RESEARCH ARTICLE

ACTA Vol. 10, No. 1, 2012, 57-67
NEUROPSYCHOLOGICA

Received: 28.10.2011
Accepted: 30.03.2012

A – Study Design
B – Data Collection
C – Statistical Analysis
D – Data Interpretation
E – Manuscript Preparation
F – Literature Search
G – Funds Collection

THE EVALUATION OF EMOTIONAL PROSODY BY STUDENTS MAJORING IN SCIENTIFIC AND HUMANISTIC SUBJECTS

Weronika D. MacQueen[1(A,B,C,D,E,F)],
Bruce D. MacQueen[2,3(A,B,C,D,E,F)]
[1] Sopot, Poland
[2] Department of Speech Therapy, University of Gdansk, Gdansk, Poland
[3] Faculty of Philology, Ateneum College, Gdansk, Poland

SUMMARY

Background: Although disturbances of prosody are known to be a common symptom of damage to the right hemisphere, little research has been done on the reception of emotional prosody in healthy subjects. The goal of this study was to examine possible differences in the reception of emotional prosody in students from different fields of study.

Material/Methods: We studied 50 students from the University of Silesia in Katowice, Poland, divided into two groups: Group H, consisting primarily of language students, and Group M, consisting of students majoring in mathematics or computer science. The test (TOPE-1) involved listening to four sentences spoken by a native speaker of Chinese, expressing fear, sadness, anger, and joy, and identifying the emotion (first part); then, the subjects were asked to match the recordings to translations of the respective texts (second part).

Results: Group M generally performed better than Group H, but the differences were statistically significant only in certain comparisons. The overall performance in the first task was much lower than in the second in both groups. Microgenetic theory can be applied to interpret the results.

Conclusions: The task of researching emotional prosody is methodologically difficult, and requires interdisciplinary cooperation. The difficulties posed by the first task for both groups strongly imply that at least some aspects of emotional prosody are language-specific.

Key words: right hemisphere, aprosody, nonverbal communication

INTRODUCTION

Despite the fact that humans are the only animal able to use language, even we, when displaying sudden or strong emotions, often revert to non-verbal signals produced by the voice. These signals consist of various kinds of sounds (for example, groans and shouts), which are generically similar to those used by the higher primates (Bostanov & Kotchoubey, 2004). The prosodic characteristics of these cries are roughly similar in humans and chimpanzees. For example, in both humans and monkeys, a cry of fear is characterized by very high sounds, which are intended to reach as far as possible. The tone is high because all the muscles in the entire body, including the speech apparatus, are tensed (Etcoff, 1989). Even lower mammals, however, can recognize and react to the prosody of a human utterance: dogs, for example, may recognize many words, but often react to them only when they are spoken with the appropriate prosody.

Unfortunately, the term "prosody" is not frequently encountered in recent neuropsychological literature. The subject is either not addressed at all, or it is discussed using other terms. For example, in a well known book by Dale G. Leathers and Michael Eaves (2008) on the subject of nonverbal communication, the word "prosody" is not used even once, but is replaced by various other terms, such as "the semantics of sounds." The elements they mention, which are in fact the constituent elements of prosody, are as follows:

- volume, measured in decibels;
- pitch, that is, the melody of speech;
- tempo;
- duration, that is, how long we speak between pauses;
- tone, which among other things enables us to distinguish one person's voice from another's;
- rhythm;
- silence, which is one of the more important aspects of speech, but one that is frequently overlooked.

According to Herzyk, emotional prosody is "the appropriate modulation of the voice, used to express the speaker's attitude towards the linguistic content of the utterance" (Herzyk, 2005:232). She mentions three functions of emotional intonation:

- to eliminate linguistic ambiguities in neutral contents;
- to reinforce meaning;
- to express irony, that is, an emotional valence contrary to the literal meaning of the utterance.

P. S. Myers (2001) also states that the prosodic characteristics of speech convey both emotional and linguistic information. Emotional prosody involves not only emotional states, but also intentions, such as sarcasm or persuasion. She points out that patients with damage to the right cerebral hemisphere often display problems in the evaluation and expression of emotional prosody. Indeed, as early as the latter half of the 19th century, John Hughlings Jackson (cited by Herzyk, 2005) observed that the right hemisphere plays the dominant role in the regulation of what he called "emotional speech." He arrived at this conclusion on the basis of observation of patients with motor aphasia, who, despite the fact that they often mispronounce words in ordinary speech, can often say the same words without difficulty when speaking under the influence of strong emotions. Jackson's observations have been confirmed by numerous studies since his time (Herzyk, 2005). Research by Haggard and Parkinson (1971) and Ley and Bryden (1982) demonstrated that the left ear is dominant for the recognition of emotional prosody.

The most systematic work in this respect, however, has been done by E. Ross (1981, 1984, 2000), who argues that prosody is organized in the right hemisphere of the brain in a manner analogous to the organization of speech in the left hemisphere. According to Ross, it is possible to distinguish specific types of prosodic disturbances, analogous to the main clinical types of aphasia: motor, sensory, transcortical motor, transcortical sensory, mixed transcortical, conduction, and anomia. Although Ross admits that the existence of the latter two forms is purely theoretical, the rest, he maintains, can be and have been identified clinically in patients with right hemisphere damage (RHD; see Gorelick & Ross, 1987). Ross also expanded the concept of prosody to include much of what other authors call "non-verbal communication."

In more recent years, however, there has been considerable critical commentary on Ross's classification of aprosodic syndromes (Bautista & Ciampetti, 2003; cf. Pachalska & MacQueen, 2006). Herzyk (2005) points out that the primary disagreement is over whether aprosody is a disturbance of a language-related system based in the right hemisphere, or a primarily emotional disturbance. In other words: do patients with RHD have difficulty with expressing or understanding emotional prosody, or with feeling the emotions themselves? Both positions have found some support in clinical research (Heilman et al., 1975; Herzyk, 2005), which makes it difficult to simply ignore Ross's theory. As is generally known, the nature of hemispheric asymmetry remains difficult to conceptualize, and continues to generate controversy among neuropsychologists.

One major problem with interpreting the clinical data on RHD patients is that relatively little is known about the reception of emotional prosody in the healthy brain. Among other things, we do not know how much of prosody is language-specific. It seems obvious that some elements of non-verbal communication are biologically conditioned, such as laughter or weeping, while other elements are most likely part of linguistic competence in a given language. If prosody involves the "melodic" or "musical" aspects of speech, it would be interesting to know if the rules for forming and interpreting the prosody of an utterance are independent of the rules of grammar and vocabulary in a given language. Are there some rules of prosody that are universal in human language?

A complete answer to that question would require very extensive interdisciplinary research, far beyond the scope of the present study. In our research, we attempted to deal with a relatively small piece of this large question. Specifically, we wanted to determine whether and how a population of healthy students representing different fields of study could interpret prosodic information when listening to a recorded utterance in a language completely unknown to them.

MATERIAL AND METHODS

Material

Our research involved 50 student volunteers recruited from the University of Silesia in Katowice, Poland. The subjects (who were guaranteed complete anonymity) were divided into two groups:

- Group H – 30 students of humanities, including classics, Polish philology, Spanish, and cultural studies;
- Group M – 20 students of mathematics and computer science.

Group H consisted of 25 women and 5 men, while Group M consisted of 14 women and 6 men. The average age of the participants was 22.96 $\pm$ 3.00 years (range 19-42). The average age in Group H was 22.23 $\pm$ 1.36 years (range 19-24), whereas the average age in Group M was 24.05 $\pm$ 4.21 years (range 21-42). The age differences between the two groups were not statistically significant (t=2.07, p=0.083). Two students from Group H were left-handed, and the rest of the participants were right-handed. Because of the small number of subjects, neither handedness nor gender could be studied as variables.

Methods

In order to evaluate the subjects' reception of emotional prosody, we used a modified version of the Polish *Test Oceny Prozodii Emocjonalnej* (TOPE-1, Test of the Evaluation of Emotional Prosody, Pachalska & MacQueen, 2008). Our modified test is based on four sentences recorded by a native speaker of Chinese, expressing four different emotions (joy, sadness, anger, fear). In the first part of the test, the subjects (who had no knowledge of Chinese and were not told what the sentences meant) were asked to listen to the recordings (audio only, to factor out the visual aspects of nonverbal communication) and identify the dominant emotion in each sentence, in a word or two. For the second part of the test they were given Polish translations of the four sentences and asked to match them to the recordings, which were played again.

The four sentences were as follows:

- sentence 1 - "Help! Don't kill me! I'm scared!";
- sentence 2 - "Oh Mom! I miss you so much!";
- sentence 3 - "Oh! I'm so glad I'm going to be a doctor";
- sentence 4 - "You cheater! I saw you! I'll kill you both!".

For purposes of quantitative analysis point scores were assigned to the answers: 2 points for a correct identification of the emotion, 1 point for an answer that was not strictly correct but indicated the correct valence (positive or negative), and 0 points for a completely wrong answer. In the second part, the participants received scores of 1 or 0.

Our statistical analysis was based on single-factor ANOVA for the first part of the test, and Kendall's correlation co-efficient (tau) for the second part. In addition, a qualitative analysis was performed. The level of statistical significance was assumed to be $p<0.05$.

RESULTS

Quantitative analysis

Table 1 presents the average scores, standard deviations, and medians for each of the two groups separately, and for the whole research group together.

A statistical analysis using ANOVA for each of the four sentences in the first part of the test is given in Table 2.

There are statistically significant differences between the two groups, in favor of Group H, in the case of sentence 1, which expresses sadness ($p=0.056$), and sentence 2, which expresses fear ($p=0.021$). In the case of sentence 3, which expresses anger, the students in Group H were more accurate, but the difference between the two groups was not statistically significant. In the case of sentence 4, which expresses joy, students of mathematics and computer science (Group M) were more accurate, and the difference was almost significant ($p=0.061$).

Since the results from the second part of the test were not parametric, the statistical analysis had to be done separately, using Kendall's correlation coefficient (tau).

Fig. 1 shows the analysis of Kendall's correlation coefficient for Group M on the second part of the test.

As can be seen in Fig. 1, on the second part of the test Group M achieved maximum scores on sentence 1, and near maximum scores on sentence 4. In sentences 2 and 3, on the other hand, they made more errors in matching the

Table 1. Mean scores, standard deviations, and medians for each group separately and both groups together, on both parts of the TOPE-1, for each of the four sentences

Group	Sentence	Part 1				Part 2			
		1	2	3	4	1	2	3	4
M	Mean	0.0	0.8	0.9	1.5	1.0	0.5	0.5	0.9
	Standard deviation	0.00	0.77	0.97	0.83	0.00	0.51	0.51	0.31
	Median	0	1	0	2	1	0	0	1
H	Mean	0.2	0.3	1.0	1.0	0.9	0.5	0.5	0.9
	Standard deviation	0.38	0.61	0.96	0.95	0.25	0.51	0.51	0.25
	Median	0	0	1	1	1	0	0	1
All	Mean	0.1	0.5	1.0	1.2	1.0	0.5	0.5	0.9
	Standard deviation	0.30	0.71	0.96	0.93	0.20	0.50	0.50	0.27
	Median	0	0	1	2	1	0	0	1

Table 2. Results of single-factor ANOVA for the four sentences in the first part of the TOPE-1

Source of variance	SS	Df	MS	F	p value	Test F
Sentence 1						
between groups	0.333333	1	0.333333	3.84	0.055864	4.042652
within groups	4.166667	48	0.086806			
Sentence 2						
between groups	2.613333	1	2.613333	5.736585	0.020565	4.042652
within groups	21.86667	48	0.455556			
Sentence 3						
between groups	0.213333	1	0.213333	0.228742	0.63463	4.042652
within groups	44.76667	48	0.932639			
Sentence 4						
between groups	3	1	3	3.692308	0.060612	4.042652
within groups	39	48	0.8125			

translation to the original Chinese sentence. The differences between the accuracy of answers in sentences 1 and 4 as compared to sentences 2 and 3 were statistically significant (Kendall's tau=0.3773; ave.r.rank=0.3445; p=0.00005).

Fig. 2 shows the analysis of Kendall's correlation coefficient for the students of humanities (Group H) on the second part of the test.

As can be seen in Fig. 2, on the second part of the test Group H achieved near maximum scores on sentences 1 and 4. In sentences 2 and 3, on the other hand, like Group M, they made more errors in matching the translation to the original Chinese sentence. The differences between the accuracy of answers in sentences 1 and 4 as compared to sentences 2 and 3 were statistically significant (Kendall's tau=0.46667; ave.r.rank=0.44828; p=0.0000).

On the second part of the test, the differences between the groups were not statistically significant.

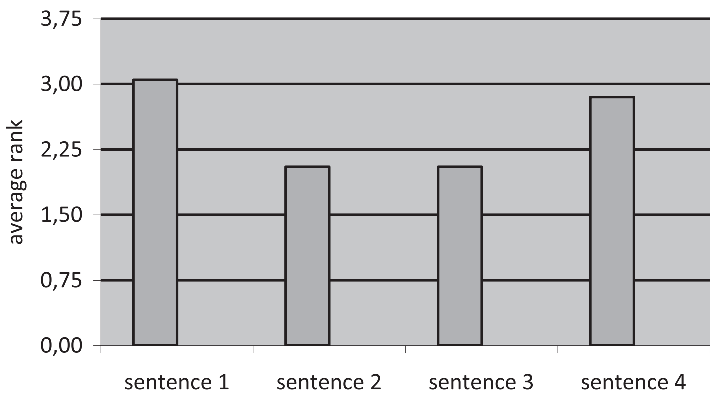

Fig. 1. Kendall's tau for Group M in the second part of the TOPE-1

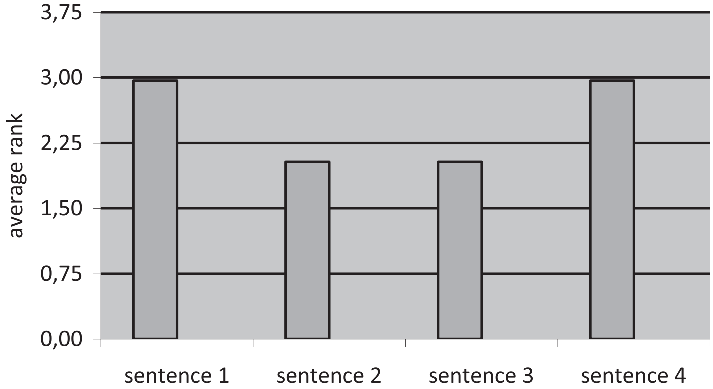

Fig. 2. Kendall's tau for Group H in the second part of the TOPE-1

Qualitative analysis

In the case of sentence 1 in the first part of the test, the students in Group M did not give the correct answer, which was "sadness," even once. The most common responses in this group were positive emotions, which means that the errors involved not only the name of the emotion, but also its valence. Among the wrong answers given by this group, the most common were "joy," "laughter," "satisfaction," and "surprise." None of these students named a negative emotion. In Group H, 8 persons mentioned "puzzlement" (1 additional person listed "amazement, embarrassment"), while 5 people answered "joy"; 4 people, "fascination"; and 2 people, "surprise." In contrast to Group M, however, 5 of the humanities students gave answers that, although they were not strictly accurate, were at least of the correct valence.

As can be seen, positive or neutral emotions dominated in both groups, while the sentence actually spoken on the recording was full of sadness.

In the case of sentence 2, the students in Group M gave 5 correct answers ("fear"). The most common incorrect answers were "upset" (5 students) and "anger" (4 students). Only 2 persons in Group H gave the correct answer, while 6 persons answered "excitement" and 5 answered "upset."

For sentence 3, 6 students from Group M gave the correct answer ("anger"). Three others identified the emotion as "fear," and another 2 thought they heard "resentment." In Group H, 4 students gave the correct answer, while 9 said the speaker was "upset," 4 identified "fear," and 2 each mentioned "sadness," "puzzlement," "resentment," and "irritation."

In the case of sentence 4, there were 14 correct answers in Group M, while 4 people answered "surprise" or "puzzlement." In Group H, 13 answered correctly. Five persons thought the emotion was "puzzlement," while 3 answered "admiration" and 3 answered "surprise." Two other persons answered "satisfaction."

The students from both groups found the second part of the test far easier than the first. The students in Group M most often switched “fear” and “anger” (5 persons). The same mistake was made by 14 people in Group H; in addition, 2 confused “sadness” with “joy,” 2 confused “fear” with “sadness” and 2 confused “joy” with “anger.”

Considering the entire study population, without regard for field of study, each successive emotion was better recognized than the preceding ones. This may result from the fact that the subjects had to get used to the speaker’s voice and the foreign language. In the second part of the test, the opposite emotions of sadness and joy were better distinguished than fear and anger, which have the same negative valence. This may have resulted, however, from the use of cognitive strategies by the students, rather than a direct reaction to the emotional prosody.

As noted above, in the first part of the test, no one from either group correctly identified the emotion expressed in sentence 1. Oddly enough, in the second part of the test, both groups did very well on this same sentence. The reason for this may be, as previously suggested, that the students in dealing with the very first task on the test were not mentally prepared for the “exotic” language they were hearing, while the second part of the test was constructed in such a way that some students could guess the correct answer using cognitive strategies. None of the students in Group M identified the emotion in sentence 1 as negative. Most of the students in Group H, on the other hand, identified the emotion as neutral, and 5 of them named a negative emotion, though it was not the correct one.

In the second part of the test, as previously noted, there were many fewer errors. Interestingly, many students in Group H (almost all of whom were language students) failed to notice that the Chinese text they were asked to match actually contained the word “mama” at the beginning. It is possible that, as students of language, they may have considered the sound of the word a mere coincidence, resisting the facile conclusion that “mama” means the same in every language.

In sentence 2, there were only 5 correct answers in Group M, but many of the wrong answers were also negative. In this case the students were probably reacting to the speaker’s loud voice and the fact that the pitch of the utterance rose considerably at the end. However, this same characteristic seems to have misled the humanities students into interpreting fear as excitement. Even so, half of the group gave an answer with the correct valence.

DISCUSSION

As previously mentioned, there is not a great deal of literature on the subject of emotional prosody. Barra et al. (2006) used recordings of several words and sentences differing in emotional expression and read by a male actor. The actor’s voice was mechanically distorted, so that the subjects could not understand the meaning of the words (the recordings were in Spanish and played for native speakers of that language). On the basis of their results, the authors stated that the most easily recognized emotion was anger, which in our study was identified moderately well in the first part of the test, and very well in the second part. The second most easily identified emotion in the Spanish study was surprise. Barra et al. classified that emotion as "complex," meaning that they did not assign it positive, negative or neutral valence. In our study surprise was mentioned very often in the first part of the test, particularly in the sentences expressing sadness and joy, even though none of the four sentences we used actually expressed surprise. Clearly there is a readiness to identify surprise, which may be biologically conditioned (surprise expressed by one member of the group may serve as an alarm signal to other members).

The third most easily identified emotion in the Spanish study was sadness, which in our study was the worst recognized in the first part, but almost perfectly recognized in the second part. Joy, on the other hand, was the hardest to recognize for the Spanish subjects, while our subjects obtained the best results for joy. These differences may result from the fact that we used a language unknown to any of our subjects, which made it unnecessary to distort the voice in order to filter out purely semantic information. At the same time it seems clear that there are significant cultural differences that have affect on emotional prosody.

The fact that many persons in the first part of the test correctly identified the valence of the emotion, even when they did not give the right answer, suggests that the emotions expressed in prosody were recognized in stages. In the first stage the subjects classified the speaker's voice to one of three categories: positive, negative, or neutral. In the second phase, the subject decided how strong the emotion was in the speaker's voice. In the final stage, the emotion was analyzed and named.

This phenomenon can easily be explained in terms of microgenetic theory (Brown, 1988; 1996; 2000). When the sound of a voice reaches the ear, the signals are first routed to the thalamus, where they trigger the primitive fight-or-flight reaction in the brainstem; in other words, the primitive affect in the receiver reflects the affect of the speaker, interpreted according to primitive categories; these, of course, are closely tied to survival, the primary criterion of value at this microgenetic level (Brown, 2005). The transition to the limbic system brings the pleasure principle into play, and the attraction-repulsion mechanism: we either like or dislike what we hear in the speaker's prosody, though at this stage we cannot give a rational account of the reasons for the liking or disliking. The task of naming the emotion that seems to be conveyed by the speaker's prosody is of course a cognitive one, which means that it belongs to the neocortex (with right hemisphere structures presumably dominant, though the actual naming is a lexical, left-hemisphere task). Language specificity is not a factor at the brainstem level, but begins to be important already at the limbic level, and is essential at the cortical level - which is consistent with our findings.

Prosody thus provides a particular interesting field for future research grounded in microgenetic theory.

CONCLUSIONS

We set out to determine, among other things, whether students of mathematics and computer science better recognized the emotions expressed by the speaker than did humanities students. This proved to be the case. The students of mathematics and computer science performed better on both parts of the test, although the differences were not statistically significant on the second part. If we look at each sentence separately, the students of mathematics and computer science outperformed the humanities students in all of the first part and in all but the last sentence of the second part. However, if we consider only valence, this difference emerges only in the case of the 2nd and 4th sentence.

We also wanted to know which emotions were more easily recognized. In our results joy was clearly easier to recognize in the first part, and both sadness and joy in the second part.

All of the subjects found the second task (matching the voice to a translation of the text) much easier than the first task, which required them to identify the emotion with no verbal cues. The relatively weak performance of all students on the first part of the test does not support the hypothesis that emotional prosody is universal, that is, independent of language and culture. However, identification of the valence of the emotion seems to be an easier task in an unknown language; but even so, the performance was surprisingly weak, especially in the case of the humanities students.

The course of our study indicates that research on emotional prosody, though it may be of crucial importance to neuropsychology (especially in the case of right hemisphere damage) is methodologically very difficult and requires interdisciplinary cooperation in both the collection and interpretation of data.

REFERENCES

Barra, R., Montero, J.M., Macias-Guarasa, J., D'Haro, L.F., San-Segundo, R. & Cordoba, R. (2006). Prosodic and segmental rubrics in emotion identification. Paper delivered during the International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, 14-19 May 2006. Available on-line at: http://www-gth.die.upm.es/research/documentation/AG-39Pro-06.pdf; DOA 24 April 2012.

Bautista, R.E.D. & Ciampetti,M.Z. (2003). Expressive aprosody and amusia as a manifestation of right-hemisphere seizure. *Epilepsia*, *44(3)*, 466-467.

Bostanov, V. & Kotchoubey, B. (2004). Recognition of affective prosody: continuous wavelet measures of event-related brain potentials to emotional exclamation. Psychophysiology, 44, 259-268.

Brown, J.W. (1988). *The life of the mind. Selected papers*. Hillsdale, New Jersey, Hove and London: Lawrence Erlbaum.

Brown, J. W. (1996). *Time, will and mental process*. New York: Plenum.

Brown, J.W. (2000). *Mind and nature: essays on time and subjectivity*. London: Whurr Publishers.

Brown, J.W. (2005). Process and the authentic life: toward a psychology of value. Heusenstamm, Germany: Ontos Verlag.

Etcoff, N.L. (1989). Asymmetries in recognition of emotion. In: F. Bolier & J. Grafman (eds.), *Handbook of Neuropsychology*, vol. 3 (363-382). Amsterdam: Elsevier.

Gorelick, P.B. & Ross, E.D. (1987). *Journal of Neurology, Neurosurgery & Psychiatry, 50,* 553-560.

Haggard, M.P. & Parkinson, A.M. (1971). Stimulus and task factors as determinants of ear advantages. Quarterly Journal of Experimental Psychology, 23(2), 168-177.

Heilman, K.M., Scholes, R. & Watson, R.T. (1975). Auditory affective agnosia: disturbed comprehension of affective speech. *Journal of Neurology, Neurosurgery & Psychiatry, 38(1)*, 69-72.

Herzyk, A. (2005). *Wprowadzenie do neuropsychologii klinicznej*, Wykłady z Psychologii. Warsaw: Scholar.

Leathers, D.G. & Eaves, M.H. (2008). *Successful nonverbal communication: principles and applications.* Boston: Allyn & Bacon.

Ley, R.G. & Bryden, M.P. (1982). A dissocation of right and left hemispheric effects for recognizing emotional tone and verbal content. *Brain & Cognition, 1(1)*, 3-9.

Myers, P.S. (2001). Communication disorders associated with right hemisphere damage. In: R. Chapey (ed.), Language intervention strategies in aphasia and related neurogenic communication disorders (pp. 809-828). Baltimore: Lippincott, Williams & Wilkins.

Pąchalska, M. & MacQueen, B.D. (2006). Komunikacja interpersonalna a poczucie tożsamości. In: B.L.J. Kaczmarek, A. Kucharski & M. Stencel (eds.), Współczesne trendy w komunikowaniu się (pp. 107-135). Lublin: UMCS.

Pąchalska, M. & MacQueen, B.D. (2008). Test Oceny Prozodii Emocjonalnej - 1 (TOPE-1). Cracow: Fundacja na Rzecz Osób z Dysfunkcjami Mózgu.

Ross, E.D. (1981). The aprosodias: functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology, 38(9)*, 561-569.

Ross, E.D. (1984). Right hemisphere's role in language, affective behaviour and emotion. *Trends in Neurosciences, 7*, 342-346.

Ross, E. (2000). *Affective prosody and the aprosodias*. In: M.-M. Mesulam (ed.), *Principles of behavioral and cognitive neurology* (pp. 316-331). Oxford: Oxford University Press.

Address for correspondence:

Prof. Bruce D. MacQueen
Katedra Logopedii
Uniwersytet Gdański
ul. Wita Stworza 58
80-956 Gdansk
e-mail: bdmacqueen@gmail.com