

# Modéliser l'interface intonosyntaxique : ratio et synchronisation entre périodes intonatives et unités illocutoires

Julie Beliao, Sylvain Kahane, Anne Lacheret

MoDyCo - UMR-7114, Université Paris-Ouest

julie@beliao.fr, sylvain@kahane.fr, anne@lacheret.com

## Abstract

The role of prosody and syntax in identifying basic discourse units is a recurring issue in the studies of spoken language. In this paper, we focus on the terminal units of both syntactic and prosodic structures, studying respectively illocutionary units and intonative periods. This study focuses on their interactions and more specifically on both the synchronization and the relative number of their boundaries. Based on the analysis of a large spoken french corpus, we identify different types of such synchronizations (total, partial or absent) and relative proportions. We interpret these results from a functional point of view as an interaction between intono-syntax and the genre of discourse. Remarkably, evidence strongly suggests a much more complex intercation between syntax and prosody than expected under intuitive assumptions. Hence, we believe that the features we propose may be interesting for the study of spoken discourse.

## 1. Introduction

Le rôle respectif de la prosodie et de la syntaxe dans l'identification des unités de discours élémentaires constitue une question récurrente dans les travaux sur l'oral. Dans la continuité des études du GARS et des travaux effectués dans le cadre du TREEBANK RHAPSODIE, un corpus échantillonné en différents genres discursifs et annoté en prosodie et en syntaxe pour modéliser l'interface intono-syntaxique en français parlé<sup>1</sup>, nous abordons cette question en nous focalisant à l'interface de la macrosyntaxe et de la prosodie.

Nous proposons une architecture structurée autour de ces deux niveaux de représentation, construits séparément l'un de l'autre, le premier qui s'appuie sur des contraintes distribu-

tionnelles et des tests syntaxiques pour identifier les unités qui le composent, indépendamment des informations prosodiques, le second, fondé sur des critères exclusivement perceptifs et acoustiques<sup>2</sup>. Cette méthode strictement modulaire ne préjuge en rien de la réalité cognitive des processus, tout au contraire : elle permet d'éviter la circularité pour mieux répondre à cette question de l'interactivité des deux composantes en situation d'interaction verbale.

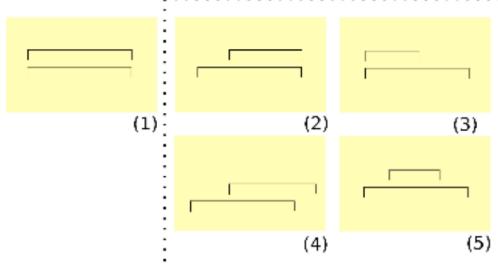
Dans cette communication, nous nous concentrerons uniquement sur les unités terminales des représentations syntaxiques et prosodiques manipulées dans le treebank RHAPSODIE, respectivement l'unité illocutoire (UI) et la période intonative (PI) telles qu'elles ont été définies par Lacheret et al. dans [Lacheret-Dujour et al., 2011]. Nous proposons une étude de leurs interactions dans le corpus RHAPSODIE et plus précisément de la coïncidence de leurs frontières.

Nous appelons frontières *synchronisées* les frontières qui sont à la fois des frontières de PI et d'UI. Les frontières sont *désynchronisées* quand frontières prosodiques et syntaxiques ne tombent pas en même temps, c'est-à-dire quand les frontières de PI ne sont pas des frontières d'UI et *vice versa*. Il est important de souligner que rien n'empêche qu'une PI dont la frontière s'aligne sur une frontière d'UI contienne plusieurs UI et inversement (voir la Fig-

1. <http://www.projet-rhapsodie.fr/>

2. Pour une méthodologie similaire voir [Degand and Simon, 2009].

ure 1). Enfin chaque PI et UI peuvent être synchronisées *totalement* lorsque leurs frontières droites et gauches sont synchronisées ou *partiellement* lorsque c'est seulement les frontières droite ou gauche qui sont synchronisées.



**FIGURE 1.** Synchronisation des frontières. (1) frontières synchronisées totalement (les deux unités s'alignent), (2) synchronisation partielle droite, (3) synchronisation partielle gauche, (4) désynchronisation chevauchante, (5) désynchronisation inclusive. En sachant que (4) et (5) sont des dérivés possibles de (2) et (3).

Notre objectif est d'abord de rendre compte statistiquement des différents types de synchronisation (totale vs partielle). Nous en proposons ensuite une lecture fonctionnelle en voyant s'il est possible de corrélérer les combinaisons intono-syntaxiques observées ( $\pm$ synchronisée) aux genres de discours à l'œuvre dans notre corpus.

## 2. Corpus

### 2.1. Corpus design

Le treebank RHAPSODIE a été créé avec comme objectif principal de proposer et de tester sur une couverture large de constructions de nouvelles méthodes d'annotation et d'analyse pour modéliser l'interface intonosyntaxique en français parlé. Un tel objectif nous a conduit à récolter à partir de données existantes (notamment les corpus présentés dans [Laks et al., 2009], [Avanzi et al., 2010] et [Branca-Rosoff et al., 2009]) et de données enregistrées pour le projet une collection d'extraits suffisamment diversifiés du point de vue de la couverture typologique (voir Table 1).

Type de parole	Monologues (M), Dialogues (D)
Situation de communication	Privée (0,1), Pro. (2)
Planification	Spontané, semi-spontané, planifié
Interactivité	Int., semi-int., non-int.
Canal	Multimédia, face-à-face
Type de séquence	Arg., descr., procédurale, ...
Tâche	cours, description, itinéraire, ...

**TABLE 1.** Variables situationnelles dans RHAPSODIE. Soit, 57 échantillons ; 52 hommes ; 35 femmes ; durée = 3h18 ; 34361 mots. Où : les monologues sont codés M, les dialogues D, la parole privée (0, ou 1), la parole publique (2). Ainsi, le fichier D2006 est un dialogue public, '006' indique le n° d'ordre.

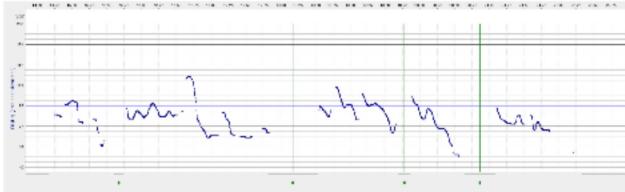
### 2.2. Annotation prosodique

La structure prosodique associée à chaque échantillon est une structure hiérarchique composée de 5 niveaux de constituante avec de bas en haut la syllabe, le pied métrique, le groupe rythmique, le paquet intonatif et la période (voir [Lacheret-Dujour et al., 2011]). Si les trois premiers et leurs différents types sont annotés sur les bases de proéminences et des disfluences perçues, la dernière, celle qui nous intéresse ici, est segmentée semi-automatiquement sur les bases de la méthode développée par Lacheret et Victorri [Lacheret and Victorri, 2002].

En pratique, la fin d'une période est détectée si et seulement si les quatre conditions suivantes sont remplies : présence d'une pause silencieuse d'au moins 300 ms, détection d'un mouvement de F0 qui atteint une certaine amplitude, définie comme la différence de hauteur entre le dernier extremum et la moyenne de F0 sur toute la portion du signal précédant la pause, détection d'un saut, défini comme étant la différence de hauteur entre le dernier extremum de F0 précédant la pause et la première valeur de F0 suivant la pause. Il convient de souligner que la décision de reconnaissance d'une rupture périodique repose sur un principe de compensation de seuils. En d'autres termes, la détection ne dépend pas des valeurs exactes des paramètres, mais de leurs seuils respectifs d'activation et des poids associés<sup>3</sup> : quand

3. Activation très forte : poids '2', forte : '1', moyenne : '0', en dessous du seuil : '-1'.

un paramètre est très légèrement au-dessous du seuil choisi, une frontière de période est détectée si les autres paramètres ont des valeurs au-dessus du seuil (ex. figure 2).



**FIGURE 2.** Segmentation par une barre verticale de 4 périodes de l'énoncé *je pense aux nombreuses victimes de la tempête [P1] et à toute leur famille [P1] endeuillée [P1] dont nous partageons la peine [P1]* [Rhap-D20004, Corpus RHAPSODIE].

### 2.3. Annotation syntaxique

Le système d'annotation développé pour le projet Rhapsodie comprend principalement un étiquetage morpho-syntaxique, une structure de dépendance fonctionnelle, et un découpage en unités macrosyntaxiques, que nous appelons unités illocutoires. C'est sur ce dernier que nous travaillons dans le cadre de cette étude. Le niveau macrosyntaxique [Berrendonner, 1990], [Cresti, 2000], [Benzitoun et al., 2010] définit la cohésion illocutoire à l'intérieur de l'énoncé. Cette structure est composée d'une unité centrale appelée « noyau », qui concentre la force illocutoire de l'énoncé et éventuellement d'unités satellites définies selon des critères strictement topologiques. On peut voir ci-dessous différentes configurations « satellites » possibles de l'unité illocutoire : l'exemple (1) est une construction en « pré-noyau – noyau », le (2) en « pré-noyau – noyau – post-noyau » et le (3) en « noyau – post-noyau »<sup>4</sup>.

(1) pour eux < c'est important // [Rhap-D0006, CFPP2000][Branca-Rosoff et al., 2009]

(2) déjà < j'ai retrouvé mes origines > quand même // [Rhap-D1003]

(3) qu'est-ce que vous en pensez > de la boule magique // [Rhap-D2011]

4. La fin des pré-noyaux est indiquée par le symbole <, le début des post-noyaux par > et l'unité illocutoire se termine toujours par //.

Les modèles macrosyntaxiques déjà existants peuvent être stratificationnels ou modulaires. Le modèle définit dans le cadre du projet se différencie des modèles stratificationnels, comme par exemple celui proposé par Berrendonner [Berrendonner, 1990] qui considère les unités maximales de la microsyntaxe, comme des points d'ancrage pour la macrosyntaxe. La macrosyntaxe de Rhapsodie s'apparente plutôt aux modèles modulaires, comme celui élaboré par l'école d'Aix-en-Provence [Blanche-Benveniste, 1990], ou par Cresti [Cresti, 2000], et qui considèrent les deux types d'organisation comme orthogonaux et donc comme pouvant opérer de concert sur les mêmes séquences. En outre, le modèle de Rhapsodie considère l'organisation macrosyntaxique comme un principe de cohésion opérant de manière indépendante de la prosodie. Au final, deux critères principaux ont été retenus pour segmenter un énoncé en unités illocutoires :

- caractériser l'organisation syntaxique indépendamment de toute organisation prosodique ou du moins de tout cadre théorique prosodique ;
- proposer des critères syntaxiques<sup>5</sup> de segmentation explicites permettant aux annotateurs d'appliquer de manière aussi formelle que possible les choix théoriques présidant à l'annotation du corpus.

Le principal critère retenu est la non-autonomie : les segments qui ne peuvent former un énoncé autonome sont considérés comme dépendants macrosyntaxiquement.

Précisons qu'une UI peut être intégrée dans une autre UI selon deux modalités : (i) les enchaînements d'UI (le discours rapporté en (5) et les greffes<sup>6</sup> en (6) notés [ ] et (ii) les insertions d'UI (les parenthèses en (7)) notées ( ).

5. La nature syntaxique de ce critère peut-être discutée notamment en regard de sa dimension pragmatique, cela-dit pour des raisons historiques nous conserverons la dénomination de « macrosyntaxe ».

6. Nous appelons greffe une UI qui vient occuper un position régie où est attendue en principe une unité lexicale. Les discours rapportés sont également des UI qui viennent occu-

- (5) Marcel Achard écrivait [ elle est très jolie // elle est même belle // elle est élégante // ] // [Rhap-D2001, Corpus Mertens][Mertens, 1987]
- (6) vous suivez la ligne du tram qui passe vers la [ je crois que c'est une ancienne caserne // ] // [Rhap-M0003, Corpus Avanzi][Avanzi, 2012]
- (7) alors que Heinze ( c'est quand même assez extraordinaire hein // ) c'est le patron de la défense // [Rhap-D2003, Rhapsodie]

### 3. Rapports et types de synchronisations intonosyntaxiques

L'objectif de cette section est de présenter l'analyse statistique conduite sur les données en termes de rapports (équation (1)), ou relation entre les fréquences des frontières droites de PI et d'UI, et de synchronisation de ces frontières.

$$\text{Ratio (échantillon)} = \log \left( \frac{\text{nombre de PI}}{\text{nombre de UI}} \right). \quad (1)$$

#### 3.1. Analyse statistique des rapports

Le corpus étudié comporte en tout 3457 UI et 2904 PI. Lorsqu'on représente pour chaque échantillon la proportion (1), on remarque que certains ressortent clairement par rapport aux autres, signifiant qu'ils présentent bien plus de PI que d'UI. Il nous est apparu de plus que ces échantillons semblaient correspondre à ceux d'un genre oratoire. Afin de vérifier statistiquement cette intuition nous avons réparti les échantillons d'après les métadonnées du corpus selon quatre genres : oratoire, procédural, argumentatif, descriptif pour lesquels il s'agit alors de vérifier si les rapports  $\frac{PI}{UI}$  sont caractéristiques. À cet effet et puisque les données ne suivent pas une loi normale, c'est-à-dire qu'elles ne se répartissent pas selon une gaussienne (un test paramétrique tel qu'une ANOVA est donc proscrit), nous avons appliqué le test non-paramétrique de Kruskal-Wallis dont le résultat est affiché en figure 3. Il apparaît que l'hypothèse nulle indiquant que les rapports  $\frac{PI}{UI}$  sont les mêmes dans tous les groupes peut être re-

per une position régie, mais à la différence des greffes, il s'agit d'une position où est attendue une telle construction.

jetée sans risque ( $p = 5 \cdot 10^{-4}$ ). Nous concluons donc que ces distributions sont significativement différentes les unes des autres et par là que le calcul du rapport  $\frac{PI}{UI}$  est pertinent si l'on souhaite classer et caractériser des types de discours.

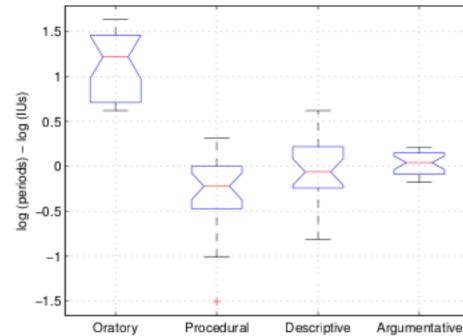


FIGURE 3. Test de Kruskal-Wallis sur les 4 groupes de sous-genres : oratoire, procédural, descriptif, argumentatif

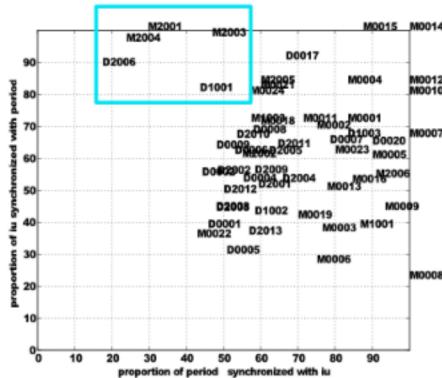
#### 3.2. Analyses des types de synchronisation

For de ces résultats préliminaires, nous avons vérifié si la synchronisation des frontières de PI et d'UI venait confirmer les résultats obtenus en calculant les rapports  $\frac{PI}{UI}$ . Pour chacun des 57 échantillons du corpus RHAPSODIE, nous avons extrait automatiquement [Belião, 2012] les frontières droites des UI et des PI ainsi que leurs positions temporelles selon différentes configurations possibles allant de la synchronisation à la désynchronisation et dont nous présentons les décomptes et pourcentages dans la table 2.

relations frontalières entre UI et PI	décompte	%
FD d'UI synchronisée avec FD de PI	1781	56
FD de PI synchronisée avec FD d'UI	1788	32
FD d'UI désynchronisées	1399	44
FD de PI désynchronisées	3799	68

TABLE 2. Décompte des relations frontalières entre PI et UI

La figure 4 présente le diagramme de dispersion des pourcentages de synchronisation des PI et des UI pour l'ensemble des échantillons.



**FIGURE 4.** Synchronisation droite partielle des UI et des périodes pour chaque échantillons du corpus

Sur cette figure, nous observons que les échantillons appartenant au genre oratoire (encadrés) se distinguent des autres venant ainsi corroborer les résultats relatifs aux ratios présentés en figure 3. S'ils présentent une forte synchronisation des frontières d'UI avec les celles de PI (plus de 80% d'occurrences), l'inverse n'est pas vrai : dans ce genre, les périodes correspondent rarement à une UI (20 à 50%), ce qui indique qu'elles se produisent souvent au sein d'une UI et montre ainsi que la planification prosodique correspond à un processus de fragmentation syntaxique qui montre les différentes étapes de programmation de l'UI (cf (2)).

- (2) lorsque vous semblez mettre en doute [PI] notre amour des libertés [PI] c'est un outrage [PI] que nous n'acceptons pas [UI-PI] nous sommes les héritiers de la tradition qui a instauré dans ce pays [PI] la démocratie politique et sociale [UI-PI] toujours [PI] toujours [PI] contre les droites coalisées [PI] nos combats pour la conquête du droit [PI] jalonnent l'histoire des deux derniers siècles [UI-PI] [Rhap-D2006, Corpus RHAPSODIE]
- (3) alors en partant de la place Paul Vallier pour aller à la place Notre-Dame [UI-PI] alors j'emprunte la rue de Strasbourg [UI-PI] je passe par la place Vaucanson [UI-PI] je prends direction Maison du tourisme [UI-PI] euh à la Maison du tourisme je contourne enfin je prends la rue de la République en remontant la rue de la République [UI-PI] je tombe sur la place Sainte-Claire on va dire là où il y a la halle [UI-PI] [Rhap-M0014, Corpus AVANZI[Avanzi, 2012]]

D'autres caractéristiques du discours pourraient être identifiées à partir de la figure 4 comme par exemple que certains échantillons situés dans le coin supérieur droit du diagramme de dispersion présentent un taux de synchronisation

quasi parfait entre frontières droites d'UI et de PI (cf exemple (3)). Les locuteurs présentent un discours relativement « canonique », c'est-à-dire un discours dont la prosodie suit la segmentation syntaxique. Dans tous les cas, il s'agit de genres descriptifs (descriptions d'itinéraires) non interactifs, sans enjeux argumentatif particulier contrairement au genre oratoire mentionné plus haut.

## 4. Conclusion

Partant de l'étude des rapports  $\frac{PI}{UI}$  (cf équation 1), les résultats observés nous ont conduit à émettre l'hypothèse qu'il s'agissait peut-être là d'un critère discriminant pour classer et caractériser les genres de discours. Pour illustrer cette hypothèse, nous nous sommes focalisés sur la question de la synchronisation des frontières droites des unités syntaxiques et prosodiques étudiées. Les points qui retiennent particulièrement notre attention sont les suivants : la différence entre la fréquence d'usage observée et la fréquence intuitive attendue qui s'illustre par une production massive de PI par rapport au nombre d'UI (3180 UI pour 5587 PI, table 2) et la relation probable entre cette répartition et les genres de discours en jeu. Si ce premier constat reste à renforcer sur du matériel plus exhaustif, et peut-être avec des descripteurs plus approfondis, il apporte une contribution significative aux études sur les genres oraux, qu'elles s'inscrivent dans le champ de la typologie textuelle ou dans celui de la phonostylistique. Il nous ouvre notamment une piste prometteuse et encore peu exploitée à l'interface de la prosodie, de la syntaxe et de la sémantique : ces traces intonosyntaxiques que les genres laissent dans le message pourraient être des marqueurs stables de la façon dont les relations de discours (contraste, concession, élaboration résumé, etc) se tissent dans les textes.

## 5. References

- [Avanzi, 2012] Avanzi, M. (2012). *L'interface prosodie/syntaxe en français. Dislocations, incises et*

asyndètes. *Bruxelles, Peter Lang*. PhD thesis, Université de Neuchâtel.

[Avanzi et al., 2010] Avanzi, M., Simon, A.-C., Goldman, J.-P., and Auchlin., A. (2010). Un corpus de français parlé annoté pour l'étude des proéminences, c-prom. *Actes des 23èmes journées d'étude sur la parole, Mons, Belgique*.

[Belião, 2012] Belião, J. (2012). Formalisation, implémentation et exploitation d'une hiérarchie objet intono-syntaxique : étude sur un treebank de français oral spontané. Master's thesis, Université Sorbonne Nouvelle.

[Benzitoun et al., 2010] Benzitoun, C., Dister, A., Gerdes, K., Kahane, S., Pietrandrea, P., and Sabio, F. (2010). Tu veux couper là faut dire pourquoi. propositions pour une segmentation syntaxique du français parlé. *Actes du Congrès Mondial de Linguistique française, La Nouvelle Orléans*.

[Berrendonner, 1990] Berrendonner, A. (1990). Pour une macro-syntaxe. *Données orales et théories linguistiques*, 21 :25–31.

[Blanche-Benveniste, 1990] Blanche-Benveniste, C. (1990). Un modèle d'analyse syntaxique 'en grilles' pour les productions orales. *Anuario de Psicología Liliane Tolchinsky (coord.) Barcelona*, vol. 47 :11–28.

[Branca-Rosoff et al., 2009] Branca-Rosoff, S., Fleury, S., Lefevre, F., and Pires, M. (2009). Discours sur la ville. Corpus de Français Parlé Parisien des années 2000.

[Cresti, 2000] Cresti, E. (2000). Corpus di italiano parlato. *Florence, Accademia della Crusca*.

[Degand and Simon, 2009] Degand, L. and Simon, A.-C. (2009). *Where prosody meets pragmatics*, chapter Mapping prosody and syntax as discourse strategies : How basic discourse units vary across genres. Emerald Group.

[Lacheret and Victorri, 2002] Lacheret, A. and Victorri, B. (2002). La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques. *Verbum*, pages 55–72.

[Lacheret-Dujour et al., 2011] Lacheret-Dujour, A., Kahane, S., Pietrandrea, P., Avanzi, M., and Victorri, B. (2011). Oui mais elle est où la coupure, là ? Quand syntaxe et prosodie s'entraident ou se complètent. *Langue française, Paris-Larousse*, 170 :61–80.

[Laks et al., 2009] Laks, B., Durand, J., and Lyche, C. (2009). Le projet PFC (Phonologie du Français Contemporain) : une source de données primaires structurées. In *Phonologie, variation et accents du français* Hermès, pages 19–6. Hermès.

[Mertens, 1987] Mertens, P. (1987). *L'intonation du français : de la description linguistique à la reconnaissance automatique*. Katholieke universiteit te Leuven.