When Prosody Helps Syntax: an HPSG Prosody-Guided Parser

Philippe Blache
2LC - CNRS
1361 route des Lucioles
F-06560 Sophia Antipolis

tél : 92.96.73.98
pb@llaor.unice.fr

Domaine d'application : Intégration parole/écrit

Résumé

L'intégration du traitement de la parole au traitement de l'écrit constitue une des grandes questions de la communication homme machine en langage naturel. Les problèmes posés sont de type linguistique (e.g. quel lien existe-t-il entre la prosodie et la syntaxe) mais concernent également la représentation des connaissances (e.g. comment représenter ces informations dans un cadre homogène).

Nous prenons dans cet article le cas de l'intonation et montrons comment ce type d'information peut aider l'analyse syntaxique, notamment en termes de désambiguisation. Après une présentation des données prosodiques et de leur lien avec la syntaxe, nous proposons une description de cette question dans le cadre de la théorie HPSG. Nous décrirons en particulier une structure de traits adaptée et contrainte par des principes. La dernière partie présente l'implantation d'un analyseur HPSG guidé par la prosodie et propose quelques résultats.

Mots clés : HPSG, prosodie, syntaxe, intégration parole/écrit

Abstract

Integrating speech and natural language processing constitutes one of the main questions for man-machine communication. There are two kind of problems : linguistic motivations (e.g. what relations between prosody and syntax) and knowledge representation (representing heterogeneous informations within a homogeneous framework).

This paper shows in what manner prosody, and more particularly intonation, may help syntax in a parsing perspective. After a presentation of the prosodic data and a description of their connection with syntax, we will propose an HPSG view of the problem. Integrating prosodic informations and representing the prosody-syntax interface comes to add new features and principles to this theory. Finally, the implementation of a prosody-guided parser and its performances are detailed and discussed.

Keywords : HPSG, prosody, syntax, SP/NLP integration

When Prosody Helps Syntax: an HPSG Prosody-Guided Parser

1 Introduction

The integration of different domains of man-machine communication such as vision, speech or natural language becomes one of the important question of this field. Concerning more particularly the integration of speech and natural language processing, there are two kind of problems. First, from a theoretical point of view, we need a formal framework in which we can represent both phonologic and syntactic information homogeneously. The two main theories used in this perspective are categorial grammars (see [Steedman90], [Steedman91] or [Prevost93]) and HPSG (see [Bird94] or [Blache93b]). Second, such an integration has to be seen in the man-machine communication perspective. In other words, the implemented mechanisms must be general enough to cover the different directions of the SP/NLP interface.

However, an observation of current works in this field underlines several problems. First of all, linguistic theories used for natural language processing still do not cover a complete phonological level. Secondly, if a lot of linguistic works are available in the question of prosody/syntax alignment, they generally stay at a descriptive level. And lastly, the scant references adressing the implementation problem are generally only situated in the text-to-speech direction. In this case, the systems consist in generating correct prosodic patterns from syntactic structures but the relations between these levels are not analysed.

This paper concerns both theoretical and implementation level in focusing on the question of the prosody/syntax interface. The first section shows in what sense we can describe relations between prosody and syntax. We propose in the second section to complete HPSG theory by integrating prosodic knowledge. The last part presents a prosody-guided parser relying on this theoretical framework.

2 From Intonation to Syntax

Intonation is one of the prosodic features which have the closest relation with syntax. More precisely, a great deal of work (see for example [Cooper77], [Pierrehumbert80], [Gee83], [Wang92] or [Mertens93][1]) has shown the possibility of aligning intonative and syntactic units. Our purpose here is to situate these results in the perspective of a given linguistic theory, HPSG (see [Pollard94]). The interest of such an approach is not only theoretical: the use of a precise formal framework allows a general, coherent and finally efficient implementation. This section presents a description of intonation concerning more particularly the construction of intonative groups. This presentation relies on the work of [Mertens93]. Let us notice that all our examples are taken from French corpora.

[1] We can found a detailed bibliography in [Hirschberg94].

2.1 Description

The following examples situate the problem and present the kind of informations provided by intonation.

(1) `(((un professeur) de français) canadien)`
A Canadian teacher of French

(2) `(un professeur) ((de français) canadien)`
A teacher of French Canadian

These examples focus on the disambiguation process relying on intonation. In case (1), the meaning would be a Canadian citizen teaching French, whereas (2) would mean a teacher of the French-Canadian language. The difference between these realizations of the same noun phrase concerns stress patterns (`/BB.HH.B-B-/` vs `/HH.BB.B-B-/`) and, as a consequence, intonative grouping (we will see in the next section the relations between stresses and groups). In (1), there is a sequence of three different groups making up a single main group. In (2), there are two distinct groups, the second being compound by two subgroups. These subdivisions of the noun phrase correspond to the two possible syntactic structures (and the two possible interpretations). In (1), the adjective *canadien* specifies the head noun *professeur* whereas in (2), *canadien* specifies the genitive *français*.

2.2 Definitions

Basically, the presentation of intonative principles requires the description features tone, stress and groups defined as follows:

- *tone*: This designates the pitch level of a syllable. There are four such levels: extra-low (denoted `b-`), low (`b`), high (`h`), and extra-high (`h+`). Minor variations are indicated by \ and /.
- *stress*: Some syllables may be stressed. We isolate three stress positions: final stress (noted `FS` and characterized by the possibility of extension and by dynamic tones), initial stress (noted `IS`) and no stressed syllables (noted `NS`). Stressed tones are indicated in uppercase. Final stresses (which in our are case the most useful) are denoted by two uppercase letters.
- *groups*: Syllables are grouped according to their possible position (i.e. their stresses)into intonative groups (denoted by parentheses in the example).

These features are the basis of grouping rules. The shemes (3) and (4) shows how the grouping proposed in (1) and (2) can be aligned with intonative patterns.

(3)
```
(((un professeur)    de français)    canadien)
b .......  b  /BB    \ b .  b  HH    b b  B-B-
```

(4)
```
(un professeur)   ((de français)   canadien)
b ....  b  /HH     b .  b  /BB     b b  B-B-
```

As shown in this example, and following the results of linguistic research, we assume that intonative groups and syntactic units can be aligned. More precisely we think that a same intonative group cannot partially recover two different syntactic contituants. We will see in the following that this result is most interesting from an implementation point of view. Let us see more precisely the mechanisms of such an intonative grouping and first of all, let us present the IG pattern as follows.

- *Intonative groups* : An intonative group follows the stress pattern $\langle$ `((NS)* IS) (NS)* FS (NS)*` $\rangle$ where NS, IS and FS represents different stresses and * represents an optional constituent.

At this point, we may introduce the notion of *package* which is a set of intonative groups. Packages are formed following a precise rule which relies to its turn on the following tone hierarchy :

$$\texttt{B} \prec \texttt{H} \prec \texttt{H+} \prec \texttt{B-}$$

Finally, the rule 1 describes how to group several IGs :

Rule 1 *Let A and B two successive IGs. Given tone(g) a function which value is the tone of the final stress of the intonative group g, then :*

- *if tone(A)* $\prec$ *tone(B) then A is embedded in B*
- *otherwise A and B are juxtaposed.*

The example (5) illustrates the formation of intonative groups and packages following this rule.

```
la lecture   n'etait pas   un niveau   auquel on s'interessait
b...b   HH   b....b   HH   b...b  HH   bb/BB  \b......b    HH
(        )   (         )   (        )  ((   )                 )
----1-----   -----2------  ----3-----  ------------4-----------
```

(5)
```
quand on faisait   une theorie   de la litterature
b...........b/BB   \b....b  HH   b........\b B-B-
(((            )              )                   )
------------------------5--------------------------
```

No one was interested in reading when writing theory was developed.

3 Prosody in HPSG

This section presents an HPSG account for prosody. The first part gives a representation of prosodical and, more generally, phonological information. The second part concerns the definition of principles governing the prosody/syntax interface.

3.1 Feature Structures

Several studies are trying to describe phonology within a given linguistic theory. It is the case for example of [Steedman90], or [Prevost93] for categorial grammars and of [Bird92], [Bird94] or [Blache93b] for HPSG. We will complete these works by the integration of prosodic information within the HPSG PHON feature.

Practically, IGs and packages representations are integrated to PHON. This is shown in the feature structure (6) which distinguishes three sub-features : LEX which is the phonetical realization of the corresponding sign, IG which indicates the corresponding intonative groups and PKG whic represents the packages. More precisely, the IG feature consists of the F-TONE feature, the value of which is the final stress of the current IG[2] and of IG-GRP which is a set of groups, each being a set of words (represented by WORD feature). In the same way, the PKG feature representing the packages of the current sign is compound with a PKG-GRP feature whose value is a set of packages, each being a set of IGs.

(6)
$$\text{PHON}\begin{bmatrix}\text{LEX}\left\langle\left[\text{WORD}\right]\right\rangle\\ \text{IG}\left\langle{}_{ig}\begin{bmatrix}\text{F-TONE}\ {}^{tone}[]\\ \text{IG-GRP}\left\langle word\right\rangle\end{bmatrix}\right\rangle\\ \text{PKG}\left\langle{}_{pkg}\left[\text{PKG-GRP}\left\langle ig\right\rangle\right]\right\rangle\end{bmatrix}$$

The WORD feature structure (7) recapitulates the phonological description of a word as described in [Blache93b]: segmentation into syllables compound by different positions (onset, nucleus, coda) and characterized by several intonative properties such as intensity, F0, duration and so forth.

(7)
$$\text{WORD}\begin{bmatrix}\text{CLITIC}\ \{+/-\ \}\\ \text{FEET}\left\langle\text{SYLLABLES}\left\langle\begin{bmatrix}\text{STRESS}\ \{+/-\ \}\\ \text{TONE}\ \{\text{b-, b, h, h+}\ \}\\ \text{DURATION}\ []\\ \text{SEGMENTATION}\begin{bmatrix}\text{ONSET}\ [\text{SEGMENTS}]\\ \text{NUCLEUS}\ [\text{SEGMENTS}\\ \text{CODA}\ [\text{SEGMENTS}]\end{bmatrix}\end{bmatrix}\right\rangle\right\rangle\end{bmatrix}$$

3.2 Principles

This section presents the main mechanisms controlling the instantiation of prosodic features and specifying relations between prosody and syntax.

The first constraint describes the particular relation between the PHON | IG | F-TONE feature and the intonative head (i.e. the word bearing the final accent) of the corresponding intonative group. This constraint simply stipulates that the F-TONE value is the tone value of the final stress of the group.

[2]The computation of packages relies on this value.

Constraint 1 *The values of the* IG | F-TONE *feature and the* IG | IG-GRP | WORD | SYL | TONE *feature of the intonative head must be token-identical. We can represent this constraint as follows :*

$$(8) \quad \left[\text{IG}[]\right] \Rightarrow \text{IG}\left[\begin{array}{l}\text{F-TONE } \boxed{1} \\ \text{IG-GRP}_{\textit{int-head}}\left[\text{WORD} \mid \text{FEET} \mid \text{SYL}_{\textit{stress-seg}}\left[\text{TONE } \boxed{1}\right]\right]\end{array}\right]$$

Let us see now more precisely the *Intonative Group Principle* formulated as follows.

Principle 1 *Each intonative group belonging to the* PHON | IG *feature of a sign must be structure-shared with the* PHON | LEX *value of an embedded sign.*

From a practical point of view, the application of this principle results in a pre-instantiation of the PHON | LEX value beginning with a word which unifies with the first index of the intonative group. This gives us an indication of phrase boundaries.

3.3 Example

The feature structure (9) describes (4) and illustrates the HPSG knowledge representation of prosodic information. Let us notice that for clarity's sake, the WORD values are replaced by their corresponding lexical realization.

$$(9) \quad \left[\begin{array}{l}\text{PHON}\left[\begin{array}{l}\text{LEX}\left[\boxed{1}\text{un } \boxed{2}\text{professeur } \boxed{3}\text{de } \boxed{4}\text{francais } \boxed{5}\text{canadien}\right] \\ \text{IG}\left\langle \begin{array}{l}\boxed{6}\left[\begin{array}{l}\text{F-TONE H} \\ \text{IG-GRP}\langle\ 1{,}2\ \rangle\end{array}\right], \boxed{7}\left[\begin{array}{l}\text{F-TONE B} \\ \text{IG-GRP}\langle\ 3{,}4\ \rangle\end{array}\right], \\ \boxed{8}\left[\begin{array}{l}\text{F-TONE B-} \\ \text{IG-GRP}\langle\ 5\ \rangle\end{array}\right]\end{array}\right\rangle \\ \text{PKG}\left\langle \boxed{9}\left[\text{PKG-GROUP}\langle\boxed{6}\rangle\right], \boxed{10}\left[\text{PKG-GROUP}\langle\boxed{7}, \boxed{8}\rangle\right]\right\rangle\end{array}\right] \\ \text{DTRS}\left[\begin{array}{l}\text{HEAD-DTR}\left[\begin{array}{l}\text{PHON}\left[\begin{array}{l}\text{LEX}\left[\boxed{1}\text{un } \boxed{2}\text{professeur}\right] \\ \text{IG}\langle\ \boxed{6}\rangle \\ \text{PKG}\langle\ \boxed{9}\ \rangle\end{array}\right] \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{MAJ} \mid \text{N} \\ \text{DTRS } \ldots\end{array}\right] \\ \text{COMP-DTRS}\left\langle\left[\begin{array}{l}\text{PHON}\left[\begin{array}{l}\text{LEX } [\boxed{3}\text{de } \boxed{4}\text{francais } \boxed{5}\text{canadien}] \\ \text{IG}\langle\ \boxed{7}, \boxed{8}\rangle \\ \text{PKG}\langle\ \boxed{10}\ \rangle\end{array}\right] \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{MAJ} \mid \text{PREP} \\ \text{DTRS} \mid \text{HEAD-DTR} \mid \text{PHON}\left[\text{LEX } [\boxed{3}\text{de}]\right] \\ \ldots\end{array}\right]\right\rangle\end{array}\right]\end{array}\right]$$

4 Implementation

There are only few references adressing the problem of the implementation of integrated parsers[3]. We can cite for example [Bear90,92], [Hunt94] or [Rowles92]. But even if some of

[3]The works of Steedman and Prevost generally focuses on generation, not on the speech-to-text sense.

them presents very good results, they do not rely on a linguistic theory. In such approaches, mechanisms cannot be very general or reusable.

The system presented here is a prosody-guided parser and implements the above-described mechanisms i.e. intonative groups and packages construction, satisfaction of prosodic constraints and prosody-syntax alignment. Its input is a sentence tagged with prosodic marks (tone and stress). The implementation of these mechanisms is finally classical whithin a constraint logic programming paradigm (see [Blache93a]). But the system also implements a prosodic heuristic concerning the attachment problem. This question is described in this section.

4.1 Heuristic

We have seen in the first section how prosody could interact with syntax. More precisely, prosody provides information that can improve parsing mechanisms in three aspects: disambiguation, control and pre-analysis. In fact, these points mainly concern the question of attachment. Indeed, the relation between intonative groups and syntactic units provides an indication in terms of constituent boundaries which can guide the parser. But prosody gives another level of information concerning attachment itself: intuitively, one intonative group is attached to another within a given package. More precisely, we can define relations between two groups as follows. Let us call a *root* the major constituent (bearing the major stress) of an intonative group. We distinguish three cases:

- Two consecutive IGs: If these IGs does not belong to different packages, then the root of the second IG is proposed to be attached to the root of the first.
- Two consecutive packages: The root of the second package is proposed for attachment to that of the first.
- A package followed by an IG: If the IG does not belong to any package, then the root of the IG is proposed for attachment to the root of the package.

4.2 Attachment Algorithm

The attachment mechanism itself consists in adding a constituent to the daughters list of another one. This is described by the *attach* procedure which stipulates that attaching g_i to g_j consists in verifying that the head of g_i is actually governed by that of g_j (i.e. belongs to its subcat list). In this case, g_i is added to the daughters list of the head of g_j'. The function $head(g)$ returns the value of the head of the corresponding phrase (i.e. the DTRS | HEAD-DTR value).

attach(g_i, g_j) :

begin
$h_1 \leftarrow head(g_i)$;
$h_2 \leftarrow head(g_j)$;
$s \leftarrow \delta(h_2,\text{SUBCAT})$;
if $h_1 \in s$ **then** $\delta(h_2,\text{DTRS}) \leftarrow \delta(h_2,\text{DTRS} \cup g_i)$
end

The following algorithm corresponds to the three cases of attachment described in the previous section. We use in the first test the function *first* returning the first argument (i.e. the first intonative group) of a package. This test verify that the token corresponds to the beginning of a package. In this case we have to try a package attachment instead of an IG one.

Input : a list of PHON feature structures (with IG and PKG features completed) S =$\text{PHON}_1 \ldots \text{PHON}_n$
Output : Partial construction of the root's SIGN value.
Step 3 :

$i \leftarrow 1$
$j \leftarrow 1$
repeat
  if $\text{IG}_i = first(\text{PKG}_j)$ **then**
    if $j > 1$ **then**
      $attach(\text{PKG}_j, \text{PKG}_{j-1})$;
      $i \leftarrow i+1; j \leftarrow j+1$
    endif
  elsif $\text{IG}_i \notin \text{PKG}_j$ **then**
    $attach(\text{IG}_i, \text{PKG}_j)$;
    $i \leftarrow i+1; j \leftarrow j+1$
  elsif $(\text{IG}_i \in \text{PKG}_j)$ **and** $(\text{IG}_{i-1} \in \text{PKG}_j$ **then**
    $attach(\text{IG}_i, \text{IG}_{i-1})$;
    i ← i+1
    endif
  endif
until false

4.3 Results

Our system is implemented in Prolog III on a SparcLX. We can note that the development of this prosody-guided parser has consisted in completing a "classical" HPSG parser (described in [Blache93a]) with the prosodic heuristic presented in this section. This remark shows incidentally the interest of using systems relying on linguistic theories : a modification of the theory can be directly implemented in the corresponding implementation.

The following table shows some results of this system. The reference time is the parsing time with a classical HPSG parser.

Test sentence	*Characteristic*	*Parsing Time*	*Reference Time*
Jean est un professeur de français canadien	intonation (3)	1"33	1"31
Jean est un professeur de français canadien	intonation (4)	1"30	1"31
Jean est un professeur de français canadien	without packages	1"38	1"31
La lecture n'était pas un niveau auquel on s'intéressait quand on faisait une théorie de la littérature	-	2"08	2"20
C'est pas grave d'être pauvre quand on est très jeune	-	1"31	1"35

These results show several properties of such a system. First of all, in terms of disambiguation, we establish that the prosody-guided parser finds the right parse using the same time than a classical parser returning the first parse. This is obviously a strong point of the method. Moreover, this approach improves the control of parsing mechanisms by providing linguistically motivated information constraining the attachment.

5 Conclusion

As conclusion, we want to highlight several points. First of all, using a linguistic theory for the integration of speech and natural language processing is not only a formal necessity. The implementation has shown that such an approach is also more efficient in terms of maintenance and reusability of parsing mechanisms. And concerning the theory itself, HPSG seems to be very well adapted for the representation of this kind of problem. From a practical point of view, the results shows that we can improve, in particular in terms of disambigation, classical parsing processes.

All these results are to be situated in the perspective of man-machine communication and more precisely whithin the problematic of the integration of different communication domains. We have shown that speech and natural language processing can enrich each other. And we think that this kind of approach can be reused in other domains such as for example the integration of vision with NLP.

References

[Bear90] Bear J. & P. Price (1990), "Prosody, Syntax and Parsing", in proceedings of the *28th annual meeting of the ACL*.

[Bear92] Bear J., J. Dowding &. E. Shriberg (1992), "Integrating Multiple Knowledge Sources for Detection and Correction of Repairs in Human-Computer Dialog", in proceedings of the *30th annual meeting of the ACL*.

[Bird92] Bird S. (1992), "Finite-state Phonology in HPSG", in proceedings of *COLING'92*.

[Bird94] Bird S. & E. Klein (1994), "Phonological Analysis in Typed Feature Systems", in *Computational Linguistics*, 20:3.

[Blache93a] Blache P. (1993), "HPSG as a Constraint Satisfaction Problem", in proceedings of *NLPRS'93*, Japan.

[Blache93b] Blache P. & J.-P. Angoujard (1993), "Phonology, HPSG and CSPs", Research Report *RR-PB4*, 2LC-CNRS.

[Blache94] Blache P. (1994), "Vers un intégration prosodie/phonologie/syntaxe", in proceedings of *IA'94*, France.

[Cooper77] Cooper W. & J. Sorensen (1977) "Fundamental Frequency Contours at Syntactic Boundaries", in *Journal of the Acoustical Society of America*, 62, 683-692.

[Gee83] Gee J. & F. Grosjean (1983) "Performance structures : A Psycholinguistic and Linguistic Appraisal", in *Cognitive Psychology*, 15.

[Hirschberg94] Hirschberg J. (1994) "Some Bibliographical References on Intonation and Intonational Meaning", Ms.

[Hunt94] Hunt A. (1994) "Improving Speech Understanding Through Integration of Prosody and Syntax", in proceedings of the *Artificial Intelligence Conference*, Australia.

[Mertens90] Mertens P. (1990), "L'intonation", in *Le Français parlé*, C. Blanche-Benveniste et al, Editions du CNRS.

[Mertens93] Mertens P. (1993), "Accentuation, intonation et morphosyntaxe", in *Travaux de Linguistique*, 26.

[Pierrehumbert80] , Pierrehumbert J. (1980), *The Phonology and Phonetics of English Inotnation*, PhD Thesis, MIT.

[Pollard87] Pollard C. & I. Sag (1987), *Information-Based Syntax and Semantics, Volume 1*, CSLI Lecture Notes.

[Pollard94] Pollard C. & I. Sag (1994), *Head-driven Phrase Structure Grammars*, CSLI Chicago University Press.

[Prevost93] Prevost S. & M. Steedman (1993) "Generating Contextually Appropriate Intonation", in proceedings of *EACL'93*.

[Rowles92] Rowles C. & X. Huang (1992), "Prosodic Aids to Syntax and Semantic Analysis of Spoken English", in proceedings of the *30th annual meeting of the ACL*.

[Selkirk84] E. Selkirk (1984) *Phonology and Syntax. The Relation between Sound and Structure*, MIT Press.

[Steedman90] M. Steedman (1990) " Structure and Intonation in Spoken Natural Language Understanding", in proceedings of *ACL'90*.

[Steedman91] M. Steedman (1991) " Parsing Spoken Language Using Combinatory Grammars", in *Current Issues in Parsing Technologies*, M. Tomita ed., Kluwer.

[Wang92] Wang M. & J. Hirschberg (1992), "Automatic Classification of Intonational Phrase Boundaries", in *Computer Speech and Language*, 6.