# Syntax and syllable count as predictors of French tonal groups: Drawing links to memory for prosody

Annie C. Gilbert & Victor J. Boucher

Université de Montréal, Montréal, Canada annie.gilbert@umontreal.ca

## Abstract

While the role and origin of prosodic structures remain unclear, there is evidence that prosody bears an intriguing relationship with serial memory processes and grouping effects. This link is seen in the fact that the recall of presented prosodic patterns and their production in speech are both restricted in term of a syllable count. The present experiment complements previous studies by examining the effects of syntactic structure as opposed to constituent length on produced tonal groups. Forty subjects produced, in quasispontaneous conditions, given utterances with differing NP, VP structures or differing lengths. The results show that constituent length is the major predictor, whereas syntactic structure appears as a secondary factor.

### 1. Introduction

Most studies of prosody attempt to model rhythm and tone by by referring to syntax, seen as either a primary factor of prosodic organization [1][2] or a basic element of prediction [3]. However, prosodic units present their own logic and are not congruent with syntactic structures. Many adjustments have to be made to derive correct prosody from bare syntax. On this problem, some authors have suggested that syntax alone is not a sufficient predictor of prosody [4] and have proposed complex algorithms to model prosodic structures (or performance structures) of sentences. Yet, as Gee and Grosjean note, these algorithms "are not in themselves interesting unless they are steps on the way to a model of the human mind" [5 p.449]. In other words, algorithms or formal model do not as such explain the role and origin of prosodic structure. Studies have also revealed that certain aspects of prosodic patterns operate irrespective of syntax.

For instance, a number of reports on both spontaneous and read speech in different languages converge on the point that stress groups or inter-stress intervals are restricted by a number of syllables that has an average upperlimit of 3.55 syllables [6][7]. Similar results were found for French by Monnin & Grosjean [4] who observed "basic prosodic groups" with an average length of 3.46 syllables. These observations of prosodic groups bear a striking similarity with known rhythmic grouping effects on serial memory. In fact, it has been established that recall of series of items is optimal when items are presented in groups of 3 or 4. [8][9][10][11][12] In fact, a direct link between such rhythmic grouping effects and stress groups in French speech was demonstrated by Boucher [13] (see also Boucher and Gilbert 2005 [14]) using a serial memory test paired with a semi-spontaneous production task.

This link with respect to memory has also been shown to operate for French tonal groups in a study by Gilbert and Boucher [15] In that report, the tonal group (or intonation group) was defined as a prosodic constituent where a pitch contour encompasses rhythmic groups and syllables. The boundary between successive tonal groups can be marked either by a change in pitch directionality (up or down) or by a pitch reset, observed by reference to fundamental frequency (F0). It was found that the maximum length of a tonal group in spontaneous French was 8 syllables (average of 5.28, see figure 1), which is consistent with the results of Sityaev [16] who showed that tonal groups in non-formal English speech generally contain two (2) stress-groups. Based on these results, a test was designed involving the recall of tonal groups, with group lengths matching those that were probable in speech. In the test, subjects were asked to reproduce sequences of repetitive non-sense syllable that had been resynthesised in order to create tonal groups. They were asked to reproduce what they heard as accurately as possible and the instructions were purposely unspecific. The results were revealing. The ability to reproduce long tonal groups (five syllables or more) was closely linked to their relative occurrence in spontaneous speech. (see figure 2)

Figure 1. % Production of tonal groups in spontaneous speech as function of length (no. syllables) and cumulative %



Figure 2. % Production of tonal groups in spontaneous speech as function of length (syllables) and memory for tonal groups performance.



At first glance the aforementioned findings on rhythm and tonal groups are difficult to interpret from a linguistic perspective, since prominent authors like Chomsky overtly excluded memory from having any role in language competence [17], which as such excludes factors of memory in linguistic models. But one can reason that learning a language requires the ability to reproduce and hence memorize and recall novel series of prosodically structured sounds constituting new words or expressions. Any restrictions on the reproducibility of heard prosody would contribute to restrict prosodic structures in one's speech to a given set of "easier" patterns. Harder to learn prosodic constituents or structures would not be learnt or used, making them less frequent in speech.

In view of the above results, and without denying the role of syntax, it appears that constraints on the number of syllable units could present an organizational factor of prosodic patterns. The present experiment was designed to verify if a syllable count constraint relying on memory capacity could better predict prosody than syntax alone. That is, the purpose was to examine if the syllable count principal, observed to apply in the case of tonal groups in French, serves to predict tonal group production better than syntax.

To test the reliability of the syllable count principle, subjects were asked to perform a semi-spontaneous or directed speech task. This kind of task is a useful way to test linguistic prosody since there is no auditory stimulus involved that could bias the subject's production. In the procedure subjects memorize a sentence that is momentarily presented on a screen and then recall the sentence upon a blanked display. It should also be noted that French offers a particular advantage in this experiment given that stress is not lexically coded but occurs at group-final positions. The aim, then, was to obtain "semi-spontaneous" speech of different phrase lengths and to compare the results to predictions by reference to the syntax and syllable count variations in the stimuli.

## 2. Method

## 2.1. Subjects

Forty (40) native speakers of French (21 males, 19 females) aged between 19 and 32 years of age (average 24.63 years) participated in the study. Subjects were required to have no

history of speech or hearing deficits. Information concerning the place of birth and level of music and dance training was obtained by an informal questionnaire but post-hoc comparisons for these factors showed no significant differences.

#### 2.2. Stimuli

Two sets of sentences were designed with a same number of syllables (12). The basic syntactic structure (SVO) was maintained within each set but unequal stress groups were used to avoid a possible bias of rhythmic symmetry.

In the first set of sentences, the independent variable was the relative location of the major syntactic boundary between the subject (NP) and the verb (VP). The direct object remained constant at 7 syllables. Table 1 presents the set of sentences used to test whether syntax or syllable count would better predict the tonal groups.

In the second set of sentences, the independent variable was the length of the whole VP. This one was always designed in order to contain the verb and its direct object, so that there were no major syntactic boundaries within the VP. The subject is always a proper noun (NP) of 3 to 5 syllables. Pretests have shown that longer subject NPs are normally produced in a tonal group of their own, so that another tonal group would be initiated at the following VP. Table 2 presents the sentences used to verify if long VPs are subject to a syllable-count principal no matter their similar syntactic structure.

Table	1. A	lumber	• of syl	lables	before	and	after	major	syntact	ic
break.										

NP	VP			
Luc (1 syll)	entretiendrait (4 syll)	les rosiers décoratifs.		
5	syll	7 syll		
Martin (2 syll)	apprécie (3 syll)	la musique instrumentale.		
5	syll	7 syll		
Dominic (3 syll)	ajoute (2 syll)	des brisures chocolatées.		
5	syll	7 syll		
Maximilien (4 syll)	vend (1 syll)	des chatons vermifugés.		
5	syll	7 syll		

Table 2 : Number of syllables in constituent.

NP	VP			
Alexandrina	regardait les candidats.			
(5 syll)	(7 syll)			
Élizabeth	psychanalysait les enfants.			
(4 syll)	(8 syll)			
Anabelle	investissait ses économies.			
(3 syll)	(9 syll)			

#### 2.3. Data collection

Subjects sat in a noise-attenuating booth. The stimuli were presented visually on a computer monitor in bold centred Arial Unicode 45 points with white characters on black background. Subjects were asked to first read the sentence aloud. Then the screen was blanked and the subject had to recall the sentence. Only the recalled version was retained in the analysis. The experiment was controlled manually from outside the booth to insure a steady pace of stimuli presentation and to verify that the subjects read and recalled the sentences without errors. When recall involved a change in constituent length (in the number of syllables), the stimulus was resubmitted (rarely more then once) until it was produced with the correct number of syllables.

Responses were digitally recorded using an AKG C477 WRL head-mounted microphone hooked up to an external sound card (CSL4400, Kay Elemetrics), providing a 16-bit digitization at 44.1 kHz.

#### 2.4. Analysis

The pitch measures were obtained using the peak-detection routines of the aforementioned analysis system (CSL4400). Three different pitch analyses were overlaid to avoid biases that could be created by certain consonants (i.e. plosives or high frequency fricatives) or an inadequate frame length or frame advance setting. The settings were as follow: length 20 ms/20 ms advance, length 25ms/20 ms advance, and length 25ms/25 ms advance. Only matching F<sub>0</sub> contours from at least two analyses served to identify tonal group boundaries. Boundaries were defined by reference to typical resets or changes in the directionality of the slope of  $F_0$ [18][19][20][21][22][23][24] as observed in a visual display of the three pitch contours overlaid on a spectrogram (using a Blackman 512 pt window). Syllable counts within the contour were performed by visual inspection of the spectrograms while simultaneously listening to portions of the signal. Figure 3 is an example of the visual display of the analysis procedure.

Figure 3: Three pitch analyses (log scale) overlaid onto spectrogram as performed during data analysis.



### 2.5. Results

For the first set of sentences, the presence of a major syntactic break was not a stable predictor of produced tonal groups Prediction by reference to syntax fluctuated (from 22.5% to 82.5%) and increased with the number of syllable contained in

the SN subject. That is, the longer the SN, the more chances of it being produced in a single tonal group. (See figure 4)

An inverted pattern of results was obtained with the second set of sentences. In this case, two tokens had to be discarded for two speakers who failed to produce the subject-SN as an independent tonal group (as required in the design). The results showed, in the second set of sentences, that the chances of the VP being produced in one single tonal group decreases with the number of syllables involved even if the basic syntactic structure remains the same. The longer VPs (longer than what is usually produced in speech) are rarely produced within one tonal group, but the shorter ones can quite easily be produced this way. (See figure 5)

Figure 4. % Correct predictionFigure 5. % Productionfrom syntax as function ofof single tone group aslength (syll)function of length (syll)



#### 3. Discussion

The data presented in Figures 4 and 5 show that, even when syntactic structure is held constant, we still get differing realisation of tonal groups based on the number of syllables. A revealing effect was that, as the length of syntactic constituents approach the average length (in no. of syllables) of tonal groups in speech, there is a greater probability that the constituent can be produced as a single tone group (see Figure 4). The relationship is further confirmed in considering the sharp drop in the production of single tonal groups for constituents of more than 8 syllables in length, and this drop corresponds to the 8-syllable limit for tonal groups observed in speech [15]. It should be noted that this correspondence cannot be associated with a tendency to create symmetrical structure. In fact, all of the sentences used as stimuli contained 12 syllables. Consequently, the length variation found between tonal groups cannot be accounted for by a tendency to create symmetry, otherwise every sentence would have been produced the same way.

Overall, the present data are consistent with earlier findings concerning French speech prosody and show the major effect of a syllable count principal over syntactic structure in predicting tonal groups. Syntax remains a minor factor in determining the prosodic structure found in speech and its potential as a predictor varies as a function of the length (syllable count) of the constituent. Syntax, however, does exert an influence, for longer constituents, in the location of the tonal group boundary.

## 4. Conclusion

These above results bear on speech contexts where subjects might produce near-spontaneous tonal-group patterns. The effects of syntactic structure on these patterns are minor compared to the effects of the number of syllables in the syntactic constituents. Such syllable-count effects may be explained by reference to constraints on memory processes, as suggested in previous work.

## 5. References

- [1] Hirst, D. (1993). Detaching intonational phrases from syntactic structure. *Linguistic Inquiry*, *24*, 781-788.
- [2] Scholes, R.J. (1971). Acoustic cues for constituent structure. The Hague, The Netherlands.: Mouton & Co.
- [3] Martin, P. (1991). A propos du statut théorique de la structure prosodique de la phrase. *Verbum*, 14(2-4), 307-310.
- [4] Monnin, P., & Grosjean, F. (1993). Les structures de performance en Français: Caractéristiques et prédiction. L'Année psychologique, 93, 9-30.
- [5] Gee, J. P., & Grosjean, F. (1983). Performance Structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15, 411-458.
- [6] Dauer, R.M. (1983). Stress-timing and syllabe-timing reanalyzed. *Journal of Phonetics*, 11, 51-62
- [7] Martin, P. (1999). Intonation of spontaneous speech in French. Paper presented at the International Congress of Phonetic Sciences, San Francisco.
- [8] Maybery, M.T., Parmentier, F.B.R. & Jones, Dylan M. (2002) Grouping of list items reflected in the timing of recall: implications for models of serial verbal memory. *Journal of Memory and Language*, 47, 360-385.
- [9] Reeves, C., Schauder, A. R., & Morris, R. K. (2000). Stress grouping improves performance on an immediate serial list recall task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1638-1654.
- [10] Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly Journal of Experimental Psychology*, 49A, 116-139.
- [11] Frankish, C. (1995). Intonation and auditory grouping in immediate serial recall. *Applied Cognitive Psychology*, 9, 5-22.
- [12] Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 469-479.

- [13] Boucher, V.J. (to appear) On the function of stress rhythms in speech: Evidence of a link with grouping effects on serial memory. *Language and Speech*.
- [14] Boucher, V.J. & Gilbert, A.C. (2005). Memory for tone contours and stress rhythms: evidence of a correspondence with the prosodic organization of meaningful speech. Proceedings of the XXVII Annual Meeting of the Cognitive Science Society, 328-330
- [15] Gilbert, A. C., & Boucher, V. J. (2005). Constraints on memory for tonal patterns : A structuring principle of speech prosody. Presented at the 9th European Congress of Psychology, Granada, Spain.
- [16] Sityaev, D. (1999). Distribution of pitch accents in intonation groups across intonational styles. Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, 1505-1508.
- [17] Chomsky, N. (1967). The formal nature of language. In E. H. Lenneberg. *Biological Foundations of Language*, p. 397-442. New York: John Wiley & Sons.
- [18] Zitter, A.D. (1992). The perceptual salience of melodic variation: contour shape and peak height. *Journal of Phonetics*, 20, 181-188.
- [19] Post, B. (1999). Solving a controversy in the analysis of French rising pitch movements. Paper presented at the International Congress of Phonetic Sciences, San Francisco.
- [20] Ladd, D.R. (1983). Peak features and overall slope. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 39-52). New-York: Springer-Verlag.
- [21] Vaissière, J. (1983). Language-independent similarities. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 53-66). New-York: Springer-Verlag.
- [22] Delattre, P. (1966). Les dix intonations de base du français. *The French Review*, 40, 1-14.
- [23] Kenning, M.-M. (1983). The tones of English and French. *Journal of Phonetics*, 13, 32-48.
- [24] Cooper, W.E., & Sorensen, J.M. (1981). Fundamental frequency in sentence production. New-York: Springer-Verlag.

## 6. Aknowledgements

This research was founded by an SSHRC graduate scholarship awarded to the first author.