



HAL
open science

De la prosodie à la syntaxe Vers une description sémantico-syntaxique des périodes intonatives binaires

Noalig Tanguy

► **To cite this version:**

Noalig Tanguy. De la prosodie à la syntaxe Vers une description sémantico-syntaxique des périodes intonatives binaires. Représentations du sens linguistique 5, 2011, France. pp.1-20. halshs-00763856

HAL Id: halshs-00763856

<https://shs.hal.science/halshs-00763856v1>

Submitted on 16 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

De la prosodie à la syntaxe

Vers une description sémantico-syntaxique des périodes intonatives binaires

Noalig TANGUY
Laboratoire LaTTiCe UMR 8094 ENS/Paris 3
noaligtanguy@gmail.com

1. Introduction

Ces dernières années, les travaux portant sur les interactions entre les niveaux syntaxiques et prosodiques se sont considérablement développés du fait de l'intérêt grandissant pour le français parlé. Cependant, les études actuelles (i) portent essentiellement sur une comparaison des niveaux (pour exemple, Avanzi 2010) ou (ii) partent de phénomènes syntaxiques tels que l'absence de subordonnant, le double marquage du sujet ou les compléments différés, que ces études expliquent, désambigüisent et/ou analysent par la prosodie.

Nous chercherons ici à procéder de manière inverse, c'est-à-dire que nous analyserons syntaxiquement des segmentations intonatives. Le présent projet, à l'interface syntaxe-prosodie, s'appuie sur les notions (i) de phrase comme prédication en acte assortie d'une modalité énonciative (Le Goffic 1993), et (ii) de période comme unité d'intégration prosodique de rang supérieur (Lacheret & Victorri 2002), comme unités de description et de segmentation.

Différentes études ont relevé trois cas de corrélation entre les unités syntaxiques et les unités prosodiques : les cas de congruence, de condensation et de dislocation, pour reprendre les termes de Degand et Simon (2008). Dans ce travail, nous nous intéresserons uniquement aux cas dits « de condensation », c'est-à-dire les cas pour lesquels une unité « période » regroupe plusieurs (au moins deux) unités « phrases ».

Le présent article constitue la première ébauche d'un programme de classement des différentes façons dont les phrases se relient les unes aux autres pour faire sens et des types de prédications mis en œuvre pour cela, dans une perspective de regroupement prosodique, mais aussi sémantico-discursif. L'objectif est de dégager les grands types de regroupements possibles des phrases au sein de périodes, selon des bases sémantiques et discursives, mais aussi par le biais d'indices syntaxiques (structures répétées, emploi de pronoms, etc.). Nous mettrons l'accent sur les relations entre les structures syntaxiques et prédictives employées et les effets produits : rapports temporels (antériorité, recouvrement, postériorité), discursifs (narration, élaboration, mise en arrière plan, annonce, explication, résultat), etc.

La question centrale sera de voir comment découper un texte oral en unités élémentaires, tout en sachant que ces unités doivent être à la fois suffisamment autonomes pour être des 'unités' délimitables, c'est-à-dire modélisables dans le cadre d'une grammaire, et suffisamment perméables et adaptables pour se relier aux autres unités dans le cadre du discours (Le Goffic 2008).

Nous présenterons ici une étude en cours sur les périodes intonatives binaires (composées de deux phrases) et réalisée sous Analec, logiciel d'annotation.

2. Cadres théoriques, corpus et méthode

2.1. Cadres théoriques

L'étude réalisée couple critères de segmentation prosodiques et critères syntaxiques, à partir des notions de période intonative et de phrase appliquée à l'oral.

La phrase comme unité de segmentation syntaxique

Notre approche syntaxique s'inscrit dans le cadre théorique développé par Le Goffic (1993), axé sur la notion de phrase comme unité de description et de segmentation, définie comme prédication en acte, assortie d'une modalité d'énonciation (Lefeuve 1999). Cela signifie qu'une phrase s'articule autour d'un prédicat, actualisé dans le discours par une modalité énonciative – assertive, interrogative ou injonctive. La notion de prédication renvoie à l'opération sémantico-syntaxique de construction d'un énoncé, c'est-à-dire à la mise en relation d'un sujet et d'un prédicat, selon plusieurs types (Tanguy 2009). Pottier (1992), dans sa rosace des possibles positionnels, distingue cinq aires événementielles, à partir desquelles se dégagent cinq types de prédications :

(i) La prédication dite « d'existence » pose une entité comme existant dans l'espace et dans le temps :

(a) *À droite, il y a le Sacré Cœur.*

(b) *À droite, le Sacré Cœur.*

(ii) La prédication d'attribution affecte une propriété à un sujet :

(c) *Ce livre est très intéressant.*

(d) *Très intéressant, ce livre.*

(iii) La prédication de localisation localise – au sens large du terme – un sujet. La localisation peut être spatiale, temporelle, notionnelle, etc.

(e) *Pierre est à Paris.*

(f) *Toujours à Paris, Pierre.*

(iv) La prédication d'activité repose sur un verbe à procès dynamique :

(g) *Pierre court.*

(v) La prédication dite « de cognitivité », en relation avec le monde subjectif, recouvre la perception par les sens et toutes les activités intellectuelles :

(h) *Pierre réfléchit.*

La phrase réalise ainsi un prédicat – d'existence, d'attribution, de localisation, d'activité ou de cognitivité – lié ou non à un sujet de manière explicite, et au(x)quel(s) se rattachent des éléments périphériques, plus ou moins extérieurs au prédicat.

Structuration syntaxique du français parlé

Les notions de phrase, de prédicat ou de prédication sont des notions très critiquées, et principalement pour la description, la segmentation et l'analyse des unités constitutives du français parlé. Les différentes théories actuelles préfèrent travailler à partir d'unités telles que l'unité macro-syntaxique (Blanche-Benveniste *et al.* 1984 et 1990), la clause et la période (Berrendonner 1990 et 1993 ; Béguelin 2002), les unités rectionnelles et illocutoires (Projet ANR Rhapsodie 2008-2012), le paragraphe oral (Morel 2003 ; Morel et Danon-Boileau 1998), l'énoncé (Cresti 1999 et 2000), etc.

Nous défendons au contraire l'idée qu'une unité de type 'phrase' existe à l'oral et que les unités du système du français parlé se reconnaissent à des critères syntaxiques, même si ceux-ci doivent être parfois associés à des indices prosodiques.

Pour décrire les différentes composantes de la phrase, nous reprenons la répartition tripartite développée principalement par l'approche aixoise en macro-syntaxe (Blanche-Benveniste 1997 et 2002 ; *et al.* 1990), mais aussi par Le Goffic au sein de la phrase (2008), à savoir une organisation en noyau et affixes. Le centre de la phrase est analysé comme un noyau : il rassemble le prédicat et le sujet si celui-ci lui est lié. Les éléments extra-prédicatifs sont appelés, selon leur place, prénoyau ou postnoyau.

La période comme unité de segmentation prosodique

Au niveau prosodique, notre étude retient comme unité d'étude la période intonative, au sens de Lacheret et Victorri (2002, Lacheret-Dujour 2003 ; Avanzi et Lacheret-Dujour 2009 ; Avanzi 2005 et 2008), définie comme l'unité d'intégration prosodique de rang supérieur et détectée par des logiciels d'étiquetage et de séquençage basés sur l'analyse prosodique du discours, tels que le logiciel Analor (Avanzi, Lacheret-Dujour et Victorri 2008).

L'analyse retient quatre paramètres : (i) la durée de la pause, et plus précisément la durée de l'intervalle entre deux portions de F0 (fréquence fondamentale), (ii) l'amplitude du « geste », c'est-à-dire la différence de hauteur entre le dernier extremum de F0 et la moyenne de F0 sur toute la portion du signal qui précède la pause, (iii) l'amplitude du « saut », c'est-à-dire la différence de hauteur entre la dernière valeur (dernier extremum) de F0 précédant la pause et la première valeur de F0 suivant la pause et (iv) la présence d'un allongement lié à une hésitation, c'est-à-dire la présence du morphème « euh », juste avant ou juste après la pause.

Le logiciel détecte une segmentation en période si quatre conditions sont réunies selon différentes variables précisées dans Lacheret et Victorri 2002 : (i) la durée de la pause dépasse un seuil de l'ordre de trois cents millisecondes, (ii) l'amplitude du « geste » dépasse un seuil de l'ordre de quatre demi-tons, (iii) l'amplitude du « saut » dépasse un seuil de l'ordre de trois demi-tons, et (iv) il n'y a pas de « euh » à la proximité immédiate de la pause.

Notre analyse convoque ainsi deux unités de niveaux différents : la phrase, pour le niveau syntaxique et la période, pour le niveau prosodique.

2.2. Corpus, méthode et segmentations

Ce premier travail à l'interface syntaxe, prosodie, mais aussi sémantique, a été réalisé à partir d'un corpus réduit de 3500 mots pour une durée totale de 19 minutes. Il rassemble quatre textes (transcriptions alignées) issus de différents corpus de français parlé existants.

<p style="text-align: center;">Texte <i>Type / genre</i> <i>Corpus</i> <i>Lieu</i> <i>Locuteur (sexe, âge)</i> <i>Données chiffrées : durée et nombre de mots</i></p>	
<p style="text-align: center;">Texte 1 « Plongée » <i>Récit de vie</i> <i>CRFP (DELIC 2004)</i> <i>Clermont-Ferrand</i> <i>Homme, 45 ans</i> <i>4'39'' – 804 mots</i></p>	<p style="text-align: center;">Texte 2 « Itinéraire 1 & 2 » <i>Indication d'itinéraire en milieu urbain</i> <i>C-PROM (Avanzi et al. 2010)</i> <i>Grenoble</i> <i>Homme, 23 ans, étudiant : 0'50'' – 133</i> <i>mots</i> <i>Femme, 34 ans, commerciale : 1'40' –</i> <i>333 mots</i></p>
<p style="text-align: center;">Texte 3 « Doreur » <i>Récit de vie</i> <i>GRE03 (Savelli, Grenoble 3, Clapi)</i> <i>Grenoble</i> <i>Homme, 46 ans, doreur sur bois</i> <i>(artisan)</i> <i>5'18'' – 1016 mots</i></p>	<p style="text-align: center;">Texte 4 « 14-01 » <i>Interview</i> <i>CFPP2000 (Branca et al. 2009)</i> <i>Paris</i> <i>Femme, 53 ans, libraire</i> <i>6'10'' – 1246 mots</i></p>

Figure 1. Corpus et métadonnées

La première étape du travail a consisté à segmenter prosodiquement différents textes oraux selon les critères de segmentation décrits ci-dessus, en combinant traitement manuel et traitement automatique à partir de logiciels d'étiquetage et de séquençage prosodique du discours. Parallèlement, nous avons réalisé une segmentation en unités « phrases ».

Nous avons classé nos résultats selon les trois cas de corrélation possibles entre unités syntaxiques et unités prosodiques (Degand et Simon 2008), à savoir (i) les cas de congruence (unité 'période' = unité 'phrase'), (ii) les cas de condensation (unité 'période' >

unité 'phrase') ou couplage (Avanzi 2010) et (iii) les cas de dislocation (unité 'période' < unité 'phrase') ou fragmentation (Avanzi 2010).

Le présent projet s'intéresse uniquement aux cas dits « de condensation », c'est-à-dire les cas pour lesquels une unité 'période' regroupe plus d'une unité 'phrase'. Ce cas représente 44% de nos résultats (Avanzi & Tanguy 2009) : une phrase correspond à une période dans 33% des cas, une période regroupe plusieurs phrases dans 44% des cas, et une phrase est 'à cheval' sur plusieurs périodes dans 23% des cas de figure. Les couplages (ou condensation) sont donc le cas le plus fréquent. La prosodie cimente ainsi plusieurs phrases : deux, trois, ou plus, jusqu'à sept phrases pour notre corpus analysé.

	Cas n°1 « congruence »	Cas n°2 « couplage »	Cas n°3 « fragmentation »
CRFP – plongée	32 %	46,5 %	21,5 %
GRE03 – doreur	17 %	50 %	33 %
C-PROM – Iti	54,5 %	36,5 %	9 %
C-PROM – Iti	28 %	61 %	11 %
CFPP – Noroy	47 %	33 %	20 %
Pourcentage	33 %	44 %	23 %

Figure 2. Répartition des cas de congruence, de couplage et fragmentation dans les quatre corpus

Les phrases sont distinctes, indépendantes les unes des autres, mais se regroupent sur des bases prosodiques, mais aussi sémantiques et informationnelles.

(i) *c'était en en hiver |¹ c'était à l'époque de Noël donc | et j'étais malade | je me souviens | j'étais malade (Texte 1 « Plongée »)*

(j) *je vous expliquerai sur mon cadre là-bas | j'ai un cadre où on voit les différentes opérations là | on voit qu'elle a fait un elle a fait un petit chiffon là | ouais (Texte 3 « Doreur »)*

Nous remarquons par exemple en (i) la répétition de la structure présentative « c'était » et l'emploi répétitif des pronoms clitiques sujets « je » qui constituent des indices syntaxiques de cohésion.

Plus spécifiquement, nous nous intéresserons aux périodes binaires, c'est-à-dire aux périodes regroupant deux unités 'phrases', du type :

(k) *y a certaines rues | elles sont une seule rue à trois sens différents (CFPP2000)*

¹ Le signe | marque la fin d'une unité 'phrase'.

3. Annotation syntaxique et sémantico-discursive

La seconde étape du travail a consisté à annoter notre corpus avec le logiciel d'annotation et d'analyse de corpus Analec (Victorri 2011). Nous avons créé deux types d'unités : la phrase et la période. Nous avons ensuite créé deux types de relations : (i) entre la période et la première phrase de la période (relation 1) et (ii) entre la période et la deuxième phrase (relation 2). À ces unités et relations, nous avons attribué différentes propriétés et valeurs.

3.1. Unités : les phrases

Nous avons défini pour les unités de type 'phrase' quatre propriétés principales précisées ensuite par différentes valeurs : (i) la prédication véhiculée, (ii) le type de prénoyau exprimé, (iii) la structure phrastique et (iv) le temps du verbe de la principale.

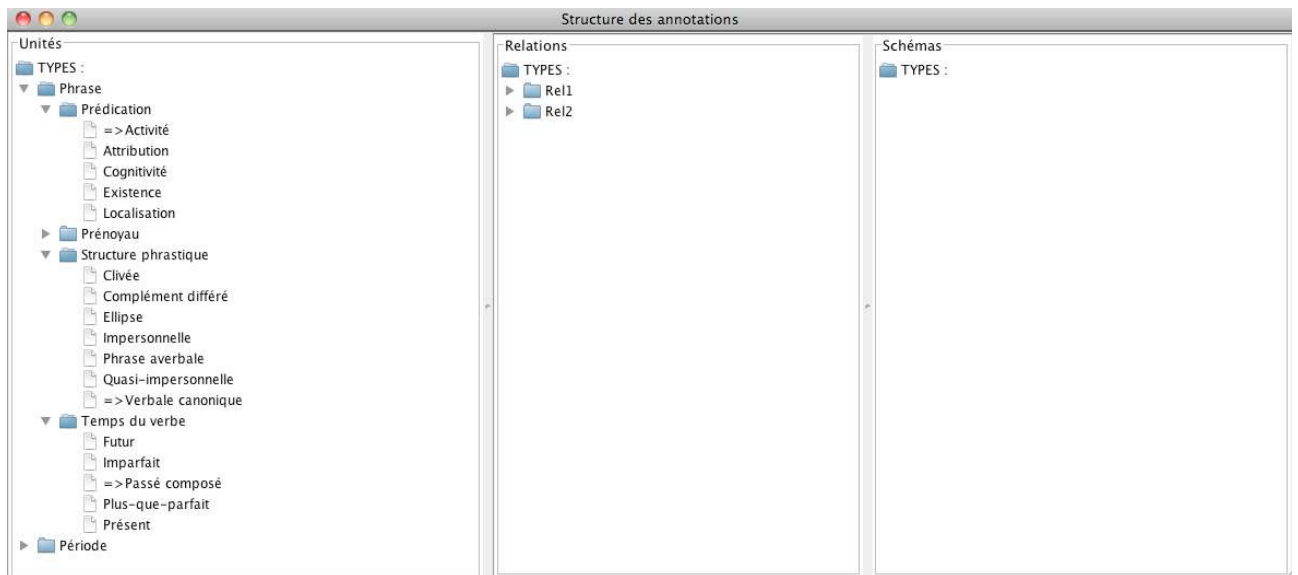


Figure 3. Gestion de la structure : unités 'phrases'

(i) Nous avons tout d'abord spécifié le type de prédication véhiculé : attribution, localisation, existence, activité, cognitivité. Nous avons classé comme attributifs des énoncés tels que :

(15) *moi j'ai quarante cinq ans + quarante six même* (Texte 3 « Doreur »)

(5) *j'étais au lit + j'avais une grosse angine* (Texte 1 « Plongée »)

(14) *il est à Parilly là + il a une écurie à Parilly* (Texte 1 « Plongée »)

(ii) Le type de prénoyau et sa réalisation apparaissent également intéressants à observer. Nous avons distingué six ensembles :

1. Les adverbiaux relatifs à l'énonciation, c'est-à-dire 'au dire' :

(42) *euh eh bien on a toujours un peu ceux-là qui vieillissent euh + y en a d'autres qui sont partis du quartier qui reviennent de temps en temps (Texte 4 « 14-01 »)*

(43) *et puis bon quand même évidemment des nouveaux +mais euh il faut quand même bien reconnaître que ils sont plus jeunes donc moins proches du livre (Texte 4 « 14-01 »)*

Parmi les adverbiaux relatifs à l'**énoncé**, c'est-à-dire 'au dit', nous distinguons :

2. Les ligateurs (Morel & Danon-Boileau 1998), autrement appelés connecteurs pragmatiques (Berrendonner 1983) :

(7) *et je savais pas nager non plus + donc j'ai jamais utilisé mon matériel (Texte 1 « Plongée »)*

(35) *et ben c'était une de mes questions + alors je vous écoute (Texte 4 « 14-01 »)*

(46) *donc euh ça c'est un petit peu pour ce qui a été des changements de du quartier + moi je suis là depuis vingt trois ans (Texte 4 « 14-01 »)*

3. Les organisateurs du discours :

(35) *et ben c'était une de mes questions + alors je vous écoute (Texte 4 « 14-01 »)*

4. Les modalisations (appréciatives ou épistémiques) :

(45) *ils restent pas longtemps sur le coin + apparemment y a pas de très grands appartements (Texte 4 « 14-01 »)*

5. Les cadratifs – ou cadres – (Charolles & Vigier 2005), adverbiaux de temps ou de lieu :

(27) *et euh après tu bifurques euh + y a une petite bifurcation euh juste avant la place du Tribunal (Texte 2 « Itinéraire 1 & 2 »)*

(28) *et quand vous serez donc à la nef Chavant alors là euh il vous vous passez à droite du cinéma + une petite rue (Texte 2 « Itinéraire 1 & 2 »)*

6. Les topiques, à savoir les participants anticipés, actants thématiques, détachés en prolepse (disloqués en tête de phrase) avec ou sans reprise anaphorique dans le noyau :

(1) *la plongée c'est euh une longue histoire + ça a commencé par une paire de palmes et d'un tuba qui ont jamais vu la mer (Texte 1 « Plongée »)*

(15) *moi j'ai quarante cinq ans + quarante six même (Texte 3 « Doreur »)*

(16) *Grenoble euh la cote a baissé maintenant + le jeune parisien aujourd'hui il va préférer Montpellier Toulouse Toulouse maintenant (Texte 3 « Doreur »)*

(iii) Nous avons spécifié le type de construction phrastique réalisé en distinguant :

1. Les constructions verbales canoniques suivant l'ordre « sujet (nominal ou pronominal – verbe – complément(s) » :

(12) *et donc là on a fait une belle croisière avec un groupe de copains + donc il y avait Dédé (Texte 1 « Plongée »)*

2. Les constructions impersonnelles, typiquement en « (il) y a » ou encore « il faut » :

(49) nous dans le quartier on a on n'a pas de lycée + déjà y a deux collèges là rue rue d'Alésia (Texte 4 « 14-01 »)

(43) et puis bon quand même évidemment des nouveaux + mais euh il faut quand même bien reconnaître que ils sont plus jeunes donc moins proches du livre (Texte 4 « 14-01 »)

3. Les constructions dites « quasi-impersonnelles » ayant pour sujet un pronom démonstratif (« ce » ou « ça ») :

(10) c'est un voyage en Egypte en Mer Rouge + c'était une croisière (Texte 1 « Plongée »)

(22) on retire les poils + ça fera des des chapeaux (Texte 3 « Doreur »)

(37) donc pour vous ce n'est pas ça + ça fait un quartier un peu mort ça (Texte 4 « 14-01 »)

4. Les constructions elliptiques, correspondant à des rectifications, corrections, ou à des compléments différés (Tanguy 2010), ajouts :

(15) moi j'ai quarante cinq ans + quarante six même (Texte 3 « Doreur »)

(39) y a bien des gens d'un certain âge maintenant qui sont très âgés voilà voilà très très proches du livre euh de la bon des gens qui à l'époque il faut quand même le dire euh avaient quand même les moyens + c'était souvent ou des gens d'une cinquantaine d'années donc euh plus d'enfants à charge (Texte 4 « 14-01 »)

(40) donc très peu regardants bon + ils venaient acheter ce qu'ils avaient envie (Texte 4 « 14-01 »)

(41) dire un petit mot + bon le côté très traditionnelle de la vie (Texte 4 « 14-01 »)

5. Les phrases averbales (Lefevre 1999) :

(43) et puis bon quand même évidemment des nouveaux + mais euh il faut quand même bien reconnaître que ils sont plus jeunes donc moins proches du livre (Texte 4 « 14-01 »)

(iv) Enfin, nous avons marqué le temps du verbe des constructions verbales : présent, passé-composé, imparfait, plus-que-parfait et futur. Nous avons annoté comme construction au futur les énoncés du type :

(18) alors là donc je suis en train de poser la feuille d'or + je vais vous montrer puisqu'il me reste un morceau à mettre là (Texte 3 « Doreur »)

(24) ça va suivre ensemble + ça va pas casser (Texte 3 « Doreur »)

(25) tu continues + tu vas arriver sur la place euh Victor Hugo (Texte 2 « Itinéraire 1 & 2 »)

3.2. De la phrase au texte : les relations

Nous avons donc délimité nos différents textes en phrases, comme unités minimales de traitement et de production à vocation phrastique, selon des critères syntaxiques et énonciatifs. À un niveau supérieur, ces unités de type 'phrase' s'organisent entre elles et entretiennent des relations de type sémantique et discursif, régies par un principe de cohérence. (Charolles 2001 ; Charolles & Combettes 1999a & 1999b ; Combettes 1983 ; Adam 1990, 1992, 1999 & 2002 ; Halliday & Hasan 1976) et appelées « relations de discours » (Sperber & Wilson 1989). Parmi les relations établies, nous avons dégagé trois types de propriétés à observer : (i) le rapport temporel entre les deux unités 'phrases' au sein de la période, (ii) leur relation de discours et (iii) la présence d'expressions coréférentielles.

3.2.1. Les relations d'ordre temporel

Nous avons dans un premier temps marqué les rapports temporels (Hobbs 1979) existant entre la phrase 1 (désormais P1) et la phrase 2 (désormais P2). Nous avons distingué (i) les rapports d'antériorité (pour lesquels P1 précède P2) :

(25) *tu continues + tu vas arriver sur la place euh Victor Hugo (Texte 2 « Itinéraire 1 & 2 »)*

(ii) les rapports de recouvrement (pour lesquels P1 = P2) :

(9) */et, il y a/ également une réserve qui est magnifique + c'est du côté de de Banyuls (Texte 1 « Plongée »)*

et (iii) les rapports de postériorité (pour lesquels P1 suit P2) :

(45) *ils restent pas longtemps sur le coin + apparemment y a pas de très grands appartements (Texte 4 « 14-01 »)*

3.2.2. Les relations de discours

Suivant les travaux de la Segmented Discourse Representation Theory (Asher et al. 1995 : 17-20 ; Lascarides & Asher 1993 : 440) sur les relations de discours, nous sommes partie du principe qu'au sein des périodes binaires, P1 et P2 entretiennent l'une envers l'autre une relation de type discursif. Nous avons ainsi distingué les relations de narration, d'élaboration, de mise en arrière plan, d'explication et de résultat.

La relation de narration

La relation est dite « de narration » lorsque l'ordre textuel est le même que celui des événements décrits et lorsque les deux phrases renvoient à un même thème (ou topique). C'est par exemple :

(26) *tu continues dans la vieille ville + tu prends la grande rue (Texte 2 « Itinéraire 1 & 2 »)*

La relation d'élaboration

Les phrases P1 et P2 entretiennent une relation d'élaboration lorsque P1 correspond à un topique développé immédiatement par P2, entraînant ainsi un rapport hiérarchique entre les deux, P1 dominant P2 :

on a changé d'avion et on a pris un Boeing sept cent quarante-sept

La relation de mise en arrière plan ou 'background'

La relation est dite « de mise en arrière plan » lorsque P2 précise le contexte de P1, ou quand P1 est un élément de premier plan et P2 est un élément d'arrière-plan (Ashler et al. 1995 : 20) :

(12) et donc là on a fait une belle croisière avec un groupe de copains + donc il y avait Dédé (Texte 1 « Plongée »)

Nous avons classé ici les énoncés de l'ordre de la reformulation, de la rectification et du développement, du type :

(10) c'est un voyage en Egypte en Mer Rouge + c'était une croisière (Texte 1 « Plongée »)

(15) moi j'ai quarante cinq ans + quarante six même (Texte 3 « Doreur »)

La relation d'explication ou 'explication'

Une relation est dite « d'explication » lorsque l'événement auquel renvoie P2 correspond à la raison de l'événement de P1 :

(5) j'étais au lit + j'avais une grosse angine (Texte 1 « Plongée »)

La relation de résultat ou 'result'

Dans une relation de type résultat, l'événement auquel renvoie P1 est la cause de celui auquel renvoie P2 :

(6) j'ai eu mon cadeau de Noël + j'étais bien content (Texte 1 « Plongée »)

L'ordre énonciatif coïncide avec l'ordre des événements.

3.2.3. Expressions coréférentielles

Nous avons enfin annoté les différentes expressions coréférentielles fonctionnant comme indices syntaxiques de cohésion.

Nous avons défini deux ensembles : (i) les cas pour lesquels le sujet de P2 était identique à celui déjà exprimé précédemment dans P1 :

(7) et je savais pas nager non plus + donc j'ai jamais utilisé mon matériel (Texte 1 « Plongée »)

(24) ça va suivre ensemble + ça va pas casser (Texte 3 « Doreur »)

(26) tu continues dans la vieille ville + tu prends la grande rue (Texte 2 « Itinéraire 1 & 2 »)

(36) alors ça c'est clair + c'est l'avenue général Leclerc (Texte 4 « 14-01 »)

(ii) et les cas pour lesquels P2 contient une expression référentielle anaphorique d'un élément mentionné dans P1 :

(1) la plongée c'est euh une longue histoire + ça a commencé par une paire de palmes et d'un tuba qui ont jamais vu la mer (Texte 1 « Plongée »)

(12) et donc là on a fait une belle croisière avec un groupe de copains + donc il y avait Dédé (Texte 1 « Plongée »)

(29) et vous remontez euh jusqu'à la à une grande place + c'est la place de Verdun (Texte 2 « Itinéraire 1 & 2 »)

(33) euh j'en ai vu vu plusieurs sur paris donc + et pis c'- c'est ici que ça c'est fait finalement (Texte 4 « 14-01 »)

(49) nous dans le quartier on a on n'a pas de lycée + déjà y deux collèges là rue rue d'Alésia (Texte 4 « 14-01 »)

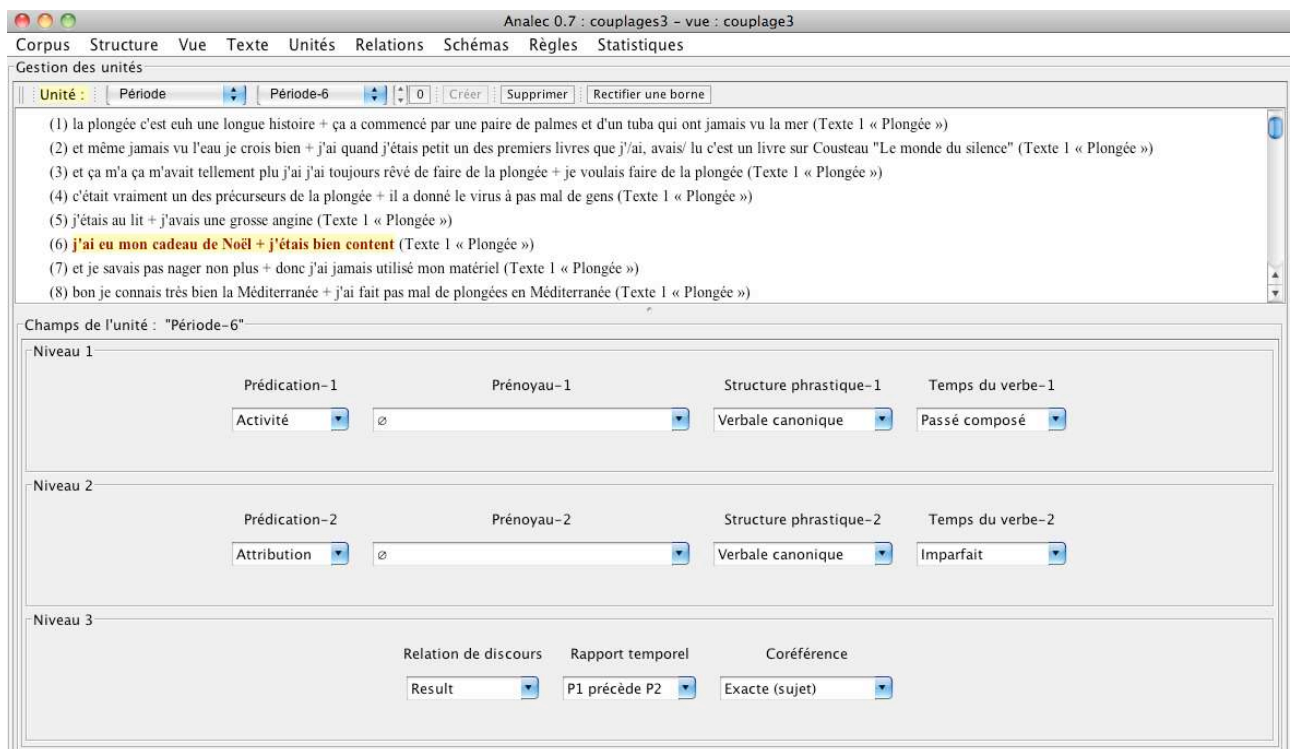


Figure 4. Annotation des propriétés des unités et des relations

4. Quelques résultats

Une fois l'annotation effectuée, le logiciel Analec permet différentes études statistiques automatiques de fréquences, de corrélations et de représentations géométriques.

4.1. Quelques fréquences

4.1.1. Sur les unités

Notre étude nous permet de vérifier que les prénoyaux apparaissent davantage dans la 1ère phrase de la période. En effet, une unité sur deux de type ‘période’ est introduite par un prénoyau, dans la première phrase (49 %). Ces prénoyaux sont majoritairement des ligateurs permettant de relier l’ensemble de la période (et non uniquement la phrase) à ce qui précède :

et, bon, donc, et/eh ben

Nous retrouvons ensuite, comme type de prénoyau, des topiques et des cadratifs. Notons qu’un cadratif de P1 porte généralement également sur P2 :

(49) nous dans le quartier on a on n'a pas de lycée + déjà y deux collèges là rue rue d'Alésia (Texte 4 « 14-01 »)

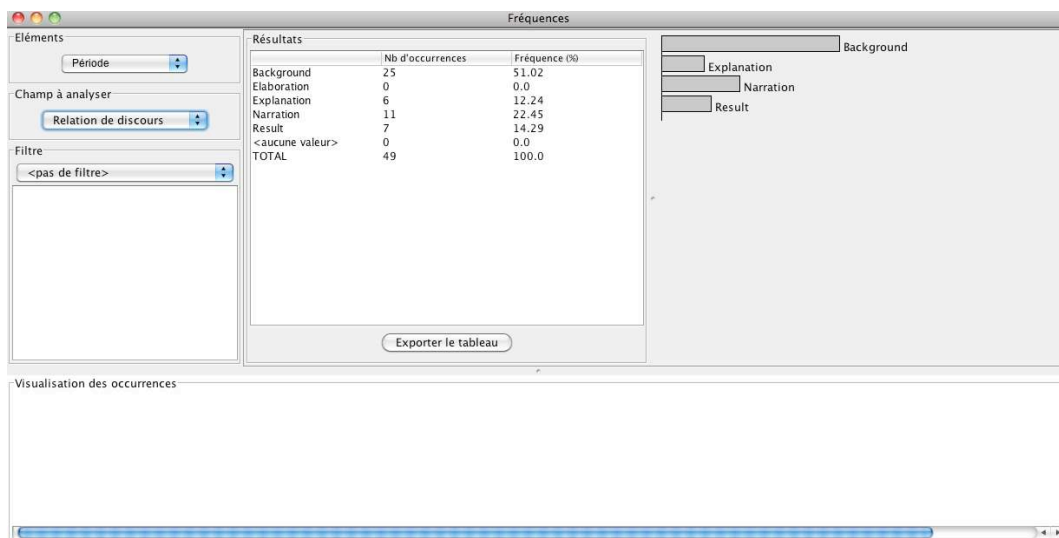
La portée n’est donc ici pas uniquement limitée à la phrase.

À opposer à ces résultats, moins de 30 % des deuxièmes phrases ont un prénoyau. Parmi ces 30 %, la moitié des prénoyau est de type ‘ligateur’ et quasiment toujours avec le coordonnant « et » :

et, et donc, et puis

4.1.2. Sur les relations

Le calcul de fréquences est intéressant pour les relations, et en particulier pour les relations de discours. Nous voyons que dans un cas sur deux, nous avons une relation de mise en arrière-plan, c’est-à-dire des cas pour lesquels la deuxième phrase de la période précise le contexte énoncé dans la première phase ou alors quand P2 constitue l’arrière plan de P1 au premier plan. Ce résultat, couplé avec les fréquences de relations de narration et de résultat, montre bien que pour 87% des périodes, on a une hiérarchie discursive entre P1 et P2, la première dominant la seconde davantage descriptive et/ou explicative.



The screenshot shows a software window titled "Fréquences". On the left, there are controls for "Éléments" (Période), "Champ à analyser" (Relation de discours), and "Filtre" (<pas de filtre>). The main area contains a table with the following data:

	Nb d'occurrences	Fréquence (%)
Background	25	51.02
Elaboration	0	0.0
Explanation	6	12.24
Narration	11	22.45
Result	7	14.29
<aucune valeur>	0	0.0
TOTAL	49	100.0

On the right, there is a legend with colored boxes for "Background", "Explanation", "Narration", and "Result". Below the table is a button labeled "Exporter le tableau". At the bottom, there is a section for "Visualisation des occurrences" which is currently empty.

Figure 5. Calcul des fréquences des relations de discours

La cohésion entre P1 et P2 est renforcée par des indices coréférentiels. 35 % des périodes ont un sujet identique, typiquement un pronom personnel. À cela, s'ajoute les cas pour lesquels la phrase 2 comporte un élément anaphorique dont l'antécédent se trouve dans la phrase 1 : 40 %. Ainsi, les trois quarts des phrases des périodes binaires comportent une relation anaphorique.

4.2. Quelques corrélations

Nous avons cherché à mettre en lien le type de prédication de P1 (champ 2, vertical) et celui de P2 (champ 1, horizontal).

	Activité	Attribution	Cognitivité	Existence	Localisation	TOTAL
Activité	7	4	1	3	1	16
Attribution	5	9	1	1	2	18
Cognitivité	2	0	2	0	0	4
Existence	0	3	0	2	1	6
Localisation	0	3	0	1	0	4
TOTAL	14	20	4	7	4	49

Figure 6. Calcul des corrélations des types de prédication

Nous remarquons qu'une activité appelle une activité, ou une attribution. Une prédication d'attribution est le plus souvent couplée d'une autre attribution. Elle peut être cependant également suivie d'une activité. Une prédication de type 'cognitivité' en P1 est suivie d'une prédication d'activité et de cognitivité. L'existentielle est associée à une attributive ou à une autre existentielle. La localisation est typiquement renforcée par une attribution.

Schématiquement, nous avons typiquement des constructions doubles identiques, c'est-à-dire des prédictions avec le même type de prédication : attribution-attribution, activité-activité, etc.

4.3. Quelques représentations géométriques

Le logiciel Analec permet enfin d'obtenir automatiquement des représentations géométriques et ainsi, de dégager des valeurs typiques. Nous en mentionnerons ici les principales.

- (i) Les périodes avec une relation de narration ont des verbes au présent marquant une prédication d'activité, avec un rapport temporel où P1 précède P2.
- (ii) Les périodes binaires construites sur le modèle « attribution-attribution » sont typiquement composées de phrases au présent, la seconde étant majoritairement quasi-

impersonnelle (c'est-à-dire ayant pour sujet un pronom démonstratif « ce » ou « ça »). La relation entre P1 et P2 est de type 'background', d'arrière-plan, avec un rapport temporel de recouvrement où P1 = P2. De plus, il existe une relation anaphorique entre un élément de P1 et de P2. Cette valeur typique peut être illustrée par la période :

(9) /et, il y a/ également une réserve qui est magnifique + c'est du côté de de Banyuls (Texte 1 « Plongée »)

(iii) Les phrases marquant une relation d'explication (explication) au sein d'une période sont liées entre elles sans prénoyau (pas de prénoyau introducteur de la deuxième phrase). Leur structure syntaxique est typiquement celle d'une phrase verbale canonique suivant l'ordre « sujet – verbe – complément(s) » avec une coréférence exacte du sujet. L'illustration type serait :

(8) bon je connais très bien la Méditerranée + j'ai fait pas mal de plongées en Méditerranée (Texte 1 « Plongée »)

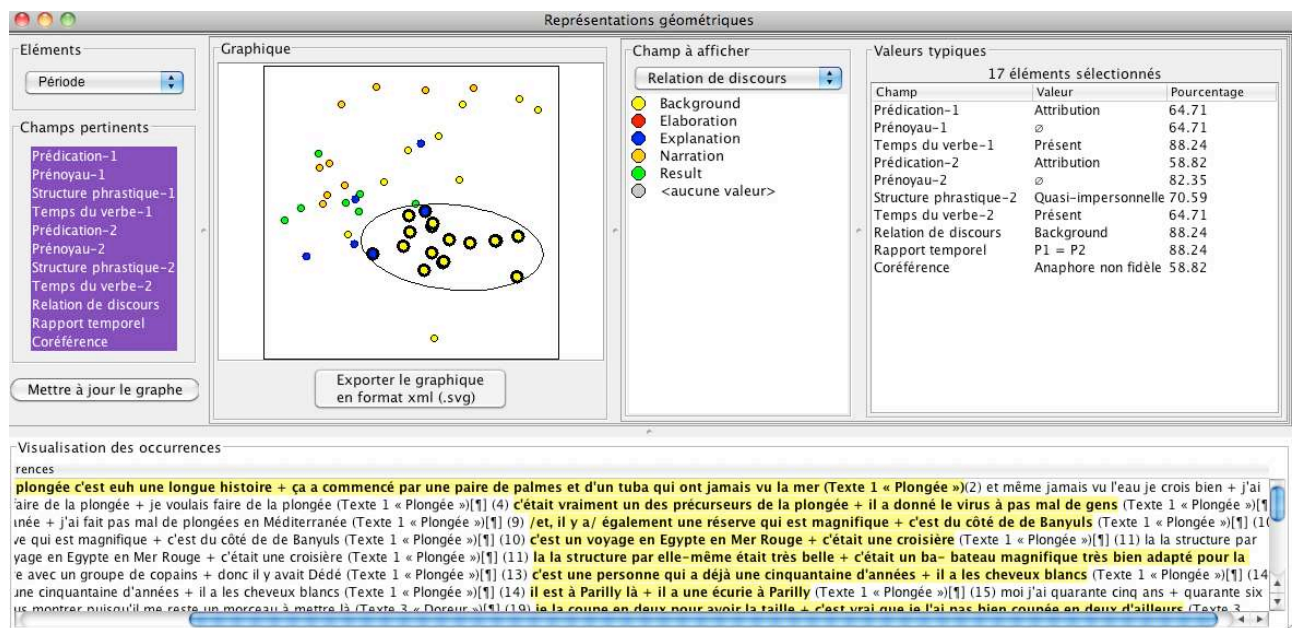


Figure 7. Calcul des représentations géométriques des relations de discours

Chaque point dans le graphique représente un exemple annoté. En fonction de leur place et de leur type, nous pouvons dégager certains regroupements et ainsi mettre en évidence les valeurs typiques des ensembles sélectionnés.

5. Conclusion

Nous espérons avec ce travail avoir posé les premières bases d'une étude à l'interface syntaxe, prosodie et sémantique. Nous avons démontré que nous avons très fréquemment en français parlé des structures binaires du type :

(4) c'était vraiment un des précurseurs de la plongée + il a donné le virus à pas mal de gens (Texte 1 « Plongée »)

combinant ici une prédication existentielle suivie d'une prédication d'activité. L'énonciateur introduit un nouveau référent dans son discours via une structure quasi-impersonnelle en « c'était » que la phrase suivante développe. L'ensemble constitue une nouvelle unité informative structurée.

Il en est de même avec des ensembles plus longs comme (1) qui constitue un ensemble cohésif :

(1) *c'est là que j'ai commencé à conduire - un camion - sans permis mais là - c'était l'exode on s'en foutait (CTFP)*

Même si l'anaphorique « là » fait référence à un cadre spatial précédemment cité, la première phrase de ce regroupement introduit dans le récit une première information que la deuxième prédication nuance via le « mais ». La dernière prédication intervient comme conclusion. Ce lien est renforcé avec l'emploi du pronom « en », anaphorique de ce qui précède.

Les regroupements textuels s'opèrent ainsi sur des bases sémantiques et discursives, avec l'appui d'indices syntaxiques (structures répétées, emploi de pronoms, etc.).

Nous devons donc maintenant étendre cette analyse à des corpus plus vastes et appliquer ce modèle d'annotation à des périodes plus importantes. Nous savons qu'une période regroupe deux à sept phrases. Nous chercherons à mettre en rapport les structures binaires, ternaires, ou plus, avec des schémas informationnels.

Références bibliographiques :

- ADAM J-M. (1990) *Éléments de linguistique textuelle*. Paris : Mardaga.
- ADAM J-M. (1992) *Les textes : types et prototypes*. Paris : Nathan. coll. : « fac ».
- ADAM J-M. (1999) *Linguistique textuelle. Des genres de discours aux textes*. Paris : Nathan.
- ADAM J-M. (2002) « De la période à la séquence. Contribution à une (trans)linguistique textuelle comparative ». In ANDERSEN H. L. & NØLKE H. (éds) *Macro-syntaxe et macro-sémantique*. Berne : Peter Lang. 167-188.
- ASCHER N. *et al.* (1995) « De l'espace temps dans l'analyse du discours ». *Sémiotiques*. 9. 13-61.
- AVANZI M. (2005) « Quelques hypothèses à propos de la structuration interne des périodes ». *Actes du symposium Interface Discours-Prosodie*. Aix-en-Provence. 8-9 sept. 05. CD-ROM.
- AVANZI M. (2008) « La différence entre micro- et macro-syntaxe est-elle marquée prosodiquement ? L'exemple des dispositifs clivés en *il y a SN qui / Ø V* ». *L'information Grammaticale*. 119. 8-13.
- AVANZI M. (2010) « Rattachement et fragmentation de la syntaxe par la prosodie ». *Travaux de linguistique*. 60. 145-166.
- AVANZI M. & LACHERET-DUJOUR A. (2010) « Micro-syntaxe, macro-syntaxe : une prosodie toujours transparente ? L'exemple des périodes asyndétiques en français

- parlé ». In BÉGUELIN M.-J., AVANZI M. & CORMINBOEUF G. (éds). *La parataxe. vol. 2 : Structures, marquages et exploitation discursive* (Actes du colloque « La parataxe », Neuchâtel, 12-15 février 2007). Bern: Peter Lang, pp. 339-364.
- AVANZI M. & TANGUY N. (2009) « Structuration syntaxique et structuration prosodique du français parlé ». Communication au séminaire LATTICE (15 mai 2009). Université Sorbonne Nouvelle Paris 3.
- AVANZI M., LACHERET-DUJOUR A. & VICTORRI B. (2008) « Analor, un outil d'aide pour la modélisation de l'interface prosodie – grammaire ». *Cahiers du CERLICO*. 21. 27-46.
- AVANZI, M., SIMON, A.C., GOLDMAN, J.-P. & AUCHLIN A. (2010) « C-PROM. Un corpus de français parlé annoté pour l'étude des prééminences ». Actes des 23èmes journées d'étude sur la parole (Mons, Belgique, 25-28 mai 2010), <https://sites.google.com/site/corpusprom/>.
- BEGUELIN M.-J. (2002) « Clause, période ou autre ? La phrase graphique et la question des niveaux d'analyse ». *Verbum*. XXIV. 1-2. 85-107.
- BERRENDONNER A. (1983) « Connecteurs pragmatiques et anaphores ». *Cahiers de linguistique française*. 5. 215-246.
- BERRENDONNER A. (1990) « Pour une macro-syntaxe ». *Travaux de linguistique*. 21. 25-36.
- BERRENDONNER A. (1993) « Périodes ». In PARRET H. (éd.) *Temps et discours*. Louvain : Presses Universitaires. 47-61.
- BLANCHE-BENVENISTE C. (1997) *Approches de la langue parlée en français*. Paris : Ophrys. coll. : « l'essentiel ».
- BLANCHE-BENVENISTE C (2002) « Macro-syntaxe et micro-syntaxe : les dispositifs de la rection verbale ». In ANDERSEN H. L. & NØLKE H. (éds). *Macro-syntaxe et macro-sémantique*. Berne : Peter Lang. 95-118.
- BLANCHE-BENVENISTE C. et al. (1984) *L'Approche pronominale et son application au français*. Paris : SELAF.
- BLANCHE-BENVENISTE C. et al. (1990) *Le Français parlé : études grammaticales*. Paris : Éditions du CNRS.
- BRANCA-ROSOFF S., FLEURY S., LEFEUVRE F. & PIRES M. (2009) « Discours sur la ville. Corpus de Français Parlé Parisien des années 2000 (CFPP2000) ». <http://ed268.univ-paris3.fr/CFPP2000/>
- CHAROLLES M. (2001) « De la phrase au discours : quelles relations ? ». In ROUSSEAU A. (éd.) *La sémantique des relations*. Université de Lille 3. Travaux et recherches. 237-260.
- CHAROLLES M. & COMBETTES B. (1999a) « Contribution pour une histoire récente de l'analyse du discours ». *Langue française*. 121. 76-116.
- CHAROLLES M. & COMBETTES B. (1999b) « De la phrase au discours : rupture et continuité ». *Langue Française*. 121. 76-116.

- CHAROLLES M. & VIGIER D. (2005) « Les adverbiaux en position préverbale : portée cadratrice et organisation des discours ». *Langue Française*. 148. 9-30.
- COMBETTES B. (1983) *Pour une grammaire textuelle. La progression thématique*. Bruxelles-paris : De Boeck-Duculot.
- CRESTI E. (1999) « Force illocutoire, articulation topic/comment et contour prosodique en italien parlé ». *Faits de Langue*. 13. 168-181
- CRESTI E. (2000) « Critère illocutoire et articulation informative ». In BILGER M. (éd.) *Corpus. Méthodologie et applications linguistiques*. Paris : Champion. 350-367.
- DELIC (2004) « Présentation du Corpus de Référence du Français Parlé ». *RSFP* 18. 11-42.
- DEGAND L. & SIMON A.-C. (2008) « Minimal discourse units in spoken French : Uncovering genre-bound segmentation strategies ». In RAMM W. & FABRICIUS-HANSEN C. (eds.) *Linearisation and Segmentation in Discourse. Multidisciplinary Approaches to Discourse 2008 (MAD 08)*, Feb 20-23 2008, Lysebu, Oslo.
- HALLIDAY M. A. K. & HASAN R. (1976) *Cohesion in English*. Longman English Language Series 9. London: Longman.
- HOBBS J. R. (1979) « Coherence and coreference ». *Cognitive Science* 3. 67-90.
- LACHERET-DUJOUR A. (2003) *La prosodie des circonstants en français parlé*. Louvain/Paris : Peeters.
- LACHERET-DUJOUR A. & VICTORRI B. (2002) « La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques ». *Verbum*. 24/1-2. 55-73.
- LASCARIDES A. & ASHER N. (1993) « Temporal Interpretation, Discourse Relations and Com- monsense Entailment ». *Linguistics and Philosophy*. 16. 437-493.
- LEFEUVRE F. (1999) *La phrase averbale en français*. Paris : L'Harmattan.
- LE GOFFIC P. (1993) *Grammaire de la Phrase Française*. Paris : Hachette Supérieur.
- LE GOFFIC P. (2008) « Phrase, séquence, période ». In VAN RAEMDONCK D. (éd.) *Modèles syntaxiques. La syntaxe à l'aube du XXI^e siècle*. Bruxelles : Peter Lang, 329-356.
- MOREL M.-A. (2003) « Phrase ? Énoncé ? Paragraphe ? Hyperparagraphe ? Quelles unités intonatives et discursives pour le dialogue oral en français ? ». *L'information grammaticale*. 98. 39-47.
- MOREL M.-A. & DANON-BOILEAU L. (1998) *Grammaire de l'intonation. L'exemple du français*. Paris : Ophrys.
- POTTIER B. (1992) *Sémantique générale*. Paris PUF.
- SPERBER D. & WILSON D. (1989) *La pertinence*. Paris : Minuit.
- TANGUY N. (2009) *Les segments averbaux, unités syntaxiques de l'oral*. Thèse de doctorat. Université Sorbonne Nouvelle Paris 3.

TANGUY N. (2010) « Étude des compléments différés à l'oral à l'interface syntaxe – prosodie ». Actes du 2^e Congrès Mondial de Linguistique Française (12-15 juillet 2010 – La Nouvelle Orléans, États-Unis). pp. 1415-1427.

VICTORRI B. (2011) « Analec : logiciel d'annotation et d'analyse de corpus écrits », logiciel téléchargeable sur : <http://www.lattice.cnrs.fr/analec>

Annexes : Relevé des périodes binaires

- (1) la plongée c'est euh une longue histoire + ça a commencé par une paire de palmes et d'un tuba qui ont jamais vu la mer (Texte 1 « Plongée »)
- (2) et même jamais vu l'eau je crois bien + j'ai quand j'étais petit un des premiers livres que j'/ai, avais/ lu c'est un livre sur Cousteau "Le monde du silence" (Texte 1 « Plongée »)
- (3) et ça m'a ça m'avait tellement plu j'ai j'ai toujours rêvé de faire de la plongée + je voulais faire de la plongée (Texte 1 « Plongée »)
- (4) c'était vraiment un des précurseurs de la plongée + il a donné le virus à pas mal de gens (Texte 1 « Plongée »)
- (5) j'étais au lit + j'avais une grosse angine (Texte 1 « Plongée »)
- (6) j'ai eu mon cadeau de Noël + j'étais bien content (Texte 1 « Plongée »)
- (7) et je savais pas nager non plus + donc j'ai jamais utilisé mon matériel (Texte 1 « Plongée »)
- (8) bon je connais très bien la Méditerranée + j'ai fait pas mal de plongées en Méditerranée (Texte 1 « Plongée »)
- (9) /et, il y a/ également une réserve qui est magnifique + c'est du côté de de Banyuls (Texte 1 « Plongée »)
- (10) c'est un voyage en Egypte en Mer Rouge + c'était une croisière (Texte 1 « Plongée »)
- (11) la la structure par elle-même était très belle + c'était un ba- bateau magnifique très bien adapté pour la plongée (Texte 1 « Plongée »)
- (12) et donc là on a fait une belle croisière avec un groupe de copains + donc il y avait Dédé (Texte 1 « Plongée »)
- (13) c'est une personne qui a déjà une cinquantaine d'années + il a les cheveux blancs (Texte 1 « Plongée »)
- (14) il est à Parilly là + il a une écurie à Parilly (Texte 1 « Plongée »)
- (15) moi j'ai quarante cinq ans + quarante six même (Texte 3 « Doreur »)
- (16) Grenoble euh la cote a baissé maintenant + le jeune parisien aujourd'hui il va préférer Montpellier Toulouse Toulouse maintenant (Texte 3 « Doreur »)
- (17) oui + d'ailleurs aujourd'hui je pense c'est c'est à Toulouse que j'irais si si si ça devait se refaire (Texte 3 « Doreur »)

- (18) alors là donc je suis en train de poser la feuille d'or + je vais vous montrer puisqu'il me reste un morceau à mettre là (Texte 3 « Doreur »)
- (19) je la coupe en deux pour avoir la taille + c'est vrai que je l'ai pas bien coupée en deux d'ailleurs (Texte 3 « Doreur »)
- (20) un pinceau en martre aussi + et je mouille (Texte 3 « Doreur »)
- (21) voilà + et en fait on va voir que la feuille elle est elle est aspirée par l'eau qui en même temps se mélange à de la colle parce que le le lien de toutes les couches qu'y a en dessous c'est la co- la colle de peau de lapin (Texte 3 « Doreur »)
- (22) on retire les poils + ça fera des des chapeaux (Texte 3 « Doreur »)
- (23) et et c'est cette co- cette peau euh + donc c'est le collagène de la peau du lapin (Texte 3 « Doreur »)
- (24) ça va suivre ensemble + ça va pas casser (Texte 3 « Doreur »)
- (25) tu continues + tu vas arriver sur la place euh Victor Hugo (Texte 2 « Itinéraire 1 & 2 »)
- (26) tu continues dans la vieille ville + tu prends la grande rue (Texte 2 « Itinéraire 1 & 2 »)
- (27) et euh après tu bifurques euh + y a une petite bifurcation euh juste avant la place du Tribunal (Texte 2 « Itinéraire 1 & 2 »)
- (28) et quand vous serez donc à la nef Chavant alors là euh il vous vous passez à droite du cinéma + une petite rue (Texte 2 « Itinéraire 1 & 2 »)
- (29) et vous remontez euh jusqu'à la à une grande place + c'est la place de Verdun (Texte 2 « Itinéraire 1 & 2 »)
- (30) et vous suivez toujours la ligne du tram + et là vous tombez donc euh à la maison du tourisme (Texte 2 « Itinéraire 1 & 2 »)
- (31) c'est pas bien compliqué à y aller hein + vous pouvez pas vous tromper (Texte 2 « Itinéraire 1 & 2 »)
- (32) donc euh dans une petite rue euh qui donne sur la place denfert rocheraud + mes parents étaient commerçant rue Daguerre (Texte 4 « 14-01 »)
- (33) euh j'en ai vu vu plusieurs sur paris donc + et pis c'- c'est ici que ça c'est fait finalement (Texte 4 « 14-01 »)
- (34) euh ça doit être en vingt euh + je crois que c'est vingt-six (Texte 4 « 14-01 »)
- (35) et ben c'était une de mes questions + alors je vous écoute (Texte 4 « 14-01 »)
- (36) alors ça c'est clair + c'est l'avenue général Leclerc (Texte 4 « 14-01 »)
- (37) donc pour vous ce n'est pas ça + ça fait un quartier un peu mort ça (Texte 4 « 14-01 »)
- (38) il s'est rajeuni + et euh donc on n'a pas euh quelque part on a on a un peu perdu la clientèle traditionnelle euh (Texte 4 « 14-01 »)
- (39) y a bien des gens d'un certain âge maintenant qui sont très âgés voilà voilà très très proches du livre euh de la bon des gens qui à l'époque il faut quand même le dire euh

- avaient quand même les moyens + c'était souvent ou des gens d'une cinquantaine d'années donc euh plus d'enfants à charge (Texte 4 « 14-01 »)
- (40) donc très peu regardants bon + ils venaient acheter ce qu'ils avaient envie (Texte 4 « 14-01 »)
- (41) dire un petit mot + bon le côté très traditionnelle de la vie (Texte 4 « 14-01 »)
- (42) euh eh bien on a toujours un peu ceux-là qui vieillissent euh + y en a d'autres qui sont partis du quartier qui reviennent de temps en temps (Texte 4 « 14-01 »)
- (43) et puis bon quand même évidemment des nouveaux +mais euh il faut quand même bien reconnaître que ils sont plus jeunes donc moins proches du livre (Texte 4 « 14-01 »)
- (44) euh ils sont en général bon pour vivre dans ce quartier + il faut quand même beaucoup d'argent parce que les loyers sont très très chers ici euh (Texte 4 « 14-01 »)
- (45) ils restent pas longtemps sur le coin + apparemment y a pas de très grands appartements (Texte 4 « 14-01 »)
- (46) donc euh ça c'est un petit peu pour ce qui a été des changements de du quartier + moi je suis là depuis vingt trois ans (Texte 4 « 14-01 »)
- (47) nous on a pas ce passage de l'avenue + les gens ne rentrent pas dans ce quartier maintenant (Texte 4 « 14-01 »)
- (48) j'avais pas prévu de vous interroger là dessus + mais du coup vous pouvez pas travailler avec les lycées et (Texte 4 « 14-01 »)
- (49) nous dans le quartier on a on n'a pas de lycée + déjà y deux collèges là rue rue d'Alésia (Texte 4 « 14-01 »)