French Listeners Counting Syllables in Read French and Russian: Implications for the Cognitive Realty of Syllable

Irina Nesterenko

Department of Phonetics
Saint-Petersburg State University, Russia
irina@in3653.spb.edu

Abstract

This paper examines the question of cognitive status of syllable as a linguistic unit. The experiment was conceived in the objective to test, in a metalinguistic task of counting the syllables, put to French speaking subjects, whether the syllable is a mere methodological concept adopted in linguistic descriptions or whether there is a cognitive reality behind it. Working with filtered stimuli permits to concentrate the attention on the role of the syllable in the prosodic organisation of the utterance.

1. Introduction

In the notorious scheme by Jakobson, representing different language functions, one of them, the metalinguistic function, is directed to the language itself. To describe different processes that take place in language, linguists introduce special operational concepts like phoneme, syllable, stress unit etc. As Otake & al.[19] note, these units are not « properties of particular language, but phonological constructs in terms of which any language can be described ». [260].

The present study deals with the cognitive value of the linguistic concept of the syllable. The syllable is a universally recognised linguistic unit. At the same time, there are a lot of discussions about its nature and status in linguistic descriptions. The concept of syllable is largely used in accounts of phonotactic constraints operating in a particular language (cf. the radical position of Pulgram [21] who considered that phonotactic rules alone determine the decomposition of an utterance into the syllables) as well as in descriptions of the prosodic level processes, for example, the tone and accent phenomena. It is a unit that does not play a role in the communication of meaning directly which makes it difficult, if not impossible, to propose a unified definition.

As an operational concept, the syllable is used in phonetic as well as phonological descriptions and in versification theory. However, as Angoujard [1] points out, the definitions of these concepts in different fields of linguistics show a remarkable discrepancy (consider as well the distinction made by Grammont [12] between a phonetic syllable and a phonological syllable).

In the framework of a phonological analysis, Blevins [4] defines the syllable as a phonological unit which organises segmental melodies in terms of sonority (cf. as well the sonorous theory of syllable). Two structural approaches are possible as well: we can define the syllable by the syllable-internal structure (cf. the branching representation of the syllable with the onset and the rhyme, the latter formed of nucleus and coda [16]), or by its role in the hierarchy of linguistic units (autosegmental approach, adopted, among others, by Kahn [15] and Clement & Keiser[6]).

At the same time, in the model of speech production proposed by Levelt [17], it is the subsyllabic components that are important at the level of phonological representations, the syllable as a whole being a minimal and invariable unit of articulation (cf. as well the different articulatory theories of syllable).

According to Blevins [4], one of the arguments in favour of the syllable as a phonological constituent is the intuition of language users who are capable to determine the number of syllables in a word, to enumerate them and to change their order (language games, cf. [7]). In this perspective the question arises whether a syllable is an operational concept, introduced by linguists and assimilated by speakers through education or there is a cognitive reality behind it. The first results obtained in psycholinguistic experiments, which followed a special syllable and phoneme detection paradigm, indicated that the syllable is the main perception unit for French speaking subjects [18]. Although the effect observed for speakers of French was not replicated for the speakers of languages differing from French in rhythmic properties (English, Japanese) [8, 19]. Consequently, the status of the syllable as a perception unit was considered to be dependent on the rhythmic properties of a given language. At the same time, it was observed that speakers of French are capable to focus their attention on a unit as small as phoneme (like the speakers of English in Cutler's experiments) but they interpret the phoneme position in terms of syllable constituents [20].

What is, then, the role of the syllable in phonological and phonetic descriptions? As Astesano points out [2], the syllable is a phonologically relevant unit insofar as it allows the association with higher level units and features (timing), but such a position does not clarify the cognitive status of the syllable. The present study was designed to verify the hypothesis of the role of the syllable in processing of prosodic information. We proposed to the subjects a metalinguistic task of counting the syllables in speech stimuli. To induce the responses based mainly on the prosodic information, filtered stimuli were used, this manipulation masking the syllable-internal structure and the articulation. At the same time, the stimuli presented to the subjects corresponded to the multisyllabic utterances so preserving the phenomena typical to the continuous speech (f.e. the reduction and syllables' elision phenomena). Two sets of stimuli (French vs. Russian stimuli) were proposed to the auditors, which permitted to compare their performance while processing the stimuli in two languages representing two rhythmic classes.

2. Perceptual experiments

The objective of the first perceptual experiment was to determine to what degree the linguistic concept of the syllable is pertinent for subjects when listening to the filtered speech stimuli.

2.1. Method

2.1.1. Material

In this experiment the subjects were presented 45 Russian stimuli (nine read utterances in the pronunciation of five native speakers of Russian, one male and four females) and 36 French stimuli (nine read utterances pronounced by four native speakers of French, two males and two females). All Russian utterances had 10 to 11 potential (i.e. determined from the orthographic version of the utterance) syllables; the French stimuli had 12 syllables, when counting all the potential syllables containing 'mute e'. The stimuli were filtered with a low-pass filter − the frequencies above 450 Hz were suppressed (this value was chosen after several preliminary trials. We kept in mind the results reported by Sluijter & van Heuven [22], where it was shown that the perceptually important intensity components, related to the presence of an accent, are located in a zone above 500 Hz. So we can consider the influence of intensity cues to be minimised in our experiment).

The stimuli were randomised and grouped in two lists. The order was kept the same for all the listeners taking part in the experiment. French and Russian stimuli were presented separately, forming two parts of the test.

2.1.2. Procedure

Ten French-speaking subjects took part in the experiment; they were tested individually. The stimuli were stored as sound files on a hard disk of the PC computer and were played back through the headphones. The subjects were seated in front of the computer and regulated the progress of the experiment themselves: to play a stimulus they needed to click on an appropriate hypertext link to the sound file in the list (Word file). The number of times the subjects could listen to the same stimulus was unlimited and uncontrolled (though this information might have been useful to evaluate the difficulty of the proposed task. It might have served as an additional parameter for the analysis as well). The questionnaire given to the subjects represented a set of lines composed of S symbols, each S standing for a 'syllable': the number of S's in a line was bigger than that of potential syllables in the stimulus. The subjects were asked to determine the number of syllables for each stimulus and to indicate all the perceived prominences.

2.2. Results and discussion

To compare the data for two sets of stimuli (Russian vs. French utterances) we decided to measure the difference between the perceived number of syllables and that determined from the transcription of the pronunciation of the utterance by a given speaker. The same utterance could contain a different number of syllables in the pronunciation of different speakers for a negligent style of pronunciation induces a greater reduction. We treat this number of syllables as objective; it was determined after performing the acoustical analysis of the phrase-stimulus (segmentation into allophones and syllables using Praat program).

For each response of each subject we have received a measure of deviation expressed in a number of syllables, this deviation ranging from −4 to +3 syllables (negative values indicate the tendency to attribute to the stimulus a number of syllables smaller than an objective one) for Russian stimuli,

and from −5 to +3 syllables for French stimuli (Figure 1 plots different observed deviations against their frequencies).
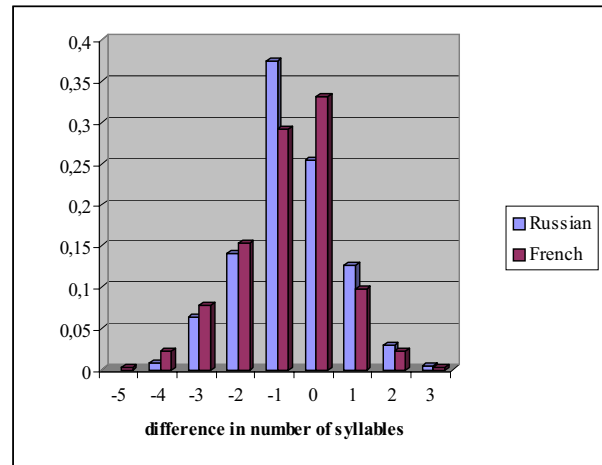

Figure 1: The frequencies of different observed deviations for Russian and French stimuli

These data show that the French-speaking subjects perceived the exact number of syllables with minimal deviation (from minus 1 to plus 1 syllable) in 75,33% of cases for Russian stimuli and in 71,99% of cases for French stimuli. These minimal deviations can be explained by the internal structure of stimuli: for example, the subjects' performance in response to the Russian stimuli can be influenced by the presence of adjacent vowels or post-accentual syllables at the end of the utterance; in French stimuli the syllables on mute e can represent the difficulty in the proposed auditory task.

We performed a one-way analysis of variance on the observed deviations testing the influence of language factor. To avoid the repeated measures design error, the mean deviations observed for each subject were used in the statistical analysis. Language yields no significant influence ($F(1,18)=0,151$; $p=0,7021$), which permits us to suppose that the model of syllable was the same when the auditors processed the stimuli in two languages. It can be supposed as well that this model forms part of the auditors' linguistic competence.

So, in the present experiment the French speaking subjects showed the ability to determine the number of syllables in the stimuli representing the entire utterances and this in the experiment situation when only the prosodic information was available. But the prosodic information in the signal is of a complex nature as it includes both temporal and melodic phenomena. After analysing the results of the first experiment, a logical question to ask would be what the role of each parameter (the duration and the melodic curve) in the constitution of the syllable as linguistic unit is. In the development of this study, the results of preliminary experiments, trying to respond on this question are presented.

2.3. Testing the impact of the duration cues

The notion of syllable is of a great importance in the rhythm studies. Di Cristo [11] proposes to define rhythm as the temporal organisation of metrically strong and metrically week syllables the former being perceptually more salient. We suppose that in the proposed task the subjects captured the

metrically strong syllables first; the following step was to measure the distance between two salient points in the signal by applying the temporal model of the syllable. If this hypothesis is true, the subjects should be sensible to syllable duration, and, as the consequence, to the total duration of the phrase-stimulus. In the second experiment, the influence of this last parameter on the performance of the subjects was tested statistically.

2.3.1. Procedure

In the description of the first experiment, the conditions applied to construction of the stimuli for our study were stated: the utterances were selected according to the potential number of syllables and were not normalised as to their absolute duration. We have chosen to subdivide our stimuli in two categories (Long vs. Short) depending on their duration, the boundary being fixed at 2,1 seconds. Table 1 presents the mean duration and the standard deviation for two categories.

Table 1: Mean duration and standard deviation per category

Category	Russian stimuli		French stimuli	
	Mean duration	Standard deviation	Mean duration	Standard deviation
Short	1,864	0,191	1,944	0,82
Long	2,506	0,297	2,387	0,18
All stimuli	2,221	0,41	2,166	0,264

The next step was to perform statistical analysis of the influence of the parameter of stimulus duration on the perceived number of syllables. Figure 2 represents the relations between the stimulus category and the mean perceived number of syllables per category for Russian and French utterances.
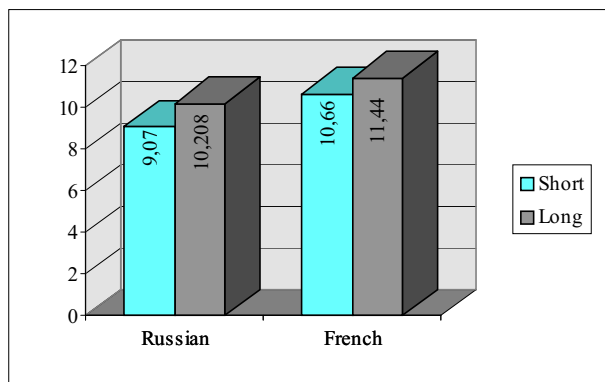

Figure 2: Mean perceived number of syllables per category for Russian and French stimuli

Two one-way ANOVAs were calculating to test the significance of the Category effect, one for the mean number of perceived syllables per category, the other for the average difference between perceived and objective number of syllables. Both return the significant effect (for Russian stimuli $F1(1,43)=40,021$; $p<0,0001$, $F2(1,43)=16,863$, $p=0.0002$; for French stimuli $F1(1,34)=14,164$; $p=0,0006$), $F2(1,34)=4,320$, $p=0,0453$).

So, it was attested that the French listeners are sensible to the stimulus duration while performing a metalinguistic task of counting the syllables in filtered speech stimuli. The statistical tests show that this influence was greater for the Russian stimuli. Two explanations can be proposed: first of all, it was proved in a number of previous research that in French, the duration of the unaccented syllable tends to be stable and represents the reference duration in the rhythmical structuring of the utterances [9, 13]. At the same time, Russian is a language with strong vowel reduction, the syllable duration being subject to a greater variance.

On the other hand, the influence of the mother tongue factor can not be excluded: it can be supposed that the listeners have applied a temporal model of the syllable in measuring the distance between two perceived prominences, and this temporal model was formed from the listener's linguistic experience with the mother tongue (cf. the Polivanov's theory of the phonological sieve which can be applied at the suprasegmental level as well as at the segmental one).

2.4. Testing the impact of the melodic curve

When processing the speech signal, the numerous speech signal analysers return a melodic curve, reflecting the changes of F0 in time, which contains two kinds of information simultaneously: the microprosodic information about the local characteristics of the signal, resulting from the interaction of adjacent sounds, and the macroprosodic information of a purely linguistic nature [10]. We decided to analyse the impact of these two components on the listeners' performance in the proposed task of counting the syllables.

2.4.1. Stimuli

To carry out this preliminary test, two utterances in the pronunciation of the male French speaker were chosen, which had served as stimuli in the previous experiment. To obtain the contrastive groups of stimuli, we applied different manipulations using the MOMEL algorithm [12]. First of all, after extracting the F0 curve, the stimuli with minimised microprosodic variations were resynthesised. The second manipulation consisted in suppressing all the melodic modulations in the utterances: we attributed the constant value of 119 Hz to all the key-points of F0 curve, obtained through MOMEL, and resynthesised the utterance. These manipulations permitted us to obtain three different categories of stimuli.

2.4.2. Procedure

The procedure and the instructions were the same as in the first experiment. Ten French speaking subjects took part in the experiment.

2.4.3. Results.

The difference in the perceived number of syllables according to the type of manipulation applied was analysed. The ANOVA returns no significant effect of the manipulation type $(F(2,27)=0,278$, $p=0,7596)$. These results permit to suppose that the melodic changes are less important for listeners when performing the task of counting the syllables. This result corroborates the conclusions made by Campbell [5] that the role of duration cues is important for indicating the boundaries and prominences in the informational structuring of the message.

3. Conclusions

The main objective of the present study was to collect the data that would permit to answer the question about the cognitive status of the syllable as a linguistic unit − whether it is a simple methodological concept or if there is a cognitive reality behind it. The main results are as follows:

1. The French speaking subjects can determine the number of syllables in speech stimuli when faced with a methalinguistic task of counting them. In our experiment the stimuli represented entire utterances and in their decisions, the subjects could rely only on prosodic cues.
2. The statistical analysis reveals that there is no difference in subjects' performance whether they are processing the French or Russian stimuli (the results are based on the measure of difference between the perceived number of syllables and the objective one). We can suppose that there is a phonological model of syllable in the linguistic competence of the speakers, which they have applied under both conditions.
3. The study revealed that the subjects are sensible to the parameter of the stimulus duration, when performing the proposed task: the longer the stimulus, the greater the number of perceived syllables. It is supposed that there is a temporal model of the syllable that the subjects applied to measure the distance between two perceived prominences.
4. The melodic changes had no significant effect on the subjects' decision of the number of syllables in the proposed task.

We have found the relation between the syllabic unit and the temporal organisation of an utterance. At the same time, in modern phonological theories it is the foot that is considered to be the basic unit governing the temporal organisation of segments (another unit, proposed by Barbosa & Bailly [3], is the inter-perceptual-center group). Yet, when applied to French, the notion of foot is isomorphous to that of syllable. So it would be interesting to replicate the experiments with the speakers of a stress−timed language (Russian or other), in which the two units will not coincide. It would also permit to test the proposed hypothesis about the listener's strategy and to evaluate the impact of mother tongue factor.

4. References

[1] Angoujard, J.-P., 1997. *Théorie de la syllabe. Rythme et qualité,* Editions CNRS, Paris.
[2] Astésano, C., 1998. Effects of prosodic constraints on the differential lengthening of syllable constituents in French: A comparison between spontaneous and read speech, *Sposs*, 24-26 septembre, Aix-en-Provence, France, 143-146.
[3] Barbosa, P. & Bailly, G., 1994. Characterisation of rhythmic patterns for text-to-speech synthesis. *Speech Communication*, 127−137
[4] Blevins, J., 1995. The syllable in phonological theory. In *J.A. Goldsmith, The Handbook of Phonological Theory.* Cambridge: Blackwell, 206-244.
[5] Campbell, W.N., 2000. Timing in Speech: a multi-level process. In *M. Horne (ed.), Prosody: Theory and Experiment.* Dordrecht, Kluwer Academic Publishers, 281−334
[6] Clements, George N. & Keyser, Samuel Jay, 1983. *CV Phonology. A Generative Theory of the Syllable.* MIT Press, Cambridge, MA, USA
[7] Content, A., Dumay, N., Frauenfelder, U.H., 1999. Segmentation syllabique chez l'enfant. *Actes des 2émes Journées d'Etudes Linguistiques - La syllabe.* Nantes: Université de Nantes, pp. 69-74
[8] Cutler, A., Mehler, J., Norris, D. & Segui, J. (1986). The syllable's different role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385-400.
[9] Duez, D., 1999. Effects of articulation rate on temporal aspects of read French speech. *Proceedings of Eurospeech,* vol.2 (1999 September 5-10 : Budapest), 715-718.
[10] Di Cristo, A. & Hirst, D.J., 1986. Modelling French Micromelody: Analysis and Synthesis. *Phonetica,* 43(1)*,* 11-30
[11] Di Cristo, A., 1999. Le cadre accentuel du français contemporain : essai de modélisation I+II. *Langues*, 2(3), 184-205 et 2(4), 258-269
[12] Grammont, M., 1933. *Traité de phonétique.* 2e Ed. 1939, Delagrave, Paris
[13] Guaïtella, I., 1988. Variations de durée en syllabe accentuée. *Travaux de l'Institut de Phonétique d'Aix.*, 12, 189-203.
[14] Hirst, D.J. & Espesser, R., 1993. Automatic modelling of fundamental frequency using a quadratic spline function. In *Travaux de l'Institut de Phonétique d'Aix*, 15, 71-85
[15] Kahn, D., 1976. *Syllable-based Generaliza-tions in English Phonology.* IULC Publications Bloomington, IN, USA
[16] Kaye, J.D. & Lowenstamm, J., 1984. De la syllabicité. In *F. Dell, D.J. Hirst, J.R.Vergnaud (eds), Forme sonore du langage.* Paris : Hermann, 123-159
[17] Levelt, Willem.J.M. (1989). *Speaking.* MIT Press, Cambridge, MA, USA
[18] Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J., 1981. The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behaviour*, 20, 298-305
[19] Otake, T., Hatano, G., Cutler, A., & Mehler, J., 1993. Mora or Syllable? Speech Segmentation in Japanese, *Journal of Memory and Language*, 32, 258-278
[20] Pallier, C., 1997. Phonemes and syllables in speech perception: size of the attentional focus in French. In *Proceedings of Eurospeech '97,* 4, 2159-2162, Sept. 22-25. Rhodes, Greece.
[21] Pulgram, E., 1970. *Syllable, Word, Nexus, Cursus.* La Haye, Pays-Bas : Mouton & Co.
[22] Slujter, Agaath M.C. & van Heuven, Vincent J., 1993. Perceptual cues of linguistic stress: intensity revisited. In *Working Papers 41*, Dept. Of Linguistics and Phonetics, Lund, Sweden.