

Speech rate in the expression of anger: a study with spontaneous speech material

Miguel Oliveira, Jr.¹, Ayane Nazarela Santos de Almeida¹, René Alain Santana de Almeida¹, Ebson Wilkerson Silva¹

¹ Faculdade de Letras, Universidade Federal de Alagoas, Brazil

miguel@fale.ufal.br, ayanesantos@hotmail.com, renealain@hotmail.com, ebswillk@gmail.com

Abstract

The study of the acoustic expression of emotion is, in general, the analysis of whether prosodic variables such as intonation (F0), speech rate, pauses, rhythm, intensity and duration, are reliable clues for the characterization of the emotional states of the speaker. The present paper aims to verify whether an association exists in Brazilian Portuguese between the basic emotion of “anger” and the prosodic variable “speech rate”, as the literature often suggests there is for other languages. The corpus consisted of fragments of spontaneous speech recorded from a radio program. The fragments were selected on the basis of a perceptual test. For the production analysis, only excerpts that were identified by more than 75% of the participants of the perceptual test as associated to the categories “anger” and “neutral” were selected. The results demonstrated that, for the data that were used for the analysis, there is a general reduction in speech rate when utterances are associated with the emotion of “anger”, if compared to utterances spoken in a “neutral” mode by the same speaker, contrary to what literature often indicates for other languages.

Index Terms: speech rate, anger, spontaneous speech

1. Introduction

1.1. Speech and emotion

There are several ways in which human emotions can be expressed at the time of communication: through the facial expression of the speaker, the verbal content of the utterance and the acoustic features of the speech, observed through the behavior of prosodic parameters such as intensity, duration, fundamental frequency (F0) and voice quality.

All these parameters collaborate to speech intelligibility and promote the understanding of the expression of emotions in human communication. [1], however, consider that the prosodic characteristics alone may warrant the recognition of emotions expressed by humans.

According to [2], real-world applications, notably those based on human-computer interaction, depend on coming to terms with the ways people express emotion. In order to contribute to the improvement of recognition and synthesis speech systems, [3] found to be of vital importance to speech technology to determine the vocal changes produced by emotional factors in various languages and cultures. Thus, [3] has analyzed the emotion and speech recognition in nine countries from three different continents and found that segmental and suprasegmental aspects, in each of the languages that were analysed, contribute significantly to the production and the perception of the emotional categories.

The classification and the concept of emotion are, however, quite controversial in the literature [4]. [5] asserts that the definitions of emotion are as varied as the researches that address it. The classification of emotions is particularly

controversial because of its spontaneous, involuntary and unpredictable feature.

Despite the lack of general agreement in the literature, there are common basic aspects that several papers share on the subject, such as the fact that emotions are almost always directed to the object, that they are activated by internal or external stimuli and that they consist of momentary states of people [5]. It also seems to be a consensus among the researchers that emotion can be broken down into three basic categories: sadness, happiness and anger ([5], [6]).

Each of the emotions, as classified by the literature, is associated with specific prosodic features that individualizes them. [5], [7] and [8], in studies of the Dutch, French and German languages respectively, sought to associate emotional patterns in these languages with acoustic parameters such as pitch levels, vocal range, global average intensity and total utterance duration. In all of these studies, the authors observed that sadness is generally characterized by a lower F0, a lower voice and a slower rate, whereas anger and joy have a higher F0, a louder voice and a faster speech rate.

[9], on the other hand, points out that not always the speakers of a language expresses emotions in the same way with the same activation levels. Then, each of the basic emotions can generate different levels of activation in the speaker, depending on its manifestation. For example, [5] observed that irritation causes a low level of activation in speaker, while fury affects the speaker more intensely.

[10] asserts that there may be very subtle differences in the way basic emotions are manifested for individual speakers. The author, however, does not make reference to different emotions, but to subtypes of the same emotion. In that way, [10] considers "fury" (hot anger) and "irritation" (cold anger), for example, as being two different manifestations of the same emotion: anger.

According to [2], there are inconsistencies for several emotions and emotion-related states that have been studied. These, however, may reflect inconsistent procedure or different interpretation of emotion categories, or variations in terms of the type of data that was used in the analysis: real or simulated. Others, though, seem likely to reflect real differences in the vocal expression of emotion, from speaker to speaker, from culture to culture, and across genders and situations. [2] points out that comparisons between languages and cultures are limited, but they suggest substantial differences.

There are still very few studies on prosodic characteristics of emotions and emotion-related events in Brazilian Portuguese. [9], for example, investigated how the expression of three primary emotions affects the characteristic melodic contours of four speech acts in Brazilian Portuguese spoken in Rio de Janeiro. In order to test that, the author used an interaction between four speech acts (assertion, question, request an order) and four emotions (three primary emotions: sadness, happiness and anger, plus the “neutral” form). She

intended to answer whether these two categories – speech acts and emotion – are really independent prosodic dimensions or whether the interaction between them causes substantial changes in intonation patterns that's been found for speech acts alone.

The author found that speech acts and emotional patterns appear to be independent categories in terms of production. With regard to the perception, however, there is some overlap between the two categories. She concludes that intonation and voice quality should be seen as complementary categories, both necessary for the recognition of emotional states of the speaker. It must be pointed out, however, that [9] analyzed utterances produced by two actors from the southeastern area of Brazil. Most of the few studies addressing the prosody of emotions in Brazilian Portuguese is based on the southern and southeastern dialects, and uses monitored utterances produced by actors.

[11] presented a method for speech expressiveness, which combines a dimensional analysis of speech expression, a Principal Component Analysis technique, as well as a multiple regression analysis. The author used as the corpus for his study recordings of a radio show called Programa do Chupim, aired by Rádio Metropolitana de São Paulo, based on a southeastern area of Brazil. He concluded that if utterances are analyzed chronologically, they reveal clear expression changes, from an automatic acoustic analysis, what implies that acoustic parameters are sufficient for the detection of emotion in speech.

[12] analyzed the prosody and expressivity of speech through emotions expressed in a poem spoken by a professional actor from southeastern Brazil. [13], investigated the melodic contour associated to the emotion of anger in theatrical speeches, based on samples of three professional actresses from southeastern Brazil, collected in laboratory condition. [14] also analysed monitored speech with the goal of improving synthesized speech produced by a TTS (text to speech) system, by adding to it emotional information from an acoustic perspective. The model used by them was based on a southeastern dialect of Brazilian Portuguese.

It is, however, unclear to what extent monitored speech truly reflects an emotion that is characterized as involuntary and unpredictable. This is, actually, one of the most important limitations in studies of emotional speech [15]. In this direction, [16] emphasized that high levels of activation are not often found in the case of elicited emotions from control conditions.

1.2. Speech rate

One of the very first references to empirical research on speech rate dates back to [17], who summarizes his findings as follows: "it takes about twice as long to read (aloud, as fast as possible) words which have no connexion as letters which make words... When a passage is read aloud at a normal rate, about the same time is taken for each word as when words having no connexion are read as fast as possible." This study dealt with several languages, such as English, French, German, Italian, Latin and Greek, as he was also concerned with the different rate employed by the same speaker while speaking a foreign language. According to [17], the rate at which someone speaks a foreign language is determined by the familiarity that the person has with the language: the more familiar a person is with it, the faster the speech.

[18] was perhaps the first to introduce objective methodology for the study of rate in speech. Using a rather

peculiar method of measurement that relied on soot marking from flames, he compared polysyllabic words with monosyllabic words, using the syllable per second unit of measure – a unit most widely employed today. In his experiment, he demonstrated that, for a given passage containing the same amount of syllables, polysyllabic words are read faster than monosyllabic words. He claims that this is due to the amount of meaning that is carried out in passages containing monosyllabic syllables: the larger the amount of meaning that is conveyed in a message, the longer the speaker will take to utter it.

Meaning is also considered to be a determining factor for the establishment of speech rate in [19]. It compared the repetition of nonsense syllables with the production of syllables articulated within words and found that nonsense syllables are often produced at a slower rate than those syllables that are part of real words. He also noted that people tend to be affected by curiosity when reading a passage of unknown content, which would result in a deliberate acceleration of the speech as a result of this curiosity.

[20] introduced developmental considerations to the study of speech rate, by investigating the speech of kindergarten children. They found that while boys tend to speak less than girls, they do so in a faster rate. [21] considered other external factors. According to these authors, emotions such as anger, fear, and indifference are closely related to fast speech rate, whereas contempt and grief are associated to slow rate.

[22] confirmed the hypothesis that meaning has a decisive influence on speech rate, by demonstrating that the more meaning there is invested in an utterance, the slower the articulation rate is. According to him, this has also to do with the emotional state of the speaker: psychological tension – a result of the demanding task of interpreting new meanings – would be the most immediate reason for speech rate variation.

On the basis of the overview presented above, it may be concluded that there is a historical tradition in the studies on speech rate to relate its observed variation to both semantic and emotional aspects associated to different speech activities. The study on speech rate, as an acoustic correlate of emotion, is still incipient for Brazilian Portuguese, as the literature review presented above indicated. Most of the few studies carried on so far focus on intonational characteristics in speech expressiveness.

One of the few studies for Brazilian Portuguese that addresses the acoustic parameter of speech rate as a correlate of emotion is that of [9]. According to her findings, there isn't a straightforward relation between duration and different basic emotions: duration may vary as a function of individual speakers. (See also [23]).

1.3. Research question

The main objective of the present study is to investigate whether speech rate is a reliable acoustic correlate, in Brazilian Portuguese, of a basic human emotion: anger. The rationale for this study lies in the scarcity of such investigation, despite the fact that the literature often relates the acoustic parameter of speech rate to human emotion ([5], [7], [8], [9]).

Contrary to most of the very few studies on the prosody of emotion in Brazilian Portuguese, the present investigation uses spontaneous material, representative of a dialect that has been neglected so far in this kind of investigation: that spoken in the northeastern area of Brazil.

2. Methods

The data used in this study consist of eighteen small fragments of recordings of spontaneous speech, extracted from a radio program called “A Hora do Mução”. It is a popular practical joke program, in which the radio broadcaster Rodrigo Vieira Emerenciano embodies the character Mução. The radio broadcaster calls people in the northeastern area of Brazil, from suggestions made by friends and relatives of the victim, in order to annoy them by constantly making reference to a physical characteristic or a nickname of the victim. One of the advantages of this kind of interaction, in radio program, is the possibility of obtaining high levels of arousal of affective responses, due to the critical events introduced by the program presenter.

In the recordings, there is a clear difference between the speech of the called parties in the first half of the phone calls and their second half, when the reason for the calls is known (i.e. the practical joke is in action). In the second half of the calls, the speech of the called party is clearly linked to the expression of anger.

The recordings of this radio program are, therefore, an excellent material for the type of analysis proposed here because they make it possible to compare, for each individual speaker, prosodic features associated with neutral speech to those associated with the expression of anger.

Furthermore, contrary to what happens with most of the data used in research about the relationship between prosody and emotion, the recordings used in the present investigation reflect the speech associated with a spontaneously occurred emotion, what makes it ecologically valid. It must also be pointed out that telephone speech is ideal for the study of vocal signs of emotion, because contrary to face to face interactions, it doesn't present any visual information that could distract the analysis ([2], [24]).

The excerpts of the recordings were selected from a perceptual test conducted with forty graduate students in Arts at Federal University of Alagoas. The data that was used in this perception test consisted of fragments of recordings available on CD “Pegadinhas do Mução” [25], which were selected based on the following criteria: (i) sound quality (intelligible and noiseless signal), (ii) sex the speaker (all excerpts were extracted from recordings featuring men only, in order to avoid the inclusion of a variable that could confound the interpretation of the results), and (iii) content of lexical information (the extract were selected by taking into account their contents: for obvious reasons, they shouldn't offer any evident clues of anger – or any other emotion, for that matter). Regarding this last criterion, however, it must be pointed out that previous studies have demonstrated that prosody processing is segregated from linguistic semantic processing, probably because both depend on partially dissociated neural mechanisms ([26], [27]).

We understand that the corpus is reduced, however, the difficulty in obtaining data from spontaneous speech justifies this limitation. Moreover, we could not use excerpts of the recordings containing insults and in the case of emotion “anger”, this occurs very often.

Transcriptions of the selected excerpts from recordings of three practical jokes, made with three different men, all speakers of northeastern dialects, along with their corresponding audio files were presented in slide-show, at random, to the participants of the perceptual test. For each excerpts, the participant had to identify an emotion related to

it. The options were the three basic emotions: “anger”, “happiness” and “sadness”, as well as “neutral”. The participants were instructed to listen to each stimulus as often as they wished before answering the form.

For the production analysis, only excerpts that were identified by more than 75% of the participants of the perceptual test as associated to the categories “anger” and “neutral” were selected.

There is a variety of units of measurement that are employed in the research on speech rate. The units range from sounds per unit of time ([22], [28], [29], [30]), to words per unit of time ([31], [32], [33], [34], [35]), to syllables per unit of time ([34], [36], [37]) and finally to beats per unit of time [35]. This plethora of units of measurements employed in the literature not only reflects a serious methodological flaw – as discussed in [37], but also makes the essential task of comparing results among various studies impossible.

The best-suited unit of measurement for speech rate, according to many authors, is the syllables per unit of time ([38], [39]). [36], for example, defines speech rate as the “rate of syllable succession.” This is the unit adopted by [34], [38], [40], [41], [42], [43], [44] and [45], to name a few. Even though, as [39] points out, this unit of measurement has also the disadvantage of not taking into consideration the processes that may result in syllable omission, that are often found in rapid speech, such as assimilation and segmental deletion, what would obviously not be covered in this unit of measurement.

Therefore, the present study opted for a measure that is mostly used in the temporal research of speech for the sake of comparability. It does recognize the pitfalls related to this choice, but assumes that they are not so serious as to invalidate the analysis. Speech rate will be interpreted in this study using the measurement of syllables per second.

3. Results

Figure 1 below shows average speech rates of the speech samples that were selected for analysis, based on the perceptual test. It is broken down by emotion (“neutral” and “anger”) and speaker (a total of three).

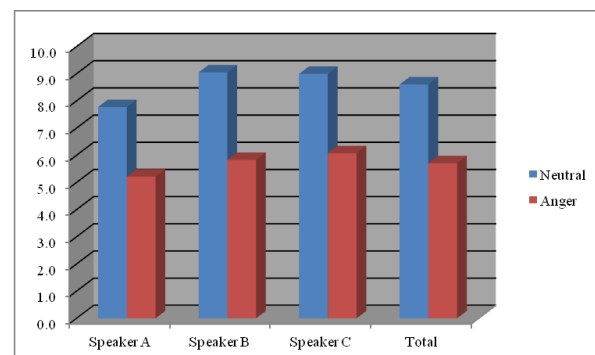


Figure 1: Average speech rate (in syllables per second) of “neutral” and “anger” utterances, broken down by speakers.

We used the mean for comparison between utterances associated with each emotion because we found that the difference between it and the median was insignificant, which allows us to state that there is no significant distortion in the mean rates found in analyzed excerpts and, therefore, can be considered a reliable measure for this data.

The results shown in Figure 1 clearly indicate that speakers employed a slower rate of speech in utterances associated to the emotion of anger, if compared to those considered to be neutral, both individually as well as a group. Speech rate of the excerpts labeled as “neutral” was, in average, 3 syllables per second faster than those that were labeled as associated to the emotion of anger.

A series of paired-samples two tailed t-tests were conducted to assess the significance of the differences in speech rate between “neutral” and “anger” expressions for each speaker and for the total sample. Table 1 below shows the results of these tests:

Table 1. Results of t-tests, broken down by speaker and by total sample.

Sample	Results
Speaker A	[t(5) = 7.02, $p < 0.0009$]
Speaker B	[t(5) = 11.08, $p < 0.0001$]
Speaker C	[t(5) = 6.77, $p < 0.0011$]
Total sample	[t(17) = 13.86, $p < 1.08E-10$]

Table 1 evidences that for all speakers, speech rates differ significantly between the two conditions: “neutral” and “anger”. It also indicates that the differences are significant when subjects are regarded as a group. Results in Table 1 show a low probability of the results being a fortuity, as indicated by the values of p .

4. Discussion

This study set out to investigate whether speech rate is a reliable acoustic correlate, in Brazilian Portuguese, of a specific basic emotion: “anger”. It followed the tradition of studies that explore how speakers use prosody to encode discrete emotions, such as happiness, anger, disgust, etc.

Emotional prosody in Brazilian Portuguese has been studied very scarcely so far. The few papers that address the issue deal almost exclusively with simulated emotion and are based on data representative of dialects spoken in the southeastern area o Brazil. In order to contribute with the still incipient research on emotional prosody in Brazilian Portuguese, the present investigation used spontaneous recordings representative of dialects spoken in northeastern Brazil.

According to the literature, the emotion of “anger” is characterized acoustically by fast speech rate, high voice intensity, high F0/pitch level, much F0/pitch variability, rising F0/pitch contour, fast voice onsets, and microstructural irregularity ([5], [7], [8], [9]).

What the general results from the analysis of speech rate reported here suggest is that there are possible dissonances between what the literature establishes as a prosodic pattern associated with the emotion of “anger” and what has been identified as a pattern for the same emotion in Brazilian Portuguese, based on spontaneous speech material. The results demonstrate that there is a general reduction in speech rate when utterances are associated with the basic emotion of “anger”, if compared to utterances spoken in a “neutral” mode by the same speaker. Statistical analysis indicate that this difference is not a fortuity or due to individual characteristics of the speakers, because the samples of each speaker were

analyzed individually, as well as a group, and very similar characteristics with regards to this speech rate’s variable in the expression of anger was found.

It is known that emotions are manifested in a continuous way, in varying dimensions, depending on the level of activation / stimulation [9]. The emotion of “anger”, for example, can be expressed on a scale ranging from a mild “irritation” to a “wrath”.

In this paper, however, we didn’t consider the varied levels of activation, but rather the basic emotion “anger”. We are of course aware that, depending on where in the continuum a given emotion is located, its corresponding acoustic characteristic may vary [46]. Notwithstanding, faster speech rate is commonly associated to any level of activation in the scale of “anger” [10].

The difference in terms of speech rate as a correlate of the emotion of “anger” reported here, as compared to the results reported elsewhere, may be interpreted as a result of the specificities of data that were used for the analysis: spontaneous material, derived from telephone speech, uttered by male speakers of scarcely studied dialects of Brazilian Portuguese. As [2] points out, variations in this kind of study may reflect differences in the vocal expression of emotion, from speaker to speaker, from culture to culture, and across genders and situations. In order to find out whether this is systematically the case, further investigation needs to be done, with a larger corpus contemplating all possible variables that are potentially important for any attempt to generalize.

5. Conclusion

According to [47], communication of emotions is crucial to social relationships and survival. It is, thus, essential to understand all the aspects related to it, including its acoustic properties. It has been suggested, however, that signs of emotion in speech is not consistent across individuals and occasions, what calls for systematic descriptions of their properties for proper comparisons between languages and cultures.

The analysis presented here allowed us to infer that in spontaneous speech data representing the variety of Portuguese spoken in northeastern Brazil, there is not an association between the basic emotion of “anger” and an increase in speech rate, as the literature often indicates for other languages. Instead, many excerpts identified as being representative of the emotion “anger” by the participants of the perception test were enunciated with a speech rate slower than that in the excerpts perceived as “neutral”.

Contrary to most studies on the prosody of emotion, the present study opted to use spontaneous material for the analysis, because, as [10], we believe that if research continues to be almost exclusively concerned with the simulation of emotion, nothing much will be gained in terms of getting to understand to what extent we are operating within a closed system of association or translation rules that may be only indirectly tied to the underlying biology of emotion.

The present paper stands as a contribution to the acoustic characterization of emotional patterns in Brazilian Portuguese. We plan for future investigation to enlarge the corpus, as to reflect other variables, and to include other prosodic parameters in the analysis, such as variation in fundamental frequency, pause and intensity. Perceptual tests with stylized samples of the corpus are also planned, as we understand that it is fundamental to describe what prosodic parameters are relevant from a perceptual perspective.

6. References

- [1] Schröder, M. "Experimental study of affect bursts". *Speech communication special issue speech and emotion*, 40: 99-116, 2003.
- [2] Douglas-Cowie, E. et al. "Emotional speech: Towards a new generation of databases". *Speech Communication*, 40(1): 33-60, 2003.
- [3] Scherer, K. "A cross-cultural investigation of emotion inferences from voice and speech: Implications for speech technology". *Proceedings of 6th ICSLP*, 2: 379-382, Beijing, 2000.
- [4] Cowie, R. and Cornelius, R. R. "Describing the emotional states that are expressed in speech". *Speech Communication*, 40(1): 5-32, 2003.
- [5] Paeschke, A. "Prosodische Analyse emotionaler Sprechweise". Berlin: Logos Verlag, 2003.
- [6] Sawamura, K., Deng, J., Akagi, M., Erickson, D., Li, A., Sakuraba, K., Minematsu, N. and Hirose, K. "Common factors in emotion perception among different cultures". *Proceedings of 16th ICPhS*, Saarbrücken, 2113-2116, 2007.
- [7] Leon, P. "Précis de phonostylistique parole et expressivité". Paris: Natan, 1993.
- [8] Bezooyen, R. "Characteristics and recognizability of vocal expression of emotion". Dordrecht: Foris, 1984.
- [9] Pereira, M. C. C. "A expressão das emoções em atos de fala no português do Brasil: produção e percepção". *Dissertação de Mestrado em Letras Vernáculas*, Universidade Federal do Rio de Janeiro, 2009. Online: <http://www.letras.ufrj.br/posverna/mestrado/PereiraMCC.pdf>, accessed on 29 Jul 2013.
- [10] Scherer, K. "Vocal affect expression: a review and a model for future research". *Psychological Bulletin*, 99: 143-165, 1986.
- [11] Barbosa, P. A. "Detecting changes in speech expressiveness in participants of a radio program". *Proceedings of Interspeech*, 2155-2158, 2009.
- [12] Santos, I. "Expressividade da Fala: o desvelar da locução de um poema a partir da Análise Acústica e da filosofia de Spinoza". *Dissertação de Mestrado em Linguística Aplicada e Estudos da Linguagem*, Pontifícia Universidade Católica de São Paulo, 2010. Online: http://www.sapientia.pucsp.br/tde_busca/arquivo.php?codArquivo=12132, accessed on 13 Aug 2013.
- [13] Vassoler, A. M. O. and Martins, M. V. M. "A entoação em falas teatrais: uma análise da raiva e da fala neutra". *Estudos Linguísticos*, 42: 9-18, 2013.
- [14] Reis, B. F. and Martins, V. V. "Síntese prosódica da fala em português do Brasil". *X SBAL*, 1185-1188, 2011.
- [15] Swerts, M. and Hirschberg, J. "Prosodic predictors of upcoming positive or negative content in spoken messages". *J. Acoust. Soc. Am.*, 128(3): 1337, 2010.
- [16] Scherer, K. R. "Vocal communication of emotion: A review of research paradigms". *Speech Communication*, 40: 227-256, 2003.
- [17] Cattell, J. M. "The time it takes to see and name objects". *Mind*, 11: 63-65, 1886.
- [18] Beer, M. "Die Abhängigkeit der Lesezeit von psychologischen und sprachlichen Faktoren". *Zeitschrift für Psychologie*, 56: 264-298, 1910.
- [19] Fröschels, E. "Untersuchungen über das Sprechtempo". *Monatsschrift für Ohrenheilkunde und Laryngo-Rhinologie*, 54: 867-871, 1920.
- [20] Olson, W. C. and Koetzle, V. S. "Amount and rate of talking of young children". *Journal of Experimental Education*, 5: 175-179, 1936.
- [21] Fairbanks, G. and Hoaglin, L. W. "An experimental study of the durational characteristics of the voice during the expression of emotion". *Speech Monographs*, 7: 85-90, 1940.
- [22] Essen, O. V. "Sprech tempo als Ausdruck psychischen Geschehens". *Zeitschrift für Phonetik*, 3: 317-341, 1949.
- [23] Moraes, J.; Rilliard, A. "Prosody and Emotion in Brazilian Portuguese", in M. Armstrong, N. Henriksen, and M. Vanrell [Eds], *Interdisciplinary approaches to intonational grammar in Ibero-Romance intonation*, *Issues in Hispanic and Lusophone Linguistics*, to appear.
- [24] Paulmann, S. and Pell, M. D. "Is there an advantage for recognizing multi-modal emotional stimuli?" *Motivation and Emotion*, 35: 192-201, 2011.
- [25] Emerenciano, R. V. "Pegadinhas do Mução". Rio de Janeiro: Polysom. 1 disco compacto: digital, stereo, 2010.
- [26] Friederici, A. D. and Alter, K. "Lateralization of auditory language functions: A dynamic dual pathway model". *Brain & Language*, 89: 267-276, 2004.
- [27] Ross, E. D. and Monnot, M. "Neurology of affective prosody and its functional-anatomic organization in right hemisphere". *Brain & Language*, 104: 51-74, 2008.
- [28] Heinitz, W. "Die Bewertung der Dauer in phonetischen Aufnahmen". *Vox* 153, 1921.
- [29] Hegedüs, L. "Sprechtempoanalysen im Ungarischen". *Zeitschrift für Phonetik*, 10: 8-20, 1957.
- [30] Fónagy, I. and Magdics, K. "Speed of utterance in phrases of different lengths". *Language and Speech*, 4: 179-192, 1960.
- [31] Brubaker, R. S. "Rate and pause characteristics of oral reading". *Journal of Psycholinguistic Research*, 1(2), 1972.
- [32] Barik, H. C. "Cross-linguistic study of temporal characteristics of different speech materials". *Language and Speech*, 20: 116-126, 1977.
- [33] Grosjean, F. and Deschamps, A. "Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomènes d'hésitation". *Phonetica*, 14: 141-148, 1975.
- [34] Grosjean, F. and Deschamps, A. "Analyse des variables temporelles du français spontané". *Phonetica*, 26: 126-156, 1972.
- [35] Scollon, R. "Tempo, Density, and Silence: Rhythms in Ordinary Talk". Center for Cross-Cultural Studies. Fairbanks, University of Alaska, 1981.
- [36] Abercrombie, D. "Elements of General Phonetics". Edinburgh: Edinburgh University Press, 1967.
- [37] Meinhold, G. "Allgemeine Probleme der Sprechgeschwindigkeit". *Zeitschrift für Phonetik*, 25: 492-505, 1972.
- [38] Uhmann, S. "Contextualizing relevance: on some forms and functions of speech rate changes in everyday conversation", in P. Auer and A. D. Luzio, *The Contextualization of Language*. Amsterdam, Benjamins: 297-336, 1992.
- [39] O'Connell, D. C. and Kowal, T. D. "Cross-linguistic pause and rate phenomena in adults and adolescents". *Journal of Psycholinguistic Research*, 1: 155-164, 1972.
- [40] Blaauw, E. "On the perceptual classification of spontaneous and read speech". Research Institute for Language and Speech, Utrecht University, 1995.
- [41] Fon, J. "Speech rate as a reflection of variance and invariance in conceptual planning in storytelling". *Proceeding of the ICPhS*, 1999.
- [42] Grosz, B. and Hirschberg, J. "Some intonational characteristics of discourse structure". *Proceeding of the International Conference on Spoken Language Processing*, Banff, 1992.
- [43] Goldman-Eisler, F. "The rate of changes in the rate of articulation". *Language and Speech*, 4: 171-174, 1961.
- [44] Van Donzel, M. "Prosodic Aspects of Information Structure in Discourse". *Faculteit der Geesteswetenschappen*, Amsterdam, University van Amsterdam: 195, 1999.
- [45] Wood, S. "Speech tempo". *Working Papers of the Phonetics Laboratory*, Lund University, 9: 99-147, 1975.
- [46] Scherer, K. "Vocal measurement of emotion". In Plutchik, R.; Kellerman, H. (Org.). *Emotion: Theory, research, and experience. The measurement of emotion*, 4: 233-260, New York: Academic Press, 1989.
- [47] Ekman, P. "An argument for basic emotions". *Cognition & Emotion*, 6: 169-200, 1992.