

Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции Диалог'2012. Вып. 11 (18). Том 1. М.: РГГУ. 2012. С. 664-673.

EXTRACTING SUBJECTIVE INFORMATION FROM SPOKEN DISCOURSE: THE CASE FOR PROSODIC EMPHASIS¹

Т.Е. Янко (tanya_yanko@list.ru)
Институт языкознания РАН, Москва, Россия

Ключевые слова: субъективная информация, интонация, звучащая речь, эмфаза, коммуникативная структура предложения

T. E. Yanko (tanya_yanko@list.ru)
The Institute for Linguistics, Russian Academy of Science, Moscow,
Russia

Key words: subjective information, spoken discourse, prosody, emphasis, communicative structure

Цель данной работы — разработка метода для извлечения субъективной информации из устного дискурса. Анализируемый тип субъективной информации — отношение говорящего к сообщаемым фактам как к нарушающим нормы жизни. Выражение говорящим сильных чувств в связи с нарушением привычного хода вещей, связано с понятием эмфазы. Эмфаза, как правило, имеет просодические средства выражения. Эмфатическая просодия означает: ‘замечательно, что он сделал это сам, так как обычно ему требуется помощь’, ‘он пришел на встречу, и это удивительно, потому что раньше он никогда к нам не приходил’, ‘их поздравил президент, что особенно почетно, потому что всегда это делает премьер-министр’. Эмфаза весьма частотна как в обыденном дискурсе, так и в речи дикторов радио и телевидения, особенно при сообщении новостей. Информация о достижениях и потерях, происшествиях, аномальных погодных явлениях, как правило, сопровождается эмфатической интонацией. В данной статье явление эмфазы иллюстрируется записями звучащей речи, для которых приводятся графики изменений частоты основного тона, полученные с помощью системы анализа устной речи Speech Analyzer.

The aim of this paper is to develop a method for extracting subjective information from spoken discourse. The example of subjective information discussed here is the attitude of the speaker to the reported events as to rare or abnormal. Disrupting a routine course of life which causes the speaker to express his/her strong feelings refers to the concept of emphasis. Emphasis basically has a prosodic expression. The emphatic prosody designates: ‘it is great that he did it by himself because generally he needs assistance’, ‘he came to the meeting which is very unusual because he never comes’, ‘it was the president who congratulated them which is highly honorable because generally it is the vice-president who does it’. Instances of emphasis are highly frequent in everyday speech as well as in mass media language, especially in breaking news. Reports of achievements and losses, accidents, severe meteorological phenomena, such as hurricanes, earthquakes, swollen streams, and numerical scores of any kind are generally made with the use of emphatic prosody. The instrumental analysis of the physical parameters of emphasis can serve as a means for extracting subjective information from spoken texts. The results presented here are exemplified by records from oral corpora. The frequency tracings are generated by the software program Speech Analyzer.

¹ This work has been supported by the Program of fundamental research “Language and literature in the context of social dynamics”, 2012—2014, section “Electronic databases and multimedia technologies in Philology”, the Russian Academy Presidium, Department of language and literature, project “The computer model and electronic data base “Meaning-Spoken Speech”” [Работа выполнена по Программе фундаментальных исследований Президиума РАН, секции языка и литературы «Язык и литература в контексте культурной динамики, 2012—2014, раздел «Электронные базы данных и мультимедийные технологии в филологических исследованиях», проект «Разработка компьютерной модели “Смысл-Звучащая речь” и электронной базы данных для ее подержки»]

The aim of this paper is to develop a method for extracting subjective information from spoken discourse. The example of subjective information discussed here is the attitude of the speaker to the reported events as to rare or abnormal. Disrupting a routine course of events which causes the speaker to be surprised and to express his/her strong feelings about it refers to the concept of emphasis. Emphasis basically has a prosodic expression. The emphatic prosody designates: 'it is great that he did it by himself because generally he needs assistance', 'he came to the meeting which is very unusual because he never comes', 'it was the president who congratulated them which is highly honorable because generally it is the vice-president who does it'. Instances of emphasis are highly frequent in everyday speech as well as in mass media language, especially in breaking news. Reports of achievements, victories, losses, accidents, severe meteorological phenomena, such as hurricanes, earthquakes, swollen streams, and numerical scores of any kind are generally made with the use of emphatic prosody. The instrumental analysis of the physical parameters of emphasis can serve, therefore, as a means for extracting subjective information from spoken texts. The results presented here are exemplified by records from the oral corpus specifically targeted at analyzing emphasis. The corpus (\approx 1 hour 15 minutes of sounding) contains records of everyday speech and breaking news from TV and radio broadcasting, as well as experimental laboratory readings with the forecasted parameters of emphasis. The waveforms and frequency tracings of examples considered below are generated by the software program Speech Analyzer.

Henceforth, the basic aim of this paper is to consider the semantic properties and the prosodic means of emphasis. The point of departure for analyzing emphasis is Russian. The contrastive analysis of Russian, English, German, French and Danish data demonstrates a significant similarity in prosodic means of emphasis: the examined languages display very similar cues for emphasis.

The semantics of emphasis is to express the speaker's attitude with respect to abnormal phenomena of life. In the question *Do you actually have a son?!*, the speaker, first, is making sure whether the hearer really has a son, and, secondly, expresses the idea that the hearer is too young to have a son and that if it is true it must make the hearer be proud of it. Emphasis, therefore, refers a current state of affairs to some ordered set where all elements are ranged in accordance with a certain standard of life. The emphasized element of the semantic structure is always the most unusual or one of the most unusual performers of the part it plays. It is placed close to the extreme of probability scale with the minimum value.

The concept of emphasis is a communicative meaning, and its basic property is the ability to combine with other communicative meanings. As a result, emphatic topics, emphatic foci, and emphatic questions arise. Therefore, emphasis does not function in a sentence independently. It can only combine with the meanings which form the components of speech acts: either statements or questions.

The principal — but not the only one — prosodic parameter of emphasis can be described as follows. The basic tone — the falling one which originally designates a focus or the rising one which designates a topic or a *yes-no*-question — in the context of emphasis associates with a tone which precedes the basic one and points in the opposite direction. Hence, with emphasis, the original fall is being preceded by a rise, while the rise is being preceded by a fall. This prosodic phenomenon might be called "emphatic tonal curving". To consider this point we need an example.

The prosodic parameters of emphasis are instantiated here by laboratory readings of the Russian word *sam* 'himself' (masculine) whose lexical semantics agrees with the concept of communicative emphasis (cf. [Yanko 2001: 287-289]), because *sam* means that what the agent unexpectedly did by himself s/he used to do with the help of his/her assistants or substitutes. In example (1) the word *sam* 'himself' is employed in a variety of communicative

functions, either composed with emphasis or not: a) *Sam?* ‘Did he do it by himself?’ (a non-emphatic *yes-no*-question), b) *Sam?!* ‘Did he really do it by himself?! Unbelievable!’ (an emphatic *yes-no*-question), c) *Sam* ‘He did it by himself’ (a non-emphatic statement), d) *Sam!* ‘He did it by himself. It is unbelievable!’ (an emphatic focus).

- (1) a) *Sam?*; b) *Sam?!*; c) *Sam*; d) *Sam!*
 ‘himself?’ ‘himself?!’ ‘himself’, ‘himself!’

In minimal pairs (a)-(b) and (c)-(d) respectively the second member of each pair is an emphatic counterpart of the first one. The pairs exemplify composing communicative meanings (or functions) — questions, topics, and foci — with emphasis. These examples demonstrate how emphasis modifies the fundamental frequency course when it combines with the communicative meanings originally designated by rises or falls of frequency. This description is obviously not complete without the pair “topic vs. emphatic topic” which is omitted on the assumption that topics and *yes-no*-questions, either composed with emphasis or not, are expressed identically. In fact, there is some difference between the expression of topics and *yes-no*-questions but here it is ignored: the analysis presumes that both topics and *yes-no*-questions in Russian are designated by the rise of the fundamental frequency which is placed on the stressed syllable of the accent-bearer of a topic or a question. The “emphatic” counterpart of a rise is therefore a rise which is preceded by an additional fall of frequency or by a low level tone, cf. frequency tracings of examples (1a) and (1b) below. Similarly, the difference between (1c) and (1d) is an additional “emphatic” rise of frequency in (1d) which precedes a fall as contrasted to (1c) where a fall designates a simple focus.

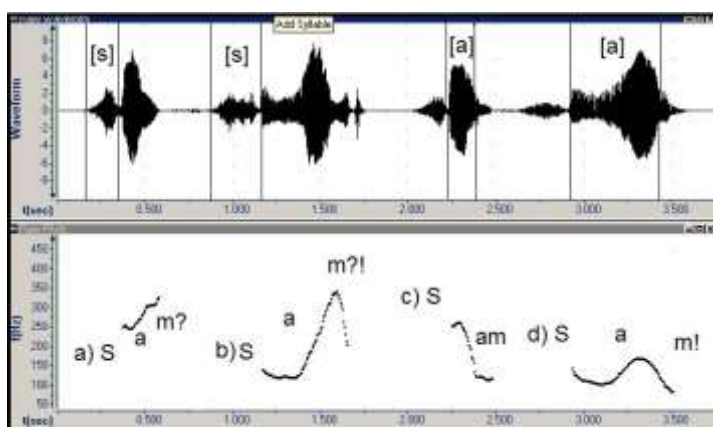


Fig. 1. *The waveform and frequency tracing of Example (1).*

The frequency tracing presented in Figure 1 on the lower panel, indeed, demonstrates “tonal curving” in (b) as contrasted to (a) and in (d) as contrasted to (c). Nevertheless, the tonal “curving” is not the only cue for emphasis. The frequency tracing shows that emphatic question (1b) has wider pitch-range than non-emphatic question (1a), namely — the rise in (1b) has a lower starting point than the initial frequency in (1a) and a higher peak frequency than the one in (1a). The accented — and the only one here — syllable of the emphatic accent-bearer has a significantly longer duration in the emphatic instances of *sam* than in the non-emphatic ones. Besides, in the context of emphasis the rise (either in a topic or in a question) or a fall (in a focus) is delayed because of a longer duration of the level onset of the syllable, especially because of the duration of the initial consonant. In example (1) it is [s]. The waveform presented by the upper panel (cf. Figure 1) shows the contrastive duration of the initial [s] in non-emphatic question (1a) and in emphatic question (1b). The intercepts of the waveform referring to [s] in (1a) and (1b) are delimited by vertical tracking.

The waveform also demonstrates the suppression of amplitude at the onset of the syllable, cf. vowel [a] in non-emphatic (1c) and emphatic (1d). Amplitude is rather low at the onset of a syllable but from approximately the middle of a syllable it grows sharply. The intercepts of the waveform referring to [a] in (1c) and (1d) are marked by vertical tracking.

An emphasized syllable gets, therefore, a binary structure: the delayed frequency peak and suppressed amplitude divide a syllable into two parts. As for the overall duration of the accented syllable of the accent-bearer, with emphasis it is about twice the corresponding non-emphatic one.

The binary structure of emphatic prosody is exemplified below by English and French emphatic sentences (5)-(6) with polysyllabic accent-bearers which in the context of emphasis get a secondary stress at the onset of a word. The components of the bidirectional emphatic frequency contours are, therefore, placed on the initial and the terminal accented syllables of the accent-bearer each. About emphasis in French cf. [Féry 2001] and [Faustova 2006].

While laboratory example (1) demonstrates the complete set of “emphatic” parameters, the examples taken from spontaneous speech do not always keep all correlations in full. It might be hypothesized that feelings of the speaker expressed by prosodic emphasis are being measured by the speaker depending on his/her agitation level. Nevertheless, all examples presented below where emphasis keeps its identity maintain the tonal curving. Tonal bending to the direction opposite to the original direction of the topical, focal, or interrogative prosodic cues may be positioned on the initial part of the accented vowel of the accent-bearer or even on a prolonged initial voiced consonant as in Russian sentence (2).

Sentence (2) exemplifies a speech of a lawyer whose client was adjudged dead by his ex-wife and was not able to retribute his rights for a year. In the emphatic topic *bolee goda* ‘more than a year’ the fall of frequency which precedes the “topical” rise is placed on the consonant [g] in the word *goda* ‘of a year’ which has a significantly longer duration than the one in the second – non-emphatic – occurrence of the same word: in order to be more convincing the lawyer pronounces the fragment *bolee goda* twice.

(2) I vot uzhe bolee goda, bolee goda...,
 And here already more than a year more than year...
 <on ne mozhet vosstanovit' svoe polozenie kak fizicheskoe litso.>
 <he is not able to retribute his position as a physical person>.

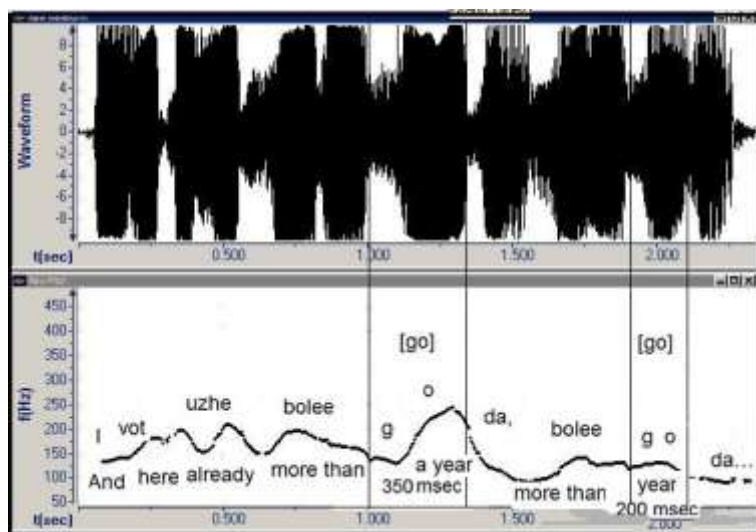


Fig. 2. The waveform and frequency tracing of Example (2).

The accented syllable [go] of the accent-bearer *goda* ‘of a year’ in the emphatic topic *I vot uzhe bolee goda* ‘And now already more than a year’ has a duration of 350 msec while the syllable [go] within the second non-emphatic occurrence of the same word only lasts 200 msec. Two occurrences of the same word successfully allow for contrasting the emphatic and the non-emphatic sounding. Emphasis in this sentence has longer duration, larger range of frequency, and syllable-initial lowering the amplitude of oscillation.

Example (3) illustrates, for comparison, the Danish emphatic question with the meaning of ‘unbelievable’. The speaker expresses his surprise with respect to the fact that the voice belongs to Jens and not to Lena: – *Who?! – Oh, Jens.* <Well, I thought it was Lena>. Consider the full text of a dialogue.

- (3) (a) – *Hej! Det er Jens i telefonen.*
 (b) – *Hvem er det?!*
 (c) – *Jens.*
 (d) – *Ah, det er Jens?! [a deɐ̯ jɛns]*
 (e) – *Jeg troede ellers, det var Lene.*

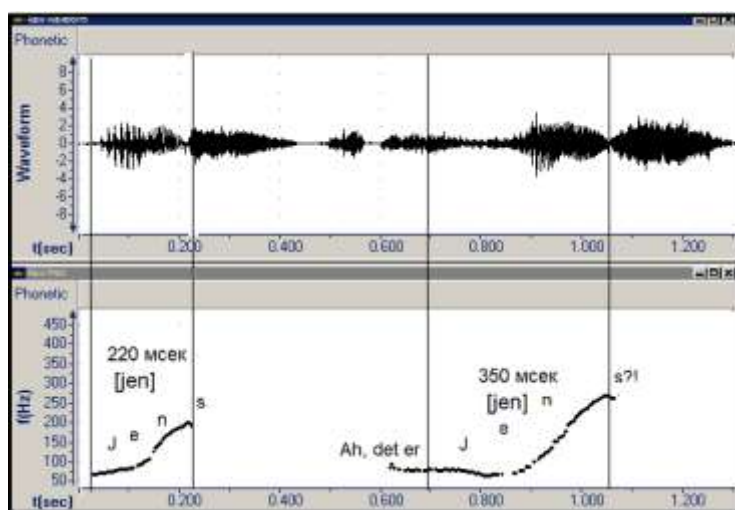


Fig. 3. *The waveform and frequency tracing of Example (3c-d).*

In example (3) the emphatic and non-emphatic rises are displayed by the same segmental material, namely — the word *Jens*. The waveform and frequency tracing of example (3c-d) demonstrates the amplitude increase, especially its higher peak value, a longer duration of the emphatic syllable, and the amplitude suppression within the consonant [j] in the emphatic sentence (3d) as compared with the non-emphatic preceding reply (3c).

German emphatic question (4) also displays all properties of emphasis. The frequency tracing demonstrates “low level - rising” course of frequency within the accent-bearer the word *Sohn* ‘son’. The waveform displays the binary structure of the vowel [o], and a longer duration of the word-initial consonant [z].

- (4) *Hast du schon einen Sohn?!*
 ‘Do you already have a son?!’

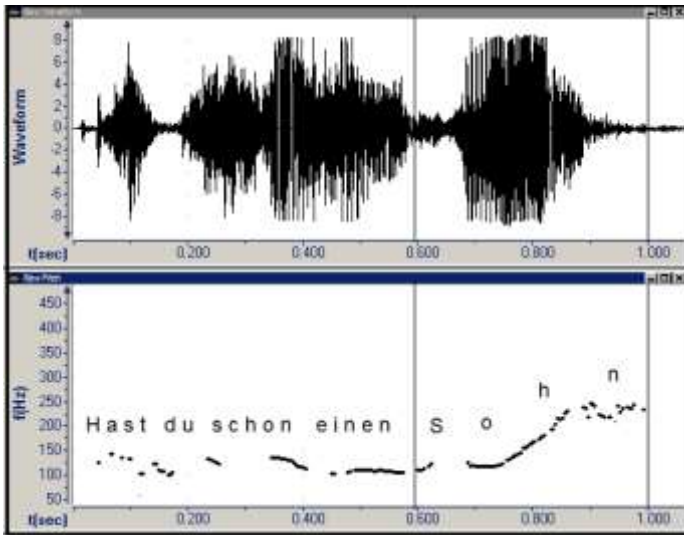


Fig. 4. The waveform and frequency tracing of Example (4).

The binary structure of the cues for emphasis allows for positioning the emphatic contour on the two syllables of a word with a secondary stress: the first component of a “curved” frequency course is placed on the first stressed syllable of a word with a secondary stress, the second — the basic — component is placed on the main word stress. Emphasis can, hence, capture two syllables of a polysyllabic word.

In English example (5) from an American TV broadcasting emphasis employs the segmental material of the numeral *fourteen*. The emphasis in the topic *fourteen years* brings the meaning ‘fourteen years is a huge span of time’. The numeral *fourteen* has two tonal peaks — the first one is positioned on the syllable *four-*, the second one — on the syllable *-teen*. The frequency tracing demonstrates a significant rise on *four-* and a fall — on *-teen*. There is a substantial pause between *four-* and *-teen* to serve as a boundary in dividing the emphatic articulation into two phases. The waveform clearly demonstrates the break of articulation between *four-* and *-teen*. On the frequency tracing the word *fourteen* is delimited by vertical tracking.

(5) *He was on death row for fourteen years!*

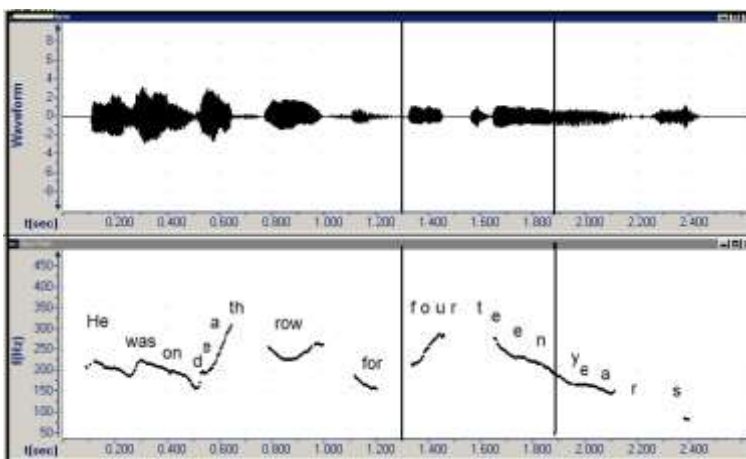


Fig. 5. The waveform and frequency tracing of Example (5).

Similarly, in example (6) taken from a French breaking news program the polysyllabic word *considérablement* ‘significantly’ displays the two components of an emphatic frequency contour: the fall on *con-* and the basic rise — on *-ment*. The rise on *considérablement*, as well as the rise on the initial topic *température*, designates the topic of sentence (5) referring to the focus *chuté ces derniers jours dans plusieurs pays d’Afrique...* with a focal fall on

chuté. Therefore, sentence (6) has two topics. The sentence-final word *Afrique* in *chuté ces derniers jours dans plusieurs pays d'Afrique...* also carries a rise which serves here as a cue for discourse continuation, i.e. it shows that the sentence (6) is not discourse-final. On the frequency tracing the emphasized word *considérablement* is delimited by vertical tracking.

(6) *Température a considérablement chuté ces derniers jours dans plusieurs pays d'Afrique...*
 ‘The temperature has severely fallen during the last days in many African countries...’

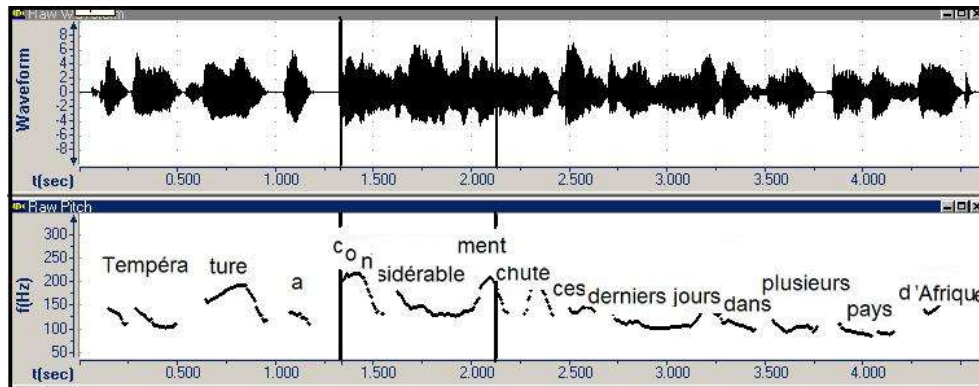


Fig. 6. The waveform and frequency tracing of Example (6).

As for the Russian language, it might be hypothesized that the so called Fifth intonational construction IK-5 as defined by E.A. Bryzgunova [Russkaia grammatika 1982: 116-117] which is characteristic of the Russian exclamations and whose frequency contour is specified by the two — the sentence-initial and the sentence-final — tonal peaks is, in fact, an instance of emphatic focus which has two accent-bearers: at the onset and at the end of a sentence. An emphatic focus can be displayed by segmental material consisting of more than one word with a sentence-initial word carrying a rise and a sentence-final word carrying a fall, and by a segmental material of only one word as well, even a monosyllabic one: *Kakoi udacha!* ‘What a good chance!’, *Skol’ko devushek!* ‘How many girls!’, *Zamechatel’no!* ‘Great!’, *Otvratitel’no!* ‘How disgusting!’, *Styd!* ‘What a shame!’. Cf. laboratory example (7) which demonstrates the parameters of Bryzgunova’s IK-5:

(7) *Kakoi nomer!*; *Zamechatel’no!*
 ‘What a show!’ ‘Great!’

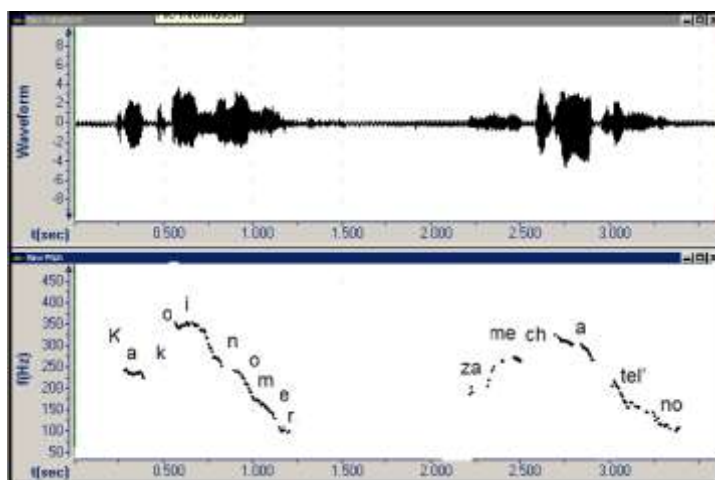


Fig. 7. The waveform and frequency tracing of Example (7).

In *Kakoi nomer!* ‘What a show!’ the stressed syllable [koj] in *kakoi* ‘what a..!’ carries a rise, while the stressed syllable [no] in *nomer* ‘show, performance’ carries a fall. *Zamechatel’no!* ‘Great!’ has two gradual rises of fundamental frequency on the initial — generally unstressed — syllables [za] and [mi] and a fall — on the stressed syllable [tʃa]. It could, therefore, be hypothesized that an emphatic focus with a monosyllabic segmental material, e.g. *Syn!* ‘Son!’ or *Styd!* ‘Shame!’, might be presented as a result of IK-5 compressing. For an emphatic topic or a question, e.g. *Syn?* ‘Son?’, which are based on the rise of frequency within the stressed syllable the pattern will be obviously different.

To conclude, from the viewpoint of semantic structure emphasis always functions in composition with the communicative parameters producing the speech acts of a statement or a question. Emphasis generates emphatic topics, emphatic foci, and emphatic questions. From the viewpoint of expression emphasis is specified by the following parameters:

- a longer overall duration of the stressed syllable in the accent-bearer of an emphatic communicative component as contrasted to the respective non-emphatic one;
- a longer duration of the accent-bearer initial consonants (if any) in comparison with the respective non-emphatic ones;
- a specific “emphatic tonal curving” when the basic tone — the falling one which originally designates a focus or the rising one which designates a topic or a *yes-no*-question — associates with a tone which precedes the basic one and points in the opposite direction;
- a syllable-initial suppression of amplitude which is followed by a jump of oscillation amplitude in an emphatic communicative component.

With the only accent-bearer of an emphatic component emphasis is specified by a heterogeneous sounding with a suppressed onset and a full quality ending of the stressed syllable. In addition, emphasis can also be displayed on various segmental material and have not only one but also two accent-bearers with two stressed syllables. With two stressed syllables emphasis is displayed with the basic — depending on the communicative function either falling or rising — frequency change course on the second stressed syllable and a frequency change course pointing in the opposite direction on the first stressed syllable. Emphasis, thus, is characterized by dividing the emphatic articulatory data into two phases and lengthening both of them as being contrasted to the respective non-emphatic topics, foci, and questions. These parameters are similar in all examined languages. The set of the distinguished prosodic parameters can serve for formalizing the process of disclosing emphasis in spoken discourse.

References

1. Faustova N.A. Emfaza i ee intonatsionnoe vyrazhenie [Emphasis and its intonational expression]. *Voprosy filologii* – Problems of philology, 2006, no. 6, pp. 255-259.
2. Féry C. The phonology of Focus in French. *Audiatum Vox Sapientiae. A Festschrift for Arnim von Stechow*. Berlin, 2001.
3. Pal’ko M.L. Emfaza v nemeckom iazyke: dialogicheskie strategii v monologicheskoi rechi [German emphasis: dialogic strategies in monologues]. *Logicheskii analiz iazyka. Mono-, dia-, polilog v raznyh iazykah i kul’turah* – Logical analysis of language. Monologue and dialogue in different languages and cultures. Moscow. Indrik Publ. 2010. 295-304.
4. Yanko T.E. *Kommunikativnye strategii russkoi rechi* [Communicative strategies of the Russian speech] Moscow. Yazyki slavianskoi kul’tury Publ. 2001. 384 pp.