

Des relations d'alignement pour décrire l'interaction des domaines linguistiques : vers des Grammaires Multimodales

Philippe Blache

Laboratoire Parole et Langage

CNRS & Université de Provence
blache@lpl-aix.fr

Résumé

Un des problèmes majeurs de la linguistique aujourd'hui réside dans la prise en compte de phénomènes relevant de domaines et de modalités différentes. Dans la littérature, la réponse consiste à représenter les relations pouvant exister entre ces domaines de façon externe, en termes de relation de structure à structure, s'appuyant donc sur une description distincte de chaque domaine ou chaque modalité. Nous proposons dans cet article une approche différente permettant représenter ces phénomènes dans un cadre formel unique, permettant de rendre compte au sein d'une même grammaire tous les phénomènes concernés. Cette représentation précise de l'interaction entre domaines et modalités s'appuie sur la définition de relations d'*alignement*.

Abstract

Linguistics is now faced with the question of representing information coming from different domains or modalities. The classical answer consists in representing separately each of these domains, building for each of them an independent structure, and then representing domain interaction in terms of relation between structures. We propose in this paper a different approach in which all information is represented within a unique and homogeneous framework, making it possible to represent into a same grammar all interaction phenomena. This precise representation of interaction relies on the definition of a new notion of *alignment* relations.

Mots-clés : *Multimodalité, interaction entre domaines, grammaire, corpus multimodaux*

Keywords: *Multimodality, domains interaction, grammar, multimodal corpora*

1 Introduction

La linguistique est confrontée à la nécessité de développer un cadre théorique prenant en compte l'ensemble de ses domaines d'information que sont la prosodie, la syntaxe, la pragmatique, etc. De nombreuses études décrivent les interactions pouvant exister entre eux, en particulier concernant l'interaction prosodie/syntaxe (voir par exemple [Hirst96], [Selkirk84], [Doetjes02], [Delais-Roussarie08], etc.). Cependant, ces études sont presque

systématiquement abordées en termes d'interaction entre structures (par exemple entre arbre syntaxique et structure prosodique) et s'appuient donc sur une conception des domaines indépendants les uns des autres. Chaque domaine dispose dans cette perspective de sa propre représentation (accessoirement de sa propre théorie et de sa propre grammaire). Or, ce type d'explication, qui s'appuie de fait sur une recherche de congruence entre structures, s'avère trop rigide lorsqu'elle est mise à l'épreuve de données réelles. La complexité du problème se renforce encore dès lors qu'on tente de décrire les interactions entre différentes *modalités* (par exemple entre le geste et la parole, voir [McNeil05]). On sait aujourd'hui que la compréhension des mécanismes d'élaboration d'un message (donc de la communication) passe par la description de ces phénomènes.

Le but de cet article est de proposer une nouvelle approche proposant de représenter ces phénomènes dans un cadre formel unique, permettant de rendre compte au sein d'une même grammaire tous les phénomènes concernés. Cette représentation précise de l'interaction entre domaines et modalités s'appuie sur la définition de relations d'*alignement*. Nous exposerons dans une première partie un ensemble de phénomènes d'alignement observés dans un corpus multimodal qui nous permettront, dans une seconde partie, de définir ces relations d'alignement. Nous en proposerons dans la dernière section une représentation sous forme de contraintes et terminerons par une discussion des résultats.

2 L'annotation de corpus multimodaux

Plusieurs projets internationaux portant sur la question des données multimodales ont été mis en place ces dernières années. Il s'agit en particulier des projets MATE (*Multilevel Annotation, Tools Engineering*, [Dybkjaer98]), ATLAS (*Architecture and Tools for Linguistic Analysis System*, [Bird00]), NITE (*Natural Interactivity Tools Engineering*), Map Task (conduit au HCRC), DAMSL (*Dialog Act Markup in Several Layers*, [Core97]) ou encore Verbmobil. Il n'existe cependant que peu d'initiatives portant sur la constitution de corpus multimodaux visant une certaine exhaustivité dans les annotations. La principale ressource développée dans ce domaine est le corpus AMI (voir [Carletta05]) qui contient 100h de réunions filmées qui ont été transcrites et annotées principalement pour les gestes et la structure de l'information.

Pour le français, le projet OTIM (*Outils de traitement d'information multimodale*)¹ se fixe pour objectif de développer une ressource de ce type mais pour laquelle l'ensemble des domaines est annoté aussi précisément que possible. Le CID (*Corpus of Interactional Data*, cf. [Bertrand09]) est la première ressource réalisée dans ce cadre. Il est composé de 8 dialogues de 1h, représentant environ 100.000 mots. Le corpus est totalement transcrit orthographiquement, phonétiquement et toutes les annotations sont alignées sur le signal acoustique, offrant ainsi la possibilité d'une étude précise des phénomènes d'interaction. Les transcriptions orthographiques et phonétiques permettent la production d'annotations de plus haut niveau (morphologie, syntaxe, prosodie, pragmatique, etc.) elles-mêmes alignées sur le signal. Il s'agit d'une ressource unique en son genre pour le français tant par sa taille que par la précision des annotations. A notre connaissance, aucune autre ressource comparable par la taille et le niveau d'information n'est disponible pour une autre langue. Il ouvre la possibilité de nouveaux types de recherches sur de la parole spontanée.

¹ Projet ANR conduit par le LPL en partenariat avec les laboratoires LIMSI, LSIS, LIA, RFC et LLING.

Les enrichissements du CID concernent à ce jour : transcription orthographique, transcription phonétique, unités interpausales, unités prosodiques, fonctions prosodiques, backchannels, unités de constructions de tours, gestes, étiquettes morpho-syntaxiques, chunks, segments syntaxiques (ou *pseudo-phrases*). Une partie du corpus (trois dialogues) a été également annotée pour ce qui concerne les phénomènes de détachement. Ces transcriptions ou annotations ont été produites automatiquement lorsque les outils existaient. Par exemple, le traitement de la transcription orthographique (segmentation, étiquetage, chunking) a été réalisé à l'aide d'outils développés au LPL (évalués dans le cadre des campagnes *Easy* et *Passage*, cf. [Blache08]). Le reste a été produit manuellement par des experts.

3 Interactions entre domaines observées dans le corpus

Nous proposons dans cette section une illustration du type d'information portant sur l'interaction de plusieurs domaines (et donc plusieurs niveaux d'annotations) qu'il devient possible d'extraire grâce à ce type de corpus. Nous nous sommes appuyés pour identifier la cooccurrence d'événements sur une synchronisation relative des frontières droites des unités annotés. Cette synchronisation a été calculée avec une marge variable de recouvrement entre unités en fonction du type de phénomène et de la façon dont il a été annoté (automatiquement, ou manuellement). Ainsi, les unités syntaxiques, générées par l'analyseur, sont alignées exactement sur le signal grâce à la synchronisation des transcriptions phonétiques et orthographiques. En revanche, les unités prosodiques ou les détachements ayant été annotés manuellement, il y a donc une marge d'erreur.

3.1 Unités prosodiques et syntaxiques

La description de la structure prosodique repose sur l'identification d'unités constituant la structure prosodique (cf. [Jun02]) : le syntagme accentuel (AP) et le syntagme intonatif (IP), caractérisés par des contours² prosodiques particuliers. Le CID a été annoté en utilisant les contours prosodiques suivants : m0 (montée mineure), m1 (autres contours mineurs), RMC (continuation majeure), ENUM (contour d'énumération), f1 (contour plat), F (contour descendant), R (contour montant), RF1 (contour montant-descendant), RF2 (contour descendant depuis la pénultième), RQ (montée de question), RT (montée terminale).

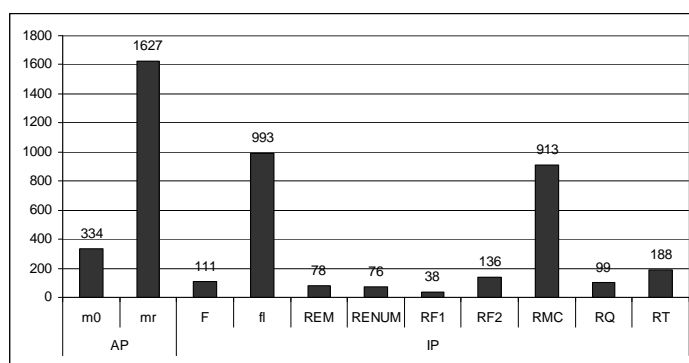


Figure 1 : Répartition des contours prosodiques par type d'unité

² Un contour est un mouvement mélodique généralement caractérisé par une direction (montant, descendant, etc.), une position (finale, pénultième, etc.) et une fonction (conclusif, continuatif).

Un des problèmes posés pour l'analyse de la structure prosodique, une fois les contours détectés, est celui de l'identification des frontières des constituants. La figure (1) illustre le type de contours observés sur une heure de corpus. Ces résultats présentent l'intérêt de porter sur des données réelles, permettant de confirmer les observations établies dans la littérature sur la base de données construites. On constate que le contour /mɾ/³ est de loin le plus utilisé pour les AP tandis que les contours des IP sont concentrés autour de deux réalisations : /fɛl/ et /RMC/. Bien entendu, le style de discours utilisé dans ce corpus (majoritairement narratif) influence également la répartition des contours réalisés. Ce schéma n'est donc pas totalement généralisable, mais suffisamment marqué pour en proposer une première relation existant entre /AP/ et /mɾ/ d'une part et /IP/ et {/RMC/, /fɛl/} d'autre part.

Nous avons ensuite recherché les relations pouvant exister entre unités prosodiques et syntaxiques. Nous avons pour cela utilisé une analyse en *chunks* réalisée par un des parseurs du LPL (cf. Blache08) et fournissant un découpage en unités syntaxiques non récursives suivant la définition proposée dans le cadre de la campagne Easy. Les groupes utilisés ici sont : GA (groupe adjectival), GN (nominal), GP (prépositionnel), GR (adverbial), NV (noyau verbal), PV (groupe verbal prépositionnel). La figure suivante présente pour chaque unité syntaxique le type d'unité prosodique correspondant. Deux groupes présentent des particularités : le GR qui est réalisé principalement comme frontière droite d'un IP et le NV qui, à la différence des autres groupes, est majoritairement relié à un AP. Cette observation confirme sur une plus grande échelle une observation sur un extrait de corpus faite dans Bertrand07 relevant une réalisation des contours mineurs en fin de construction verbale, ou entre deux constituants régis par le verbe (sujet ou complément). Une analyse de la répartition des contours par chunks, présentée dans la figure suivante, montre une relation forte entre le /NV/ et le contour /mɾ/ de même qu'entre /GR/ et le contour /fɛl/. En revanche, les autres contours semblent proportionnellement mieux répartis pour les groupes.

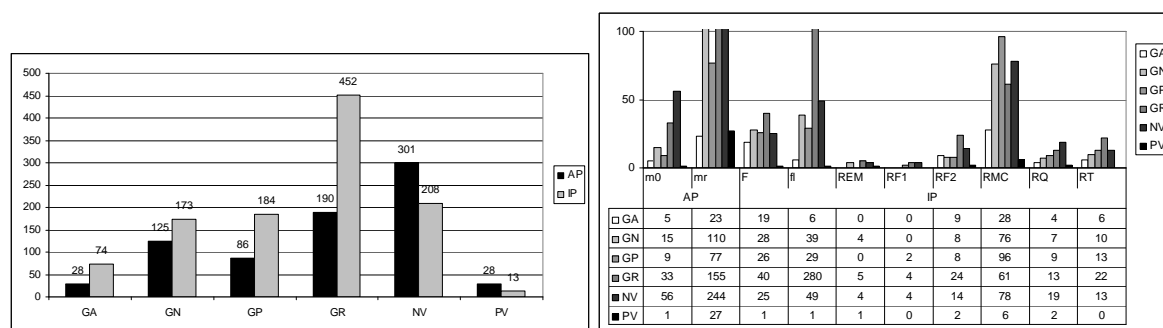


Figure 2 : Répartition des contours prosodiques par unités syntaxiques

Nous n'entrons pas ici dans l'interprétation des relations observées, ce travail faisant l'objet d'une étude linguistique en cours. En revanche, il est utile, notamment dans une perspective de catégorisation ou de désambiguïsation lors du processus d'analyse, de disposer de ces informations : la connaissance du contour prosodiques permettra de favoriser certaines catégories au détriment d'autres.

³ On note dans le reste de l'article les étiquettes entre barres obliques.

3.2 Deux phénomènes : *détachements* et *backchannels*

Nous présentons dans cette section deux types d'annotations de plus haut niveau, effectuée manuellement par des experts⁴ et permettant d'illustrer l'intérêt de ce type de ressources. Ces annotations portent sur les *constructions détachées* (phénomènes consistant à détacher un élément de la phrase et le réaliser généralement avant elle) et les *backchannels* (phénomènes discursifs, vocaux ou gestuels, de réaction de l'interlocuteur au discours d'un locuteur).

L'annotation des structures détachées a porté sur les phénomènes suivants : SbjD (sujet lexicalisé disloqué), SbjProD (pronom sujet disloqué), DirObjD (objet direct lexicalisé disloqué), DirObjProD (pronom objet disloqué), IndObjD (objet indirect lexicalisé disloqué), IndObjProD (pronom objet indirect disloqué). De plus, les sujets lexicalisés ont également été repérés (étiquetés /Sbj/, sujet lexicaux), dans la perspective d'une étude comparative sur le redoublement clitique.

Le premier tableau de la figure 3 indique la répartition des contours utilisés sur ces constructions, tous types confondus. Il est important de souligner que le niveau d'information est ici très général dans la mesure où il cumule tous les types de détachement. Il n'est donc pas possible ici de rechercher des relations directes qui permettraient d'associer un type de contour à un type de détachement. Par ailleurs, il faut rappeler que les unités syntaxiques générées sont des *chunks* non récursifs. L'alignement, obtenu automatiquement, indique les relations entre l'élément disloqué (nous y reviendrons plus loin) et le dernier chunk le composant (qui n'est pas nécessairement l'unité syntaxique de plus haut niveau qui correspondrait au segment étudié, mais simplement la dernière), expliquant par exemple la proportion importante de GR reportée dans le tableau. Nous sommes cependant dans cet article essentiellement intéressés par les aspects quantitatifs de la corrélation entre domaines. Les informations fournies sont de ce point de vue fiables.

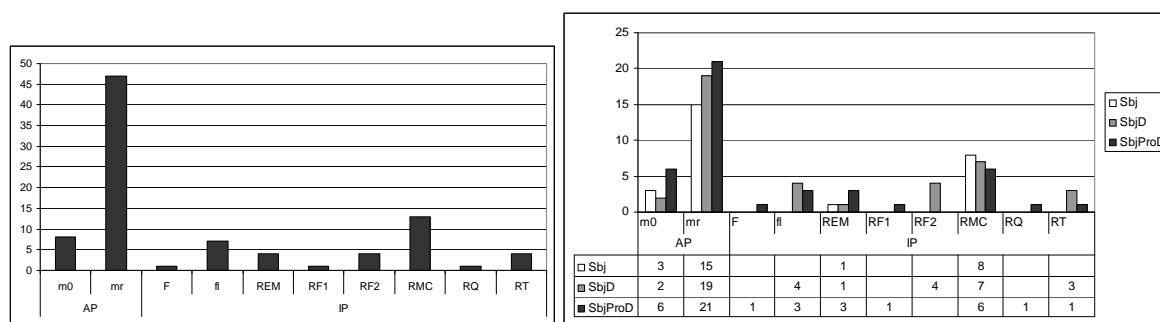


Figure 3 : Répartition des contours prosodiques par types de détachement

Du point de vue de la répartition globale des contours, on constate pour les détachements une répartition identique à celle des données globales de la figure 1. En particulier, le contour le plus fréquent est /mr/. On constate en revanche la faible proportion de contours d'IP, révélant une relation syntaxique forte entre l'élément détaché et le reste de la phrase. Les données plus précises du second tableau indiquent quant à elles un fonctionnement très homogène des types concernés, y compris pour les sujets lexicalisés⁵. Il s'agit là d'une information intéressante

⁴ Roxane Bertrand (LPL), Gaëlle Ferré (LLING) et Lisa Brunetti (LPL) ont effectué ces annotations.

⁵ Il s'agit de sujets construits de façon canonique comme dans la phrase « *Le livre est intéressant* », à la différence de sujets disloqués comme dans « *Le livre, il est intéressant* ». On constate dans nos corpus, comme le tableau 3 l'indique, une proportion relativement faible de ce type de sujets dans les corpus oraux.

tendant à montrer que ce type de sujets a un fonctionnement marqué prosodiquement et proche de celui des sujets disloqués.

Le second phénomène d'alignement que nous avons choisi concerne les backchannels. Il s'agit plus précisément de signaux vocaux ou gestuels émis par l'interlocuteur pour montrer son écoute, sa compréhension, son accord, etc. Les éléments fournis ici sont repris de l'étude présentée dans Bertrand09 qui a montré que ces phénomènes étaient favorisés par des événements prosodiques ou discursifs. En effet, nous avons montré que les backchannels interviennent de façon tout à fait prépondérante après des IP (en particulier après des contours plats / $\text{f}1$ /). De même, du point de vue discursif, trois types de tours conversationnels ont été annotés : les tours ayant un caractère complet du point de vue syntaxique, sémantique et pragmatiques (appelés tours *finaux* et notés / f /), les composants incomplets (notés / nf /) et les cas où le narrateur atteint un point de complétude de son discours et marque une volonté de passer le tour de parole (appelés tours de continuation et notés / cont /). De ce point de vue, les backchannels apparaissent préférentiellement après des tours / f /. Le tableau de la figure 4 récapitule ces différentes informations et indique (en pourcentage) les proportions d'objets après lesquels des backchannels apparaissent.

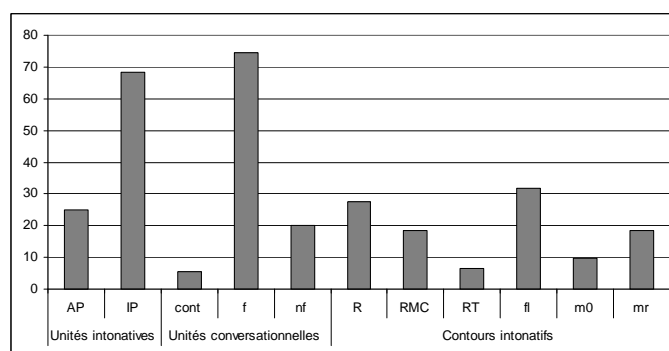


Figure 4 : Répartition des backchannels

4 Informations nécessaires à la représentation de l'alignement

Nous revenons dans cette section sur les observations proposées dans la section précédente. Il s'agit ici de déterminer les informations qu'il sera nécessaire d'utiliser pour décrire les phénomènes d'alignement dans une grammaire multimodale.

Concrètement, ces relations se manifestent par une synchronisation des événements. Bien entendu, la synchronisation de deux événements ne signifie pas nécessairement leur interaction. Elle est cependant une contrainte forte dans le cas où l'interaction est avérée. C'est vrai à l'intérieur d'une même modalité (par exemple pour les interactions entre tons et contours intonatifs, relevant tous deux d'informations véhiculées par le signal acoustique), mais également entre différentes modalités. L'interaction geste/parole est par exemple synchronisée lorsqu'elle est réalisée dans le cadre d'une expression référentielle. Dans ce cas, un geste déictique pourra être synchronisé par exemple avec un pronom. L'alignement temporel dépend par ailleurs du type d'interaction observé. Nous trouverons en effet dans certains cas (généralement à l'intérieur d'une même modalité) des synchronisations strictes, imposant des frontières temporelles identiques (ou très proches) pour les objets en interaction. Dans d'autres cas en revanche (par exemple pour l'interaction geste/parole), une simple intersection temporelle sera suffisante. Il est ainsi nécessaire de représenter par la suite le type d'alignement (*strict* ou *lâche*) de la relation.

Par ailleurs, deux grands types de relation d'interactions peuvent être distingués. Le premier concerne des relations exprimant un caractère obligatoire et constitutif de la construction analysée. Par exemple, les constructions énumératives ont une double contrainte syntaxique et prosodique requérant généralement la reprise d'un constituant syntaxique et d'un contour mélodique de même type. De même, les constructions interrogatives non marquées syntaxiquement requièrent un contour mélodique ascendant. En revanche, certaines relations d'alignement sont indépendantes d'une construction spécifique. C'est le cas par exemple des backchannels vocaux qui apparaissent toujours après des IPs. De même, l'observation faite dans la section précédente sur l'interaction entre contours intonatifs et catégories syntaxiques montre que le contour /mɪr/ apparaît de façon privilégiée avec certaines catégories syntaxiques comme /NV/ ou, dans une moindre mesure, /GR/. Ces secondes relations d'interaction sont plus générales dans le sens où elles portent sur des objets de différents domaines, indépendamment d'un type de construction ou d'un phénomène particulier. On remarquera par ailleurs que les relations peuvent également être négatives. Par exemple, les backchannels n'apparaissent pas après les connecteurs. Nous proposons donc de distinguer deux types de contraintes d'alignement :

- Contraintes d'alignement **locales** : caractéristiques d'une construction particulière, elles ont un caractère fort, voire obligatoire
- Contraintes d'alignement **globales** : indiquant une relation entre deux objets, indépendamment d'une construction précise

En termes de représentation, dès lors qu'il est possible selon le formalisme choisi d'associer une contrainte à une construction, ces deux types de contraintes d'alignement ne différeront donc que par l'identification de la construction ainsi que par le caractère plus ou moins impératif de la relation. Pour récapituler, voici les informations caractéristiques d'une contrainte d'alignement :

- **Type** : indique le caractère nécessaire ou interdit de la relation
- **Niveau** : contrainte globale ou locale à une construction
- **Objets** : spécification des objets en interaction
- **Construction** : spécification dans le cas de contrainte locale de la construction concernée
- **Poids** : degré d'obligation de la contrainte. A minima, distinction entre contrainte impérative (*hard*) et relâchable (*soft*)
- **Synchronisation** : alignement strict ou lâche des frontières temporelles des objets

5 Proposition de représentation

Les Grammaires de Propriétés (dorénavant GP) se prêtent facilement à la représentation de ce type d'information. Nous allons dans ce qui suit poursuivre la description en utilisant cette formalisation. Sans entrer dans les détails (on pourra pour cela se reporter à [Blache01]), rappelons qu'en GP, toutes les informations sont représentées sous la forme de contraintes. Une construction, qui peut être un objet linguistique de n'importe quel niveau (syntagme, proposition, phénomène particulier), est définie par un ensemble de contraintes. La particularité des GP est que les contraintes contiennent la totalité de l'information. Elles sont de fait indépendantes les unes des autres et il est possible de les évaluer séparément. Il est donc tout aussi possible d'ajouter dans la description d'une construction des contraintes relevant de domaines différents.

Prenons l'exemple des contraintes globales citées plus haut. La première stipule que les contours de type /mɪr/ correspondent en général à une unité accentuelle /AP/. Cette relation est

orientée dans la mesure où la réciproque n'est pas vraie. Dans la terminologie des GP, il s'agit donc d'une contrainte d'exigence qu'il est possible de représenter de la façon suivante :

$$P1_{[GLOBAL]} : mr_{[PROSO]} \rightarrow_{[soft, strict]} AP_{[PROSO]}$$

On retrouve dans cette contrainte les différentes informations stipulées précédemment : il s'agit d'une contrainte globale, reliant les objets /mr/ et /AP/ (relevant du domaine prosodique) par une contrainte d'alignement faible (dans la mesure où elle n'est pas impérative) et synchronisée strictement.

La contrainte P2 stipule quant à elle une relation non impérative et synchronisée de façon stricte indiquant une réalisation préférentielle de contour /mr/ avec des /NV/. On remarque au passage l'intérêt potentiel de ce type de contrainte dans une perspective de désambiguïsation.

$$P2_{[GLOBAL]} : NV_{[SYNT]} \rightarrow_{[soft, strict]} mr_{[PROSO]}$$

Le troisième exemple de contrainte globale indique que les backchannels vocaux sont réalisés préférentiellement après des IPs. La contrainte P4, plus spécifique, indique les contours prosodiques favorisant leur apparition. La contrainte P5, enfin, décrit la relation entre backchannel et tour final. Ces trois contraintes reposent sur une relation de précédence. Il s'agit là encore d'une contrainte non impérative (*soft*) et pour laquelle la synchronisation peut être lâche (*fuzzy*).

$$\begin{aligned} P3_{[GLOBAL]} &: VocBC_{[DISC]} >_{[soft, fuzzy]} IP_{[PROSO]} \\ P4_{[GLOBAL]} &: VocBC_{[DISC]} >_{[soft, fuzzy]} \{RF1, RF2, RT\}_{[PROSO]} \\ P5_{[GLOBAL]} &: VocBC_{[DISC]} >_{[soft, fuzzy]} TCU_f_{[DISC]} \end{aligned}$$

Prenons maintenant des exemples de contraintes locales, qui sont donc rattachées à des constructions. Pour illustrer plus précisément la représentation des constructions, l'exemple suivant décrit le cas de la dislocation⁶.

$$\begin{aligned} P6_{[DISLOC]} &: Const = \{S1_{[SYNT]}, S2_{[SYNT]}, mr_{[PROSO]}\} \\ P7_{[DISLOC]} &: S1_{[SYNT]} > S2_{[SYNT]} \\ P8_{[DISLOC]} &: S1_{[SYNT]} \Rightarrow S2_{[SYNT]} \\ P9_{[DISLOC]} &: S1_{[SYNT]} \rightarrow\rightarrow S2_{[SYNT]} // Clit \\ P10_{[DISLOC]} &: S1_{[SYNT]} \rightarrow_{[soft, strict]} mr_{[PROSO]} \end{aligned}$$

Cet ensemble de contraintes permet de décrire les informations de base d'un type simple de dislocation. Cette construction est formée de deux unités syntaxiques S1 et S2, qui apparaissent dans cet ordre (contrainte P7) et qui sont obligatoirement cooccurrentes (P8). De plus, S1 est en relation anaphorique avec un clitique réalisé dans S2⁷. Cette description est complétée par une contrainte d'alignement (P9) stipulant une réalisation de S1 préférentiellement associée à un contour /mr/. Dans le même ordre d'idée, la construction énumérative peut être décrite comme la répétition de plusieurs syntagmes de même type, tous alignés (P12) avec un contour spécifique.

⁶ Les dislocations sont ces constructions détachant un élément (souvent un SN) à gauche de la phrase (quelquefois en fin) avec une possibilité de reprise anaphorique de cet élément par un clitique dans la phrase. Exemple : « *Le chocolat, j'aime ça* ».

⁷ On utilise l'opérateur // (emprunté à XPATH) pour indiquer un élément situé à un niveau quelconque à l'intérieur d'un autre.

$$\begin{aligned} P11_{[ENUM]} &: \text{Const} = \{SX^*_{[SYNT]}, ENUM_{[PROSO]}\} \\ P12_{[ENUM]} &: SX_{[SYNT]} \rightarrow_{[soft, strict]} ENUM_{[PROSO]} \end{aligned}$$

L'exemple suivant propose une contrainte isolée permettant la description d'une question formée d'une phrase sans marque interrogative particulière (donc à l'affirmatif) alignée avec un contour montant spécifique à l'interrogation. Nous sommes ainsi en mesure de décrire une tournure particulière dont l'interprétation ne peut provenir que de la réalisation conjointe de ces deux éléments. Il s'agit d'une contrainte d'alignement impérative, l'absence de réalisation du contour /RQ/ empêchant en effet l'interprétation interrogative

$$P13_{[QUEST]} : P_{[SYNT]} \rightarrow_{[hard, strict]} RQ_{[PROSO]}$$

6 Discussion

Les relations d'alignement entre les différents domaines sont d'une importance fondamentale. Elles permettent une description très précise du niveau de l'interaction, mais également de son type. Elles permettent de plus de décrire des relations entre des types d'objets de *granularité* différentes pouvant être généraux (des constructions entières) ou au contraire spécifiques (des sous-catégories, du type consonne occlusive sourde ou verbe à l'infinitif). Il est ainsi possible de décrire des phénomènes d'interaction ponctuels, ou à l'inverse des relations entre des catégories de plus haut niveau voire des méta-catégories. Par ailleurs, la distinction de *niveau* entre relations locales et globales offre une grande souplesse de description, rendant possible la représentation de relations très générales (éventuellement universelles), ou au contraire spécifiques à une construction.

Une des difficultés dans la représentation de ces informations réside dans la grande variabilité des phénomènes observés, par exemple pour ce qui concerne la prosodie, allant jusqu'à faire coexister des phénomènes pouvant apparaître comme contradictoires. Nous avons proposé ailleurs (cf. [Blache04]) une explication de la variabilité s'appuyant sur la quantité d'information véhiculée par chaque domaine linguistique. En tout état de cause, il est nécessaire d'indiquer le *poids* de la relation pour représenter son degré d'exigence : certaines relations sont impératives, d'autres le sont moins. De même, le type de *synchronisation* (stricte ou lâche) est spécifique à chaque relation. Il est donc nécessaire de proposer une représentation permettant de préciser des informations spécifiques à chaque contrainte.

Une approche totalement basée sur les contraintes du type de celle proposée ici répond à ces besoins. Chaque relation d'alignement est représentée dans la grammaire par une contrainte spécifique. De plus, respectant ainsi l'idée de base des GP, toutes ces contraintes peuvent être évaluées séparément, sans avoir besoin au préalable de construire une quelconque structure. Nous disposons ainsi d'un cadre formel nous permettant de nous affranchir d'une explication de l'interaction entre domaines linguistiques en termes de superposition de structures (par exemple insertion d'événements prosodiques à l'intérieur d'arbres syntaxiques), requérant la construction préalable des structures de chaque domaine avant d'en pouvoir déterminer les interactions. Ce type d'explication est la plupart du temps utilisé dans la littérature. Elle ne rend cependant pas compte de la réalité du phénomène, ou de façon très limitée. En revanche, la possibilité de décrire dans la grammaire des relations pouvant être générales ou spécifiques, impératives ou permissives, permet de représenter avec précision l'interaction. Nous sommes ainsi en mesure de proposer un cadre permettant le développement de véritables *grammaires multimodales*.

Références

- Bertrand R., Ferré G., Blache P., Espesser R. & Rauzy S. (2007) «Backchannels revisited from a multimodal perspective», *Proceedings of Auditory-visual Speech Processing* Hilvarenbeek, The Netherlands.
- Bertrand R., P. Blache, R. Espesser, G. Ferré, C. Meunier, B. Priego-Valverde, S. Rauzy (2009) Le Cid - Corpus of Interactional Data - Annotation et Exploitation Multimodale de Parole Conversationnelle, in revue TAL 49:3
- Bertrand R., Portes C. & Sabio F. (2007) «Distribution syntaxique, discursive et interactionnelle des contours intonatifs du français dans un corpus de conversation», *TRANEL*, 47, 2007
- Bird S., Day D., Garofolo J., Henderson J., Laprun C. & Liberman M. (2000) «ATLAS : A Flexible and Extensible Architecture for Linguistic Annotation», *Proceedings of the Second International Conference on Language Resources and Evaluation*
- Blache P. (2001) *Les Grammaires de Propriétés : Des contraintes pour le traitement automatique des langues naturelles*, Hermès.
- Blache P. A. Di Cristo (2002), « Variabilité et dépendances des composants linguistiques », in *Actes de TALN'02*.
- Blache P. & Rauzy S. (2008) «Influence de la qualité de l'étiquetage sur le chunking : une corrélation dépendant de la taille des chunks», *Actes de TALN08*
- Carletta J. et al. (2005) «The AMI Meetings Corpus», *Proceedings of the Measuring Behavior 2005 symposium on Annotating and measuring Meeting Behavior*,
- Carletta, J. (2006) Announcing the AMI Meeting Corpus. The ELRA Newsletter 11(1)
- de la Clergerie E., Ayache C., de Chalendar G., Francopoulo G., Gardent C. & Paroubek P. (2008) «Large Scale Production of Syntactic Annotations for French», *Proceedings of the International Workshop on Automated Syntactic Annotations for Interoperable Language Ressources*.
- Delais-Roussarie, E et B. Post (2008) « Unités prosodiques et grammaire de l'intonation : vers une nouvelle approche », in *actes des XXVIIème Journées d'Études sur la Parole*
- Doetjes J., E. Delais-Roussarie & P. Sleeman (2002) "The Prosody of Left Detached Constituent in French", in *proceedings of Speech Prosody 2002*.
- Gendner V., Illouz G., Jardino M., Monceaux L., Paroubek P., Robba I. & Vilnat A. (2003) «PEAS, the first instantiation of a comparative framework for evaluating parsers of French», *Research Notes of EACL 2003*
- Hirst D. & Di Cristo A., *Intonation Systems*, Cambridge University Press, 1998.
- Jun S.-A. & Fougeron C. (2002) «Realizations of accentual phrase in French intonation», *Probus* 14
- Kipp M. (2006) *Anvil 4.0. Annotation of Video and Spoken Language*. (<http://www.dfki.de/~kipp/anvil>)
- McNeill D. (2005) *Gesture and Thought*. University of Chicago Press
- MUMIN : A Nordic Network for MultiModal Interfaces (<http://www.cst.dk/mumin/>)
- NITE : Natural Interactivity Tools Engineering (<http://nite.nis.sdu.dk/>)
- Paroubek P., Robba I., Vilnat A. & Ayache C. (2006) «Data Annotations and Measures in EASY the Evaluation Campaign for Parsers in French», *Proceedings of the 5th International Conference on Language Resources and Evaluation*
- Portes C., Bertrand R. & Espesser R. (2007) «Contribution to a grammar of intonation in French. Form and function of three rising patterns», *Nouveaux Cahiers de Linguistique Française*, 28
- Rauzy S. & Blache P. (2007) «Un lexique syntaxique des verbes du français : VfrLPL», *Rapport de recherche RAU-3055*, Laboratoire Parole et Langage.
- Selkirk E. (1984) *Phonology and Syntax: the Relation between Sound and Structure*, MIT Press.
- VERBMOBIL : Alexandersson J., Buschbeck-Wolf B., Fujinami T., Kipp M., Koch S., Maier E., Reithinger N., Schmitz B. & Siegel M. *Dialogue Acts in VERBMOBIL-2 - Second Edition*. Rapport interne, 1998, DFKI GmbH.