

Prosodic Characteristics Associated with The Clause in Venezuelan Spanish: an Acoustic, Audio-Perceptive Analysis

Elsa Mora and Hernán Martínez

Universidad de Los Andes. Mérida, Venezuela
hjmart@gmail.com

Abstract

This study presents a series of analyses related to the prosodic characteristics of clauses in the Mérida dialect of Venezuelan Spanish. We based this study on two analyses: one acoustic and one audio-perceptive. The parameters we analysed were the distribution and nature of pauses, the variations on fundamental frequency at the beginning and at the end of clauses, the different types of accents in both stressed initial and final syllables and the duration of stress syllables both in initial and final positions. The results suggest that clauses have certain prosodic features that identify them. Our data reveals that prosody leads to the recognition of syntactic-discursive structures in the spoken language.

1. Introduction¹

Speech perception is a very complex phenomenon involving a series of events such as sound, word and utterance identification. In this process, one important factor is the perception of acoustic-prosodic parameters, i.e., the prominences and the boundaries marked by accents, tones and intonation. However, this is not as simple as it seems. Prosody has a essential role in speech perception, structuring of speech, lexical retrieval, comprehension of oral messages and its syntactic structuring, expressing of emotions and in other functions associated with linguistic and communicative competence.

To perceive the spoken language, listeners perform a segmentation process, which is the most important task to decode a message. This segmentation process is also associated with the boundaries found in the speech chain, which in turn are associated with syntactic and prosodic units.

The speech flow does not indicate --in most cases-- word boundaries. The speaker must use prosodic elements to mark them. Wingfield & Titone [1] claim that during segmentation, speakers have to group words into clauses, i.e., they must mark the syntactic boundaries. Next, they have to identify larger units in the paragraph or semantic structures. Moreover, they have to perceive, based on the syntactic-prosodic structures, the organization of the structure according to the interaction, or turn taking, and the situational context [2]. Prosody helps to carry out all these segmentation tasks. In the different language levels, "listeners bring forward their linguistic experience to heuristically perform the segmentation according to the limits that their native language imposes on words, clause and sentences" [1].

Therefore, we can infer that at the perceptive level there exists a close link between prosody and syntax in the segmentation of spoken language. There is a perceptive correspondence between the melody and rhythm of a language and its syntactic and sense groups. From the perceptive point of view, prosody comes to be the phonetic realization of syntactic structures.

In the case of the listener's task of segmentation, we depend on pauses (real or virtual) to perceive and segment a message. Virtual pauses are those which are not marked by silence but where there is a change in fundamental frequency or the lengthening of end of the syntactic or prosodic unit. Blondet [3] notes that these pauses have a syntactic function. On the perceptive level, pauses are not the only features that mark the limits of information units or clauses, but this is also done by other prosodic elements [4].

These information units do not quite correspond with traditional grammar sentences, since they not always have a subject, a verb and an object. Prosody is very important to determine these units, usually called *clauses*. For Halliday [5] and Chafe [6], the clause coincides with the information unit that is realized in the same intonation contour. This implies that, in spoken language, "it is plausible that these units have different characteristics depending on the information they transmit and the kind of relationship there is among them" [7]. Quillis [8] notes that "intonation is [...] the most common and elemental resource in an utterance: there can be utterances without grammatical form, but none without intonation." The effect of all fundamental frequency indices plus duration allows listeners to identify information units [4]. We propose that the clause is the unit in which information is presented and developed in the spoken language. Moreover, we can say that what characterises the clause in spoken language is that it has a distinctive intonation contour, is found between pauses, has syntactic independence, is semantically self-sufficient and lacks a strict correspondence with the /subject+predicate/ structure (in this characterization, mainly three types of elements intervene: meaning content, form and physical characteristics). The clause is informative, communicative and pragmatic. The different prosodic signals define the number of clauses and their boundaries in spoken language.

According to Bradlow & Pisoni [9], syntax and prosody affect each other to achieve meaning effects. Therefore, syntactic and semantic structures have substantial effects on the configuration of the fundamental frequency, on the duration patterns and on the intensity of sounds. According Schuetze-Coburn [10] the prosodic signals (fundamental frequency, pauses, stress and duration) are used by listeners to define and segment the units in the spoken language.

Some studies that analyse the interrelation that exists between prosody and syntax follow an experimental approach since it can be argued that prosody play a major

¹ This study was financed by the CDCHT-ULA with the project H-942-06-06-B.

role in the identification of discursive units. For instance, Hirschberg & Litman (1993. In [11]) found that “intonational phrasing and pitch accent play a role in disambiguating cue phrases, and hence in helping determine discourse structure”. According to Ferreira (1993, 2002. In [12]) “durational differences signal syntactic structure even in the absence of other prosodic markers, and are correlated with other prosodic cues. Moreover, comprehension studies find that listeners can use durational cues to correctly interpret an ambiguous syntactic structure even in the absence of a disambiguating pitch contour”.

From the perceptive point of view, some studies concluded that prosody can provide listeners useful cues during the processing of ambiguous utterances. According Clifton, Carlson & Frazier [13] the production of a prosodic chunk (e.g. pauses) in spoken language could affect its comprehension in an independent context. Likewise, there is recent evidence that indicates that listeners use global prosodic representations during in the processing of spoken language (Schafer, 1997. In [13]).

According to Doncel [14] in Dutch these “boundaries” can be prosodically marked by the speaker using, for example, a pause, along with a falling f_0 movement. Other possible marks are the lengthening of the final syllable and a variation in intensity. Van Donzel adds that it is the listeners who control the status of the information received to be segmented. According to this author’s analysis, listeners perceive two types of syntactic structures (clauses): ‘non-final’ and ‘utterance final’. The data from his study prove that the clause is the unit for the processing of spontaneous speech. These units coincide with those a listener can perceptively process during a communicative exchange. Similar results are reported for Swedish in the studies carried out by Strangert & Strangert [15].

For Spanish, Mora & Cavé [4], applying the metric autosegmental theory, report that the intonation contour in the case of continuative or non-terminal statements, the most frequent tone is L-H%, and only seldom H-H%. According to these authors, it is the tonal level what determines the intonation unit [4]. The configuration of the intonation contour, the tonal level and the final lengthening are key factors in determining the syntactic boundaries of terminal and non-terminal intonation units in spoken discourse.

On the other hand, Lacheret-Dujour & Beaugendre [16] suggest that only the semantic function (resolving structural ambiguities when the syntactic indicators are not clear) has and effect on intonation structures. Intonation guides the listener to a possible segmentation and consequently -- beyond syntax -- to a semantic interpretation. Intonation has the fundamental function of eliminating ambiguities, as well as supporting and reinforcing the syntactic organization.

2. The study

Based on acoustical and perceptive analyses, this study aims to determine the prosodic characteristics of clauses in the Mérida dialect of Venezuelan Spanish. Specifically, we focus on the combination of various prosodic aspects, such as the type and distribution of pauses, changes in F_0 at the beginning and end of clauses, types of accents found on initial or final stressed syllables, and duration of stressed and unstressed syllables, both in initial and final position. The purpose of this analysis was to determine how these parameters change according to the clause structure. Based on these data, we draw a conclusion about the prosodic classification of discourse and its conditions.

3. experimental methods and procedures

3.1. Corpus

The corpus used in the study consists of 6 samples of spontaneous speech taken from the *Corpus lingüístico del habla de Mérida* [17] which contains recordings of conversations between Merida Spanish speakers belonging to different socioeconomic and age groups. These samples had previously been transcribed orthographically and segmented into clauses.

3.2. Audio-perceptive analysis

The 6 audio recordings and their orthographic transcriptions (with no punctuation marks) were presented to 12 listeners, who were asked to mark the pauses they heard in the recordings. A total of 683 clauses was obtained from this experiment. This number of clauses corresponded almost exactly (96.8%) with the segmentation previously done

3.3. Acoustic analysis

An analysis of the main acoustic properties of the clause units segmented by the participants in the perceptive analysis was carried out using the computer program *Speech Analyzer*. The acoustical parameters measured were duration and fundamental frequency. The unit of analysis was the syllable. The duration of external boundaries of the intonation units was taken from the duration of the first unstressed and stressed syllables as well as from the duration of the last stressed syllable and the last syllable in unit. In the analysis of the fundamental frequency, values of f_0 were taken in the middle of the first unstressed syllable and in the middle of the first accented syllable, as well as in the middle of the last unstressed syllable and of the last syllable in the unit. Also the duration of initial and final pauses in each clause was measured. Then it was determined whether the characteristics acoustic-prosodic coincided in all of the corpus clauses.

4. Results and discussion

4.1. Pauses

From the different acoustic analyses carried out, we obtained the following results. Firstly, we should mention that the initial pauses in the clauses we chose were grouped according to their duration and to their syntactic structures. The differences between the different types of initial pauses was significant (.0001):

1. (IP1) Initial pauses longer than 0.9 seconds, which seem to precede other clauses in which the processing or planning of what is to be said takes place. That is, these are pauses that precede clauses presenting the information or theme. Discursive markers are present in the clauses preceded by this type of pauses. The average duration of initial pauses is 1.40 seconds.
2. (IP2) these are initial pauses with a duration between 0.1 and 0.5 seconds (the average duration of these pauses is 0.3 sec.) which seem to precede clauses that develop the information previously introduced. In the clauses preceded by these pauses we find descriptions, enumerations or narrations, and they usually present discursive markers such as *y* (and) or *y entonces* (and so).
3. (IP3) Finally, there are initial pauses lasting between 0.6 and 0.9 seconds. These pauses seem to mark the beginning of a clause closing the development of information in the major syntactic-semantic unit. The clauses marked by these pauses seem to complete the meaning of the larger unit. This corresponds with the data obtained from the perception test, in which the listeners

recognized the clauses marked with pauses as final units of the larger sense unit. The clauses that have these pauses also present discursive markers such as *¿no?*, *¿ves?*, *¿me entiendes?*, *¿me sigues?* [18]. The average duration of these initial pauses is 0.78 seg.

Based on this data, it can be said that in the spoken language there is larger syntactic-semantic unit coded in the clauses and that these have specific prosodic marks. Specifically, clauses are found in larger sense units according to the duration of the preceding pause.

On the other hand, we found a significant difference (.0027) between to types of final pauses also related to syntactic structures. These are:

1. (FP1) Final pauses lasting between 0.1 and 0.5 seconds, which seem to precede the clauses in which the information is developed. That is, the pauses lasting between 0.1 and 0.5 seconds are those that follow clauses with initial pauses of 0.3 seconds.
2. (FP2) Final pauses lasting more than 0.9 which usually follow both syntactic constructions in which the planning takes place and clauses beginning a major syntactic-semantic unit.

For instance,

(PI1) *segundo y tercer año lo estudié en La Inmaculada*(PF1)

(PI2) *allá éramos un poquito más traviesas*(PF1)

(PI2) *era de tirar los pipotes de basura por las escaleras y todo*(PF1)

(PI2) *y a la final nadie era*(PF1)

(PI2) *todo mundo tenía que ir...*(PF1)

(PI2) *todo mundo paraba en el patio castigado porque nadie...*(PF1)

(PI2) *nadie era culpable de lo que se hacía*(PF1)

(PI3) *todo el mundo se defendía.*(PF1) (md1fa)

These pauses are used to segment the discourse. This is based in the fact that, at a perceptive level, only a small part of the information stored in our brain can be activated and processed at a time; consequently, the flow of speech is segmented by pauses and organized into clauses that include just one item of information at a time [6].

Table 1. Types of pauses with duration averages and types of clauses related to them. (IP= Initial Pause; FP= Final Pause).

Type of pause	Duration	Related Clause
IP1	1,40	Introduce information or theme
IP2	0,3	Develop initial information
IP3	0,78	Close development of information
FP1	0,3	Develop information
FP2	0,9	Introduce information

4.2. Syllable duration

As for the duration of the initial syllables, we found that the average duration of the first unstressed syllable in the clauses analysed is 0.137 sec. and that of the first stressed syllable was 0.141 sec. Likewise, the average duration of the last stressed syllable is 0.164, while the average duration of the last syllable of the clause was 0.183 sec. The data shows that there is a significant difference between the two initial syllables (.0007) and the two final syllables (.0025), and between initial and final syllables (.0001).

Table 2. Duration averages for initial and final unstressed and stressed syllables (IUS= initial unstressed syllable; ISS=

initial stressed syllable; FUS= final unstressed syllable; FSS= final stressed syllable).

Type of syllble	Duration
IUS	0,137
ISS	0,141
UFS	0,164
SFS	0,183

We can conclude from these data that the duration of final syllables, both stressed and unstressed, is usually longer that the duration of the first syllable of the clause, which indicates that the final part of the clause is longer that the initial part. The temporal characteristic of intonation units (information units) is a beginning having two short syllables, associated to the first syllable and the first stressed syllable and the ending of two long syllables corresponding to the last accent in the intonation or information unit and to the last syllable in this unit, showing a typical final lengthening. Chafe [6] has called this pattern acceleration plus deceleration tempo. We can say that intonation units show a compression of initial syllables and a lengthening of final syllables. According to Mora & Cavé [4], this fact is typical in declarative clauses, which show initial temporal compression and final lengthening. This is a very common pattern: some phonetic components are longer before a boundary, as in the end of an utterance or an intonation phrase, than in other positions.

4.3. Fundamental frequency

On the other hand, the average fundamental frequency of the first unstressed syllable is 187.4 Hz, and the average of the first stressed syllable is 189.8 Hz. The averages of fundamental frequency of both last unstressed (177.6 Hz.) and last stressed syllables (184.2 Hz) are higher than those of the initial syllables. These results seem to correspond with the duration of clausal segments. Based on these data, we can say that the movement of f0 starts with a rise up to the first stressed syllable, where a gradual fall begins, which can continue falling, or be followed by a plateau or a rise, depending on the type of clause. As for the values of f0, there is a significant difference between the first unstressed syllable and the first stressed syllable (.0034), and the last stressed syllable and the last unstressed syllable (.0001).

Table 3. Fundamental frequency averages for syllables.

Syllable type	Fundamental frequency f0 (Hz)
SII	187,4
SIA	189,8
SFI	177,6
SFA	184,2

We also noticed that the f0 values at the beginning of the first clause of the text tend to be higher than the values corresponding the beginning of clauses that develop and end the text. Besides, the different positions that each clause takes can be clearly seen by the highest f0 initial value and by the lowest f0 initial value. Therefore, the intonational levels vary according to the "purpose" clauses have.

According to Kraljic & Brennan [12], prosody marks syntactic boundaries in several ways: "the final word of a phrase and the pause immediately following it may be lengthened; such durational differences (and the phonetic variations that accompany them) are correlated with suprasegmental markers such as pitch accents and boundary tones. Syntactic boundaries can be signalled via pitch contour variation such as a pre-boundary fall-rise or rise or

by a decrease in amplitude on the pre-boundary syllable. In short, a variety of phonological and intonational cues may be used in combination to mark syntactic structure”.

4.4. Perception

The data obtained from the acoustic analysis greatly coincide (96.8%) with the data from the perception test in which the subjects had to identify the pauses. The subjects indicated the final boundaries of all clauses, whether they introduce, develop or complete information. In other words, the pauses, as well as the configurations of the fundamental frequency and the duration of initial and final syllables, are used by the speaker to signal the segmentation of the information, i.e., to indicate the boundaries of the information units.

Based on this, we can say that there is some regularity in the grouping that takes place during speech perception. Also, speech perception requires to carry out several psychological operations such as to separate and segment words, phrases and larger units, as well as giving them a meaning. In other words, in order to organize the aural input, we have to be able to assign it an adequate linguistic structure with all its syntactic, prosodic, lexical and pragmatic characteristics.

The data presented here allow us to say that clauses are group into larger discursive information and processing units (or paragraphs). We say this based on the different f0 configurations, the duration of different types of the pauses analyzed, the syllabic lengthening, and the perceptive segmentation carried out by the listener. We suggest, along Garrido et al. [19], that there is a textual superstructure in spoken Spanish, which usually shows a tendency to declination. This has been observed in English, Danish and Dutch.

5. Conclusions

Studies on the segmentation marks have helped us to determine how speakers signal or mark the beginning and the end of long and short discursive units and how listeners recognize such units. This is fundamental in discourse analysis.

To sum up, we can consider that speech is structured into intonation units that correlate with information and syntactic units which are different to those established in classical grammar. These units can be identified by certain suprasegmental features: pauses at the beginning and at the end of the unit, an f0 contour rising at the beginning and falling at the end, a decrease in timbre and in clear phonetic configurations. Likewise, these units are characterized by a particular lexical-syntactic structure that makes them different.

Therefore, prosody groups the units that constitute the phonic flow and that present information: we could say that the prosodic contrasts have a grammatical value in the syntax of a language. As we have argued, listeners perceive sound signals that "pack" the information [6], [20]. Therefore, pauses, a long with the movement of f0 and duration, make possible the recognition of the syntactic-discursive structure of speech.

It is clear that speakers know a lot about the structure of their native language. Understanding the clauses present in the spoken language involves an amazingly active process, even if this is normally done quickly and without a conscious effort. Therefore, the processing of connected speech implies a continuous analysis of the aural input in order to identify the structure and meaning of utterances as they are heard. In the speech perception, all the prosodic and syntactic knowledge of the language are used together.

The results presented in this study show that there is a close relationship between syntax and prosody. Undoubtedly, prosody reflexes syntax and the syntactic hierarchy of

language. However, we do not claim that this is a one to one relationship; both disciplines are united at the textual level; however, semantic and, above all, pragmatic factors intervene in the production and processing of spoken language. Certainly, the semantic pragmatic information has much influence in the prosodic syntactic configuration of speech.

6. References

- [1] Wingfield, A.; Titone, D., 1999. Procesamiento de oraciones en *Psicolingüística*, Brenstein, N.; Berko Gleason, J.; Narasimhan, B. Madrid: MacGraw Hill.
- [2] Schegloff, E. 1998., Reflections on studying prosody in talk-in-interaction. *Language and speech*, 41 (3-4), 235-263.
- [3] Blondet, M., 1999. *Estudio acústico-prosódico de los fenómenos de hesitación: análisis contrastivo entre los dialectos andino y central*. Mérida: Universidad de Los Andes. Tesis de Maestría.
- [4] Mora, E.; Cavé, C., 2000. Fronteras internas y externas: límites de palabras y de unidades entonativas. *Entre lenguas*, 5 (2), 118-127.
- [5] Halliday, M.A.K., 1990. *Language, context and text: aspects of language in a social-semiotic perspective*. Oxford: University press.
- [6] Chafe, W., 1994. *Discourse, consciousness and time*. Chicago: Chicago University Press.
- [7] Elorrieta, G.; Romera, M., 2004. *Estudio experimental de las unidades prosódicas del discurso y sus funciones*. Círculo de Lingüística Aplicada a la Comunicación, 18. Disponible en: <http://www.ucm.es/info/circulo/no18/elorrome.htm>
- [8] Quillis, A., 1981. *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- [9] Bradlow, A.; Pisoni, D., 1998. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 12. 255-272.
- [10] Schuetze-Coburn, S., 1992. *Information flow in discourse*. PhD dissertation, UCLA.
- [11] Jurafsky, D.; Shriberg, E.; Fox, B.; Curl, T., 2000. Lexical, prosodic, and syntactic cues for dialog acts. *SRI Internacional*. University of Colorado. 114-120.
- [12] Kraljic, T.; Brennan, S., 2004. Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive psychology*, 50, 194-231.
- [13] Clifton, C.; Carlson, K.; Frazier, L., 2002. Informative prosodic boundaries. *Language and Speech*, 45, 87-114.
- [14] van Donzel, M., 1997. Perception of discourse boundaries and prominence in spontaneous Dutch speech. *Working Papers*, 46, 5-23.
- [15] Strangert, E.; Strangert, B., 1993. Prosody in the perception of syntactic boundaries. *EUROSPEECH'93*. 1209-1210.
- [16] Lacheret-Dujour, A.; Beaugendre, F., 1999. *La prosodie du français*. París: CNRS Editions.
- [17] Domínguez, C.; Mora, E., 1998. *El habla de Mérida*. Mérida: Universidad de Los Andes.
- [18] Martínez, H.; Domínguez, C., 2006. Análisis prosódico de algunos marcadores discursivos en el habla de Mérida, Venezuela. *Lingüística Española actual*, 28 (2). 1-18.
- [19] Garrido, J.; Llisterri, J.; de la Mota, C.; Ríos, A., 1995. Estudio comparado de las características prosódicas de la oración simple en español en dos modalidades de lectura en *Phonetica. Trabajos de fonética experimental*. Elejabeitia, A. & Iribar, A. (Eds.). Bilbao: Universidad de Deusto. 173-194.
- [20] Obediente, E. 1998. *Fonética y fonología*. Mérida: Universidad de Los Andes.