Prosody and discourse structure in Buenos Aires Spanish

Leopoldo Omar Labastía
leolabas@hotmail.com

Alejandra Eva Dabrowski
Master candidate in Linguistics. Facultad de Lenguas, Universidad Nacional del Comahue, Argentina.
aledabro@gmail.com

María de los Ángeles Guglielmone
Master candidate in Linguistics. Facultad de Lenguas, Universidad Nacional del Comahue, Argentina.
angeles_guglielmone@yahoo.com

Abstract: In previous work, we examined the role of some prosodic features in structuring spoken discourse. In this paper, we explore monologue in political discourse, and attempt to link these prosodic cues to the speaker’s argumentative goals. The prosodic analysis is carried out in the Autosegmental-Metrical framework applied to Buenos Aires Spanish. The pragmatic analysis follows Relevance Theory. Tonal configurations guide the hearers to those parts of the discourse the speaker wants the audience to consider highly relevant.

Keywords: Discourse structure; prosody; relevance

Resumo: Em trabalhos anteriores, estudamos o papel de alguns recursos prostom nuclear, de fronteira e a altura tonal na estruturação do discurso espontâneo em unidades de processamento, para marcar a informação como fundo ou primeiro plano, indicando a relevância das partes dos enunciados. Neste trabalho, exploramos o monólogo político, e tentamos relacionar estas pistas prosódicas com as metas argumentativas do falante. A análise prosódica realiza-se dentro do marco do modelo Métrico-Autosegmental aplicado ao espanhol de Buenos Aires. A análise pragmática baseia-se na Teoria da Relevância. As configurações tonais guiam à audiência aquelas partes do discurso que o emissor quer que se considerem como muito relevantes.

Palavras chaves: Estrutura do discurso; prosódia; relevância
1. Introduction

In previous work (LABASTÍA, 2011; LABASTÍA AND DABROWSKI, 2011; DABROWSKI AND LABASTÍA, 2013), we analysed the role of intonation in Buenos Aires Spanish in structuring spoken discourse in terms of a contrast between suspensive or rising nuclear tones, on the one hand, and falling nuclear tones, on the other. We interpreted this contrast in procedural terms, as processing instructions to indicate to the hearer how to derive the speaker’s intended cognitive effects with the least processing effort. Specifically, we proposed that suspensive and rising nuclear configurations signal to the hearer that they should process the information as background, that is, as context in which forthcoming foreground information will achieve the speaker’s intended effects. In contrast, falling nuclear configurations mark the information as foreground, that is, as relevant in its own right. Suspensive and rising configurations suggest to the hearer that he should hold off deriving cognitive effects until a falling configuration is heard. Instead, a falling configuration encourages the hearer to process the accumulated intonational phrases with suspensive or rising configurations together with the one(s) with falling intonation. The audience can now derive richer cognitive effects than they would have done, had they processed each intonational phrase separately. To sum up, nuclear tone choice organises spoken discourse in processing units (PUs), indicating which information is to be treated as background, and which as foreground (HOUSE, 1990), and the joint processing of background and foreground lead to rich cognitive effects.

The overall pitch range, or tonal envelope, also contributes to signalling the limits of PUs. The beginning is marked by an upwards pitch resetting, then pitch fluctuates between a mid-to-high level, and it reaches the speaker’s baseline at the end. In other

\[1\text{ It should be noted that background does not necessarily mean given information. Background information is information which contributes only indirectly to relevance. Nor should foreground be interpreted as new information. Foreground information is information that is relevant in its own right by having contextual effects (SPERBER and WILSON, 1995, p. 217).} \]
words, PUs constitute a declination domain, and pitch range is exploited to signal topic cohesion\(^2\) (LADD, 2008, p. 304-309).

In previous work, we analysed TV and radio interviews. On this occasion, we focus on monologue: we analyse a piece of political discourse and attempt to link nuclear tone choice and declination to the speaker’s argumentative goals. Gutiérrez (1997) analyses some features of political discourse which are worth considering for the analysis of our data: it is strategic, since it defines goals, means and opponents; it has performative features, because it produces a speech act and adopts a position; it is essentially polemic, as it defines an adversary and it counters real or anticipated arguments by that adversary; it is argumentative, insomuch as it presents a thesis and arguments supporting it.

The paper is organised as follows: first, we introduce the theoretical framework in which we carry out our research. Next, we describe the data and methodology, and we analyse extracts from the monologue. We link the prosodic choices with the speaker’s argumentative goals at different stages in the discourse. Finally, we draw some conclusions about the study of discourse and the study of intonation from a pragmatic point of view.

**Theoretical framework**

The prosodic analysis of the data is carried out in the Autosegmental-Metrical (A-M) framework of intonational phonology (PIERREHUMBERT, 1980; LADD, 2008). Intonational contours are interpreted in terms of high (H) and low (L) tonal targets, and make up two types of units: pitch accents (T\(^*\)), associated with prominent syllables, and boundary tones (T%) at the edge of intonational phrases (IPs).\(^3\) An application of the A-M framework to the analysis of intonation in Buenos Aires
Spanish, using the Sp-ToBI transcription system, can be found in Gabriel et al (2010). Particularly important for our analysis is the nuclear tone, the last pitch accent in the IP, whose location and configuration affects the pragmatic value of the whole intonational phrase.

The pragmatic analysis follows Relevance Theory (RT) (SPERBER and WILSON, 1995; WILSON and SPERBER, 2004). Information is relevant when it connects with other (background) information the hearer has in a productive way, yielding cognitive effects, or improvements in the hearer’s representation of the world. These effects are mainly of three types: contextual implications (conclusions), and revisions: strengthening or abandonment of background information. Relevance can be assessed in terms of the impact of new information in the context of already existing information. The more cognitive effects information triggers, the greater the relevance. But processing information involves cognitive effort of perception, memory and inference. The greater the effort, the less relevant the information will appear to be; the lesser the effort, the more relevant it will appear to be. Greater effort will be acceptable only when it leads to greater cognitive gain. Languages contribute two types of input to the process of pragmatic interpretation: conceptual and procedural. Procedural devices (discourse connectives, pronouns, determiners, etc.) (WILSON AND SPERBER, 1993; LEONETTI AND ESCANDELL-VIDAL, 2004) encode processing instructions which reduce cognitive effort by guiding the hearer to the speaker’s intended interpretation. Intonation has also been analysed as a procedural device with a pragmatic import (WILSON AND WHARTON, 2006; HOUSE, 2006; ESCANDELL-VIDAL, 2011a).

4 Sp-ToBI stands for Tone and Break Indices in Spanish. ToBI is the transcription system which implements the tenets of the A-M approach to intonation in different languages.

5 GABRIEL et al (2010) include a mid-pitch boundary tone, M%, in which the voice neither falls to a low pitch nor rises to a high pitch. This boundary tone plays a crucial role in our analysis of spontaneous spoken discourse.
Data and methodology

The data consists in a press conference given by the Argentinian Minister of Foreign Affairs, Mr. Héctor Marcos Timerman, at the United Nations (UN) in New York on February 12th, 2012. The 22-minute press conference video, available on Youtube at http://www.youtube.com/watch?v=e2C_tx3OyqY, was turned into a 22,050 Hz, 16-bit mono audio file. It was transcribed auditorily, separating the text into intonation phrases and marking the nuclear tone configurations. Pitch declination throughout tone units was also taken into account, especially low pitch and upward reset, to segment speech into PUs. The authors transcribed the recording separately, and then checked their transcriptions for reliability. Then key instances of nuclear tone configurations were analysed using PRAAT (BOERSMA and WEENICK, 2010), to corroborate auditory impressions. Table 1 shows the pitch accent and boundary tone contrasts. Arrows are used to show pitch direction for those readers who are not familiar with the A-M framework and Sp-ToBI transcription system. The procedural interpretation of the configurations is also included.

Table 1: Nuclear configurations and their procedural interpretation. Nuclear pitch accents and boundary tones are based on Gabriel et al, 2010.
Analysis of the data
Prosody and discourse organisation

If we concentrate on the opening section of the press conference, we will notice that the speaker’s voice fluctuates between approximately 200 Hz (12.02 st\textsubscript{100} and 100 Hz (0.03 st\textsubscript{100}). Taking into account the pitch height of the nuclear tone configurations in the intonational phrases, we locate the first configuration at 180 Hz (10.14 st\textsubscript{100}). The nuclear tone configurations then fluctuate between approximately 180 Hz (10.14 st\textsubscript{100}) and 140 Hz (6.60 st\textsubscript{100}) (occasionally reaching 223 Hz, or 13.92 st\textsubscript{100}, on the prenuclear pitch accent on “no” in the phrase “al no uso”), and at the end of the section, the voice reaches 97 Hz (-0.57 st\textsubscript{100}) on the nuclear syllable of a low-falling configuration (Figure 1). This section is devoted to explaining the purpose of coming to the UN: the Argentinean search for a peaceful settlement based on dialogue and diplomacy. The nuclear tone configuration in the following intonational phrase is located at 191 Hz (11.18 st\textsubscript{100}) and a new downward trend sets in. This new section is used for a different purpose: to thank those (Latin American) diplomatic representatives present at the conference for their support of Argentina’s bid to keep peace in the Southern Atlantic area and to regain sovereignty of the ‘Malvinas’ Islands. We thus conclude that the first section constitutes a PU, a declination domain employed to signal topic cohesion.

Fig. 1: Waveform, spectrogram and F0 trace of the end of the opening section. The PRAAT graph shows the pitch fluctuations ending in a very low pitch, and thus marking the end of the PU.
This opening section consists mainly of rising-suspensive nuclear tone configurations \((L+H^* M\%\), example [1a], figure 2 below). The speaker is creating the context /background to introduce the most relevant information, and suggesting that the hearer should wait until that information is uttered to process the accumulated material as a whole (capital letters identify the nuclear syllable):

[1a]

\[y\text{ no}\text{ SOtros\text{ somos\text{ FIRmes\text{ YENtes\text{ en que los proBLEmas entre los}\text{ del DIÁlogo/ más en PAZ/}}}}\]

Towards the end, the speaker uses a high-falling configuration \((H+L^* L\%)\) and a rising-falling configuration \((L+H^* L L\%,\) example [1b], figure 2 below), which suggest that the material in these units should be taken as foreground, as information relevant in its own right, and should be processed in the context of the preceding tone units to obtain rich cognitive effects:

[1b]

\[de\text{ un DIÁlogo/ que compro\text{ a TODos los involu\text{ al NO Uso/}}\]

\[de\text{ las ARmas\text{ a la diplo\text{ Macia/}}\]

\[de\text{ las ARmas\text{ a al Uso de la raZÓN/ y la diplo\text{ Macia/}}\]

leitura • maceió, n.52, p. 227-247, jul./dez. 2013
The use of these falling nuclear configurations in these intonational phrases is in keeping with the speaker’s goals: to get Great Britain to commit itself to not using weapons. In fact, the whole of the press conference will be about the U.K.’s increased introduction of sophisticated, high-tech weaponry in the area of the ‘Malvinas’ archipelago. The relevance of this part of the message is further emphasized by the fact that, when mentioning the “non-use” of weapons, Timerman expands his pitch range, and the two prominent syllables in the intonational phrase reach the highest pitch in the section:

\[1b\] / al NO Uso /
\[1b\] L+H* H+L* L%
223 127 Hz

Throughout the conference, Foreign Affairs Minister Timerman shows different slides with information about Great Britain’s ‘militarisation’ of the ‘Malvinas’ area: a map of Britain’s military bases in the Southern Atlantic area, cutting-edge submarines, warships, warplanes and missiles, and different military installations, such as a shooting range, airfields and communications networks. He organises his speech as PUs on the basis of declination. These PUs are composed of background intonational phrases (with a suspensive or rising/falling-rising configuration) and foreground phrases (with falling/rising-falling configuration), and they all end on a low-pitched falling configuration (L* L%, figure 2 below). Then pitch is reset at the beginning of the following intonational phrase, and a new declination domain starts.
When the speaker shows a new slide with the different weapons and devices the United Kingdom has introduced, he explains to the audience what is displayed in it, and elaborates on the impact of the introduction of that weapon or device in the area, and the danger it involves for all the neighbouring countries. Then he sums up the main points of that part of the presentation.

**Two organisational patterns in PUs**

It is interesting to notice that the nuclear tone configurations have a different distribution at different points in the presentation and discussion of each slide, and thus background-foreground relations vary in different PUs. The distribution seems to be cut basically to two patterns. One pattern appears at the presentation stage of each slide, when the speaker provides factual information – often reading from notes which he uses to guide his presentation – and when he sums up the main points in that slide, before passing on to the next. The other surfaces after the presentation, when the speaker assesses the consequences of the facts he has just stated for the countries in the area.
The first pattern consists of a series of non-falling nuclear tone configurations, and a low-falling tone at the end of PUs. Examples [2a] and [2b] below are a case in point. In these two PUs, the speaker first introduces the context (background) with respect to which the information in the low-falling intonational phrase at the very end operates as a signal to process the series of intonational phrases in the PU, thus achieving the desired cognitive effects:

[2a]
/ ÉSte es el misil TAUrus //<
L* L%
this is the Taurus missile

[2b]
/ Esta discarnado / para destruir / blancos / atacar / lugares protegidos / y varios / declarar / L+H* H%        L+H* M%        L+H* M%  
It has been designed to destroy targets highly protected and valuable

/ como centros de comando / de infraestructura TAUr //
L+H* M%        L+H* M%        L* L%
such as command centres radar stations infrastructure

Fig. 3: Waveform, spectrogram and F0 trace of example [2b] above. The graph shows the sequence of rising-suspensive tones ending in a low-falling tone, pronounced as a declination domain or PU.

This pattern also occurs in PUs where the speaker presents a conclusion to the points he has made, as is shown in examples [3a] and [3b] below. Timerman has just shown a map of the Southern Atlantic Ocean, with the distribution of British military bases in the area, and pointed out the extent to which the British control access
to the Atlantic and the Pacific oceans. He sums up his point saying the following:

[3a]
/a\Sí que\Ésta es la situación /mili\TAR que hay /Hoy en el Atlántico \SUR/
H+L* L%                               L+H’ M%                     L+H’ H%               L* L%            So this is the situation military at present in the South Atlantic

[3b]
/usTEdes PUE\eden \VER / quién \ES / la po\TEN\cia / mili\TAR por exce\LCia //
L+H’ M%                               H’ M%                           L+H’ M%               L* L%            You can see which is the power military par excellence

In both PUs ([3a] and [3b]), the speaker introduces the background by means of suspensive or rising tonal configurations, and ends the unit with a low-falling configuration, thus instructing the hearer to process the unit as a whole. PUs are coextensive with sentences.

In the second pattern, the nuclear tone configurations are distributed in a different way. The speaker does not wait until the end of the processing units to present the most relevant information and to instruct the hearer to process the whole. He uses high-falling or rising-falling tones in the middle of the PU to indicate that the information he is presenting is crucial to draw the intended conclusions. The hearer is to assess the impact of these tone units immediately, and not wait until the end of the PU. In the PUs in examples [4a] and [4b] below, Timerman shows a photograph of a cutting-edge naval vessel, a destroyer, which poses a serious danger to the South Atlantic area. The speaker uses the rising-falling configuration to signal the high relevance of the information he is conveying, and does not wait until the end of the PU to do this. This configuration is accompanied by an increase in pitch range, which makes the configuration more noticeable. The first syllable in the term cuadruplicado (increased fourfold) is also made prominent by means of a pitch accent on the first syllable
which is not normally used there. This concept is repeated three times, always with the same tonal configuration. The speaker wants to point out the importance of the facts he is stating, and expects his audience to derive rich contextual implications about the danger the whole area is exposed to, and the possible consequences for each of the neighbouring countries whose representatives are present.

[4a] / con él la incorporación de este barco Gran Bretaña / with the incorporation of this warship Great Britain

has quadrupled its naval power in the South Atlantic

[4b] / ha multiplicado por cuatro Gran Bretaña dice que no está inten
tando militarizar el Atlántico Sur / Great Britain claims not to be trying to 'militarize' the South Atlantic but it has increased fourfold its naval power

Fig. 4: Waveform, spectrogram and F0 trace of example [4a] above. The graph shows the sequence of rising-falling tones ending in a low-falling tone, pronounced in a wider pitch range and as a declination domain or PU.
Further evidence of these two patterns can be found in lists, as evidenced by examples [5a] and [5b] below. In [5a], as the Argentinian foreign minister shows a map of the South Atlantic area, he first points to the different countries in the area, using a suspensive configuration for all except for the final, low-falling nuclear tone, which signals the end of list and the PU. This can be considered to be one of the possible standard ways of intoning a list in Spanish. However, in [5b] he immediately afterwards points to the different military bases Great Britain has in the area. He wants to show how powerful the British military forces are in the area, and the extent to which they control access to the oceans, sea and air traffic. He mentions each location with a falling configuration. Each base is relevant in itself, and adds up to the list, as if the speaker were saying “not only X but also Y and Z and…”, creating a completely different effect: Great Britain appears to be the single most powerful force, and poses a real, immediate threat to the countries in the region, whose diplomatic representatives are attending the press conference.

[5a]

/ está Bra→SIL / Uru→GUAY / Argen→TIna / África del →SUR / An→GOla /
L+H' M% L+H' M% L+H' M% L+H' M% L+H' M%
There is Brazil Uruguay Argentina Southern Africa Angola

/ disTINtos paÍses en →África / y en →MÉrica / del →SUR /
L+H' M% L+H' M% L' L%
different countries in Africa and America in the south

[5b]

/ Asen→CIÓN / Santa E→LEna / TrisTAN de →CUnha / ISlas Mal→VInas /
L+H'+L L% L+H'+L L% L+H'+L L% L+H'+L L%
Ascension Saint Helena Tristan da Cunha Malvinas Islands

/ ISlas GOERgias y SANDwich del →SUR / y el TerriTOrio An→TÁrtico //
L+H' H% L' L%
South Georgia and South Sandwich and Antarctica
It is in this discourse pattern that another phenomenon shows up: deaccenting in the intonational phrase. Most nuclear tone configurations in Spanish occur at the end of the IP, on the last content word. Displacing the nuclear accent to an earlier position in the IP is a marked option: the deaccented chunk after the nuclear accent is somehow background information. Besides, deaccenting may be related to a contrastive or emphatic meaning (ZUBIZARRETA, 1998). In the following example [6], Timerman claims to know that Great Britain has sent nuclear weapons to the South Atlantic, and they refuse to admit that they have done so. In his speech, he anticipates a question the audience may ask:

[6]  
Tal vez ustedes me preguntan, ¿Por qué ahora?  
Perhaps you might ask me, why now?

and he himself answers:  
/ no no Ynes ahora / no Ynes la primera vez que han introducido armas nucleares /  
L+H*+L L% L+H*+L  
No, it’s not (just) now; it’s not the first time they have introduced nuclear weapons.

Then he goes on to relate another occasion when it leaked out that UK troops had nearly had an accident transferring nuclear weapons from one ship to...
another. Deaccenting, then, appears as a characteristic of this second prosodic pattern, where highly relevant information, marked by the use of a falling tonal configuration, appears early in PUs.

It is worth pointing out that the three features discussed above – increased range, double stress and deaccenting – do not occur in any of the PUs with the first pattern (a sequence of suspensive, rising or falling-rising configurations).

We consider that these two prosodic patterns of organization of PUs reflect the speaker’s strategies in pursuit of different communicative goals at different stages in the discourse. These goals can be linked to different types of argumentation in this type of discourse. It is this aspect that we turn to in the next section.

Discussion

The first pattern discussed in the previous section, a sequence of non-falling nuclear configurations which ends in a low-falling configuration, uttered on a declination domain, that is within a single PU, is perhaps the easiest way to present information to facilitate processing. The information in the non-falling nuclear configurations provides a context in which to process the most relevant information, marked with falling intonation at the end. As Sperber and Wilson (1995: 203) express it, “it is natural for given information to come before new... as this in some way facilitates comprehension.” It is more natural for the background to come before the foreground in the same way that it is easier to communicate a piece of complex information step by step.

This prosodic pattern is used by our speaker in order to present ‘factual’ information, or at least information which is meant to be taken as such. At the beginning of each slide, when Mr. Timerman introduces
a new topic, he describes the slide, and presents details about the item in question. The PUs are coextensive with the sentence, simple, compound or complex. The information is expressed in present tense and indicative mood, and is not modalized in any other way. No doubt, the speaker intends his audience to take this information as objective and true to life. These sections qualify as expository discourse. This prosodic pattern also tends to appear in the concluding part of each topic, where the speaker also seems to appear as objective and unbiased.

The second prosodic pattern discussed above, where falling tonal configurations occur earlier in PUs, and where foreground may precede background, reveals a different aspect of political discourse. The objective information provided in the introductory section of each topic, which was processed at that point as foreground, has become the background against which new information can be processed (SPERBER AND WILSON, 1995: 118). On the basis of that background, the speaker can now achieve another goal, which is characteristic of political discourse. He can engage in polemic, refuting claims by his adversary, the United Kingdom, either real or possible. Falling intonation is used to highlight those aspects of the polemic which he wants his audience to reflect on. These aspects are the ones which are meant to lead to rich cognitive effects in the minds of the members of his audience. They concern the security of all the countries in the area, of which they are the diplomatic representatives. The security of these countries is supposed to be endangered by the mere presence of cutting-edge, high-tech weapons introduced by the United Kingdom into the 'Malvinas' Islands.

Foreign minister Timerman concludes the press conference by accusing the United Kingdom of using the right of the inhabitants of the islands to self-determination in order to establish a powerful military
base in the area in the service of its strategic interests, and insists on Argentina’s decision to strive for sovereignty over the islands by peaceful means. Finally, he urges the U.K. to obey the resolutions of the UN to discuss sovereignty and stop the military escalation. Interestingly, he uses the more objective prosodic pattern which we have characterised as typical of expository discourse in this concluding section. Clearly, his ultimate purpose has also been not only to present counter-arguments against the U.K., but also to muster the support of Latin American countries by showing them that they, too, may be affected by the increasing ‘militarisation’ of the ‘Malvinas’ Islands, whilst appearing to be fair and objective in the end.

To sum up, we propose that prosody is used to organize this type of discourse into units which reflect the speaker’s varying strategic goals as his speech progresses. It enables him to signal which bits are to be treated as expository discourse, and present objective, irrefutable information, and which bits are to be treated as polemic with respect to the contrary opinions of an opponent.

**Conclusions**

In this paper we have tried to show the way intonation contributes to the interpretation of a piece of political discourse in Buenos Aires Spanish: how the speaker organises speech as processing units through declination, and how he signals background and foreground by means of contrasting nuclear tone configurations. The analysis of prosody in our data reveals two different but complementary aspects of political discourse: it is both expository and polemic.

It is interesting to point out that the nuclear tone configurations found in this piece are the same as the ones found in the other type of discourse we have analysed, interviews and, we dare say, in everyday conversation.
As we pointed out in previous work (DABROWSKI AND LABASTÍA, 2013), the instruction to assess relevance or to postpone assessing it seems to be encoded not in the whole nuclear configuration, but in the boundary tone: L% operates as a signal to evaluate the relevance immediately, whereas M% and H% (the final part of suspensive and rising configurations) suggest postponing it until an L% is found. A configurational analysis needs to be complemented by a levels analysis (PIERREHUMBERT AND HIRSCHBERG, 1990), and it is also necessary to identify the meaning of the nuclear tones independently of that of boundary tones. We are confident that the procedural view proposed by RT can help us to achieve that aim, too.

Needless to say, a more complete prosodic analysis of spoken data would need to include other aspects, such as segmentation into prosodic units at different levels (intermediate phrases and intonational phrases), pauses, duration and tempo variations. Once again, we trust that the inferential framework provided by RT offers useful tools to make progress in this area (ESCANDELL-VIDAL, 2011a, b).

**Acknowledgements**

We would like to thank our colleagues María Palmira Massi and Bettiana Blazquez for useful advice and for reading and commenting on our paper, and Alejandro Renato for helping us with the acoustic measurements.
References


[Recebido em 30 de junho de 2013
e aceito para publicação em 15 de novembro de 2013]

PROSODY AND DISCOURSE STRUCTURE IN BUENOS AIRES SPANISH