

Prosody and Discourse: a Multi-linear Analysis.

C. Portes, E. Rami, C. Auran, A. Di Cristo

Institut de Phonétique / Laboratoire Parole et Langage CNRS UMR 6057

Université de Provence, Aix-en-Provence

cristel.portes@lpl.univ-aix.fr ; emmanuelle.rami@free.fr
cauran@wanadoo.fr ; albert.dicristo@lpl.univ-aix.fr

Abstract

The aim of this paper is to present the multi-linear approach being developed among the ProDiGE research group at the Institut de Phonétique of the Laboratoire Parole & Langage (Université de Provence) in Aix-en-Provence. We first introduce the conceptual bases of this approach, before focusing our attention on the prosodic levels which constitute one of the lines in our formal and functional representation. In the last part, we present a sample analysis grid of part of our spontaneous speech corpus, detailing each level in turn.

1. Introduction

This work constitutes an illustration of the analysis framework which we collectively develop among the ProDiGE research group; this framework aims at describing and modeling the prosody of spontaneous spoken discourse. The word discourse will hereafter be considered in a wide sense as concerning all forms of speech production in a given communication situation. “spontaneous” speech is simply used here as opposed to “read” speech.

As Whichmann [1] emphasizes, three approaches to discourse, in which prosody may play a major role, can be distinguished: the propositional approach, which aims at differentiating discourse segments and specifying the rhetorical relations which link them; the cognitive approach, according to which text structure would depend on several cognitive aspects, more specifically planning activities; and finally, the formal approach, more centered on the hearer and his or her real-time interpretation of the text structure.

2. Conceptual bases

The perspective adopted here partakes of these three approaches which, for us, do not seem to be mutually exclusive, but rather complementary, even inseparable. Indeed, we consider that discourse activity, for the speaker, consists in organizing diverse semiotic hierarchies into linear discourse segments which depend on distinct structure levels. For the hearer, discourse activity consists in inferring the heterogeneous semiotic hierarchies which have been organized linearly, in order to achieve the interpretation of the speaker’s informative and communicative intentions.

In this perspective, we consider the production of discourse to consist in the transmutation of heterogeneous conceptual configurations into a sequence of linguistic units. This transmutation is operated by semantic-pragmatic operators; these operators constantly select (and simultaneously adapt to) the relevant linguistic elements: firstly to instantiate discourse coherence (notably the

appropriateness of utterance acts and utterances to their contexts), and secondly to achieve the speaker’s informative and communicative goals.

The linguistic sequence, configured by conceptual operators, consists of two parallel layers: the verbal (lexical-syntactical) sequence, and the prosodic sequence. However, the components of the linguistic sequence do obey intrinsic constraints (syntactical and metrical constraints for instance), so that the emergent forms of this “surface” linguistic sequence in discourse simultaneously reflect both these intrinsic constraints and those imposed by the semantic-pragmatic operators (for instance, the operators governing the building of the focus scale of the utterances). Moreover, “surface” prosodic forms can also display aspects of discourse contextualization which depend on the speaker, such as propositional attitude and emphasis. In this perspective, we consider that the “surface” prosodic forms and the prosodic features which are associated with them in discourse are only interpretable in a constraints based framework (cf. [2]) which would not be restricted to phonological and syntactical constraints only.

3. Prosody and Discourse

From our point of view, the prosodic sequence instantiates three structural orders which we term tonal, metrical and temporal respectively. This multi-systemic prosodic sequence is simultaneously a local mark indicator (discrete entities, such as the presence of an accent or boundary) and a local and/or global quantifying vector — a scale — (register, span, etc.). These two aspects (discrete and gradual) are complementary and indispensable in order to interpret the role played by prosody in discourse. They must therefore be taken into account in the coding and transcription of both the phonetic formal and functional marks associated with the representation of this component of the linguistic sequence.

The preceding remarks lead us to propose a tool under the form of a multi-linear analysis grid the purpose of which is to systematically link the components of the linguistic sequence (syntax and prosody) with the elements which partake of discourse organization and activity. More precisely, the purpose of this grid is to generate and test hypotheses about the very nature of the (certainly not bijective) relations which exist between discourse functions and the linguistic material which they contribute to configure. This grid also materializes a heuristic approach designed to emphasize the functional heterogeneity of prosody and discourse (cf. [3]), in the framework of an investigation of “speech in action” (cf. [4]).

The following analysis exposes the development of our research and explains the (temporary) choice of the different levels which constitute the multi-linear grid.

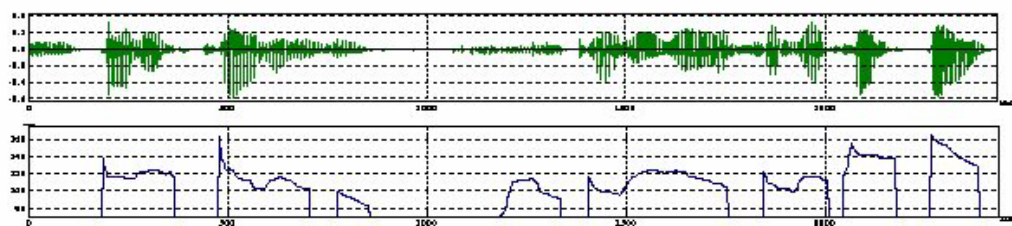
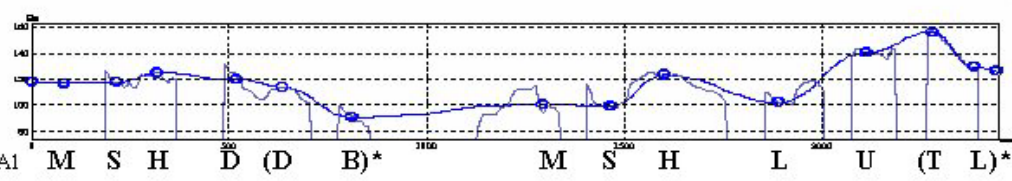
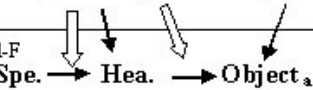
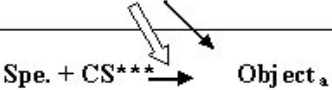
	Vous parlez d(e) charisme euh le charisme ne se décrète pas	
Prosody		
		
	A1 M S H D (D B)* M S H L U (T L)**	
	A2 (N)	
	A3 (N) (Ex)	
	B1 R N L N XL* L R N L N N N N XL**	
	B2 P ^{ss} P ^H P ^{ss}	
	C 1 1 2 1 3* 1 1 2 1 1 1 1 3**	
	F NU AC NU BO* BO** EM EM -T +T	
Syntax	A1 [(NP)(VP)] [(NP)(VP)]	
	A2 [Acc. Reflexive]***	
	F /	
Semantics and Pragmatics	Ref-A (Content _a) (Content _a)***	
	Ref-F Re-Init. _a Continue. _a	
	Pol-A [Point. Speech Vb (Point.) Point.]	
	 Pol-F Spe. → Hea. → Object _a	
	 Spe. + CS*** → Object _a	
	Rhe-A (Clause) ₁ (Clause) ₂	
	Rhe-F Satellite ← ANTITHESIS → Nucleus	
Pro-A2 : Normal / Raised / Lowered Pro-A3 : Normal / Extended / Reduced Pro-B1 : Normal / Reduced / Lengthened / eXtra-Lengthened Pro-B2 : P^s : silent pause / P^h : hesitation pause / P^b : breathpause Pro-F : A Ccent / N Ucleus / B Oundary / E Mphasis / +/- Terminal Syn-A1 : Noun Phrase / Verb Phrase / Adjective Phrase / Adverb Phrase / Preposition Phrase / [Proposition] Syn-F : / : Independence / P1 > P2 : P1 dominates P2 / P1 < P2 : P1 depends on P2		
SPr-Ref-A : Content : Semantical Content SPr-Ref-F : Initialize / Re-Initialize / Continue / End SPr-Pol-A : Pointer / Speech Verb SPr-Pol-F : Speaker / Hearer / Object		

Figure 1: Sample analysis grid

4. A case study

This sample grid (previous page, figure 1) constitutes the analysis of the first part of the following extract:

(1) “*Vous parlez de charisme / le charisme ne se décrète pas // vous parlez d'éloignement / l'éloignement des politiques a tourné au dédain // vous parlez de mystère / le mystère en politique a trop souvent tourné à l'opacité y compris dans les systèmes de financement // »* («*You talk about charisma / charisma is not something you can decree // you talk about distance / the politicians' distance has turned into disdain // you talk about mystery / mystery in politics has turned into opacity including in the financing systems // »*)

This passage is an extract from a radio debate on the issue of new conceptions of politics. Here, one participant exposes his arguments referring to the previous speaker's turn taking. This excerpt comprises three main parts (separated with //) which display the same configuration, each one containing two members (separated with /). Each time, the first member relays the opponent's argument, while the second consists in the current speaker giving his own interpretation of the same topic.

The analysis presented in this grid aims at describing and emphasizing some of the processes set up by the speaker in order to achieve his argumentative purpose.

The uppermost part introduces the prosodic description which is to be connected with the other levels of the analysis. At each level, we have tried to obey the form-function distinction principle, the forms of each level being, as much as possible, defined in terms of intrinsic criteria. The analysis of this example will show that this methodological constraint is far from being so straightforward as it may seem.

The very top of the grid displays the unpunctuated orthographic transcription of the passage under scrutiny. This method allows an easier reading of the analysis and avoids the potential bias induced by punctuation marks.

The part devoted to our prosodic description consists of five tiers:

- Pro-S: graphical representations of the acoustic level, itself divided in two parts, namely the speech wave representation (top) and the F0 curve obtained through AMDF detection with the Phonedit software (bottom);
- Pro-A: phonetic and formal coding of the tonal component, comprising three sub-levels: MOMEL targets and INTSINT (cf. [5]) alphabet coding (A1), pitch range register coding (A2: Normal/Raised/Lowered) and pitch range span coding (A3: Normal/Extended/Reduced);
- Pro-B: formal coding of the temporal component, comprising two sub-levels: syllable length variation coding (B1: Normal/Reduced/Lengthened/eXtra-Lengthened) and pause coding (B2: P^S=silent pause/P^H=hesitation pause/P^B=breath pause);
- Pro-C: formal metrical coding (three levels coded 1, 2 and 3);
- Pro-F: coding of prosodic functions (ACcent, NUcleus, BOundary, EMphasis, +/- Terminal).

The second level is devoted to our syntactic analysis, here divided in three tiers:

- Syn-A1: formal coding of syntactical phrases (Noun/Verb/Adjective/Adverb/Preposition Phrase, grouped into clauses, marked with []);
- Syn-A2: formal coding of a specific type of verb construction (accusative reflexive, which plays a major role in the speaker's argumentation, as will be detailed later on in this paper);
- Syn-F: coding of syntactic functions (≠Independency, P1>P2=P1 dominates P2, P1<P2=P1 depends on P2).

The third part of this grid is devoted to the semantic-pragmatic level of our analysis and can be divided up in three sub-levels (SPR-Ref, SPR-Pol and SPR-Rhe), each one observing the form-function divide (six lines in total):

- SPR-Ref-A: formal coding of relevant semantic contents (Content);
- SPR-Ref-F: coding of semantic functional processes (Initialize/Re-Initialize/Continue/End);
- SPR-Pol-A: formal coding of relevant polyphonic markers (Pointer/Speech Verb);
- SPR-Pol-F: coding of polyphonic functions (Speaker/Hearer/Object; → = Points at);
- SPR-Rhe-A: formal coding of rhetorical units (clauses marked with ()_i);
- SPR-Rhe-F: coding of rhetorical relations (cf. [6]).

The grid representation embodies (even visually) the common contribution of different levels to the argumentative strategy of the speaker.

On the prosodic level, several forms jointly contribute to the definition of a boundary function (Pro-F: BO) between the two members of the utterance under analysis: the DDB pattern (Pro-A1), the final lengthening (Pro-B1:XL), the silent pause (Pro-B2:P^S) and the maximum local metrical degree (Pro-C:3).

This separation is simultaneously marked by cues on other levels of the analysis:

- the absence of syntactic dependency markers (Syn-A1) which induces a paratactic relation between the two clauses (Syn-F: /);
- the polyphonic difference between the two members of the utterance, the authoring of the topic being attributed to the opponent in the first member, and personally assumed by the current speaker in the second (SPR-Pol);
- and, on the rhetorical level, the use of an antithesis relation between the two members (SPR-Rhe).

However, if forms from different levels sometimes collaborate to render a common function, there remain times when different levels give competitive (sometimes even opposite) data. For instance, the selection of the falling DDB (Pro-A1) variant of continuative pattern (as opposed to the more frequent “rising” pattern) implies (cf. [7]: p. 265) some sort of junction of the first member with the second; this configuration is systematic in our example (same pattern for

the first members of each main part of the utterance) and can thus be regarded as giving cues which may seem compatible with semantic continuity (SPr-Ref) but somehow incompatible with the paratactic configuration mentioned earlier (Syn).

The apparent clash between these competing data can certainly be accounted for if one takes the rhetorical level of our analysis into account; indeed, the very notion of antithesis (SPr-Rhe-F) implies, at the same time, both the similarity and the opposition which exist between its two constituents; prosodic dependency and syntactic independency can then be considered as materializing the two opposing forces which are at stake in this specific rhetorical function, competing cues thus playing a major role in the pragmatic interpretation of the entire excerpt.

As we have already mentioned, such a multi-linear representation makes it possible to factor out the formal elements which do play an active part in the marking of certain functions. For instance, the proverbial (or “common sense”) value of the second member of our example seems not to be marked on the prosodic level. Indeed, the generalized truth effect associated with this statement relies on formal markers on the syntactic and semantic levels only: the accusative reflexive verb construction (Syn-A2: Acc. Reflexive) combined with the inanimate feature of the semantic component “charisme” (SPr-Ref-A: Content_a) jointly induce this common sense (CS) effect (SPr-Pol-F: +CS). This contributes to strengthen the legitimacy of the speaker’s argument, thus giving more weight to the nucleus, as opposed to the satellite inside the antithesis relation (SPr-Rhe).

The emphatic pitch accent (Pro-F: EM) on the final negative particle (“*ne se décrète pas*”), eventually, corresponds to a “climax” (cf. [8]), a cumulative pitch accent where nucleus accent and emphasis accent merge. The function of this climax remains highly ambiguous; indeed, following Coleman [9], it may receive two possible interpretations: either an intensifying or a contrastive one. In this precise case, the climax concerns a negatively oriented expression in two parts out of the three which constitute the speaker’s argumentation (1). This could reinforce the contrastive interpretation, contrast thus being inferable from context and situation and not distinctively marked by prosody alone. Nonetheless, this hypothesis has to be confirmed through systematic verification in a much wider corpus.

5. Conclusion

As a conclusion, we wish to emphasize the fact that such a multi-linear grid, although it is based on strong theoretical conceptions, is but a heuristic procedure designed to display the complex relations which exist between forms and functions at different levels of the analysis. More specifically, it allows us to emphasize the part played by prosody in both the design and the layout of semantic-pragmatic meaning in spontaneous speech.

This proposition of a multi-linear analysis (and representation system) has thus shown that it is possible to decompose the semiotic syncretism inherent in spontaneous speech into distinct levels in order to better describe and explain the otherwise intricate complex relations which define discourse.

We can regard this grid as the current version of the graphical representation of the prototype (very much in the car

industry sense of the word) of the model it contributes to develop.

This implies two remarks. Firstly, partly being but a graphical representation, the grid is independent of the model, which implies that other representations (such as graphs, for instance) could very well be chosen to assume the same functions (and even make room for an automated treatment of the encoded data). Secondly, the model being at an early developmental stage, numerous questions remain to be asked (and potentially answered) about its very structure; for example, the question of a typology of discursive functions (e.g. specific vs. shared) seems to be an interesting avenue, which we aim at exploring in forthcoming research; similarly, we plan to develop the description of prosodic global markers such as pitch range (register and span), speech rate and downdrift and their relations with discourse planning and structure. The presence of psycho-linguists inside the ProDiGE group allows us to consider validating the hypotheses produced by our approach through perceptive experiments. We have used such an experimental procedure about the issue of terminality in a previous communication (cf. [10]). These results, together with the present analysis, induce us to regard this terminality function as the emergent product of the combination of markers on different levels rather than that of the prosodic level alone. The same questioning holds for other functions, such as emphasis, and will constitute objectives for further research.

References

- [1] Wichmann, A., 2000. *Intonation and text in discourse*. London: Longman.
- [2] Di Cristo, A., 2000. Une Grammaire Écologique comme cadre interprétatif de la Prosodie. *Actes du Congrès International de Sémiotique*, juin 2000, Imatra.
- [3] Bertrand, R., 1998. *De l'hétérogénéité de la parole : analyse énonciative de phénomènes prosodiques et kinésiques dans l'interaction interindividuelle*. Thèse de doctorat, Université de Provence.
- [4] Edwards, D., 1997. *Discourse and Cognition*. London: Sage Publications.
- [5] Hirst, D.; Di Cristo, A., 1998. *Intonation Systems, a Survey of Twenty Languages*. Cambridge, U.K: Cambridge University Press.
- [6] Mann, W.C.; Thompson, S.A., 1988. Rhetorical Structure Theory: toward a functional theory of text organization. *Text* 8(3): 243-281.
- [7] Martin, P., 1981. Pour une théorie de l'intonation. In *L'intonation, de l'acoustique à la sémantique*, M. Rossi et al. Paris: Klincksieck, 234-271.
- [8] Bolinger, D.L., 1989. *Intonation and Its Uses: Melody in Grammar and Discourse*. London : Edward Arnold.
- [9] Coleman, H.O., 1914. Intonation and emphasis. *Miscellanea Phonetica* 1, 6-26.
- [10] Auran, C.; Portes, C.; Rami, E.; Rigaud, N., à paraître. La distinction entre frontières conclusives et continuatives est-elle pertinente dans le discours spontané ? *Actes des Journées Prosodie 2001*, octobre 2001, Grenoble.