Prosodic Phrasing of SVO Sentences in French

Mathieu Avanzi[1], *George Christodoulides*[2], *Elisabeth Delais-Roussarie*[1]

[1]Laboratoire de Linguistique Formelle, UMR 7110, University Paris Diderot, France
[2]Institut Langage & Communication, Centre VALIBEL, UCLouvain, Belgium

mathieu.avanzi@gmail.com, george@mycontent.gr, elisabeth.roussarie@wanadoo.fr

10.21437/SpeechProsody.2014-129

Abstract

In the literature on prosody/syntax interface, syntactic information is usually considered as playing an important role in deriving the prosodic phrasing of an utterance. NP subjects, for instance, have often been claimed to phrase independently from the VP. It has nevertheless been shown that metrical factors could have an impact on phrasing, and that NPs could be phrased in the same prosodic phrase as the VP, or that the verb could be phrased with the subject. Several methods were used to measure metrical weight: number of syllables, of prosodic words, syntactic branchingness, etc. In order to determine which factors are more important, and how the all interact, we evaluate the weight that different metrical predictors have on prosodic phrasing. This is done by analyzing the phrasing of SVO structures in 200 sentences extracted from various French corpora. From the observation of the data that were semi-automatically annotated, it appears that subjects can be phrased independently or in the same prosodic phrase as the VP, and that objects are rarely isolated from the verb. The analysis reveals interesting results regarding the effect of articulation rate and number of syllables, whereas syntactic-branchingness didn't show any effect.

Index Terms: Prosody, metrical structure, phrasing, interface prosody-syntax, articulation rate

1. Introduction

Several theoretical and experimental studies have shown that multiple factors constrain prosodic phrasing, among which we may mention metrical constraints. Nevertheless, syntactic information are often considered as playing the most fundamental role, regardless of the approach used to account for the mapping between the prosodic structure and the syntactic one. In [1], for instance, it is assumed that edges of prosodic pharses should coincide with edges of syntactic phrase, whereas [2] insists on head-complement relations in determining the prosodic phrases (see for a review of the various approaches [3] and [4]).

Whatever approach is taken into account (end-based or relational), one would predict for French SVO sentences that NP subjects form an independent prosodic phrase, separated from the rest of the utterance, and more specifically the VP (see, for instance, [5] and [6]). On this basis, the sentence in (1) would be phrased in two prosodic groups, as shown in (2):

(1) Le gardien a vu le beau chien de ma voisine

(2) [Le gardien] [a vu le beau chien de ma voisine]

Nevertheless, it has been shown that mapping rules were not sufficient to predict phrasing. Factors such as length of the syntactic phrases could play a crucial role in the prosodic parsing of an utterance. As for SVO structure in French, [7] and [8] showed that NP subjects could be phrased in the same prosodic phrase as the VP when they are short (i.e. when they contain a small number of syllables), and that objects could be separated from the verbal head V to obtain well-balanced phrases. As a consequence the parsing in (3) is completely acceptable for the sentence (1), and maybe even better than the one proposed in (2):

(3) [Le gardien a vu] [le beau chien de ma voisine]

In the literature, the length of the constituents have been evaluated by several means: syntactic branchingness or complexity ([2]), number of prosodic units (generally number of prosodic words, as in [9]) or number of syllables ([10] and [8] among others). The way certain of these factors interact have been analyzed in details in some experimental studies (see [11] and [12] among others). In [12], where four different Romance languages were compared, it has been shown that SVO structures do phrase differently in the four Romance languages under investigation. Nevertheless, the (S)(VO) pattern was the most common for Catalan, Spanish and Italian, while (SVO) is the unmarked phrasing for European Portuguese. In addition, the authors showed that syntactic branchingness, number of prosodic words, and number of syllables had a different impact on phrasing decisions, and interacted in a different fashion across the four languages, leading for example to the production of patterns such as (SV)(O) in Catalan under certain length condition, a pattern which never was found for European Portuguese. A systematic study of the impact of these different factors has never been achieved for French. The goal of this paper is thus to investigate how all these factors interact in French. Due to a limited number of utterances and to the important number of parameters at play, the results presented here are only preliminary.

2. Methods

2.1. Data and participants

The analysis presented here is based on corpus data. We explored a total of 90 minutes of speech recorded by 18 speakers in four different speaking conditions. Among these 90 minutes, 20 minutes consist in the reading of fairy tales by professionals, 20 minutes in political speeches, and the remaining 50 minutes consist in different types of spoken data produced by 8 Parisian speakers recorded within the PFC project [13]. Following the data collection protocol used for the project, each speaker was asked to read carefully at a normal speech rate a journalistic text including 22 sentences (398 words), and then to converse freely in pair for 20 minutes. Among these data, the reading of the text and 3.30 minutes of conversation were taken into consideration for each of the 8 speakers in the present studies. Table 1 gives an overview of the number of speakers and of the duration of the samples.

Table 1: *Composition of the data set*

Speaking Style	Duration (min.)	Nb. of speakers
Fairy tales	19'28	6
Political speeches	21'08	4
Reading (PFC)	18'53	8
Spontaneous (PFC)	30'10	

We controlled for the sex of the speakers, selecting an equal number of male and female for each sub-corpus; however age was not a controlled variable. All participants speak a standard variety of French.

2.2. Labeling of the data and prosodic structure

The recordings were transcribed within Praat [14], and aligned in phones, syllables and words with EasyAlign [15]. All alignments were manually checked and corrected by hand by one of the authors. Then, the orthographic transcription of the data was annotated in PoS with the Dismo software [16]. This allows assigning a phonological status to the different words with respect to their ability to be accented (see among others [17] and [18]), and then segmenting the data in Phonological Words (henceforth PW).

The strength of the prosodic boundaries occurring at the end of prosodic words depends on their phonetic realization. When the last metrical syllable of a prosodic word is associated with a pitch movement, is lengthened, and even followed by a pause, it is considered as corresponding to a major prosodic boundary that may correspond to a major phrase or an IP boundary. In order to analyze intonational phrasing without going into details as for the distinction between major phrase or intermediate phrase on the one hand, and intonational phrase on the second (see [19] among others), the two levels have been grouped together and referred to as *major prosodic phrases* (MaP). The strength of the boundary occurring at the end of each prosodic word has been automatically detected with the Analor tool [20]. On the basis of four automatically measured acoustic parameters (relative syllabic duration, relative f0 average, slope contour amplitude and presence of an adjacent silent pause), the software estimates a degree of strength for the last syllable of each PW on a scale from 0 to 10 (from the least to the most prominent). The calculations rely on two fundamental principles. The first is a quantity principle: the greater the number of acoustic parameters involved in the identification of a prominence and the distance from predetermined thresholds, the stronger the prominence is perceived. The second is a compensation principle, which stipulates that if one of the classic parameters involved in the perception of prominence in French presents a low value and another presents a high value, there will be the same feeling of prominence as if the two parameters involved both presented a medium score. We considered that the right edge of a PW coincides with a major prosodic phrase boundary when the degree of strength associated to the last syllable of the PW reached a score of 4/10.

2.3. Extraction and prosodic encoding of the SVO structure

Using a manual annotation, we extracted from the entire corpus all sentences having a SVO pattern, where S, V and O were not separated by any other kind of linguistic material (appositive clauses, adverbs and other parentheticals), and where S and O consisted of a SN. In addition we restricted ourself to main clauses, leaving aside embedded SVO structure. In total, 198 sentences were extracted. Table 2 gives the distribution among the sub-corpus of the extracted sentences.

Table 2: *Nb of SVO sentences extracted from each sub-corpus*

Speaking Style	Nb. of sentences
Fairy tales	85
Political	28
Reading	67
Spontaneous	29

On the basis of the annotation described in section 2.2, it was possible to assign to each syntactic units of the SVO structure prosodic labels/ features(including F0 patterns, number of syllables, and number of PWs). Due to the uncontrolled character of the data (corpus data), the length of the various units vary greatly, as shown in **Erreur ! Source du renvoi introuvable.**.

Table 3: *Range and mean (standard deviation) nb. of syllables and of PWs for each syntactic constituent*

Synt. Const.	Nb. of Syll.		Nb. of PW	
S	Min.-Max Mean (SD)	1-17 4.8 (2.6)	Min.-Max Mean (SD)	1-4 1.7 (0.9)
V	Min.-Max Mean (SD)	1-8 3.1 (1.6)	Min.-Max Mean (SD)	1-3 1.1 (0.5)
O	Min.-Max. Mean (SD)	1-15 4.7 (2.7)	Min.-Max. Mean (SD)	1-4 1.7 (0.8)

In addition, the syntactic complexity of each S and O (simple nominal head or nominal head with branching complements) was manually coded. The articulation rate of the entire sentence was also automatically calculated by dividing the duration of the entire sentence (and by excluding the time of silent pauses) by the total number of syllables (it is therefore expressed in ms/syll, (see [21]).

3. Results

The analysis of the data, and more specifically of the boundary strength associated at the end of S, V, and O respectively, showed that S and O could be phrased independently from V, or phrased in the same MaP than V. It thus leads to four possible combinations, which were all obtained in our data. Yet, as table 4 shows, the four combinations do not occur with the same frequency.

Table 4: *Distribution of the phrasing obtained for the SVO structures*

Pattern	Count	%
(S)(V)(O)	16	8.1
(S)(VO)	92	46.5
(SV)(O)	4	2
(SVO)	86	43.3

The (S)(VO) pattern is the most frequently observed, followed by the (SVO) pattern. Taken together, these two patterns represent approximately 90% of the data. Cases where both S and O are isolated from V by a major prosodic phrase boundary are quiet rare, and cases where S and V are phrased in a single IP while O is in a distinct IP are even rarer (these two patterns taken together account for 10% of all SVO occurrences)

Due to the small amount of cases for the two last patterns, and in order to simplify the analysis and presentation of the results, we separated the data according to the way S and O were phrased in regards of V. Essentially, we compared the group of sentences where S was phrased with V (N= 90) and the group of sentences where S was phrased in a different IP than V (N=108), *cf*. §3.1; and we compared the group of sentences where O was phrased with V (N= 178) and the group of sentences where O was phrased in a different IP than V (N = 20), *cf*. §3.2.

Data were analyzed by mean of Nominal Logistic Regression within SPSS (v. 21.1). Due to the fact that the number of syllables and the number of PWs were strongly correlated for S ($r = 0.808$, $p < 0.01$), for V ($r = 0.718$, $p < 0.01$) and for O ($r = 0.786$, $p < 0.01$), and in order to avoid effects of co-variables, two different models were run to test the effect of prosodic branchingness. Note that due to the small amount of data, we did not include speaking style as a predictor.

3.1. S prosodic phrasing

First, a model was run with S phrasing status as the dependent variables (possible value: within an independent MaP/within the same MaP as V), and the following variables as predictors: syntactic branching of the element (branching/non-branching), number of syllables in S, number of syllables in V, number of syllables in O, the interaction between the number of syllables in S and the number of syllables in V, and the interaction between the number of syllables in V and the number of syllables in O.

Apart from syntactic constraints, results of the model revealed that the articulation rate has a significant impact on the phrasing of S ($\beta = 0.21$, $z = 20.114$, $p < 0.001$). As it can be seen in Figure 1 below, syllabic mean duration of the SVO sentence is greater (articulation rate slower) in cases where S is phrased as an independent MaP.

An effect of the number of syllables in the subject phrase was also found ($\beta = 0.404$, $z = 4.507$, $p < 0.05$). This shows that the more syllables are contained in S, the bigger the chances for S to be phrased as an independent MaP. In other words, an S containing 12 syllables stands in greater chance to be phrased in a separate MaP, separated from the verb, in comparison with an S containing 2 syllables. Prosodically embedded S constituents contain fewer syllables than prosodically independent S constituents; this can be seen by comparing the mean number of syllables in S when it is phrased as an independent MaPP and when it is phrased in the same MaPP as V (see Figure 2).

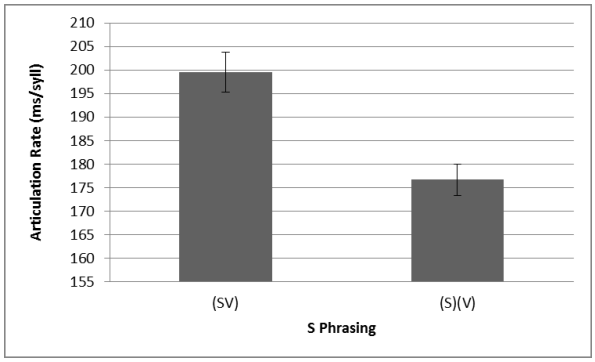

Figure 1: *Articulation rate (in ms/syll) by mean of S phrasing. Errors bars are standard errors from the mean.*

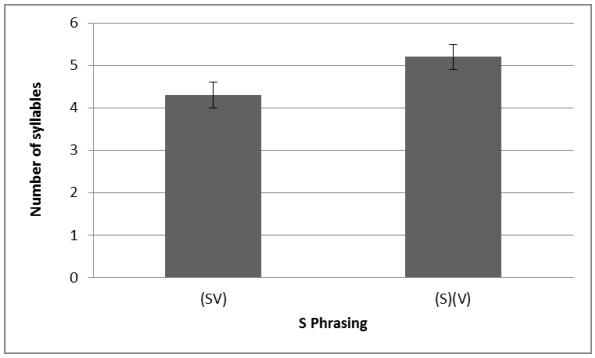

Figure 2: *Number of syllables by mean of S phrasing. Errors bars are standard errors from the mean.*

A marginal effect of the number of syllables in V was also observed ($\beta = 0.582$, $z = 2.896$, $p = 0.079$), revealing that the more syllables V contains, the greater the chances for S to be phrased in a single MaP. Nevertheless, it appeared that the number of syllables in S and the number of syllables in V interact ($\beta = 0.109$, $z = 6.543$, $p < 0.01$). The effect shows that when S is long and V is short, the chances to find a Major Prosodic Phrase boundary after S are stronger than when S is short and V is long. In the latter case, there are smaller chances for S to be followed by an MaP boundary.

The same model with the number of PWs instead of the number of syllables per constituent show similar results, with an effect of articulation rate ($\beta = 0.011$, $z = 6.087$, $p < 0.05$), an effect of the number of PWs in S ($\beta = 1.335$, $z = 6.520$, $p < 0.05$).

3.2. O prosodic phrasing

A first model was run with O phrasing status as the dependent variable (possible values: within a single MaP/within the same MaP than V), and the following variables as predictors: syntactic branchingness of the element (branching/non-branching), number of syllables in S, number of syllables in V, number of syllables in O, the interaction between the number of syllables in S and the number of syllables in V, and the

interaction between the number of syllables in V and the number of syllables in O.

Only the articulation rate appeared to be a significant predictor in this model. Articulation Rate had impact on O phrasing ($\beta = 0.21$, $z = 119.745$, $p < 0.001$), and, as can be seen in Figure 3 below, in the cases where O is phrased as an independent MaP, syllabic mean duration of the SVO sentence is greater (articulation rate slower).

The same model with the number of PWs instead of the number of syllables per constituent produced similar results, again showing solely an effect of articulation rate ($\beta = 0.021$, $z = 13.554$, $p < 0.001$).

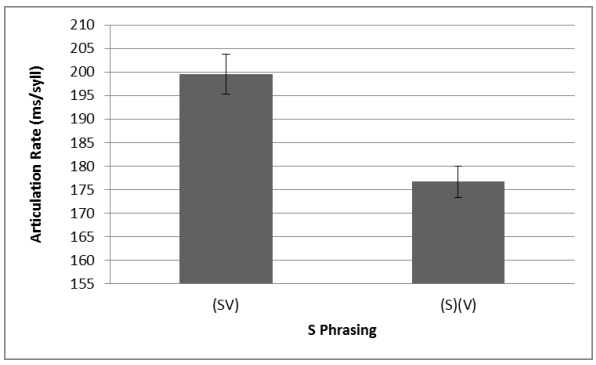

Figure 3: *articulation rate (in ms/syll) by mean of O phrasing. Errors bars are standard errors from the mean.*

4. Discussion

The analysis revealed interesting results regarding the phrasing pattern for SVO, and the effect of the different non syntactic factors constraining prosodic phrasing.

Regarding phrasing patterns, it appeared that S can be phrased with or without V with almost the same probability (in 54.1% of the data, S is phrased as an independent MaP, whereas in 45.6% of the data, it is in the same MaP than V). This is an important result, which indicates that the phrasing of S in a single MaP is not the default pattern in French, contrary to what has been claimed in the literature. In addition, this distribution of the various phrasing patterns shows that French differs from other Romance Language, such as Spanish and Catalan, where S is usually phrased in an independent MaP (see [11] and [12]). On the other hand, Object NP constituents seldom form a single MaPP: they are phrased with the V in more than 90% of the cases.

Regarding the factors constraining phrasing, our results show that syntactic complexity does not seem to be an important parameter to predict the phrasing of SVO structures. Indeed, the effect of syntactic branchingness of either O or S was never significant in our data: it does not help to explain why in some cases S or O are phrased separately or in the same MaP than the V. Articulation rate has a significant effect on phrasing of S and of O, supporting what has been found in the literature regarding the effects of articulation rate on phrasing (see [22] and [23]), i.e.: the faster a speaker articulates, the greater the chances for him to obliterate prosodic boundaries, and therefore to phrase the subject in the same MaP than the V. The number of syllables does not have any effect on the way the object NP is phrased: short or long Os are phrased in the same MaP or in an independent MaP because of different reasons. However, it clearly has an effect on the way the NP subject is phrased, supporting the idea the heavy NP subjects tend to be phrased in independent prosodic phrases, and light NPs to be embedded in one MaP along with the verb that follows (see [8]). It appeared that the number of syllables in V interacted with the number of syllables in S, confirming the idea that balance effects also have an importance on phrasing (see [8]). When calculated in term of prosodic branchingness, i.e. in terms of number of PWs, the lengths of V and O have not been observed to have any effect on phrasing.

It is necessary to conduct further analyses on a larger set of data in order to determine which parameters are the most robust for evaluating prosodic weight in French (number of syllables or number of prosodic words), in particular when analyzing phrasing patterns in SVO structure. According to [11], the choice of on parameter over the other is language-specific.

5. Conclusions

The aim of this paper was to test the impact of different phonological and syntactic factors on prosodic phrasing of French SVO structures. A set of approximately 200 sentences were extracted from 4 corpora of different speaking styles, and were semi-automatically annotated to study prosodic phrasing at the level of the Major Prosodic Phrase. The results have shown that NP subjects could be phrased within the same MaP than the verb, or in an independent MaP. The decision for one phrasing over the other is mainly correlated with articulation rate (the faster the speaker articulates, the greater the chances for the subject NPs to be phrased in an independent MaP) and on prosodic weight, be it calculated in terms of number of syllables or in terms of number of PWs in S (the longer the S, the greater the chances for S to be phrased as an independent MaP). Regarding the phrasing of O, results indicate that this constituent was rarely phrased autonomously. Furthermore, whenever it was realized as an independent MaP, the phrasing obtained was motivated by a slow articulation rate, and not by metrical properties (expressed in terms of number of syllables or in terms of number of PWs). Finally, for both S and O, no effect of syntactic branching was found. Further developments are needed to confirm these results, in particular in enlarging the set of analyzed data.

6. Acknowledgements

This paper is related to the work package "Prosodic phrasing and prosodic hierarchy: a data driven approach" of the Labex Empirical Foundations of Linguistics and is supported by the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083). It is also partly achieved within the research project ANR-CONTINT 2011 PHOREVOX funded by ANR/CGI.

7. References

[1] Selkirk, E., "On derived Domains in Sentence Phonology ", Phonology 3, pp. 371-405, 1986.

[2] Nespor, M., & Vogel, I., Prosodic Phonology, Foris Publication, 1986.

[3] Inkelas, S., & Zec D., 'Syntax-phonology interface', in J. Goldsmith (ed), The Handbook of phonological theory, Blackwell Publishers, pp. 535-549, 1995.

[4] Selkirk, E., 'The Syntax-Phonology Interface', in J. Goldsmith, J. Riggle & A. Yu (Eds.), The Handbook of Phonological Theory, Blackwell Publishing, pp. 435-484, 2011.
[5] Delais-Roussarie, E., Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française, Thèse de Doctorat, Université de Toulouse-Le Mirail, 1995.
[6] Rossi, M., L'intonation, le système du Français : description et modélisation, Ophrys, 1999.
[7] Dell, F., 'L'accentuation dans les phrases en français', in F. Dell, D. Hirst & J.R. Vergnaud (Eds.), Forme sonore du langage: structure des représentations en phonologie, Hermann, pp. 65-122, 1984.
[8] Delais-Roussarie, E., "Phonological phrasing and accentuation in French", in M. Nespor & N. Smith (Eds.), Dam phonology : HIL phonology papers II, Holland Academic Graphics, pp. 1-38, 1996.
[9] Ghini, M., "Φ-formation in Italian: A new proposal", Toronto Working Papers in Linguistics 12, pp. 41-78, 1993.
[10] Martin, P., Prosodic and Rhythmic Structures in French. Linguistics, 25, 925-949, 1987.
[11] Elordieta, G., Frota, S. & Vigário, M., 'Subjects, objects and intonational phrasing in Spanish and Portuguese', Studia Linguistica 59, 110-143, 2005.
[12] D'Imperio, M., Elordieta, G., Frota, S., Prieto, P., and Vigário, M., "Intonational Phrasing in Romance: The role of prosodic and syntactic structure", in S. Frota, M. Vigário & M.-J. Freitas (eds), Prosodies, Mouton de Gruyter, 59-97, 2005.
[13] Durand, J., Laks, B., & Lyche, C. (Eds.), Phonologie, variation et accents du français, Hermes, 2009.
[14] Boersma, P., & Weenink, D., Praat (Version 5.3), Retrieved from http://www.fon.hum.uva.nl/praat/, 1995-2013.
[15] Goldman, J.-P. EasyAlign: an Automatic Phonetic Alignment Tool under Praat. Proceedings of Interspeech.
[16] Christodoulides, G., Avanzi, M., Goldman, J.-P. 'DisMo: A Morphosyntactic, Disfluency and Multi-Word Unit Annotator. An Evaluation on a Corpus of French Spontaneous and Read Speech'. Proceedings of LREC, 2014.
[17] Mertens, P., Goldman, J.-P., Wehrli, E., & Gaudinat, A., 'La synthèse de l'intonation à partir de structures syntaxiques riches', Traitement Automatique des Langues, 42, pp. 145-192, 2001.
[18] Goldman, J.-P., Auchlin, A., Roekhaut, S., Simon, A. C., & Avanzi, M., 'Prominence Perception and Accent Detection in French. A Corpus-based Account' in Proceedings of Speech Prosody, 2010.
[19] Frota, S., "Prosodic structure, constituents and their representations", in A. Cohn, C. Fougeron & M. Huffman (eds), The Oxford Handbook of Laboratory Phonology Oxford, Oxford University Press, Chapter 11, pp. 255-265, 2012.
[20] Avanzi, M., Obin, N., Lacheret-Dujour, A., & Victorri, B. (2011). Toward a Continuous Modeling of French Prosodic Structure: Using Acoustic Features to Predict Prominence Location and Prominence Degree. Paper presented at the Proceedings of Interspeech.
[21] Miller, J.L., Grosjean, F., Lomato, C., "Articulation rate and its variability in spontaneous speech: A reanalysis and some implications". Phonetica, 41, 215-225, 1984.
[22] Fougeron, C., & Jun, S. A. (1998). Rate Effects on French Intonation: Prosodic Organization and Phonetic Realization. Journal of Phonetics, 26, 45-69.
[23] Post, B. (2011). The multi-facetted relation between phrasing and intonation contours in French. In C. Gabriel & C. Lleó (Eds.), Intonational Phrasing in Romance and Germanic: Cross-linguistic and bilingual studies (pp. 43–74). Amsterdam: John Benjamins.