accountable for all aspects of the research. Second author: designing, constructing, and validating the Persian emotional speech database; writing the second part of the article; revising and correcting the entire article.

Funding

The German Research Foundation (DFG) awarded a scholarship to the second author for designing the Persian language emotional speech database.

Acknowledgments

Thanks to Shahla Raghibdoust for her assistance in writing this paper's the first part (review).

Conflicts of interest

The authors have no conflict of interest.

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

مجله تازه‌های علوم شناختی
دوره ۲۶، شماره ۱، بهار ۱۴۰۳
صفحات ۱۰۸-۹۱
مقاله پژوهشی

مروری بر پژوهش‌های نوایِ گفتار عاطفی انجام شده با استفاده از تکنیک پتانسیل وابسته به رویداد و تصویربرداری تشدید مغناطیسی کارکردی و معرفی پایگاه دادهٔ گفتارِ عاطفی زبان فارسی (Persian ESD)

نیره جودی[1*]، نیلوفر کشتیاری[2]

۱. کاندید دکتری زبان‌شناسی، دانشگاه علامه طباطبائی، تهران، ایران
۲. دکتری تخصصی روان‌شناسی زبان، دانشگاه آزاد برلین، برلین، آلمان

دریافت: ۱۴۰۲/۰۸/۱۸
اصلاح نهایی: ۱۴۰۳/۰۱/۱۹
پذیرش: ۱۴۰۳/۰۳/۱۷

واژه‌های کلیدی
نوای گفتار عاطفی
پتانسیل وابسته به رویداد (ERP)
تصویربرداری تشدید مغناطیسی کارکردی (fMRI)
پایگاه دادهٔ گفتار عاطفی زبان فارسی (Persian ESD)

نویسنده مسئول
نیره جودی، کاندید دکتری زبان‌شناسی، دانشگاه علامه طباطبائی، تهران، ایران
ایمیل: Nayerehjoodi4@gmail.com

doi.org/10.30514/icss.26.1.91

چکیده

مقدمه: نوای گفتار ابزارِ انتقال اطلاعات زبانی و عاطفی است که بدون آنها ارتباط کلامیِ مؤثر امکان‌پذیر نیست. پژوهش حاضر می‌کوشد با مرور پژوهش‌های انجام شده در زمینهٔ نوای عاطفی با به کارگیری روش‌های تصویربرداری عصبی و ثبت سیگنال‌های مغزی وابسته به رویداد و با توجه به یافته‌های این پژوهش‌ها گزارش کند که هنگام پردازش نوای عاطفی کدام مناطق مغز فعال هستند و سازوکار درک نوای عاطفی چگونه است.

روش کار: برای دست‌یابی به این هدف، مطالعاتی که از سال ۲۰۰۵ تا ۲۰۲۱ در پایگاه‌های اطلاعاتی معتبر PubMed و Google Scholar منتشر شده بودند و بر پایهٔ دو روش نام برده و نیز با استفاده از جملات به عنوان محرک، آزمودنی‌های سالم را بررسی کرده بودند، انتخاب و به صورت نظام‌مند مرور شدند. در بخش دوم این نوشتار برای نخستین بار به زبان فارسی مراحل طراحی، ساخت و رواسازی «پایگاه دادهٔ گفتارِ عاطفی زبان فارسی (Persian ESD)» شرح داده شده است.

یافته‌ها: فعالیت مغز هنگام پردازش نوای گفتار به منطقه‌ای خاص در نیمکرهٔ راست مغز محدود نمی‌شود. بلکه در نواحی مشخصی فعالیت همزمان دو نیمکره مشاهده می‌گردد. پردازش نوای گفتار زبان‌شناختی و عاطفی سازوکارهای متفاوتی دارند.

نتیجه‌گیری: سازوکار عصبی مغز در هنگام پردازش نوای گفتار به عواملی مانند درجهٔ دشواری تکلیف آزمایش، کیفیت محرک‌ها و چگونگی طراحی و اجرای آزمایش بستگی دارد. دسترسی به ابزارهای پژوهشی رواسازی شده برای ارزیابی نوای گفتار از ملزومات مطالعهٔ عصب‌شناختی گفتار عاطفی است.

مقدمه

بررسی سازوکار پردازش نوای گفتار در مغز، اطلاعات جدیدی دربارهٔ رابطهٔ زبان و مغز به پژوهشگران می‌دهد. نوای گفتار دو کاربرد دارد: ۱) انتقال اطلاعات زبان‌شناختی و ۲) انتقال اطلاعات فرا زبانی (مانند عواطف). مطالعات فراوانی با استفاده از مغزنگاری الکتریکی الکتروانسفالوگرافی (Electroencephalography (EEG))، ثبت پتانسیل وابسته به رویداد (Event-related potential (ERP)) و تصویربرداری تشدید مغناطیسی کارکردی (functional magnetic resonance imaging (fMRI)) در پی مشخص کردن نقش نوای گفتار زبان‌شناختی/نوای گفتار (Linguistic prosody) و نوای گفتار عاطفی (Emotional prosody) در درک زبان و مناطق فعالِ مغز افراد سالم هنگام پردازش آنها بوده‌اند. پژوهش حاضر با مرورِ نظام‌مند این مطالعات و جمع‌بندی نتایج آنها می‌کوشد به دو پرسش زیر پاسخ دهد:

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

۱) فرایند پردازش نوای عاطفی در مغز چگونه است؟
۲) در پردازش نوای عاطفی کدام مناطق مغز فعالیت دارند؟

در بخش پایانیِ این مقاله به شرح مراحل طراحی، ساخت و رواسازی پایگاه دادهٔ گفتار عاطفی زبان فارسی (Persian ESD) پرداخته‌ایم. این ابزار که برای استفاده در مطالعات حوزه‌های گوناگون از جمله عصب‌شناسی زبان، روان‌شناسی زبان، زبان‌شناسی بالینی، گفتار درمانی و نیز سنتز گفتار تهیه شده است، این امکان را به پژوهشگران می‌دهد که تأثیر نوای عاطفی گفتار، متن گفتار (جملات) و نیز تعامل این دو را در درک گفتار عاطفی به طور جداگانه بررسی کنند.

پیشینه و چارچوب

پیش از پرداختن به روش گردآوری مطالعات مرتبط با موضوع این پژوهش و چگونگی دسته‌بندی آنها بیان مفاهیم اصلیِ مطرح در پژوهش‌های عصب‌شناختی در زمینهٔ نوای عاطفی ضروری به نظر می‌رسد که در ادامه به آن پرداخته می‌شود.

پتانسیل وابسته به رویداد (ERP) یا تصویربرداری تشدید مغناطیسی کارکردی (fMRI)

عواطف یا هیجان‌ها (Emotions) راهی برای انتقال اطلاعات هستند و نقشی حیاتی در بقایِ بشر و ارتباطات اجتماعی دارند. بنابراین پردازش مغزی آنها توجه پژوهشگران در حوزه‌های متفاوتی را به خود جلب کرده است (۱). کاربرد روش ثبت پتانسیل وابسته به رویداد (ERP) و تصویربرداری تشدید مغناطیسی کارکردی (fMRI) در عصب‌روان‌شناسی سبب شده است ارزیابی پاسخ‌های عصبی مغز به محرک‌های عاطفی با دقت هزارم‌ثانیه و شفافیت مکانی بالا امکان‌پذیر باشد. در دو دههٔ اخیر با استفاده از روش fMRI شبکه‌های فعال مغز در پردازش عواطف مشخص و معرفی شده‌اند. بدیهی است که با توجه به ماهیت چندبُعدی عواطف در انسان و بسته به اهدف و محرک‌های هر پژوهش، نتایج تصویربرداری عصبی بخشی از موضوع را روشن می‌کند (۲-۴). بسیاری از پژوهش‌های انجام شده با هدف بررسی فعالیت مغز به صورت عمیق‌تر و با دقت مکانی مناسب که ویژهٔ این روش است، فرضیهٔ پردازش انحصاریِ نوای عاطفی در نیمکرهٔ راست مغز را به چالش کشیده‌اند (۵).

با به کارگیری روش ERP اطلاعات ارزشمندی دربارهٔ زمان مورد نیاز برای پردازش عواطف در مغز به دست آمده است. برای مثال Begleiter و همکاران در آزمایشی بر پایهٔ روش ERP نشان‌دادند که پردازش عواطف متفاوت سبب بروز مؤلفه‌هایی با دامنه‌های مختلف می‌شوند (۶). به بیان دیگر محرک‌های شامل واژه‌های ناخوشایند نسبت به محرک‌هایی که واژه‌هایِ خنثی یا خوشایند داشتند سبب بروز مؤلفه‌هایی با دامنهٔ بیشتر می‌شوند. در سال‌های اخیر بررسی چگونگی پردازش اطلاعات عاطفی با استفاده از ERP به معرفی مؤلفه‌های مختلفی از این روش (LPC، LPP، P300، N300، N250، N170، P100، N100) منجر شده است (۷).

در پژوهش‌های بر پایهٔ ERP روش‌های متفاوتی به کار گرفته شده است بنابراین نتایج این پژوهش‌ها همواره با هم همسو نبوده‌اند. برای نمونه نتایج Liu و Pell (۲۰۱۲) (۸) و نتایج پژوهش Kotz و همکاران (۲۰۰۶) (۹) کاملاً متفاوت گزارش شده است. همچنین نتایج Kotz و همکاران (۹) با نتایج پژوهش Chen و Yang (۲۰۱۲) (۱۰) تفاوت دارد. استفاده از تکلیف‌ها و محرک‌های متفاوت و همچنین زبان‌های متفاوت را می‌توان از دلایل چنین ناهماهنگی‌هایی دانست. در مطالعهٔ Schirmer و Kotz (۲۰۰۶) (۱۱) محرک‌های مطالعه به زبان آلمانی بودند در صورتی که Chen و Yang (۲۰۱۲) (۱۰) محرک‌های زبان چینی را مطالعه کردند که برخلاف زبان آلمانی، یک زبانی نواختی (Tone language) است. این تفاوت‌های زبانی سبب تفاوت گسترهٔ فعالیت مغزی آزمودنی‌های این دو پژوهش در پردازش نوای عاطفی گفتار شده است. همان‌طور که Chen و Yang (۲۰۱۲) (۱۰) نشان داده‌اند گویشوران زبان چینی از مشخصه‌های صوت‌شناختی متفاوتی برای بازشناسی نوای عاطفی نسبت به انگلیسی زبان‌ها بهره می‌برند (۱۰). همچنین در توضیح چنین تفاوت‌هایی نقش خطاهای آماری را نیز نمی‌توان نادیده گرفت. با توجه به ماهیت روش ERP تعداد محرک‌ها و میزان دشواری تکلیف آزمایش، در کنارِ عوامل دیگری مانند وضعیت جسمی و ذهنی آزمودنی‌ها و مکان آزمایش نیز می‌تواند در بروز اشکالات آماری و به پیروی از آن اختلال در نتایج مؤثر باشد. بنابراین برای دست‌یابی به نتایج دقیق باید در تحلیل سیگنال‌های مغزی و استخراج مؤلفهٔ مرتبط با موضوع آزمایش تا آن جا که ممکن است به این عوامل توجه گردد (۱۲).

به طور کلی در هر دو روش توجه به انتخاب محرک‌های مناسب از مهمترین و حساس‌ترین مراحل کار است. بررسی محرک‌های این پژوهش‌ها نشان می‌دهد، با توجه به پرسش و هدف هر پژوهش واحدهای مختلف زبانی شامل آواها، هجاها، واژه‌ها، شبه واژه‌ها و جمله‌ها و شبه جملات به عنوان محرک به کار رفته‌اند. اما با توجه به ماهیت نوای گفتار عاطفی به نظر می‌رسد استفاده از «جمله» به عنوان محرک، به پژوهشگر کمک می‌کند بافت عاطفی را آسان‌تر بازسازی و بررسی کند. هرچند شایسته است به این نکته اشاره شود که استفاده از جمله مشکلات به همراه خواهد داشت. از جمله این که آنگاه تأثیر معنای عبارات و معنای گزاره‌ای و تلویحی جمله‌ها را باید در محرک‌ها

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

کنترل یا همسان کرد. در برخی از آزمایش‌های مرتبط با سنجش تأثیر جنبه‌های نوایی بر رفتار یا تصویربرداری عصبی محتوای واژگانی برای حل چنین چالش‌هایی عناصر زبانی از محرک‌ها حذف و صرفاً مشخصه‌های نوایی به شرکت‌کننده‌ها ارائه می‌شود. بنابراین یکسان نگه داشتن محرک‌ها از نظر معنایی، نحوی و طول جمله‌ها عوامل مهم دیگری هستند که در طراحی و اجرا و نتایج آزمایش‌های عصب‌شناسی زبان تاثیر به سزایی دارند. در بیشتر پژوهش‌های معرفی‌شده در جدول ۱ و ۲ پژوهشگران از پیکره‌های گفتاری استفاده کردند و پیش از انجام آزمایش اصلی، شدت و نوع نوای عاطفی محرک‌ها را در آزمونی رفتاری تعیین و آنها را رواسازی کرده‌اند (۱۳).

نوای گفتار عاطفی

در برقراری ارتباط کلامی مؤلفه‌هایی مانند: زیروبمی، بلندی، سرعت («سرعت» گفتار خود می‌تواند دو ویژگی را شامل شود: سرعت فراگویی (سرعت تولید آواهای زبانی (Speech rate) با حذف مکث‌ها برحسب هجا بر ثانیه یا ضرباهنگ گفتار (Tempo) تعداد واحدهای زبانی تولید شده در واحد زمان بدون حذف مکث برحسب واژه بر دقیقه)، و کیفیت صدا نقش دارند. استفادهٔ به‌جا از این مؤلفه‌ها به گوینده کمک می‌کند تا منظورش را به درستی به شنونده منتقل کند. این سازوکار نه تنها در درک نوای عاطفی (شادی، غم، خشم و...) بلکه در پردازش اطلاعات غیرعاطفیِ زبان (اطلاعات زبان‌شناختی) نیز کاربرد دارد. برای مثال در زبان انگلیسی آهنگ خیزان یا افتان یک جمله یا برجستگی بخشی از آن تعیین‌کنندهٔ ساخت نحوی، معنا و اطلاعات گفتمانی آن است. بنابراین عناصر زبرزنجیری (Suprasegmental) یا نوایی (Prosodic) دو نوع اطلاعات زبان‌شناختی و عاطفی را منتقل می‌کند (۱۴). جداکردن این اطلاعات و معرفی مؤلفه‌های صوت‌شناختی (Acoustic) هریک به هرچه دقیق‌تر شدن پژوهش‌های آتی دربارهٔ رابطهٔ زبان و مغز خواهد انجامید (۵).

برای پرداختن به مفهوم نوای عاطفی ابتدا باید عاطفه یا هیجان را از دیدگاه علمی تعریف کنیم. «عاطفه» را می‌توان پیوستاری از دو مفهوم ظرفیت (Valence) و برانگیختگی (Arousal) در نظر گرفت. در این تعریف منظور از «ظرفیت» میزانِ خوشایندی یا ناخوشایندی است و «انگیختگی» به حس درونی هر فرد از میزان فعال بودن و غیرفعال بودن وی بر می‌گردد. با توجه به این تعاریف و پژوهش‌های بینافرهنگی شش حس شادی، غم، ترس، خشم، چندش و شگفت‌زدگی به عنوان عواطف پایه (Basic emotions) شناخته می‌شوند. طبق نظر برخی از روان‌شناسان عواطف پایه بدون در نظر گرفتن تفاوت‌های فرهنگی و زبانی میان تمامی افراد بشر مشترک هستند (۱۵). همچنین گروه دیگری از عواطف نیز هستند که نقشی حیاتی در تعاملات اجتماعی دارند. این دسته شامل: غرور، گناه، حسادت، شرم‌ساری، خجالت هستند که پس از شش حس پایه و بیشتر در بافت اجتماعی رشد می‌کنند (۱).

روش‌های ثبت سیگنال و تصویربرداری عصبی به چگونگی انتخاب محرک حساس هستند. از همین رو، در آن دسته از پژوهش‌های عصب‌شناسیِ زبان که بر چگونگی پردازش و درک عواطف تمرکز دارند، دسته‌بندی دقیق عواطف ضروری است. مثلاً Haworth (۲۰۱۸) با به کارگیری روشی صوت‌شناختی ابتدا انواع نوای عاطفی گفتار و ویژگی‌های صوت‌شناختی آنها را دسته‌بندی کرده و سپس آزمایش ثبت سیگنال مغزی را با توجه به همین دسته‌بندی اجرا کرده است. نقش نوای گفتار در کنار دیگر ابزارهای انتقال پیام (مانند: نشانه‌های بافتی، زبان‌شناختی، و حالات چهره) در ایجاد ارتباط مؤثر و انتقال و درک عواطف اهمیت ویژه‌ای دارد. عواطف موجود در نوای گفتار را می‌توان به صورت زیر دسته‌بندی کرد (۱۳):

نوای نگرشی (Attitudinal prosody): شواهد نشان می‌دهد نگرش ما در نوای گفتار نمود پیدا می‌کند. برای مثال وقتی صادقانه مطلبی را بیان می‌کنیم سرعت کلام کم می‌شود، زیروبمی افزایش پیدا می‌کند و شدت کاهش می‌یابد (برای جزییات (۱۰) را ببینید).

نوای گفتار انگیزشی/انگیزه‌ای (Motivational prosody): با نوای گفتار عاطفی تفاوت چندانی ندارد و مطالعهٔ آن به پژوهش‌های سال‌های اخیر بازمی‌گردد (۱۸-۱۶). نوای گفتار انگیزشی شبیه نوای گفتار نگرشی، ابزاری است برای انتقال منظور گویشوران در گفتگو. با توجه به بافت می‌توان نوای انگیزشی را به دو دسته تقسیم کرد: ۱) نوای کنترل‌کننده و ۲) نوای استقلالی/حمایتی (Controlling and autonomy supportive) که هر کدام از این دو با توجه به رویکرد گوینده و شنونده در بیان و درک مفهومی خاص، تعریف می‌شود (۱۳).

در بخش بعدی چگونگی و معیارهای انتخاب پژوهش‌های عصب‌شناختی برای مرور نظام‌مند شرح داده شده است.

روش کار

روش اجرا

در این پژوهش با مراجعه به پایگاه‌های معتبر PubMed و Google Scholar و جستجوی کلید واژه‌هایِ نوای عاطفی، پردازش گفتار، پتانسیل وابسته به رویداد (ERP) و تصویربرداری تشدید مغناطیسی کارکردی (fMRI)، پژوهش‌های منتشر شده از سال ۲۰۰۵ تا ۲۰۲۱ دربارهٔ پردازش نوای گفتار در مغز آزمودنی‌های سالم و بزرگسال

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

بررسی شد. با توجه به اهمیت پردازش نوای عاطفی به عنوان بخشی از پردازش و درک عواطف نتایج جستجوی این کلید واژه‌ها بسیار گسترده و بیش از ۱۰۰ مقالهٔ مرتبط بود. با مطالعهٔ چکیدهٔ آنها مواردی که معیارهای مورد نظر را داشتند انتخاب شد. پژوهش‌های انجام شده دربارهٔ رابطهٔ نوای گفتار با استفاده از رویکردهای رفتاری، مواردی که روش‌های مورد نظر (ERP و fMRI) در آنها به کار نرفته بود و پژوهش‌هایی که روی مغز بیماران، نوزادان و کودکان انجام شده بودند با توجه به محدودیت‌های این پژوهش از تحلیل کنار گذاشته شده‌اند. در همهٔ پژوهش‌های انتخاب شده آزمودنی‌ها از نظر شنوایی، بینایی و عصبی_روانی، سالم و راست‌دست بوده‌اند، همچنین فقط پژوهش‌هایی بررسی شدند که محرک آنها جمله است. در دو جدول ۱ و ۲ تنها اطلاعات مرتبط با پرسش‌های پژوهش حاضر گنجانده شده است و نگارنده از بیان جزئیات فنی مرتبط با چگونگی تصویربرداری و ثبت سیگنال مغزی و پردازش تصاویر و سیگنال‌های و مسائل مربوط به مراحل پردازش آماری نتایج هر پژوهش اجتناب کرده است.

جدول ۱. پژوهش‌های انجام شده بر پایهٔ روش fMRI

	آزمودنی	محرک	تکلیف	هدف پژوهش	منطقهٔ فعال، نتایج	پژوهشگران
۱	۱۲ آلمانی‌زبان ۸ زن، ۴ مرد	۱۰۸ جمله خبری آلمانی متن جملات: از نظر معنایی خنثی نوای گفتار: مثبت، منفی و خنثی	تشخیص نوای عاطفی مثبت (شادی) منفی (خشم) و خنثی با دادن امتیاز از ۱ تا ۵	- بررسی سو برتری نیمکرهٔ راست در پردازش نوای عاطفی گفتار - مشخص شدن نقش عقده‌ها (هسته‌های) قاعده‌ایِ (basal ganglia) مغز در درک نوای عاطفی	۱) فعالیت دو نیمکره در مناطق مرکزی زبانی اطراف و برخی مناطق زیر (perisylvian) شیار سیلویان قشری و بخش پیشانی مغز در پردازش گفتار فعالیت در دو نیمکره به صورت دو طرفه ۲) فعالیت در دو نیمکره در منطقه درپوش پیشانی در هنگام پردازش نوای گفتار مثبت (opercular) ۳) فعالیت دو نیمکره در منطقهٔ پیشانی_گیجگاهی و زیرقشری (bilateral fronto-temporal) در پردازش نوای گفتار منفی (subcortical)	Kotz و همکاران (۲۰۰۳) (۴۳)
۲	۱۰ آلمانی‌زبان ۵ زن، ۵ مرد	۱۰۰ جملهٔ خبری آلمانی معنادار متن جملات: خنثی نوای گفتار: شادی، خشم، ترس، نفرت و غم	تشخیص نوای عاطفی با نام‌گذاری عواطف تشخیص و بیان واکه‌ها	- بررسی مناطق فعال در پردازش و شناسایی نوای عاطفی و واکه‌ها (آواها)	۱) درک نوای عاطفی وابسته به الگوی خاصی از (PSTS)/ فعالیت نیمکرهٔ راست شامل فعالیت در (posterior superior) منطقهٔ برادمن ۲۲ و منطقهٔ پشتی_جانبی temporal sulcus برادمن ۴۵/۴۴ و اطراف عقده (dorsolateral) مناطق پیشانی (۴۷B47) قاعده‌ای ۲) هنگام انجام تکلیف تشخیص آوایی فعالیت در قسمت‌های چپ مناطق گفتاری مشاهده شد	(Wildgruber) و همکاران (۲۰۰۵) (۱۹)
۳	آزمایش با دو گروه آلمانی زبان: ۸ نفر ۴ مرد و ۴) (زن ۱۰ نفر ۵ مرد و ۵) (زن	جملهٔ خبری آلمانی معنادار متن جملات: خنثی نوای گفتار: شادی، خشم، ترس	تشخیص نوای عاطفی و نام‌گذاری آنها تکلیف تشخیص و بیان واکه‌ها	شناسایی شبکه مغزی مسئول پردازش نوای عاطفی شناسایی سه مرحلهٔ پردازش و درک نوای گفتار	۱) فعالیت همزمان دو نیمکره در هنگام ارزیابی و پردازش نوای عاطفی ۲) فعالیت مناطق زیر قشری در تشخیص و پردازش در (amygdala) عواطف خاص مانند فعالیت بادامه پاسخ به محرک احساس ترس ۳) پردازش عناصر زبرزنجیری (مشخصه‌های صوت شناختی)، نوای گفتار زبان‌شناختی و عاطفی فعالیت مشابهی در نیمکرهٔ راست ۴) فعالیت مناطق زبانی در نیمکرهٔ چپ در هنگام ارزیابی آشکار و پردازش جنبه‌های زبان‌شناختی نوای گفتار	(Wildgruber) و همکاران (۲۰۰۶) (۲۰)

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

	آزمودنی	محرک	تکلیف	هدف پژوهش	منطقهٔ فعال، نتایج	پژوهشگران
۴	۲۲ آلمانی‌زبان (۱۲ و ۱۲ مرد)	۳۰ جملهٔ معنادار خبری آلمانی متن جملات: خنثی نوای گفتار: شاد، خشمگین و خنثی	تشخیص نوای عاطفی در جمله‌های دستکاری شده	بررسی تسلط نیمکرهٔ راست در پردازش نوای عاطفی با توجه به نقش ضمنی چگونگی تکلیف	۱) فعالیت هر دو نیمکره در مناطق پیشانی (fronto- striato- temporal) –جسم مخطط – گیجگاهی بدون مشاهدهٔ تسلط آشکار نیمکرهٔ راست در پردازش نوای عاطفی گفتار ۲) چگونگی تکلیف آزمایش تأثیری در سوبرتری ندارد	Kotz و همکاران (۲۰۰۶) (۴۴)
۵	۲۴ ژاپنی زبان: (۱۲ مرد و ۱۲ زن)	۸۰ عبارت (۴۰ عبارت با بار معنای مثبت و ۴۰ عبارت با بار معنایی منفی) زبان ژاپنی (لهجهٔ توکیو) متن عبارت: منفی و مثبت نوای گفتار: مثبت و منفی و نوای عاطفی محبت‌آمیز یا سرد و سنگ‌دلانه	تشخیص معنای زبان‌شناختی و رویکرد عاطفی گوینده	بررسی چگونگی فرایند درک ذهن گوینده در فعالیت کلامی (قصد گوینده)	۱) فعالیت عصبی مسئول درک قصد و منظور: (superior temporal sulci ((STS گوینده شیار گیجگاهی فوقانی (inferior frontal regions) بخش‌های تحتانی پیشانی مناطق میانی پیشانی و مناطق پشتی مخچه	Imaizumi و همکاران (۲۰۰۶) (۴)
۶	۶ چینی زبان (۲ مرد و ۴ زن)	۱۶ جملهٔ معنادار پرسشی برگرفته از اخبار متن جملات: خنثی نوای گفتار: خنثی	گوش دادن و پاسخ دادن به پرسش‌ها (با پاسخ درست/نادرست)	- بررسی سازوکار پردازش نوای عاطفی انسان در طراحی عواطف هوش مصنوعی - نقش مشخصه‌های صوت‌شناختی در درک نوای عاطفی گفتار (رابطهٔ انسان و هوش مصنوعی)	۱) فعالیت مناطق گیجگاهی نیمکرهٔ راست و مناطق پیشانی در پردازش نوای عاطفی مستقل از دسته‌بندی احساس خاص ۲) کمک به پردازش عواطف توسط ربات‌ها (هوش عاطفی) و قابلیت تحلیل عواطف در مواقع ضروری	Wang و همکاران (۲۰۱۴) (۲)
۷	۲۰ انگلیسی زبان (۱۲ زن و ۸ مرد) انگلیسی آمریکایی	۵ شرایط مختلف آزمایش جمله‌های خبری معنادار متن جملات: خنثی نوای گفتار: شاد، غمگین و خشمگین	گوش دادن به محرک‌های شنیداری و تشخیص نوای گفتار با توجه به محرک‌های دیداری	- مقایسهٔ فعالیت مغز در پردازش و درک جمله‌ها با پردازش و درک نوای عاطفی	۱) فعالیت مناطق پیشانی_گیجگاهی نیمکرهٔ راست در پردازش نوای عاطفی همراه با فعالیت (بی ای ۴۷)، بادامه و اینسولا پیشینی/قدامی	Seydell-Greenwal و همکاران (۲۰۲۰) (۳)
۸	۱۶ زن (به زبان آزمودنی‌ها اشاره نشده است)	۳۵ ارائهٔ ۱ دقیقه‌ای انگلیسی معنادار متن جملات: روایت بازنمایی‌کنندهٔ خشم، ترس، چندش، شادی، غم، شگفت‌زدگی: شگفت‌زدگی) و وضعیت خنثی: نوای گفتار خنثی	گوش‌دادن به روایتی همزمان با شنیدن صحبت‌های یک دوست	- بررسی سازوکار پردازش نوای عاطفی انسان در طراحی عواطف هوش مصنوعی - نقش مشخصه‌های صوت‌شناختی در درک نوای عاطفی گفتار (رابطهٔ انسان و هوش مصنوعی)	۱) فعالیت مناطق گیجگاهی نیمکرهٔ راست و مناطق پیشانی در پردازش نوای عاطفی مستقل از دسته‌بندی احساس خاص ۲) کمک به پردازش عواطف توسط ربات‌ها (هوش عاطفی) و قابلیت تحلیل عواطف در مواقع ضروری	Wang و همکاران (۲۰۱۴) (۲)

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

	آزمودنی	محرک	تکلیف	هدف پژوهش	منطقهٔ فعال، نتایج	پژوهشگران
۹	۱۵ فرانسوی‌زبان (۸ زن ۷ مرد)	۲ شبه جمله متن جملات: بی‌معنا نوای گفتار: خنثی، شاد و خشمگین	گوش دادن به صدای زن و مرد بازیگر و فشار دادن دکمه درصورتی شنیدن دو جملهٔ یکسان پشت سرهم	- بررسی نقش مناطق زیرقشری مانند عقده‌های قاعده‌ای و مخچه - درگیری کارکردی این مناطق در پردازش نوای عاطفی گفتار	۱) فعالیت به صورت ارتباطات مستقیم بین (putamen) عقده‌های قاعده‌ای به ویژه پوتامن با (globus pallidus) و گولیبوس پالیدوس هسته‌های زیرتالاموسی در مخچه ۲) تأکید بر نقش حیاتی این مناطق در رمزگشایی و پاسخ عواطف کلامی (نوای عاطفی)	Ceravolo و همکاران (۲۰۲۱) (۴۶)

جدول ۲. پژوهش‌های انجام شده بر پایهٔ تکنیک ERP

	آزمودنی	محرک	تکلیف	هدف پژوهش	منطقهٔ فعال، نتایج	پژوهشگران
۱	۳۴ آلمانی زبان (۱۸ زن و ۱۶ مرد)	زبان آلمانی متن و نوای گفتار (شادی): ۳۰ جمله متن و نوای گفتار (خشم): ۲۷ جمله متن و نوای گفتار (خنثی): ۳۰ جمله	تشخیص درست بودن یا نبودن نوای گفتار و دسته‌بندی عواطف	- سازوکار پردازش نوای عاطفی/پردازش معنایی - تشخیص عواطف با توجه به نوای گفتار	۱) تخطی از نوای گفتار سبب بروز مؤلفه‌ای مثبت ۲) تخطی معنایی_نوایی سبب بروز مؤلفه‌ای منفی ۳) دسته‌بندی عواطف	Kotz و Paulmann (۲۰۰۷) (۳۴) و Kotz و همکاران (۲۰۰۶) (۴۴)
۲	۱۵ انگلیسی استرالیایی زبان (۹ مرد و ۶ زن)	۱۵۰ جملهٔ معنادار خبری انگلیسی متن جملات: با بار معنایی خنثی و خشمگین و کنایه‌آمیز نوای گفتار: ایجاد تخطی نوایی با روش بریدن و چسباندن ابتدای یک جمله خنثی از نظر نوایی به جملاتی با نوای خنثی و خشمگین و کنایه‌آمیز	بررسی چگونگی تأثیر نوع تکلیف در پردازش دو آزمایش قضاوت دربارهٔ درست بودن یا نبودن نوای گفتار و دسته‌بندی عواطف	- بررسی چگونگی تأثیر نوع تکلیف در پردازش - بررسی سازوکار پردازش برخط نوای گفتار غیر کلامی و مؤلفه‌های عاطفی گفتار (تحلیل خشم و نوای کنایه‌آمیز)	۱) سازوکار پردازش خشم و کنایه از نظر زمانی و گستردگی فعالیت مغزی مشابه است ۲) چگونگی تکلیف آزمایش در مراحل ابتدایی تفاوت نشان داد ۳) همگام با پیشینهٔ پژوهش مؤلفه‌ای مثبت در بخش‌های پیشانی و زودهنگام ثبت شد ۴) چگونگی پردازش مشخصه‌های صوت‌شناختی نوای عاطفی بررسی شد	Wickens و Perry (۲۰۱۵) (۱۲)
۳	۲۰ چینی زبان (۱۰ زن و ۱۰ مرد)	۳۶۰ جملهٔ معنادار انتخاب شده از پایگاه دادگان (emotional speech database of Taiyuan University of Technology) زبان چینی ۹۰ نمونه برای هریک از ۴ دستهٔ: غم، شادی، خشم، شگفت‌زدگی جملات معنادار و متن: بار معنایی شادی، خشم، غم، شگفت‌زدگی نوای گفتار: نمایش رادیویی با نوای گفتار شادی، خشم، غم و شگفت‌زدگی	فقط گوش دادن به محرک‌ها	- بررسی نقش دیرش محرک‌های شنیداری بر مؤلفه‌های ERP پژوهش‌های نوای عاطفی	۱) محرک عاطفی کوتاه‌تر در تحلیل مؤلفه‌های مربوط به عواطف (مانند N100; P200; N300) مؤثرتر است	Chang و همکاران (۲۰۱۸) (۲۲)

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

	آزمودنی	محرک	تکلیف	هدف پژوهش	منطقهٔ فعال، نتایج	پژوهشگران
۴	۳۸ انگلیسی‌زبان (۱۷ زن و ۲۱ مرد)	۱۲۸ جملهٔ پرسشی معنادار متن جملات: خنثی نوای گفتار: (۱۶ جمله با نوای خشمگین ۱۶ جملهٔ کنترل، ۳۲ جملهٔ خودمختاری، ۳۲ جملهٔ لذت و ۳۲ خنثی)	فقط گوش دادن به محرک‌ها	- بررسی شباهت و تفاوت پردازش نوای انگیزشی و عاطفی از نظر زمان	۱) پردازش نوای عاطفی و انگیزشی گفتار از نظر سازوکار مغزی به صورت مشابهی پردازش می‌شوند ۲) اما با توجه به تفاوت در چگونگی استفاده از سرنخ‌های صوتی از هم متمایز می‌شوند	Haworth و همکاران (۲۰۱۸) (۱۳)
۵	۲ انگلیسی زبان (تعداد زنان و مردان گزارش نشده است)	۱۹۸ جملهٔ کوتاه انگلیسی ۹۹ جمله با ظرفیت (متن و نوای گفتار) مثبت (شادی)، ۹۹ جملهٔ منفی (خشم)، ۲۵ جملهٔ خنثی به عنوان پرکننده	گوش دادن و تشخیص اسم‌های اول جمله	- بررسی فعالیت بیوالکتریکی مغز هنگام پردازش ظرفیت عواطف	۱) سه مؤلفهٔ P300 و N400 و LP (late Positivity) در قسمت‌های پیشینی مغز پردازش ظرفیت عاطفی را بازنمایی می‌کند. ۲) LP و P450 در پردازش ظرفیت مثبت و N400 در پردازش ظرفیت منفی مشاهده شد. ۳) P300 در پردازش مثبت و منفی در ((SFG، MFG) superior and middle frontal gyri) بازنمایی‌کنندهٔ پردازش معنایی و حافظهٔ فعال	Proverbio و همکاران (۲۰۲۰) (۴۵)

یافته‌ها

بخش اول

در ادامه با استفاده از نتایج و شواهد ارائه شده در جدول ۱ و ۲، به دو پرسش پژوهشِ حاضر پاسخ داده می‌شود.

مناطق فعال مغز در پردازش نوای عاطفی گفتار

بررسی نتایج مطالعات معرفی شده در جدول ۱ و ۲ نشان می‌دهد که فعالیت مغز هنگام پردازش نوای گفتار به منطقه‌ای خاص در نیمکرهٔ راست مغز محدود نمی‌شود (۹). بلکه در نواحی مشخصی (عقدهٔ عصبی پایه‌ای مغز، بادامه، مناطق پیشانی (۴-۲) فعالیت همزمان دو نیمکره مشاهده می‌گردد. البته با توجه به هدف و روش هر پژوهش این یافته‌ها متفاوت هستند (۱۹، ۲۰). پردازش نوای گفتار زبان‌شناختی و عاطفی سازوکارهای متفاوتی دارند. پردازش نوای گفتار زبان‌شناختی سبب ظهور فعالیت نیمکرهٔ چپ می‌گردد و پردازش نوای گفتار عاطفی عمدتاً در نیمکرهٔ راست و همراه با فعالیت در مناطقی دیگری از مغز مشاهده می‌شود (۱۴). همچنین نتایج این پژوهش‌ها نشان می‌دهند که اگرچه سازوکار پردازش عواطف در مغز یکسان است (۱۳)، اما هر کدام از عواطف سبب ظهور الگوی فعالیت مغزیِ خاصی می‌شوند (۱۲، ۲۱). تشخیص نوای شادی، خشم و غم (عواطف پایه) آسان‌تر از تشخیص نوای ترس و چندش است (۱۹). به طور کلی نتایج این پژوهش‌ها بر نقش محرک و حساسیت نتایج به چگونگی آنها تأکید می‌کنند (۲۲).

چگونگی و مراحل پردازش و درک نوای عاطفی

پژوهش‌های انجام شده دربارهٔ پردازش نوای عاطفی، الگوهایی ارائه کردند که مراحل پردازش نوای عاطفی را صورت‌بندی می‌کند. مدل Schirmer و Kotz (۲۰۰۶) (۱۱) یکی از این الگوهاست که در پژوهش‌های پس از آن نیز مورد توجه قرار گرفته است و زمینه‌ساز دستیابی به جزئیات بیشتری دربارهٔ مناطق فعال مغز هنگام پردازش نوای گفتار عاطفی شده است (۲، ۱۹). در این مدل برای پردازش نوای گفتار و رمزگشایی از جریان اطلاعات شنیداری در گوش، ساقهٔ مغز، تالاموس و قشر شنوایی اولیه سه مرحله در نظر گرفته می‌شود:

اول: استخراج مشخصه‌های صوت‌شناختی از سرنخ‌های نوایی در نواحی راست قشر شنوایی.

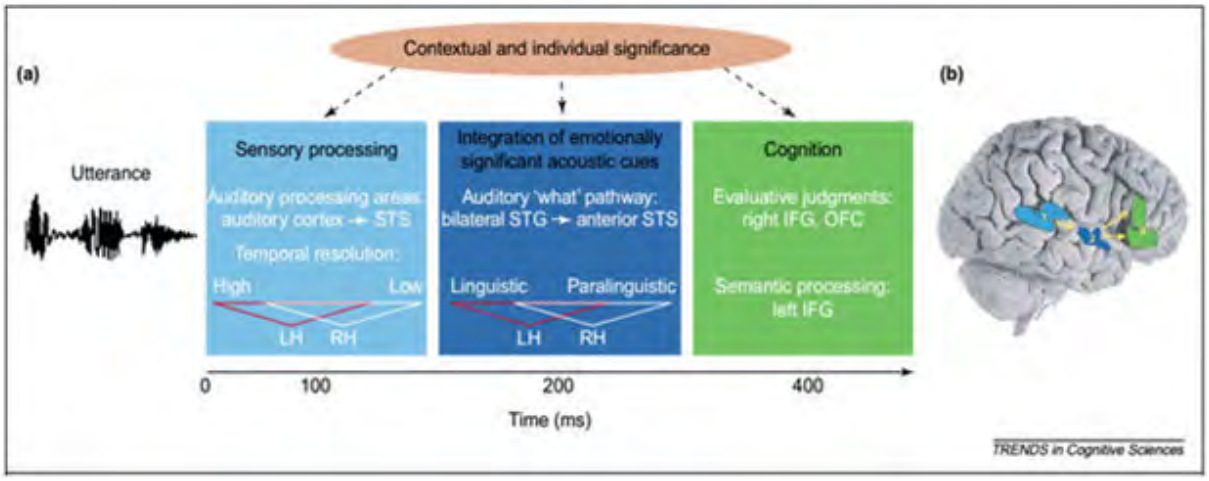
دوم: شناسایی عواطف در گفتار، بازنمایی توالی‌های معنادار عناصر صوت‌شناختی زبرزنجیری در بخش‌های خلفی ((STS) Superior temporal sulcus) نیمکرهٔ راست.

سوم: ارزیابی آشکار و تفسیرِ شناختیِ عواطف بیان شده، ارزیابی آشکار نوای عاطفی و فعالیت همزمان دو نیمکره در مناطق تحتانی قشرپیشانی.

شکل ۱ و ۲ این مراحل را نشان می‌دهند.

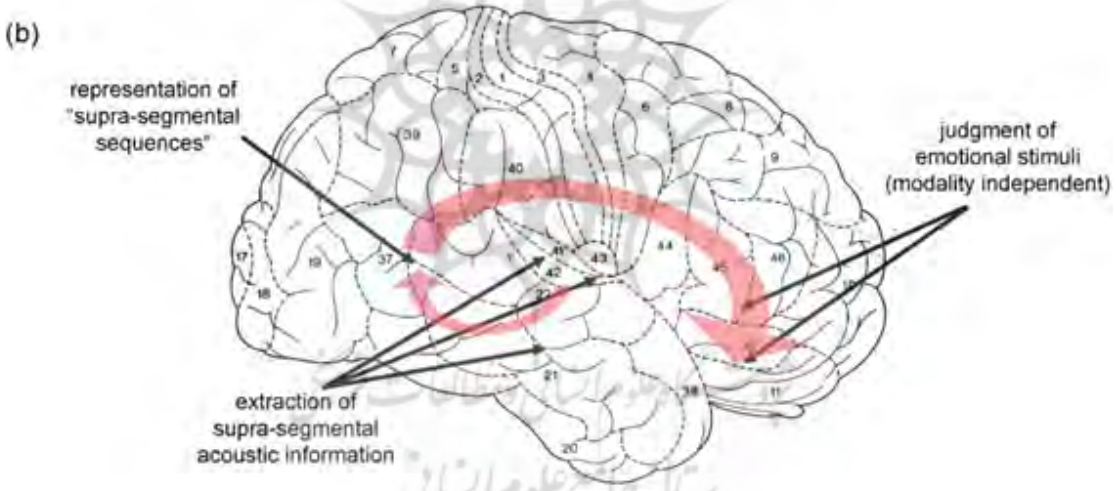
[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License



شکل ۱. انگارهٔ پردازش نوای عاطفی (۱۱، صفحه ۲۵)

شایسته توجه این که گویشوران با توجه به خصوصیات فرهنگی_اجتماعی زبان ویژهٔ خود از این مشخصه‌های صوت‌شناختی (مانند بسامد پایه، شدت یا بلندی صدا، سرعت گفتار و کیفیت صدا (نفسی، نجوایی، با تندی، غر زدن و...)) برای تولید و درک نوای عاطفی و نگرشی استفاده می‌کنند (۵، ۲۴، ۲۵).



شکل ۲. گام‌های پردازش نوای عاطفی (۲۰، صفحه ۲۶۱)

بنابراین، مشاهده می‌شود که تنها یک نیمکرهٔ مغز در پردازش نوای گفتار نقش ندارد. همچنین سازوکار عصبی مغز به عواملی مانند درجهٔ دشواری تکلیف آزمایش، کیفیت محرک‌ها و چگونگی طراحی و اجرای آزمایش بستگی دارد. با توجه به پیچیدگی کاربردی نوای گفتار زبان‌شناختی به نظر می‌رسد دست‌یابی به تصویری از ساختار عصبی نوای گفتار عاطفی آسان‌تر است. اگرچه نتایج پژوهش‌های بالینی و عصب‌شناختی نشان می‌دهند که هر کدام از این دو در بخش‌های متفاوتی از مغز پردازش می‌شوند. بنابراین نوای گفتار نه به عنوان مفهومی کلی بلکه باید با توجه به تقسیم‌بندی چند سطحی و کاربرد آن بررسی شود (۵).

در ادامه مراحل ساخت و رواسازی «پایگاه‌دادهٔ گفتار عاطفی زبان فارسی» را توضیح می‌دهیم (۲۶). پژوهشگران علاقه‌مند به بررسی چگونگی پردازش و ادراک نوای عاطفی گفتار فارسی در مغز گویشوران این زبان می‌توانند از این پایگاه‌داده برای طراحی و انجام آزمایش‌های عصب‌شناختی بر پایهٔ تکنیک‌های ERP یا fMRI بهره‌مند شوند.

معرفی پایگاه دادهٔ گفتارِ عاطفی زبان فارسی (Persian ESD)

در این بخش از مقاله به شرح مراحل طراحی، ساخت و رواسازی «پایگاه دادهٔ گفتارِ عاطفی زبان فارسی (Persian ESD)» می‌پردازیم. این پایگاه داده پیش‌تر در دانشگاه Freie Universitat Berlin با

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

همکاری گروهی پژوهشگر (۲۶) طراحی و تولید شده است. اما اطلاعات تا پیش از این فقط به زبان انگلیسی در دسترس بوده است.

ساخت متن جملات

پژوهشگرانی که تا به امروز گفتار عاطفی را مطالعه و بررسی کرده‌اند عموماً بر نقش نوای گفتار تمرکز کرده‌اند و نقش متن جملات را نادیده گرفته‌اند (۲۷). آنها اغلب تعدادی جمله ویژهٔ پژوهش خود تهیه کرده‌اند و بدون رواسازی متن جملات از آنها در مطالعهٔ گفتار عاطفی استفاده کرده‌اند (۲۸، ۲۹). اما برای بررسی تأثیر «متن جملات» و «نوای عاطفی» در پردازش گفتار عاطفی به مجموعه‌ای از جمله‌های معتبر نیاز داریم (۲۷). در این پژوهش با هدف ساخت یک مجموعه جملهٔ رواسازی شده برای زبان فارسی محاوره‌ای سه مطالعهٔ رفتاری جداگانه انجام شد. شرکت‌کننده‌هایی به صورت داوطلبانه و برخط در هر سه مطالعه شرکت کردند. همهٔ شرکت‌کننده‌ها در هر سه آزمایش از نظر عصبی روانی سالم بودند و غالب بودن زبان فارسی در آنها با استفاده از پرسش‌هایی تأیید می‌شد.

به عنوان اولین گام ۲۵۲ جملهٔ خبری ساخته شد. هر یک از این جملات بیانگر فقط یک حس خاص و یا حالت خنثی بودند (۳۶ جمله برای هر یک از حس‌های مورد نظر به اضافهٔ ۳۶ جملهٔ خنثی). تمام این جملات بر اساس یک ساختار سادهٔ دستوری (فاعل+مفعول+گروه حرف اضافه‌ای+فعل) ساخته شدند. اسامی خاص زنان و مردان به عنوان فاعل جمله و به تعداد برابر در جمله‌ها استفاده شد تا از اثر جنسیت اجتناب شود. تمام این ۲۵۲ جمله توسط چهار گویشور فارسی زبان (دو روان‌شناس، دو زبان‌شناس) از نظر طبیعی و روان بودن بررسی شدند.

رواسازی متن جملات

برای اطمینان از این که هر جمله فقط یک حس خاص را القا می‌کند و یا هیچ حسی را برنمی‌انگیزد (حالت خنثی) دو مطالعهٔ رفتاری جداگانه برای رواسازی این ۲۵۲ جمله انجام شد. شیوهٔ انتخاب گزینهٔ درست در مطالعات تشخیص احساس بر اعتبار نتایج بسیار اثرگذار است. اگر تعداد محدودی گزینه (هرکدام با نام یک حس) در اختیار آزمودنی‌ها قرار بگیرد درصد تشخیص هر حس افزایش پیدا می‌کند ولی در عین حال میزان اثرات تصنعی (Artifacts) هم زیاد می‌شود (۳۰). از سوی دیگر اگر آزمودنی‌ها گزینه‌های زیادی برای علامت زدن در اختیار داشته باشند، میزان تشخیص کاهش می‌یابد (۳۰، ۳۱). برای حل این مشکل می‌توان گزینهٔ «هیچ کدام از موارد بالا» را نیز به گزینه‌ها اضافه کرد (۳۲). بنابراین در تمام بخش‌های این پژوهش نام حس‌های مورد نظر به همراه گزینهٔ «هیچ یک از موارد بالا» برای تشخیص در اختیار آزمودنی‌ها قرار گرفت.

مطالعهٔ اول

در این آزمایش ۱۱۲۶ شرکت‌کننده شرکت کردند. در یک پرسشنامهٔ برخط ۲۵۲ جمله به شرکت‌کننده‌ها ارائه شد و از آنها خواسته شد تا هر جمله را بخوانند و موقعیت شرح داده شده در هر جمله (یا سناریو) را تصور کنند و حسی که همان جمله برمی‌انگیزد را علامت بزنند. در این مطالعه هشت گزینهٔ (خشم، چندش، ترس، غم، شادی، شگفت‌زدگی، خنثی و «هیچ یک از موارد بالا») برای تشخیص هر جمله به آزمودنی‌ها ارائه شد (سطح شانس ۱۲/۵ درصد). برای جلوگیری از اثر توالی، جمله‌ها در سه بخش به صورت کاملاً تصادفی به نمایش در می‌آمدند. بر اساس معیارهای آماری معرفی شده در پژوهش‌های پیشین (۳۳) پنج برابر سطح شانس (یعنی با در نظر گرفتن هشت گزینه عدد ۶۲/۵ درصد) به عنوان میزان تشخیص گفتار عاطفی در نظر گرفته شد و ۱۰۲ جمله (خشم: ۱۸ جمله؛ چندش: ۲۳؛ ترس: ۱۷؛ شادی: ۲۱؛ شگفت‌زدگی: ۰ و غم: ۲۳) و ۲۱ جملهٔ خنثی معیارهای لازم را کسب کردند. درصد تشخیص حس شگفت‌زدگی بسیار پایین بود (بین ۱۶/۹ درصد و ۵۰/۵ درصد، میانگین ۳۹/۴ درصد) و هیچ‌کدام از جملات مربوط به این حس به حد نصاب نرسیدند. مطالعات پیشین هم حاکی از تشخیص اشتباه حس شگفت‌زدگی و دسته‌بندی آن به عنوان شادی است (۳۳، ۳۴، ۴۳، ۴۴) که می‌تواند ناشی از هم‌پوشانی حس شادی و شگفت‌زدگی باشد. تحلیل نظرات شرکت‌کننده‌ها نشان‌دهندهٔ مواردی از سوگیری به دلیل استفاده از اسامی خاص به عنوان فاعل جملات بود. بنابراین برای حذف اثر سوگیری مطالعهٔ دیگری ترتیب داده شد.

مطالعهٔ دوم

برای حذف این اثر اسامی خاص با مشاغل (مانند «خانم معلم» و «آقای کشاورز») جایگزین شدند. استفاده از مشاغل به عنوان فاعل جمله علاوه بر حذف اثر سوگیری اطلاعات بافتی بیشتری به خواننده/شنونده می‌دهد که می‌تواند در تجسم سناریوی هر جمله مفید باشد. برای اجتناب از اثر جنسیت نیز به تعداد یکسان از عنوان آقا و خانم به عنوان فاعل جمله استفاده شد. سپس مطالعهٔ دوم با در نظر گرفتن موارد بیان‌شده (یعنی حذف حس شگفت‌زدگی از فهرست حس‌های مورد بررسی و جایگزینی اسامی خاص با عنوان مشاغل) انجام شد.

در مجموع ۷۱۶ آزمودنی در این مطالعه شرکت کردند. تعداد ۱۲۳ جمله در یک پرسشنامهٔ برخط با مشابه شیوهٔ اجرای مطالعهٔ قبلی به آزمودنی‌ها ارائه شد. البته با توجه حذف گزینهٔ شگفت‌زدگی این بار هفت گزینه (خشم، چندش، ترس، غم، شادی، خنثی و هیچ یک از موارد بالا) برای تشخیص هر جمله به صورت تصادفی به آزمودنی‌ها

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

ارائه شد (سطح شانس ۳/۱۴ درصد) در این مطالعه حد نصاب پنج برابر سطح شانس (یعنی ۵/۷۱ درصد) به عنوان معیار در نظر گرفته شد. بر این اساس ۹۰ جملهٔ رواسازی شده به دست آمد که یکی از حس‌های مورد نظر (یعنی خشم ۱۷ جمله، چندش ۱۵، ترس ۱۵، غم ۱۴ و شادی ۱۵ و خنثی ۱۴ جمله) را القا می‌کنند. «شدت» حس منتقل شده از طریق متن جمله می‌تواند بر تشخیص حس مورد نظر تأثیر بگذارد.

مطالعهٔ سوم

در گام بعدی در یک مطالعهٔ رفتاری شدت عواطف مورد نظر در متن هر یک از ۹۰ جملهٔ رواسازی شده بررسی گردید. در این مطالعه که روالی مانند دو مطالعهٔ قبلی داشت پاسخ‌های ۲۰۰ شرکت‌کنندهٔ فارسی زبان بررسی شد. در پرسشنامهٔ برخط ۹۰ جملهٔ رواسازی شده با ترتیب تصادفی به هر یک از شرکت‌کننده‌ها ارائه شد. در انتهای هر جمله حسی که آن جمله در خواننده ایجاد می‌کند در یک پرانتز نوشته شده بود. شرکت‌کننده‌ها شدت آن حس را با شش گزینه از مقیاس پنج امتیازی لیکرت، که روشی معتبر برای ارزیابی شدت یک ویژگیِ (۳۵) است، مشخص می‌کردند: خیلی کم، کم، متوسط، زیاد، خیلی زیاد و به هیچ وجه. آنها علاوه بر این پنج گزینه اگر با حسی که در انتهای هر جمله مشخص شده بود موافق نبودند می‌توانستند از «به هیچ وجه» نیز استفاده کنند.

با بررسی نتایج، شدت حسی که هر جمله در خواننده ایجاد می‌کند مشخص شد و مجموعه‌ای معتبر از جملات عاطفی زبان فارسی شامل ۹۰ جمله به دست آمد که هر کدام تنها یک حس به خصوص را به خواننده القا می‌کند و برای ساخت پایگاه دادهٔ گفتار عاطفی زبان فارسی این ۹۰ جمله به عنوان متن گفتار استفاده شدند.

تولید و ضبط گفتار عاطفی

هدف نهایی آزمایش (۲۶) تولید و رواسازی پایگاه‌داده‌ای برای گفتار عاطفی زبان فارسی بوده است. بنابراین در گام بعدی هر ۹۰ جملهٔ رواسازی شده توسط دو گویشور فارسی زبان و میان‌سال (آقا ۵۰ ساله و خانم ۴۹ ساله) که دوره‌های بازیگری را گذرانده بودند اما بازیگر حرفه‌ای نبودند (چند سال پیش از این مطالعه برای مدتی به این حرفه مشغول بوده‌اند)، در پنج حس و حالت خنثی اجرا شدند و در یک استودیوی تخصصی ضبط صدا در شهر برلین در آلمان ضبط شدند. هر ۹۰ جمله‌ای که در مرحلهٔ رواسازی متن جملات انتخاب شده بودند (خشم ۱۷ جمله، چندش ۱۵، ترس ۱۵، غم ۱۴ و شادی ۱۵ و خنثی ۱۴ جمله) در این مرحله توسط هر دو گوینده در سه دسته اجرا شدند: (۱) همگون: جملات با متن عاطفی و نوای عاطفی همگون (۷۶ جمله)، (۲) ناهمگون: جملات با متن عاطفی و نوای گفتار خنثی (۷۰ جمله) و (۳) حالت پایه: متن جملات عاطفی و نیز خنثی با نوای گفتار خنثی (۹۰ جمله). به این ترتیب در مجموع ۴۷۲ جملهٔ صوتی تولید شد. صدای گوینده‌ها به صورت مجزا در استودیویی در برلین با نظارت یک زبان‌شناس و یک مهندس آکوستیک طی شش جلسه ضبط شدند (برای اطلاع از جزئیات فنی به (۲۶) مراجعه کنید).

رواسازی جملات صوتی

پژوهش‌های مشابهی که تاکنون در زمینهٔ گفتار عاطفی انجام شده است بین ۱۰ تا ۲۴ آزمودنی برای رواسازی جملات ستفاده کرده‌اند (برای مثال ۲۴). اما در پژوهش (۲۶) از ۳۴ شرکت‌کننده برای رواسازی گفتار عاطفی (گوش کردن و تشخیص حس مورد نظر) استفاده شد (۱۷ مرد و ۱۷ زن با میانگین سنی ۲۶/۳ سال). همهٔ شرکت‌کننده‌ها دانشجویان و فارغ‌التحصیلان ایرانی ساکن شهر برلین بودند و فارسی را از بدو تولد فرا گرفته بودند و کمتر از ۳ سال بود که از ایران دور بودند. همهٔ آنها از نظر روانی_عصبی، شنوایی و بینایی سالم بودند. با استفاده از پرسشنامهٔ مفصل بومی و غالب‌بودن زبان فارسی زبان در آنها تأیید می‌شد.

محرک‌های این مطالعهٔ رفتاری در مجموع ۴۷۲ جملهٔ صوتی (همهٔ جملات عاطفی و خنثی) که شامل سه دستهٔ همگون (۷۶ جمله)، ناهمگون (۷۰ جمله) و حالت پایه (۹۰ جمله) بودند. از شرکت‌کنندگان خواسته شده بود به جملات که برای جلوگیری از اثر توالی در چهار بخش و به صورت کاملاً تصادفی برای هر آزمودنی پخش می‌شد، گوش دهند و بدون توجه به متن جملات نوای عاطفی آن را شناسایی کنند. برای این کار آزمودنی‌ها هفت گزینه در اختیار داشتند: خشم، چندش، ترس، شادی، غم، خنثی و گزینهٔ آخر هیچ یک از موارد بالا. زمان آزمایش برای هر شرکت‌کننده ۹۰ دقیقه بود. برای کاهش اثر خستگی و بهبود توجه بعد از هر ۲۰ دقیقه هر آزمودنی ۵ دقیقه استراحت می‌کرد.

درصد تشخیص هر جملهٔ صوتی محاسبه شد. در نتیجه در مجموع ۴۶۸ جملهٔ صوتی که میزان تشخیص آنها بالاتر از حد نصاب مشخص شده بود به عنوان جملات معتبر در نظر گرفته شدند. یافته‌های این مطالعه نشان می‌دهد که همهٔ جملات عاطفی (با دقت بین ۵/۹۰ درصد، ۹۰/۵ درصد و ۹۸ درصد) بازشناسایی شده‌اند. مقایسهٔ میزان تشخیص پنج حس مورد بررسی نشان می‌دهد که تشخیص حس چندش دشوارتر از باقی حس‌ها بوده است (۹۰/۵ درصد در دستهٔ ناهمگون و ۹۵/۶۵ درصد در دستهٔ همگون). جالب این که در پیشینهٔ پژوهش نیز دشواری‌هایی در شناسایی حس چندش گزارش شده است (۳۳). شرکت‌کنندگان در بازشناسی بازنمودهای ترس در مقایسه با دیگر بازنمودهای عاطفی نیز

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

دشواری داشتند (۹۴/۵ درصد در دستهٔ ناهمگون و ۹۷/۷ درصد در دستهٔ همگون).

سپس این ۴۶۸ جمله صوتی از لحاظ صوت‌شناختی بررسی شدند. نتایج نمایان‌گر تفاوت معناداری در شدت، زیر و بمی صدا و کشش (دیرش) و نقش آنها در تشخیص دقیق حس در پنج آهنگ عاطفی مورد نظر بود (برای آگاهی از شرح مفصل آزمون‌های آماری این بخش (۲۶) را ببینید.

بحث

هدف از پژوهش (۲۶) ساخت یک پایگاه دادهٔ معتبر برای گفتار عاطفی زبان فارسی بود. این پایگاه داده شامل مجموعه‌ای است متشکل از ۹۰ جملهٔ رواسازی شدهٔ فارسی که در پنج حس (خشم، چندش، ترس، شادی و غم) و نیز حالت خنثی طبقه‌بندی شده‌اند. این جمله‌ها در دو مطالعهٔ رفتاری جداگانه توسط ۱۱۲۶ گویشور فارسی زبان رواسازی شده‌اند. دو بازیگر (یک زن، یک مرد) فارسی زبان این جمله‌ها را سه دسته اجرا کرده‌اند. این سه دسته عبارت‌اند از: ۱) همگون: جملات با متن عاطفی و نوای عاطفی همگون، ۲) ناهمگون: جملات با متن عاطفی و نوای خنثی، ۳) حالت پایه: متن جملات عاطفی و نیز خنثی با نوای گفتار خنثی. حاصل این کار ۴۷۲ جمله صوتی با آهنگ‌های عاطفی متفاوت است. روایی محتوایی این جملات صوتی در یک مطالعهٔ رفتاری توسط ۳۴ آزمودنی فارسی‌زبان ارزیابی شد و ۴۶۸ جمله صوتی که درصد تشخیصشان بالای ۷۱/۴۲ درصد بود به عنوان جملات معتبر (رواسازی شده) در نظر گرفته شده‌اند. همچنین تحلیل صوت‌شناختی این ۴۶۸ جمله صوتی نمایان‌گر تفاوت معناداری در زمینهٔ شدت، زیروبمی صدا، و کشش (دیرش) بیان جملات در پنج آهنگ عاطفی مورد مطالعه است. این ویژگی‌های صوت‌شناختی می‌توانند برای تشخیص حس مورد نظر به شنونده کمک کند.

نتیجه‌گیری

در پایان با توجه به ماهیت پیچیده و چند وجهی زبان به نظر می‌رسد که بررسی چگونگی پردازش نوای عاطفی گفتار به صورت مجزا اگرچه جزئیات دقیقی از چگونگی ایجاد ارتباط کلامی_عاطفی به دست می‌دهد، توجه به تأثیر سطوح نحو، معنا و کاربردشناسی و تعامل این سطوح با نوای گفتار (عاطفی و زبان‌شناختی) نیز ضروری است. در مطالعهٔ مجزای سطوح زبان به ویژه در مطالعات عصب‌شناختی با توجه به محدودیت‌های فنی آنها داده‌های زبانی به صورت بافت زدوده بررسی می‌شوند و برای به دست آوردن شواهد دقیق از چگونگی پردازش و درک زبان، توجه به طراحی تکلیف مناسب و دقت در انتخاب و تدوین محرک‌های مناسب از مهمترین مراحل انجام این آزمایش‌ها است. بنابراین دسترسی به پیکره‌ای استاندارد از نوای عاطفی گفتارِ هر زبان، پیش نیاز پژوهش دربارهٔ نوای عاطفی گفتار است. (۸). همچنین با توجه به این که در این آزمایش‌ها معمولاً از محرک دستکاری شده استفاده می‌شود لازم است ویژگی‌های صوت‌شناختیِ مؤثر (بسامد پایه، دیرش، شدت و...) در تداعی نوای عاطفی، نگرشی و انگیزشی در هر زبان با استفاده از پژوهش‌های ادراکی و رفتاری مناسب مشخص گردد.

پایگاه دادهٔ گفتار عاطفی زبان فارسی (Persian ESD) یک ابزار پژوهشی معتبر (متنی و گفتاری) برای پژوهش‌های بینا-زبانی و بینا-فرهنگی در زمینهٔ گفتار عاطفی است (برای آشنایی با محدودیت‌های این پژوهش و یافته‌های آماری آن به (۲۶) مراجعه کنید). این ابزار که برای استفاده در مطالعات حوزه‌های گوناگون از جمله عصب‌شناسی زبان، روان‌شناسی زبان، زبان‌شناسی بالینی، گفتار درمانی و نیز سنتز گفتار تهیه شده است، این امکان را به پژوهشگران می‌دهد که تأثیر نوای عاطفی گفتار، متن گفتار (جملات) و نیز تعامل این دو را در درک گفتار عاطفی به طور جداگانه بررسی کنند. این پایگاه داده تاکنون در مطالعات گوناگونی در حیطه جامعه‌شناسی زبان (برای مثال ۳۶، ۳۷) زبان‌شناسی بالینی و گفتار درمانی (برای مثال ۳۸) و سنتز گفتار (برای مثال (۳۹-۴۲) به عنوان ابزار پژوهش استفاده شده است.

ملاحظات اخلاقی

پیروی از اصول اخلاق در پژوهش

نگاشتهٔ حاضر ماهیت مروری دارد. همهٔ مطالعات مرور شده در این پژوهش در مجلاتی چاپ شده‌اند که به اصول اخلاق در پژوهش پایبند بوده‌اند.

مشارکت نویسندگان

نویسندهٔ اول: تهیه پیش‌نویس مقاله، بازبینی و اصلاح آن، پاسخ‌گو بودن نسبت به تمام جنبه‌های پژوهش. نویسندهٔ دوم: طراحی ساخت و رواسازی پایگاه داده گفتار احساسی زبان فارسی، نگارش بخش دوم مقاله، بازبینی و تصحیح کل مقاله.

منابع مالی

بورسیه بنیاد تحقیقات آلمان (DFG) به نویسنده دوم برای طراحی، ساخت و رواسازی پایگاه داده گفتار احساسی زبان فارسی.

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

تشکر و قدردانی

از شهلا رقیب‌دوست برای راهنمایی‌های ایشان در نگارش بخش ابتدایی (مروری) این نگاشته سپاسگزاریم.

تعارض منافع

نویسندگان هیچ‌گونه تعارض منافع احتمالی را اعلام نمی‌کنند.

References

1. Ding R, Li P, Wang W, Luo W. Emotion processing by ERP combined with development and plasticity. *Neural Plasticity*. 2017;2017:5282670.

2. Wang Y, Hu X, Dai W, Zhou J, Kuo T. Vocal emotion of humanoid robots: A study from brain mechanism. *The Scientific World Journal*. 2014;2014:216341.

3. Seydell-Greenwald A, Chambers CE, Ferrara K, Newport EL. What you say versus how you say it: Comparing sentence comprehension and emotional prosody processing using fMRI. *NeuroImage*. 2020;209:116509.

4. Imaizumi S, Noguch1i Y, Homma M, Yamasaki K. Development of the brain mechanism for understanding speakers' intents from speech. *Age*. 2006;7(14):15.

5. Paulmann S. The neurocognition of prosody. In: Hickok G, Small S, editors. Neurobiology of language. Amsterdam:Elsevier;2016. pp. 1109-1120.

6. Begleiter H, Gross MM, Kissin B. Evoked cortical responses to affective visual stimuli. *Psychophysiology*. 1967;3(4):336-344.

7. Wiens S, Sand A, Olofsson JK. Nonemotional features suppress early and enhance late emotional electrocortical responses to negative pictures. *Biological Psychology*. 2011;86(1):83-89.

8. Liu P, Pell MD. Recognizing vocal emotions in Mandarin Chinese: A validated database of Chinese vocal emotional stimuli. *Behavior Research Methods*. 2012;44:1042-1051.

9. Kotz SA, Meyer M, Paulmann S. Lateralization of emotional prosody in the brain: An overview and synopsis on the impact of study design. *Progress in Brain Research*. 2006;156:285-294.

10. Chen X, Yang Y. When brain differentiates happy from neutral in prosody. In Speech prosody, 6th International Conference. 2012 May 22-25; Shanghai, China;2012.

11. Schirmer A, Kotz SA. Sex differences in vocal emotional processing. *Trends in Cognitive Sciences*. 2006;1(10):24-30.

12. Wickens S, Perry C. What do you mean by that?! An electrophysiological study of emotional and attitudinal prosody. *Plos One*. 2015;10(7):e0132947.

13. Haworth S. It's how you said it and what I heard: A comparison of motivational and emotional tone of voice [PhD Dissertation]. Colchester, England:University of Essex;2018.

14. Friederici AD, Alter K. Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*. 2004;89(2):267-276.

15. Ekman P. Basic emotions. In: Dalgleish T, Power MJ, editors. Basic emotions. The handbook of cognition and emotion. 1st ed. New York:John Wiley & Sons;1999. pp. 45-60.

16. Weinstein N, Ryan RM. When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient. *Journal of Personality and Social Psychology*. 2010;98(2):222-224.

17. Weinstein N, Zougkou K, Paulmann S. Differences between the acoustic typology of autonomy-supportive and controlling sentences. In: Proceedings of the 7th Conference on Speech Prosody. 2014 May 20-23; Dublin, Ireland;2014.

18. Weinstein N, Zougkou K, Paulmann S. You 'have'to hear this: Using tone of voice to motivate others. *Journal of Experimental Psychology: Human Perception and Performance*. 2018;44(6):898-913.

19. Wildgruber D, Riecker A, Hertrich I, Erb M, Grodd W,

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Ethofer T, Ackermann H. Identification of emotional intonation evaluated by fMRI. *Neuroimage*. 2005;24(4):1233-1241.

20. Wildgruber D, Ackermann H, Kreifelts B, Ethofer T. Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*. 2006;156:249-268.

21. Saarimaki H, Glerean E, Smirnov D, Mynttinen H, Jaaskelainen IP, Sams M, et al. Classification of emotions based on functional connectivity patterns of the human brain. *bioRxiv*. 2020:2020-01.

22. Chang J, Zhang X, Zhang Q, Sun Y. Investigating duration effects of emotional speech stimuli in a tonal language by using event-related potentials. *IEEE Access*. 2018;6:13541-13554.

23. Pell MD. Judging emotion and attitudes from prosody following brain damage. *Progress in Brain Research*. 2006;156:303-317.

24. Pell MD, Paulmann S, Dara C, Alasseri A, Kotz SA. Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics*. 2009;37(4):417-435.

25. Paulmann S, Uskul AK. Cross-cultural emotional prosody recognition: Evidence from Chinese and British listeners. *Cognition & Emotion*. 2014;28(2):230-244.

26. Keshtiari N, Kuhlmann M, Eslami M, Klann-Delius G. Recognizing emotional speech in Persian: A validated database of Persian emotional speech (Persian ESD). *Behavior Research Methods*. 2015;47:275-294.

27. Ben-David BM, Van Lieshout PH, Leszcz T. A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Injury*. 2011;25(2):206-220.

28. Luo X, Fu QJ, Galvin III JJ. Cochlear implants special issue article: Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends in Amplification*. 2007;11(4):301-315.

29. Maurage P, Joassin F, Philippot P, Campanella S. A validated battery of vocal emotional expressions. *Neuropsychological Trends*. 2007;2(1):63-74.

30. Russell JA. Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*. 1994;115(1):102-141.

31. Banse R, Scherer KR. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*. 1996;70(3):614-636.

32. Frank MG, Stennett J. The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology*. 2001;80(1):75-85.

33. Scherer KR, Banse R, Wallbott HG, Goldbeck T. Vocal cues in emotion encoding and decoding. *Motivation and Emotion*. 1991;15:123-148.

34. Kotz SA, Paulmann S. When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research*. 2007;1151:107-118.

35. Calder J. Survey research methods. *Medical Education*. 1998;32(6):636-652.

36. Keshtiari N, Kuhlmann M. The effects of culture and gender on the recognition of emotional speech: Evidence from Persian speakers living in a collectivist society. *International Journal of Society, Culture & Language*. 2016;4(2):71-86.

37. Keshtiari N. Gender effects on the recognition of emotional prosody: Evidence from Persian language. *Journal of Sociolinguistics*. 2016;1(1):87-99. (Persian)

38. Arefi HN, Sameni SJ, Jalilvand H, Kamali M. Effect of hearing aid amplitude compression on emotional speech recognition. *Auditory and Vestibular Research*. 2017;26(4):223-230.

39. Bastanfard A, Abbasian A. Speech emotion recognition in Persian based on stacked autoencoder by comparing local and global features. *Multimedia Tools and Applications*. 2023;82:36413-36430.

40. Naderi N, Nasersharif B. Cross corpus speech emotion recognition using transfer learning and attention-based fusion of Wav2Vec2 and prosody features. *Knowledge-Based Systems*. 2023;277:110814.

41. Pourebrahim Y, Razzazi F, Sameti H. Semi-supervised parallel shared encoders for speech emotion recognition. *Digital Signal Processing*. 2021;118:103205.

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

42. Shirani A, Nilchi AR. Speech emotion recognition based on SVM as both feature selector and classifier. *International Journal of Image, Graphics and Signal Processing*. 2016;8(4):39-45.

43. Kotz SA, Meyer M, Alter K, Besson M, Von Cramon DY, Friederici AD. On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language*. 2003;86(3):366-376.

44. Kotz SA, Paulmann S, Raettig T. efMRI evidence for implicit emotional prosodic processing. In: Processing Speech Prosody. 2006 May 2-5; Dresden, Germany;2006.

45. Proverbio AM, Santoni S, Adorni R. ERP markers of valence coding in emotional speech processing. *Iscience*. 2020;23(3):100933.

46. Ceravolo L, Fruhholz S, Pierce J, Grandjean D, Peron J. Basal ganglia and cerebellum contributions to vocal emotion processing as revealed by high-resolution fMRI. *Scientific Re-*

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

[DOI: 10.30514/icss.26.1.91]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License