MODELS OF ANALYSIS AND PROSODIC LABELING SYSTEMS

Philippe Martin
UFR Linguistique, Université Paris 7 Denis Diderot
philippe.martin@linguist.jussieu.fr

1. INTRODUCTION

The underlying principle used in this paper to give an appropriate account to phonetic variations of prosody is based on the classical (structural) view of phonology, where phonetic variants of units appear as specific realizations of phonemes, which adequately contrast with each other inside the system. For example, although vocalic system in Italian can be considered the same in all regions where Italian is spoken, the phonetic realization of each vowel can be different (i.e. its articulatory, acoustic and contextual properties) in various parts of Italy.

Applied to sentence intonation, this view assumes that a phonological model is available, so that phonetic variations of prosodic entities can be described by reference to the units assumed by the model. In our approach, we will use a model for Italian described earlier (Martin, 2004), which basically poses the existence of a prosodic structure, independent but associated with the sentence text through a set of universal constraints (syntactic clash condition, eurhythmicity, stress clash node condition, stress rule …). The prosodic markers and their grammar which indicate the prosodic structure are a priori specific to each language. In Italian, these markers use the prosodic features of Height, Rise, Amplitude, Duration, etc.

Assuming these hypotheses, the coding of prosodic variations involves a somewhat detailed phonetic description of the phonological prosodic markers. They should use phonetic features such as amplitude of fundamental frequency variation, melodic slope, syllable duration, and any prosodic detail which appears to be specific to the variant studied. These phonetic features nevertheless ensure the appropriate contrasts between contours to encode the prosodic structure.

2. FRAMEWORK

2.1 Autosegmental-Metrical vs. Phonosyntactic Approach

There is a general agreement to look on accented (stressed) syllables or around the accented syllables to describe prosodic phenomena. Indeed, perceptually the stressed syllables are the most prominent, and a sentence reduced to only one word with only one syllable carries a stress (minimal condition). Minimal units are also similar for both approaches. They contain one (lexical) stress and, for French for example, one optional initial stress. They are called Accentual Phrases (AP) in the Autosegmental-Metrical (AM) framework and prosodic words (PW) (or stress groups) in the phonosyntax (PY) approach.

2.2 Content words and function words

In both theories, a minimum prosodic unit (AP or PW) contains one or more content word (open class word), and optional grammatical (closed class) words, but AM proceeds directly
with this definition whereas in the case of PY it results from a seven unstressed syllables rule (Wioland, 1985).

AM: one or more content word and grammatical words, with one final stress and an optional secondary stress (Jun and Fougeron, 2002);

PY: one content word forms a group with grammatical words through dependency relations with one stressable last syllable. Depending on the speech rate, stressable syllables are effectively stressed (with a final – primary - stress). If two groups have few syllables (in the order of 2 or 3) they can form a larger group with its final syllable stressed. If the group has a large number of syllables (say > 7), it will receive a secondary stress (Martin, 2004).

Figure 1: Various configurations of stress groups with 2 to 8 syllables, showing the dependency relation (arrows and double arrows) between components, and the resulting stressed syllables(s), underlined in red.

2.3 Prosodic structure

Differences appear in AM and PY in the definition of the prosodic structure which organizes hierarchically the AP and PW. In AM, the structure has only 2 levels, and has the form \{IP IP … IP\}, with IP = Intonation Phrase, formed with a sequence of Accentual Phrases AP: [AP AP … AP]. This arrangement is governed by the Strict Layer Hypothesis where every AP is completely contained in an IP (as it is the case in any hierarchy…), and a single AP can constitute a complete IP (this is a reminiscence of a property of syntactic structures).

In PY, the prosodic structure is not level limited, and PW form larger prosodic units in a hierarchy that can also be represented by a tree. Here a Prosodic Word is simply part of a larger prosodic unit, except when the prosodic structure has only one prosodic word.
2.4 Prosodic structure properties

Larger differences between AM and PY emerge essentially in the properties of the prosodic structure.

In AM, AP are normally described in French by the sequence /LHiLH*/; the Hi part corresponding to the optional initial (secondary) accent, and the H* part to the (primary) final stress. The L and H symbols refer to the ToBI transcription of raw fundamental frequency acoustical data, and the standard pattern implies that all melodic contours are High, except the last one in a declarative sentence, which is Low (and transcribed L%) and H% in an interrogative sentence (Jun and Fougeron, 2002).

In PY, prosodic words have no pre-established standard pattern, as their melodic characteristics depend on the application of 2 rules (Germain et Martin, 2005):

   a. IMS: Inversion of Melodic Slope rule
   b. AMV: Amplitude of Melodic Variation rule

The description of the final accent of a PW results from the application of these rules for a specific prosodic structure, and usually uses phonetic features such as Length (i.e. syllable duration), melodic Rise or Fall, Amplitude of melodic variation, etc. (Martin, 1987, 2004). Initial (secondary) accents do not play a role in the marking of the prosodic structure in PY, and are therefore always normally described with a melodic rise. Their role is only to ensure the presence of at least one stress in sequences of 7 consecutive syllables.

2.5 Prosodic and syntactic structures

The relation between the prosodic and syntactic structures is defined by alignment rules in AM, and by “sovereignty-association” in PY.

In AM, IP are basically aligned with major syntactic constituents and their hierarchy is predefined by the sequence {IP IP … IP} with IP = AP AP …AP.
The concept is very different in PY: in general more than one prosodic structure can be associated with a given syntactic structure, and every prosodic structure complies with the following constrains:

a. Planarity
b. Connexity
c. Syntactic clash condition
d. Eurhythmic restructuration
e. Prosodic 7 syllables condition
f. Prosodic word stress clash condition

2.6 Stress clash

Stress clash may occur when 2 consecutive syllables are stressed, but only in specific conditions when the implied units are dominated by the same node in the syntactic structure (Martin, 1987). In this case, the first stressed syllable in clash moves to the left and becomes a secondary stress, located on the first or on the penultimate syllable *(hippopotame gris vs. hippopotame gris)*.

![Diagram](image1)

**Figure 3:** An example of stress clash with no stress shift. The clashing units are not dominated by the same node in the syntactic structure. Case of an answer to a question such as “Comment Julien adore-t-il son café ?”.

![Diagram](image2)

**Figure 4:** An example of stress clash with stress shift. The clashing units are dominated by the same node in the syntactic structure and the shifted accent becomes a secondary accent. Case of an answer to a question such as “Qu’est-ce que Julien adore ?”.

46
2.7 Rhythmic restructuration

Figure 5: An example of rhythmic restructuration. The prosodic structure groups [Julien adore] and [le café chaud] instead of [Julien] [[adore][le café chaud]] to achieve a more balanced number of syllables in the first level of the structure (eurythmicity).

In PY, rhythmic restructuring can take place to ensure a more balanced number of syllables in the main (top) prosodic groups of the prosodic structure. If not, for instance to maintain congruence with the syntactic structure, a pause may be inserted by the speaker in an attempt to equilibrate syntactic groups duration.

2.8 Syntactic clash

Not all prosodic structures can be associated with a given syntactic structure if a syntactic clash condition prevails, when the prosodic structure assembles syntactic units dominated by distinct nodes as shown in Fig. 7. No syntactic clash condition occurs however if the prosodic structure separates units dominated by the same node in the syntactic structure.

Figure 6: The prosodic structure separates syntactic units dominated by the same node “adore” and “vraiment”. There is no syntactic clash.
2.9 French vs. Italian

Figure 8 summarizes the theoretical differences found in the encoding process of the prosodic structure by prosodic markers, i.e. by specific prosodic events located on stressed and large syntactic group boundaries.

The main differences pertain to the first stressed syllable, which (almost) always has a high and rising melodic contour, and to the PS first level final contour, located on the last stressed syllable of the corresponding syntactic group. This contour can either rise on the stressed syllable and of various slope on the last, or slightly fall on the stressed syllable and rise on the last syllable. The particular realization of either one of these phonological variants seems to be
conditioned by the complexity of the prosodic structure: simple rises on the stressed syllable for PS with few PW, and complex contour for PS involving more prosodic units. Indeed in the latter case, more contrasts are needed to encode a more complex PS.

3. CONTRASTIVE ANALYSIS OF FIVE SELECTED ITALIAN VARIETIES

Given the differences between autosegmental and phonosyntactic approaches to describe sentence intonation, the best candidates for comparison are the sentence final contour (SFC) and the so-called “major” continuation contour (MCC) (continuation majeure in the French tradition), having the advantage to be somewhat theoretically neutral.

In Italian, as marker of final first level groups in the prosodic structure PS, the MCC manifests itself under various forms, depending on the location of the lexical stress on the last text unit (word) and the complexity, in terms of stress groups and levels in the PS. Indeed, complex PS require a larger number of contrasts between prosodic contours in order to be properly encoded, and the MCC is these cases exhibits a complex melodic variation, completely associated with the final syllable of the group if stressed, and spread over the stressed syllable and the last syllable if the final syllable is not stressed. In both cases, the first part of the contour is flat or slightly falling, whereas the second part is rising. These movements occur both on the final syllable if stressed and in two steps if the final syllable does not bear the lexical stress: the flat or slightly falling melodic movement is placed on the stressed syllable, and the rising part of the prosodic marker is associated with the final (thus unstressed) syllable. This corresponds to the “delayed rise” in AM approaches.

If the PS does not require a complex network of contrasts between prosodic markers, the contour is simply rising on the lexically stressed syllable of the last word, regardless of its final or not final position.

In the declarative case, the SFC is simply falling on the last stressed syllable of the sentence, and usually reaches its lowest melodic level. When the sentence has an interrogative modality with a non neutralized melodic marker, the rising melodic slope is located on the last syllable, regardless of the last syllable being stressed or not.

3.1 Different phonological systems or variants?

In order to have different phonological prosodic systems, both the system of contrasts between contours AND the grammar of prosodic contours must be different. In our view, the grammar of prosodic contours gives an account of dependency relations existing between prosodic contours to encode the PS. In French for example, prosodic markers always depend on, the SFC being at the highest level of the PS. The mechanism is different in Italian. For example, the first stressed syllable melodic contour functions as another head of the prosodic structure, independently from the modality terminal contour. Therefore, Italian and French prosodic grammars are clearly different.

On the other hand, prosodic realizations of MCC contours observed on Belgian, Parisian, or Swiss speakers show phonetic variations in the details of the pitch movements of these contours.
Looking at the examples according to the principles outlined above, the following preliminary quantitative observations can be made:

**Sentence final contour (SFC):** described as simple for declaratives, and simple or complex for interrogatives.

*Simple:* described in terms of slope (in Hz/s), starting and ending f₀ for a declarative contour and for an interrogative contour with a final stressed syllable;

*Complex:* Part1: slope [Hz/s], starting and ending f₀ [Hz];
Part2: slope [Hz/s], starting and ending f₀ [Hz];
Difference of f₀ between the ending point of part 1 and the starting point of part 2 [Hz];

**Major continuation contour (MCC):** described as simple or complex.

*Simple:* slope [Hz/s], starting and ending f₀ [Hz];

*Complex:* Part1: slope [Hz/s], starting and ending f₀ [Hz];
Part2: slope [Hz/s], starting and ending f₀ [Hz];
Difference of f₀ between the ending point of part 1 and the starting point of part 2 [Hz];
A complex contour can occur on a final stressed syllable (depending on the complexity of the prosodic structure);

**Rhythm:** described here by the average absolute value of CMM and SFC durations [ms].

Spontaneous material should only be analyzed after a somewhat valid model of sentence intonation has been established by analyzing laboratory (read) data, unless one has a clear understanding of macro syntax, which best describes spontaneous material.

Three types of *corpora* that complement each other:

1. Laboratory speech, i.e. read speech of carefully designed sentences according to specific hypotheses (for example with syntactic expansions of syntactic categories such as NP, VP, PP, ...);
2. Semi spontaneous speech, recorded in good technical conditions, such as map tasking;
3. Spontaneous speech, recorded in various discourse productions settings (monologs, family dialogs, etc.).
Validation of contour phonetic description could easily be made through re-synthesis (using the PSOLA prosodic morphing for instance). However, it has been shown elsewhere that, if the proper intonation pattern was necessary for listeners to identify a specific variety of say French or Arabic, together with the corresponding segmental features, it is not at all the case if the intonation was ported on another variety, without changes to the segmental features pertaining to this variety.

3.2 Some experimental results

The analyzed corpus contains five Map Task dialogues collected in different Italian cities. Pitch curves and spectrograms were obtained with the software tool WinPitch (WinPitch, 2006). Relevant pitch movements (initial, simple and complex continuations, and final contours were highlighted in specific colours, as shown in figure 10.

![Figure 9: Examples analysis made with the software tool WinPitch.](image)

![Figure 10: Initial (blue), simple (green), complex (red), and final (purple) contours highlighted in specific colors.](image)
3.2.1 Final contours

Figure 11 shows various realizations of final prosodic contours in the 5 varieties of Italian analyzed. Even without statistical analysis, examples taken from Turin contrast with the other 4 by a final rise on the last syllable of the sentence. Likewise, final contour from Naples appear much sharper (i.e. with larger range of frequency fall in a shorter time) than the other varieties.

Figure 11: Examples of final declarative contour in 5 varieties of Italian.
3.2.2 Complex continuation contours

Figure 12 shows various realizations of complex continuation contours in the 5 varieties of Italian analyzed. Even without statistical analysis, Turin and Naples contours appear much sharper (i.e. with larger range of frequency rise in a shorter time) than the other varieties.

Figure 12: Examples of complex continuation contour in 5 varieties of Italian.
3.2.3 Simple continuation contours

Figure 13 shows various realizations of simple continuation contours in the 5 varieties of Italian analyzed. The number of measured cases is somewhat smaller than for the complex variety, and tendencies are harder to observe.

Figure 13: Examples of simple continuation contour in 5 varieties of Italian.
3.2.4 Initial contours

Figure 14 shows various realizations of sentence initial contours in the 5 varieties of Italian analyzed. In this case, a tendency to realize a rise-fall pattern can be observed for all varieties.

Figure 14: Examples of initial contour in 5 varieties of Italian.
Finally, figure 15 gives a summary of realizations of initial, simple and complex continuation and final melodic contours for Rome, Turin, Palermo, Naples and Florence.

Figure 15: Examples of initial, simple, complex continuation and final contours in 5 varieties of Italian.

4. CONCLUSION

Although only 2 speakers were considered for each regional variant, clear regularities appeared for the five varieties from Rome, Florence, Palermo, Turin and Naples. More details phonetic analysis, this time including rhythm, should be conducted to better describe differences and similarities in the realization of prosodic markers in Italian.

5. REFERENCES


