Revista CEFAC

Original articles

Emotional prosody recognition using pseudowords from the Hoosier Vocal Emotions Collection

Paulo Vitor de Oliveira Santos Matos[1] (ID)

Rayssa Silva Santos Andrade[1] (ID)

Maria Inês Beltrati Cornacchioni Rehder[2] (ID)

Raphaela Barroso Guedes-Granzotti[1] (ID)

Kelly da Silva[3] (ID)

Carla Patrícia Hernandez Alves Ribeiro César[1] (ID)

[1] Universidade Federal de Sergipe, São Cristóvão, Sergipe, Brasil.

[2] CEFAC, Rio Claro, São Paulo, Brasil.

[3] Universidade Federal de Sergipe, Lagarto, Sergipe, Brasil.

ABSTRACT

Purpose: to verify whether the Hoosier Vocal Emotions Collection *corpus* allows the identification of different emotional prosodies in Brazilian adults.

Methods: 60 healthy adults equally distributed by sex, aged between 18 and 42 years, participated in the Mini-Mental State Examination and subtests related to prosody (Montreal communication battery and those from the Hoosier Vocal Emotions Collection *corpus*, with 73 pseudowords produced by two different actresses). The results were analyzed using descriptive statistics and the Chi-square test, which had a significance of 5%.

Results: in general, the emotional prosodies from the Hoosier Vocal Emotions Collection were identified with an average accuracy of 43.63%, with the highest hits, in descending order, for neutrality, sadness, happiness, disgust, anger, and fear. As for sex, there were statistically significant differences regarding the correct answers in the neutrality and disgust prosodies for males, while for females, there were differences in happiness and anger prosodies. Both sexes had more incredible difficulty in identifying prosody related to fear.

Conclusion: the Hoosier Vocal Emotions Collection *corpus* allowed the identification of the emotional prosodies tested in the studied sample, with sexual dysmorphism to emotional prosodic identification being found.

Keywords: Emotions; Voice Recognition; Speech, Language and Hearing Sciences

INTRODUCTION

The expression of emotion is multimodal (face, voice, and body) and needs to be minimally coherent between the resources used so that the emotional prosody can be understood, depending on how the sender interacts with a given situation[1]. Regarding voice, the subject of this study, emotional prosodic differences allow the identification and discrimination of distinct emotional states. Several aspects can be analyzed, such as pitch (the subjective sensation of frequency), loudness (the personal sense of intensity), duration, and speed of speech. Furthermore, it is also possible to distinguish between simulated and non-simulated voices with adequate professional training. Identifying the sender's emotional state in most circumstances is plausible, thus increasing the degree of relevance for their interpretation in different fields of knowledge[2].

However, the variations involve the larynx and the entire vocal tract. For example, in the prosody of sadness, the vocal tract tends to be less open for low vowels. In contrast, in the prosody of happiness, the vocal tract is significantly shorter than in anger and sadness, in most cases[3].

Some factors can impair the recognition of emotional prosody, such as auditory[4] and neurological[5] disorders, psychological disorders[6], and disabilities in executive functions[7]. Given the above, researchers[2] suggested paying attention to these aspects when analyzing emotional prosody.

According to the literature[8], validating corpus containing emotional prosodic stimuli, such as those from the Hoosier Vocal Emotions Collection, can facilitate the understanding of prosodic use and identification by researchers and clinicians. Thus, the use of adequately calibrated instruments would make it possible to investigate the processing of emotions in individuals with psychopathic traits, aphasia, schizophrenia, and other mental disorders, bilingual individuals, or non-native speakers of English, or even it could be used in the training of automatic emotion recognition algorithms, justifying the carrying out of this research on a Brazilian sample.

Therefore, this research aimed to verify whether the Hoosier Vocal Emotions Collection corpus allows the identification of emotional prosodies (happiness, sadness, anger, disgust, fear, and neutrality) in Brazilian adults and whether the respective identification is the same between the sexes.

METHODS

This research was initiated after approval by the Research Ethics Committee (CEP) of the Federal University of Sergipe, Brazil, under CAAE nº 59618322.0.0000.5546 and opinion nº 5,539,794, following the ethical research recommendations described by Council Resolution 466/12 National Health.

This cross-sectional, descriptive, observational study was conducted with a meticulously selected convenience sample. The research took place at a Brazilian university, in a controlled environment, on days and times agreed upon with the research participants. The sample consisted of healthy Brazilian adults, carefully chosen for their lack of neurological disorders, as confirmed by their reports.

With a solid commitment to the rights and safety of the research participants, we ensured that the explanatory letter and the Free and Informed Consent Form were thoroughly read and signed. This process guaranteed the participants' right to privacy, secrecy, confidentiality, and anonymity of personal data. They were also assured the right to obtain information about the results of the tests applied and compensation for any signs of damage during and after the research.

Participants were recruited through an oral invitation, totaling 60 subjects distributed equally between the sexes. Ages ranged between 18 and 42 years old (average: 23.15±5.17), with 46 young adults between 18 and 24 and 14 adults between 25 and 42. Regarding education, they had between eight and 25 years of study (average: 16.15±4.06). According to the literature[9], age and education may affect the interpretation of results, justifying the choice of literate adults.

The inclusion criteria were complete primary education, age between 18 and 42 years old, negative screening for hearing loss, prosodic difficulties (comprehension and production), and cognitive changes. The exclusion criteria were: positive history regarding the use of drugs or medications that act on the central nervous system, neurological, psychic, and mental disorders, as well as the presence of visual difficulties (except those duly corrected).

Regarding the research procedures, participants carried out:

- Anamnesis to collect information on identification data, socioeconomic data, and data relating to possible auditory, neurological, and comprehension complaints.

- The Mini-Mental State Examination (MMSE)[10], translated and validated into Portuguese (BR)[11], was thoughtfully used to ensure the inclusion of all participants, regardless of cognitive abilities. This approach was taken to avoid the exclusion of individuals with possible cognitive impairment, which could make the test difficult to understand. The Mini-Mental State Examination mean was 29.03 ± 0.89, a result compatible with the participants' educational level, demonstrating the study's inclusivity.

- Part of the adaptation of the "Protocole Montréal d'Évaluation de la Communication- Protocole MEC" (Montréal Communication Assessment Battery - MAC Battery), validated for Brazilian Portuguese[12]. This battery comprises nine tests: the questionnaire on awareness of difficulties, proof of conversational and narrative speeches, interpretation of metaphors, lexical evocation, linguistic and emotional prosody, indirect speech acts, and semantic judgment. As a screening for prosodic difficulties, only the conversational speech tests (through spontaneous conversation with possible topics: family, work, leisure, and current news, for four minutes) and linguistic and emotional prosodic skills (comprehension and production) were applied. The participant must present intact pragmatic, lexico-semantic, discursive, and prosodic aspects to be able to participate in the research. Those who obtained an adequate score for their age group and years of schooling, as proposed by the assessment instrument[12], were included in the study. The average number of correct answers regarding the emotional prosody of the Montreal drums was 9.4 ± 1.69.

- Assessment of emotional prosody recognition: the Hoosier Vocal Emotions Collection corpus is composed of 73 disyllabic pseudowords pronounced in English by two actresses (AG and KM, as described in the original research), tested and validated by the authors of the respective collection8. The pseudowords were phonetically balanced using the International Phonetic Alphabet (IPA). The American actresses emitted each pseudoword in six different ways (each pseudoword was emitted twice by each actress): with happiness, sadness, fear, anger, disgust, and

neutrality, totaling 1,763 audio files divided into four lists (1 and 2 were produced by one actress, with 438 sounds each and 3 and 4 by another, with list 3 presenting 443 and list 4, 444 sounds). To use the pseudowords with the highest number of correct answers per tested emotion uttered by the actresses and reduce the application time, twelve pseudowords were selected from these lists for the prosodies of happiness, sadness, fear, anger, and disgust (totaling 50 pseudowords). There were thirteen with neutral prosody, so the total number of pseudowords tested for the Brazilian sample was 73. To this end, the pseudoword selection criterion was the highest percentage of correct answers based on the results of the original research8, as can be seen in Table 1. Therefore, the lists were offered in a randomized manner (randomization carried out using the Excel spreadsheet from the Microsoft Office® package) among the participants; with every 50 stimuli, there was a pause in order not to cause fatigue, thus avoiding possible errors. The final list used in the research is detailed in Table 1. The answers were written down on a sheet with six answer sheets. The accuracy rates for identifying emotions were analyzed based on the responses obtained. It is worth noting that the authors allowed using prosodic collections of pseudowords.

The previously selected pseudowords were presented on a pre-scheduled day and time with the participants in an air-conditioned room, using the Audacity® software, AKG K72 headphones, and Dell Intel core i5 computer. There was prior instruction so that the participant paid attention to the pseudoword offered and, later, marked on a specific sheet which emotional prosody corresponded to the one uttered. Two sentences were added to each pseudoword, created by a native Brazilian speech therapist, without changes in speech, namely: "I say" and "I say again". Thus, each pseudoword was presented twice immediately following the sentence issued in Portuguese. For example: "I say < pseudoword of a certain emotional prosody" and "I say again <the same pseudoword emitted previously". To insert these excerpts in Portuguese (Brazil) into the pseudowords in the Hoosier Vocal Emotions Collection, the software Audacity® was used.

Table 1. Selection of 73 pseudowords from the Hoosier Vocal Emotional Collection corpus for each actress's emissions with the average number of hits and standard deviation per emotion tested

Emotions	Corpus 1 – Actress AG			Corpus 2 – Actress KM		
	Hit Range (%)	Mean	SD	Hit Range (%)	Mean	SD
Sadness	100 to 79.2%	87.74	11.9	95.7 to 56.5%	78.73	11.13
Fear	87.5 to 56%	75.37	10.34	87 to 56.5%	71.28	9.01
Happiness	96 to 52%	75.75	12.93	100 to 52.2%	66.15	14.67
Anger	80 to 58.3%	73.9	5.58	87.5 to 58.3%	71.57	7.86
Disgust	96 to 60%	74.7	12.22	91.7 to 62.5%	82.89	7.95
Neutrality	80 to 50%	61.31	10.36	87.5 to 54.2%	73.95	10.40

Captions: SD = standard deviation; % = percentage.

Chart 1. List of 73 pseudowords selected, by Emotions, related to American actresses, used in the present research

Item	Orthographic Representation	Corpus Actress AG/ Emotions	Corpus Actress KM/ Emotions
1	**Ner**vack	Sadness	Disgust
2	**Lor**ack	Sadness	Disgust
3	**Lai**ret	Sadness	Disgust
4	**Vok**ered	Fear	Happiness
5	**Tair**ack	Sadness	Anger
6	**Dom**ner	Sadness	Neutrality
7	**Nam**my	Sadness	Disgust
8	**Tan**nock	Sadness	Neutrality
9	**A**gerth	Happiness	Disgust
10	**Arm**idge	Fear	Neutrality
11	**Bu**rish	Sadness	Neutrality
12	**Der**nom	Sadness	Neutrality
13	**Re**vo	Sadness	Disgust
14	**Fin**gill	Sadness	Disgust
15	**Jou**less	Sadness	Sadness
16	**Leb**by	Happiness	Happiness
17	**Low**men	Neutrality	Disgust
18	**Mad**age	Happiness	Neutrality
19	**Men**no	Anger	Neutrality
20	**Mer**rus	Fear	Disgust
21	**Mow**an	Anger	Anger
22	**Na**bick	Fear	Disgust
23	**Nem**my	Anger	Neutrality
24	**Nid**der	Fear	Anger
25	**Nil**len	Anger	Disgust
26	**Nom**el	Anger	Fear
27	**Nom**ey	Fear	Anger
28	**Ram**idge	Fear	Neutrality
29	**Sha**vil	Happiness	Neutrality
30	**Shi**bur	Happiness	Neutrality
31	**Slo**ver	Sadness	Anger
32	**Ter**rel	Sadness	Anger
33	**Tha**ger	Anger	Anger
34	**Tho**mer	Sadness	Fear
35	**Val**ish	Neutrality	Anger
36	**Ven**ner	Anger	Anger

Item	Orthographic Representation	Corpus Actress AG/ Emotions	Corpus Actress KM/ Emotions
37	**Ver**ney	Disgust	Neutrality
38	**Vig**ging	Disgust	Sadness
39	**Vok**er	Disgust	Anger
40	**Vokered**	Sadness	Fear
41	**Vol**ers	Fear	Happiness
42	**Win**nith	Fear	Sadness
43	**Zid**dy	Fear	Fear
44	**Zil**ard	Happiness	Happiness
45	**Ver**coed	Happiness	Fear
46	**For**ny	Disgust	Fear
47	**Ad**mage	Neutrality	Happiness
48	**Aff**ning	Neutrality	Fear
49	**El**by	Neutrality	Fear
50	**Er**vy	Disgust	Fear
51	**Inf**ess	Disgust	Fear
52	**You**ssle	Anger	Sadness
53	**Ker**vo	Disgust	Sadness
54	**Ker**voed	Disgust	Fear
55	**Lar**py	Disgust	Happiness
56	**Lek**nodge	Neutrality	Sadness
57	**Mod**ner	Disgust	Sadness
58	**Mok**ers	Anger	Happiness
59	**Mus**ser	Sadness	Sadness
60	**Naf**fing	Anger	Sadness
61	**Nif**ish	Anger	Anger
62	**Niph**er	Anger	Sadness
63	**Oth**ening	Sadness	Fear
64	**Rack**ies	Fear	Neutrality
65	**Sco**pies	Happiness	Happiness
66	**Shi**fin	Neutrality	Disgust
67	**Vack**ner	Disgust	Anger
68	**Va**shil	Happiness	Happiness
69	**Vi**shal	Fear	Happiness
70	**Wed**ick	Fear	Happiness
71	**Win**thy	Fear	Sadness
72	**You**shing	Sadness	Sadness
73	**Zu**ber	Happiness	Happiness

Note: The boldface in pseudowords refers to the stressed syllable.

At the end of the collection, the data were tabulated in Microsoft Office Excel 2013 spreadsheets. The results were analyzed using descriptive statistics, such as frequency, mean, and standard deviation measurements, and inferential statistics, such as the test Chi-square, considering a significance level of 5%, using the JAMOVI software.

RESULTS

Our research, conducted with a comprehensive sample encompassing all the prosodies in the Hoosier Vocal Emotions Collection, yielded a significant finding. In general, the correct answers corresponded to 43.63%. The detailed results obtained by emotional prosody can be seen in Table 2, providing a thorough and reliable data analysis.

Table 2. Number and percentage of correct answers from the 60 participants (30 for each corpus) for the prosodies tested, using pseudowords from the Hoosier Vocal Emotional Collection

Prosody	CORPUS 1 – ACTRESS AG				CORPUS 2 – ACTRESS KM			
	Pseudow	IPA	Hits	%	Pseudow	IPA	Hits	%
NEUTRALITY	shiffin	/ˈʃɪfɪn/	22	73.33%	armidge	/ˈɑɹmɪdʒ/	24	80.00%
	lowmen	/ˈloʊmən/	19	63.33%	ramidge	/ˈɹæmɪdʒ/	21	70.00%
	dernom	/ˈdɜɹnəm/	18	60.00%	dernom	/ˈdɜɹnəm/	20	66.67%
	admage	/ˈædmɪdʒ/	18	60.00%	shavil	/ˈʃævɪl/	19	63.33%
	armidge	/ˈɑɹmɪdʒ/	17	56.67%	menno	/ˈmɛnoʊ/	18	60.00%
	agerth	/ˈægəɹθ/	15	50.00%	nemmy	/ˈnɛmi/	18	60.00%
	leknodge	/ˈlɛknədʒ/	15	50.00%	domner	/ˈdɑmnəɹ/	16	53.33%
	fingill	/ˈfɪŋgəl/	14	46.67%	burish	/ˈbʊɹɪʃ/	15	50.00%
	elby	/ˈɛlbi/	12	40.00%	rackies	/ˈɹækiːz/	15	50.00%
	othening	/ˈɔθ(ə)nɪŋ/	12	40.00%	tannock	/ˈtænək/	15	50.00%
	merrus	/ˈmɛɹəs/	6	20.00%	madage	/ˈmædədʒ/	14	46.67%
	affning	/ˈɑfnɪŋ/	5	16.67%	shibur	/ˈʃɪbəɹ/	13	43.33%
	valish	/ˈvælɪʃ/	1	3.33%	verney	/ˈvɜɹni/	12	40.00%
	Mean AG		13.38	44.62%	Mean KM		16.92	56.41%
	Mean				15.15±5.19 (50.51%)			
	P=0.603							
DISGUST	infess	/ˈɪnfɛs/	23	76.67%	nabick	/ˈnæbɪk/	25	83.33%
	verney	/ˈvɜɹni/	22	73.33%	nervack	/ˈnɜɹvæk/	18	60.00%
	vackner	/ˈvæknəɹ/	22	73.33%	nammy	/ˈnæmi/	17	56.67%
	ervy	/ˈɜɹvi/	20	66.67%	agerth	/ˈægəɹθ/	15	50.00%
	kervo	/ˈkɜɹvoʊ/	20	66.67%	lorack	/ˈlɔɹæk/	13	43.33%
	vigging	/ˈvɪgɪŋ/	14	46.67%	merrus	/ˈmɛɹəs/	13	43.33%
	kervoed	/ˈkɜɹvoʊd/	13	43.33%	nillen	/ˈnɪlən/	12	40.00%
	forny	/ˈfɔɹni/	9	30.00%	shiffin	/ˈʃɪfɪn/	11	36.67%
	modner	/ˈmɑdnəɹ/	7	23.33%	fingill	/ˈfɪŋgəl/	10	33.33%
	ziddy	/ˈzɪdi/	7	23.33%	revo	/ˈɹɛvoʊ/	6	20.00%
	voker	/ˈvoʊkəɹ/	6	20.00%	lairet	/ˈlɛɹət/	6	20.00%
	larpy	/ˈlɑɹpi/	5	16.67%	lowmen	/ˈloʊmən/	5	16.67%
	Mean AG		14	46.67%	Mean KM		12	41.67%
	Mean				13.25±6.31 (44.17%)			
	P=0.819							
ANGER	mokers	/ˈmoʊkəɹs/	20	66.67%	voker	/ˈvoʊkəɹ/	21	70.00%
	nipher	/ˈnɪfəɹ/	19	63.33%	tairack	/ˈtɛɹæk/	20	66.67%
	nomel	/ˈnoməl/	12	40.00%	nifish	/ˈnɪfɪʃ/	18	60.00%
	mowan	/ˈmoʊwən/	10	33.33%	venner	/ˈvɛnəɹ/	17	56.67%
	thager	/ˈθægəɹ/	10	33.33%	slover	/ˈsloʊvəɹ/	16	53.33%
	youssle	/ˈjusəl/	10	33.33%	vackner	/ˈvæknəɹ/	14	46.67%
	nifish	/ˈnɪfɪʃ/	9	30.00%	mowan	/ˈmoʊwən/	14	46.67%
	naffing	/ˈnæfɪŋ/	8	26.67%	valish	/ˈvælɪʃ/	14	46.67%
	nillen	/ˈnɪlən/	3	10.00%	nomey	/ˈnoʊmi/	11	36.67%
	venner	/ˈvɛnəɹ/	2	6.67%	terrel	/ˈtɛɹəl/	9	30.00%
	nemmy	/ˈnɛmi/	1	3.33%	thager	/ˈθægəɹ/	7	23.33%
	menno	/ˈmɛnoʊ/	0	0.00%	nidder	/ˈnɪdəɹ/	4	13.33%
	Mean AG		8.67	28.90%	Mean KM		13.75	45.83%
	Mean				11.21±6.30 (37.37%)			
	P=0.251							

Prosody	CORPUS 1 – ACTRESS AG				CORPUS 2 – ACTRESS KM			
	Pseudow	**IPA**	**Hits**	**%**	**Pseudow**	**IPA**	**Hits**	**%**
HAPPINESS	**lebby**	/ˈlɛbi/	30	100%	**admage**	/ˈædmɪdʒ/	26	86.67%
	zilard	/ˈzɪləɹd/	24	80.00%	**vashil**	/ˈvæʃɪl/	21	70.00%
	nervack	/ˈnɜɹvæk/	24	80.00%	**scopies**	/ˈskoʊpiːz/	20	66.67%
	madage	/ˈmædədʒ/	22	73.33%	**lebby**	/ˈlɛbi/	16	53.33%
	jouless	/ˈdʒoʊlɛs/	20	66.67%	**larpy**	/ˈlɑɹpi/	11	36.67%
	shibur	/ˈʃɪbəɹ/	19	63.33%	**mokers**	/ˈmoʊkəɹs/	10	33.33%
	scopies	/ˈskoʊpiːz/	16	53.33%	**vokered**	/ˈvoʊkəɹd/	7	23.33%
	vercoed	/ˈvɜɹkoʊd/	15	50.00%	**zilard**	/ˈzɪləɹd/	7	23.33%
	vashil	/ˈvæʃɪl/	12	40.00%	**volers**	/ˈvoʊləɹs/	6	20.00%
	shavil	/ˈʃævɪl/	9	30.00%	**wedick**	/ˈwɛdɪk/	6	20.00%
	burish	/ˈbʊɹɪʃ/	9	30.00%	**vishal**	/ˈvɪʃəl/	5	16.67%
	zuber	/ˈzubəɹ/	4	13.33%	**zuber**	/ˈzubəɹ/	4	13.33%
	Mean AG		17	56.67%	Mean KM		11.58	38.60%
	Mean				14.29±7.81 (47.63%)			
	P=0.298							
FEAR	**nomey**	/ˈnoʊmi/	11	36.67%	**forny**	/ˈfoɹni/	30	100%
	nidder	/ˈnɪdəɹ/	9	30.00%	**ziddy**	/ˈzɪdi/	24	80.00%
	winthy	/ˈwɪnθi/	7	23.33%	**vercoed**	/ˈvɜɹkoʊd/	24	80.00%
	winnith	/ˈwɪnɪθ/	6	20.00%	**ervy**	/ˈɜɹvi/	24	80.00%
	vokered	/ˈvoʊkəɹd/	6	20.00%	**othening**	/ˈɔθ(ə)nɪŋ/	21	70.00%
	revo	/ˈɹɛvoʊ/	2	6.67%	**affning**	/ˈɑfnɪŋ/	14	46.67%
	nabick	/ˈnæbɪk/	1	3.33%	**elby**	/ˈɛlbi/	14	46.67%
	volers	/ˈvoʊləɹs/	1	3.33%	**infess**	/ˈɪnfɛs/	11	36.67%
	rackies	/ˈɹækiːz/	1	3.33%	**nomel**	/ˈnɔməl/	7	23.33%
	wedick	/ˈwɛdɪk/	1	3.33%	**vokered**	/ˈvoʊkəɹd/	7	23.33%
	ramidge	/ˈɹæmɪdʒ/	0	0%	**thomer**	/ˈθoʊməɹ/	5	16.67%
	vishal	/ˈvɪʃəl/	0	0%	**kervoed**	/ˈkɜɹvoʊd/	5	16.67%
	Mean AG		3.75	12.45%	Mean KM		15.5	51.66%
	Mean				9.63±8.94 (32.1%)			
	P=0.007*							
SADNESS	**tairack**	/ˈtɛɹæk/	24	80.00%	**youshing**	/ˈjuːʃɪŋ/	24	80.00%
	slover	/ˈsloʊvəɹ/	24	80.00%	**jouless**	/ˈdʒoʊlɛs/	21	70.00%
	lorack	/ˈloɹæk/	22	73.33%	**musser**	/ˈmʌsəɹ/	21	70.00%
	lairet	/ˈlɛɹət/	18	60.00%	**naffing**	/ˈnæfɪŋ/	21	70.00%
	domner	/ˈdɑmnəɹ/	18	60.00%	**modner**	/ˈmɔdnəɹ/	19	63.33%
	youshing	/ˈjuːʃɪŋ/	18	60.00%	**winthy**	/ˈwɪnθi/	19	63.33%
	musser	/ˈmʌsəɹ/	17	56.67%	**winnith**	/ˈwɪnɪθ/	13	43.33%
	nammy	/ˈnæmi/	11	36.67%	**nipher**	/ˈnɪfəɹ/	11	36.67%
	vokered	/ˈvoʊkəɹd/	11	36.67%	**leknodge**	/ˈlɛknədʒ/	9	30.00%
	tannock	/ˈtænək/	9	30.00%	**vigging**	/ˈvɪgɪŋ/	5	16.67%
	thomer	/ˈθoʊməɹ/	7	23.33%	**kervo**	/ˈkɜɹvoʊ/	5	16.67%
	terrel	/ˈtɛɹəl/	4	13.33%	**youssle**	/ˈjusəl/	4	13.33%
	Means		15.25	50.84%			14.33	47.77%
	Mean				14.79±6.89 (49.30%)			
	P=0.157							
	Overall Average Percentage with all tested prosodies: 43.63%							

Captions: Chi-square test; N = absolute frequency; % = relative frequency percentage; * = indicates values with statistically significant difference; Pseudow = pseudo-words; IPA = International Phonetic Alphabet

When comparing the responses about gender, statistically significant differences were observed for the pseudowords that expressed the emotions of neutrality (p=0.015), with more excellent correct answers for male individuals (average number of correct answers=50.76%), disgust (p=0.042), as well as happiness (p=<0.001), with more excellent correct answers for female individuals (average number of correct answers=49.72%) and anger (p= 0.002), as can be seen in Table 3. No significant differences were found about sex for the emotions of sadness or fear.

Table 3. Number and percentage of male and female correct answers for the emotional prosodies tested using pseudowords from the Hoosier Vocal Emotional Collection

| Prosodies | Males | | | | Females | | | | P-Value |
| | Hits | | Misses | | Hits | | Misses | | |
	N	%	N	%	N	%	N	%	
Neutrality	198	50.76%	192	49.23%	193	49.48%	197	50.51%	P= 0.015
Happiness	164	45.55%	196	54.44%	179	49.72%	181	50.27%	P< 0.001
Anger	122	33.88%	238	66.11%	146	40.55%	214	66.94%	P= 0.002
Disgust	179	49.72%	181	50.27%	140	38.88%	220	61.11%	P= 0.042
Sadness	157	43.62%	203	56.39%	190	52.78%	170	47.23%	P= 0.504
Fear	110	30.55%	250	69.44%	123	34.16%	237	65.83%	P= 0.269

Captions: Chi-square test; N = absolute frequency; % = relative frequency percentage

DISCUSSION

The objective of this research was to verify whether the Hoosier Vocal Emotions Collection corpus allows the identification of different emotional prosodies (happiness, sadness, anger, disgust, fear, and neutrality) in Brazilian adults, bearing in mind that, if possible, the results can be compared with other inter-national research, with the instrument already validated. Generally, the tested prosodies were identified with an accuracy of 43.63%, similar to that obtained by the original authors[8], whose percentage was 45%. In this study, the entire corpus was not applied, and the pseudowords with the highest percentages of correct answers were selected from the collection of pseudo-words[8]. This choice was made due to the tiredness and difficulties reported by the participants since the Hoosier Vocal prosodies are emitted with medium/average prosodic intensity, different from the prosody of the MAC[12] battery, which can be considered as strong, and for agility in the procedure (the original corpus of the Hoosier Vocal Emotions Collection has 1,763 files). However, 73 pseudowords were used for each corpus.

The prosodies that obtained above-average hits were neutrality, sadness, happiness, and disgust. Only neutral and sad results were presented in these prosodies, similar to those in the literature[8]. Anger and fear were below average. In the study by Darcy

and Fontaine[8], the emotional prosody of rage had the lowest identification rates. The prosody related to fear was among the prosodies with good percentages of correct answers. One study[13] used functional near-infrared spectroscopy (fNIRS) with three disyllabic pseudowords ("minad," "lagod," "namil") emitted by four subjects (two of each sex) with different prosodies (happiness, sadness, fear, anger and neutrality). To this end, 28 healthy volunteers participated in the study, with a mean age of 26.44±4.7 years. The authors found that participants were faster in discriminating than in naming the prosodies tested and in processing the linguistic content than in emotional prosodies, especially in angry, fearful, and neutral prosodies. There was modulation of oxyhemoglobin changes in the inferior frontal gyrus depending on the condition, task, emotional prosody tested, and cerebral hemisphere. For fear prosody, they verified the involvement of the right hemisphere and, for anger, both hemispheres. Given the above, it can be inferred that the cognitive activity to identify fear and anger prosodies can be justified by the tasks implying greater neuronal load, resulting in more significant difficulties for their identi-fication in the present study. In addition to the above, differences can be justified by the different prosodic use between countries, languages, sexes, and individuals[14].

An important consideration to be made concerns the differences obtained in the prosodic identification of the

MAC battery (94% correct) compared to the pseudo-words from the Hoosier Vocal Emotions Collection (43.63%) selected in this study. In the MAC battery, sentences with a solid prosodic load are used. In contrast, in the present study, pseudowords with medium/average load were used, making the identification task much more difficult, as the literature points out[8]. Furthermore, the literature[15] confirms the above, considering that listeners' accuracy in identifying specific emotional prosody increases according to its emotional intensity, such as, for example, anger. Therefore, when this emotion is transmitted in an average way or with a weak emotional charge, it is more likely to be misinterpreted[15].

No statistically significant differences were found for identifying pseudowords about the presented corpus (described by the original authors as the AG corpus and the KM corpus), except the fear prosody, in which there was a more significant number of correct answers in the KM actress corpus. This difference did not occur in the original corpus and must have happened due to the prior selection used in the present study. Based on the results obtained, it is suggested that the pseudowords with the highest percentages of correct answers selected below be used to screen or evaluate the prosodies related to happiness, sadness, fear, disgust, anger, and neutrality (Table 2 – supplementary material), expanding, thus, the current evaluation options, since in Brazil there is only adaptation and validation of the MAC battery with three prosodies: happiness, anger, and sadness. Future research may clarify the use of pseudowords from the Hoosier Vocal Emotions Collection in different age groups and clinical conditions.

Chart 2. Suggested list of pseudowords for screening/evaluating emotional prosody in young and adult Brazilians, indicating the file to be used

Emotional Prosody	Corpus Actress AG	File	Corpus Actress KM	File
Neutrality	shiffin	n_AG_66_1.wav	armidge	n_KM_10_2.wav
	lowmen	n_AG_17_1.wav	ramidge	n_KM_28_1.wav
	dernom	n_AG_12_2.wav	dernom	n_KM_12_1.wav
Disgust	infess	d_AG_51_1.wav	nabick	d_KM_22_2.wav
	verney	d_AG_37_2.wav	nervack	d_KM_01_1.wav
Anger	mokers	a_AG_58_2.wav	voker	a_KM_39_2.wav
	nipher	a_AG_62_1.wav	tairack	a_KM_05_1.wav
Happiness	lebby	h_AG_16_1.wav	admage	h_KM_47_1.wav
	zilard	h_AG_44_1.wav	vashil	h_KM_68_1.wav
Fear	nomey	f_AG_27_1.wav	forny	f_KM_46_1.wav
	nidder	f_AG_24_2.wav	ziddy	f_KM_43_1.wav
Sadness	tairack	s_AG_05_1.wav	youshing	s_KM_72_1.wav
	slover	s_AG_31_1.wav	jouless	s_KM_15_1.wav

Some prosodies are considered strong emotional activators (such as anger, fear, and happiness), while others show weak activation (such as sadness, boredom, and tenderness). In those with strong activation, the characteristics of acoustics are an increase in fundamental frequency, pitch, and speech speed. In those with weak emotional activation, the opposite occurs[16], and it is worth investigating whether there are differences between the sexes in this identification.

In this sense, in the present study, differences were found between the sexes about the prosodic identification of happiness and disgust (more excellent hits for women) and neutrality (men), and it is not possible to compare the results obtained with the original study[8], since this analysis was not carried out. However, researchers[17] did not find differences in identifying emotional prosody about the sexes. A literature review study[18] showed differences in emotional prosodic identification between the sexes, and, according to the authors, differences in this sense may show that the processing of information between men and women occurs differently, both due to faster female temporal processing and due to the social role

played by women in most cultures. A systematic review study with meta-analysis[19] confirmed sexual dimorphism related to emotional reactivity in the activation of different brain areas, concluding that it is essential to consider sex in research involving emotion.

The importance of screening and evaluating prosody concerns the possibility of early diagnosis of mild cognitive disorders[20,21], allowing early intervention in these clinical conditions. In autism spectrum disorder (ASD), there may be difficulties in recognition and identification and in the use of emotional prosody[6,22], depending on the number of response options offered, the emotion tested, and the patient's verbal and cognitive skills6. In psychopathies, such difficulties can also occur, and it is even possible to observe them in children at high risk of developing future criminal behavior[23]. Furthermore, changes in comprehension and prosodic production may highlight a neurological disorder that needs to be investigated[5], justifying the research effort in the area, mainly due to the insufficient quantity of materials validated for use in Brazil.

As mentioned by Darcy and Fontaine[8], this research's limitation lies in the production of emotional prosodies uttered exclusively by two female people. This prevents the comparison of their identification in relation to the prosodic production emitted by male people, which is a gap for future investigations. However, there are reports in the literature that prosodic recognition is facilitated when the sender is female[24].

As suggestions for further research, the application of the synthesized corpus with the best percentages of correct answers (both for actress AG and KM) in different age groups as an instrument to obtain correct scores and, as an example of research[25] on the differences between the emotional prosodic recognition of young and older adults, verifying whether the scores differ between age groups; apply the synthetic corpus to different conditions such as mild cognitive disorders, Parkinson's and Alzheimer's diseases, depressive conditions and psychopathies, for example, is suggested.

CONCLUSION

The findings showed that the Hoosier Vocal Emotions Collection corpus effectively identified emotional prosodies (happiness, sadness, anger, disgust, fear, and neutrality) in the study sample. Among the emotional prosodies tested, the most easily identified, in a descending order of correct answers, were neutrality, sadness, happiness, disgust, anger, and fear. Notably, statistically significant differences were found in the identification of neutrality and disgust for males and happiness and anger for females, indicating sexual dimorphism in emotional prosodic identification.

ACKNOWLEDGEMENTS

REFERENCES

1. Sznycer D, Cohen AS. Are emotions natural kinds after all? Rethinking the issue of response coherence. Evol. psychol. 2021;19(2):14747049211016009. https//:doi.org/10.1177/14747049211016009 PMID: 34060370.

2. César CPHAR, Pellicani AD, Farias IS, Reis LF, Santos L. The identification of voice in emotions: A narrative review of the literature. In: Pereira S, César CPHAR, Landy N, editors. Face Summit 2021. Porto: FeeLab; 2022. p. 51-67.

3. Kim J, Toutios A, Lee S, Narayanan SS. Vocal tract shaping of emotional speech. Comput. speech lang. 2020;64:101100. https//:doi.org/10.1016/j.csl.2020.101100 PMID: 32523241.

4. Yeshoda K, Raveendran R, Konadath S. Perception of vocal emotional prosody in children with hearing impairment. Int. j. pediatr. otorhinolaryngol. 2020;137:110252. https://doi.org/10.1016/j.ijporl.2020.110252 PMID: 32896359.

5. Coulombe V, Happinessal M, Martel-Sauvageau V, Monetta L. Affective prosody disorders in adults with neurological conditions: A scoping review. International Int. J. Lang. Commun. Disord. 2023;58(6):1939-54. https://doi.org/10.1111/1460-6984.12909 PMID: 37212522.

6. Zhang M, Xu S, Chen Y, Lin Y, Ding H, Zhang Y. Recognition of affective prosody in autism spectrum conditions: A systematic review and meta-analysis. Autism. 2022;26(4):798-813. https://doi.org/10.1177/1362361321995725 PMID: 33722094.

7. Ikeda S. Overcoming lexical bias in the judgment of emotion in speech: Role of executive function and usefulness understanding in young children. J. genet. psychol. 2022;183(3):211-21. https://doi.org/10.1080/00221325.2022.2037499 PMID: 35132942.

8. Darcy I, Fontaine NMG. The Hoosier vocal emotions corpus: a validated set of north american English pseudowords for evaluating emotion processing. Behav. Res. Meth. 2020;52(2):901-17. https://doi.org/10.3758/s13428-019-01288-0 PMID: 31485866.

9. Kerr MS, Pagliarin KC, Mineiro A, Ferré P, Joanette Y, Fonseca RP. Montreal Communication Evaluation Battery - Portuguese version: Age and education effects. CoDAS. 2015;27(6):550-6. https://doi.org/10.1590/2317-1782/20152015029 PMID: 26691619.

10. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: A practical method for grading the cognitive state of patients for clinician. J. Psychiatr. Res. 1975;12:189-98. https://doi.org/10.1016/0022-3956(75)90026-6 PMID: 1202204.

11. Bertolucci PHF, Brucki SMD, Campacci SR, Juliano Y. O mini-exame do estado mental em uma população geral: impacto da escolaridade. Arq. Neuro-Psiquiatr. 1994;52(1):1-7. https://doi.org/10.1590/S0004-282X1994000100001 PMID: 1202204.

12. Fonseca RP, Parente MAMP, Côté H, Joanette Y. Adaptation process to Brazilian Portuguese of the Montreal communication evaluation battery: MAC battery. Psicol. Reflex. Crit. 2007;20:259-67. https://doi.org/10.1590/S0102-79722007000200012

13. Gruber T, Debracque C, Ceravolo L, Igloi K, Bosch BM, Frühholz S et al. Human discrimination and categorization of emotions in voices: A functional near-infrared spectroscopy (fNIRS) study. Front. Neurosci. 2020;14:570. https://doi.org/10.3389/fnins.2020.00570. eCollection2020 PMID: 32581695.

14. van Rijn P, Larrouy-Maestri P. Modelling individual and cross-cultural variation in the mapping of emotions to speech prosody. Nat. Hum. Behav. 2023;7(3):386-96. https://doi.org/10.1038/s41562-022-01505-5 PMID: 36646838.

15. Morningstar M, Gilbert AC, Burdo J, Leis M, Dirks MA. Recognition of vocal socioemotional expressions at varying levels of emotional intensity. Emotion. 2021;21(7):1570-5. https://doi.org/10.1037/emo0001024 PMID: 34570558.

16. Scherer KR. Vocal communication of emotion: A review of research paradigms. Speech Commun. 2003;40(1-2):227-56. https://doi.org/10.1016/S0167-6393(02)00084-5

17. Ertürk A, Gürses E, Kulak Kayıkcı ME. Sex related differences in the perception and production of emotional prosody in adults. Psychol. Res. 2024;88(2):449-57. https://doi.org/10.1007/s00426-023-01865-1 PMID: 37542581.

18. Lin Y, Ding H, Zhang Y. Unisensory and multisensory stroop effects modulate gender differences in verbal and nonverbal emotion perception. J. Speech Lang. Hear. Res. 2021;64(11):4439-57. https://doi.org/10.1044/2021_JSLHR-20-00338 PMID: 34469179.

19. Filkowski MM, Olsen RM, Duda B, Wanger TJ, Sabatinelli D. Sex differences in emotional perception: Meta analysis of divergent activation. Neuroimage. 2017;147:925-33. https://doi.org/10.1016/j.neuroimage.2016.12.016 PMID: 27988321.

20. Themistocleous C, Eckerström M, Kokkinakis D. Identification of mild cognitive impairment from speech in Swedish using deep sequential neural networks. Front. Neurol. 2018;9:975. https://doi.org/10.3389/fneur.2018.00975 PMID: 30498472.

21. Gosztolya G, Vincze V, Tóth L, Pákáski M, Kálmán J, Hoffmann I. Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. Comput. Speech Lang. 2019;53:181-97. https://doi.org/10.1016/j.csl.2018.07.007

22. Zuanetti PA, Silva K, Pontes-Fernandes ÂC, Dornelas R, Fukuda MTH. Characteristics of the emissive prosody of children with Autism Spectrum Disorder. Rev. CEFAC. 2018;20(5):565-72. https://doi.org/10.1590/1982-021620182051718

23. Van Zonneveld L, De Sonneville L, Van Goozen S, Swaab H. Recognition of facial emotion and affective prosody in children at high risk of criminal behavior. J. Int. Neuropsychol. Soc. 2019;25(1):57-64. https://doi.org/10.1017/S1355617718000796 PMID: 30394247.

24. Eskritt M, Zupan B. Emotion perception from vocal cues: Testing the influence of emotion intensity and sex on in-group advantage. Can. J. Exp. Psychol. 2023;77(3):202-11. https://doi.org/10.1037/cep0000310 PMID: 37535514.

25. Martzoukou M, Nasios G, Kosmidis MH, Papadopoulou D. Aging and the perception of affective and linguistic prosody. J. Psycholinguist Res. 2022;51(5):1001-21. https://doi.org/10.1007/s10936-022-09875-7 PMID: 35441951.

Author contributions:

PVOSM, RSSA: Data curation; Formal analysis; Investgation; Methodology; Software development, implementation, and testing; Writing - Original draft; Writing – Review and editing.

MIBCR, RBGG, KS: Writing - Original draft; Writing – Review and editing.

CPHARC: Conceptualization, Data curation, Data Analysis, Research, Methodology; Project administration; Software development, implementation, and testing; Supervision; Writing - Original draft; Writing – Review and editing.

Data sharing statement:

The individual data of de-identified participants (gender and age) can be shared upon request to the corresponding author by email. However, those using the shared data must commit to citing both the original authors of the Hoosier Vocal Emotions Collection and those of the present study.