

## **Quelques principes pour une grammaire multimodale non-modulaire du français**

Marie-Laure Guénot & Emmanuel Bellengier  
Laboratoire Parole et Langage  
CNRS – Université de Provence  
29 avenue Robert Schuman 13621 Aix-en-Provence cedex 1  
{mlg, bellengier}@lpl.univ-aix.fr

### **Mots-clefs – Keywords**

Analyse multimodale – Linguistique formelle – Développement de grammaire – Grammaires de Propriétés

*Multimodal analysis – Formal linguistics – Grammar development – Property Grammars*

### **Résumé - Abstract**

Dans cet article, nous introduisons une approche de la représentation et de l'analyse des discours multimodaux, basée sur un traitement unimodulaire par contraintes. Le but de cet article est de présenter (i) un système de représentation des données et (ii) une méthode d'analyse, permettant une interaction simplifiée entre les différentes modalités de communication. L'avantage de cette méthode est qu'elle permet la prise en compte rigoureuse d'informations communicatives de natures diverses en un traitement unique, grâce à une représentation homogène des objets, de leurs relations, et de leur méthode d'analyse, selon le modèle des Grammaires de Propriétés.

*In this paper, we introduce an approach to multimodal discourse representation and analysis, based on a unimodular constraint-based treatment. The aim of this paper is to present (i) a data representation system, and (ii) an analysis method, which allow a simple interaction between the various modalities of communication. The advantage of this method is that it allows to rigorously take into account multi-character information within a single treatment, thanks to a homogeneous representation of the objects, of their relations, and of their analysis method, according to the Property Grammars formalism.*

# 1 Introduction

La communication humaine est une activité qui s'effectue sur différents canaux ; son caractère multimodal n'est plus à prouver aujourd'hui<sup>1</sup>. Chaque modalité participe à l'élaboration du message, et joue un rôle dans son interprétation, rôle qui peut être de renforcer ou de compléter les autres modalités. La prise en compte d'informations communicatives non-linguistiques revêt donc une importance non négligeable dans la construction de la signification<sup>2</sup> d'un énoncé, puisque ne pas en tenir compte conduit inmanquablement à passer à côté d'une partie du message, et mène donc à des erreurs et des ambiguïtés superflues.

Parmi les approches multimodales suivies actuellement en TALN, on peut retenir deux façons générales de procéder. La première, présentée par exemple dans (Landragin, 1999), consiste à effectuer des analyses séparées (parallèles ou successives) d'une même entrée, correspondant à autant de modalités, niveaux ou domaines, avant d'en unifier les résultats au cours d'une ultime étape<sup>3</sup>. Autrement dit, une partie de la construction de la signification du message est effectuée avant même que les modalités ne soient mises en relation, ce qui exclut toute possibilité d'interaction multimodale de bas niveau.

La seconde approche, proposée notamment dans (Pineda & Garza, 2000) et (Godbert, 2002), consiste à traduire dans un premier temps les données multimodales en une entrée unimodale linguistique, avant d'en effectuer l'analyse. Cette perspective implique d'adopter une position théorique selon laquelle pour toute production non-linguistique et au-delà pour toute production multimodale, il existe toujours une production unimodale linguistique strictement équivalente.

De plus, les deux procédés ne semblent pas s'appuyer sur une justification cognitive théorique : il paraît peu probable que l'interprétation naturelle d'un message (par un humain) s'effectue en phases successives et indépendantes, ou par une traduction préalable des différents canaux en un seul (cf. par exemple, à ce propos, (Langacker, 1999) ou (Croft & Cruse, 2003)).

Dans cet article nous proposons un extrait de grammaire et la méthode d'analyse correspondante, permettant de traiter en un seul passage une entrée multimodale, en se référant à une unique grammaire au sein de laquelle toutes les informations communicatives interagissent librement, fournissant en sortie un réseau de relations entretenues par les divers éléments constitutifs du message. Nous appelons cette proposition une **grammaire multimodale non modulaire**<sup>4</sup> ; nous en illustrons les caractéristiques à travers un exemple concret, et détaillons la représentation formelle de ce type de grammaire, représentation chargée d'une part de mettre en évidence l'interaction des modalités, et d'autre part, plus généralement, de faire le lien entre analyse théorique et traitement automatique.

---

<sup>1</sup>On retrouve cette multimodalité aussi bien dans la communication *inter-humaine orale* (canaux verbal et gestuel) ou *écrite* (chaîne linguistique, mise en page), que dans le *dialogue homme-machine* (clavier, souris, son, image, voire parole, clignement des yeux, etc.).

<sup>2</sup>Nous entendons ici par "signification" une représentation assez générale de l'ensemble des éléments pouvant servir à l'interprétation d'un message. Nous nous plaçons ainsi dans un paradigme proche de celui des Grammaires de Construction ou d'HPSG.

<sup>3</sup>Il est intéressant de remarquer que ce type de procédé, **modulaire**, est pratiqué non seulement dans des perspectives multimodales, mais aussi dans les traitements multi-domaines ou multi-niveaux de l'analyse linguistique, par exemple dans (Bear & Prince, 1990), (Hirst, 1993) ou (Pavia, 2000). De la même façon, on peut alors considérer que les quelques principes que nous proposons ici peuvent aussi bien s'appliquer à ces cas de figure.

<sup>4</sup>Ce type d'approche a été également introduit dans (Bellengier, 2003), article dans lequel les fondements d'une Grammaire de Propriété Multimodale sont proposés ainsi qu'une réflexion sur la nécessité de prendre en compte des données hétérogènes et de qualifier les relations entre les niveaux.

## 2 Transcription segmentée de corpus audiovisuel

Pour analyser un corpus, il est utile et traditionnel en linguistique comme en TALN de l’avoir préalablement transcrit et segmenté. Nous utilisons donc comme entrée de nos analyses un document qui contient non seulement la représentation écrite des données, découpées en unités atomiques de chaque modalité, mais également l’affectation à celles-ci d’un certain nombre de caractéristiques propres. Il est cependant important de souligner que ce “document” que nous créons et manipulons représente une abstraction par rapport à la langue elle-même, contrairement à ce qui est généré dans l’approche de (Pineda & Garza, 2000). Il s’agit d’une transcription (et non d’une traduction) qui est à un message de langue, ce qu’un fichier audiovisuel informatique est à la scène réelle correspondante.

Les corpus que nous avons utilisés pour notre présent exemple sont de courts extraits de dialogues libres authentiques, enregistrés à l’Université de Provence. Le protocole d’expérimentation imposait aux sujets de choisir un thème de discussion parmi trois et d’en débattre immédiatement pendant dix minutes.

Chaque unité découpée, indexée et représentée en référence aux données audiovisuelles<sup>5</sup> correspond à un élément atomique appelé **objet**. La figure 1 illustre sous forme de structures de traits, l’objet linguistique *ouais* tiré de l’exemple 1, et un objet gestuel simultanée (un hochement de tête).

- (1) Loc.A : ça fait partie des 10 minutes le choix du sujet ou  
 Loc.B : *ouais* ben je sais pas quel quel sujet toi tu as

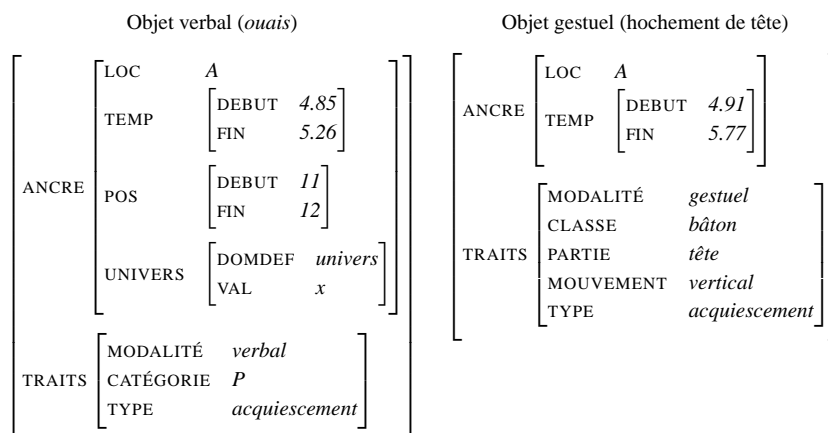


FIG. 1 – Structures des objets *ouais* et *hochement de tête* de l’exemple (1).

On peut voir dans cette figure que chaque description d’objet est constituée du même ensemble d’informations, quelle qu’en soit la modalité d’appartenance : la structure générale d’un objet est composée d’une **ancre** à localisation multiple et d’un ensemble de **traits** de spécification catégorielle.

<sup>5</sup>Ces opérations ont été effectuées à l’aide de TASX-ANNOTATOR, qui fait partie d’une suite d’outils logiciels développés à l’Université de Bielefeld en Allemagne (<http://tasxforce.lili.uni-bielefeld.de>)

## 2.1 Ancre

Plusieurs types de références ou d'index sont utilisés simultanément et optionnellement pour localiser un même objet, et sont rassemblés au sein de l'ANCRE (cf. figure 1). Ainsi, une unité identifiée dans une entrée acoustique pourra être référencée non seulement par sa position temporelle dans le signal (TEMP dans l'illustration), mais également par son index dans la chaîne linguistique correspondante (POS ; cf. (Bird & Liberman, 1999)). S'il s'agit en outre d'un objet référentiel, on peut y ajouter un troisième attribut s'appuyant sur l'univers du discours (UNIVERS ; cf. (Kamp & Reyle, 1993)). Enfin, pour toutes les unités on ajoutera l'identification du locuteur qui l'a produite (LOC). Cette technique d'ancrage multiple permet de mettre en évidence des relations, pouvant être de natures diverses, entre des objets de domaines distincts, repérés sur des ensembles d'axes différents en fonction de la nature de leurs modes de production.

Les attributs sont dotés de valeurs appartenant à leurs domaines de définition respectifs. La définition du domaine est essentiellement utile à préciser l'univers du discours utilisé dans le point d'ancrage correspondant. Elle permet de tirer parti de différents types d'information. L'apparition de catégories aura en effet pour conséquence de compléter les domaines de définition correspondants. Par exemple, en rencontrant un nouveau référent, celui-ci est ajouté au domaine UNIVERS de l'ancre. En revanche, l'analyse d'un déictique ou plus généralement d'une quantification permettra de préciser voire de spécifier une seule valeur dans un contexte donné en associant une formule (ou une valeur) à un objet préexistant.

## 2.2 Traits

On regroupe dans cette structure l'ensemble des informations définitives caractéristiques des objets segmentés. Pour illustrer notre exemple et compte tenu de sa nature (dialogue inter-humain oral), nous avons référencé des objets **linguistiques** et des objets **gestuels** (cf. l'attribut MODALITÉ dans la figure 1). Concernant les objets linguistiques, les informations contenues dans les TRAITS représentent des informations morphologiques, lexicales et sémantiques assez classiques. Les objets gestuels quant à eux ont été identifiés en référence à la classification des gestes et des postures que l'on peut trouver dans (McNeill, 1992).

Après cette transcription-segmentation, toutes les unités sont donc représentées de façon homogène, ce qui participe amplement à leurs possibilités d'interaction dès le début de l'analyse. A l'issue de celle-ci, notre structure complexe sera complétée par une sous-structure supplémentaire, que nous appellerons **caractérisation**, et qui représentera l'ensemble des relations entretenues par les différents objets, atomiques ou complexes. C'est ce que nous allons voir dans la partie suivante.

# 3 La grammaire

## 3.1 Analyse des structures brèves d'acquiescement

Le fragment de grammaire du français que nous proposons afin d'illustrer notre approche traite des structures brèves d'acquiescement (ci-après SBA) dans un dialogue oral. Nous motivons le

choix de ce phénomène d'abord par le fait qu'il est très productif en français. De plus, il présente un certain nombre de régularités de production qui nous permet d'opérer sur des situations relativement faciles à capter. Enfin, nous avons remarqué que les cas de non-congruence entre l'apport fourni par les gestes et l'apport fourni par l'entrée lexicale permettaient de différencier très nettement les structures d'acquiescement.

Nous ne faisons pas de différence *a priori* entre les acquiescements ou retours positifs. En apparence, les acquiescements sont des structures qui doivent marquer l'acceptation par un agent, du discours d'un autre agent impliqué dans le processus discursif. L'intérêt du phénomène est qu'il marque profondément le caractère coopératif de la communication, et nous nous intéressons à la façon dont ces éléments peuvent être rattachés à un discours antécédent. Nous réinvestissons la classification proposée par (Muller & Prevot, 2003) et distinguons à leur suite deux types principaux de SBA, caractérisés par un tissu d'informations issues de modalités différentes :

- les **acquiescements faibles** marquent que le discours a bien été perçu par leur producteur, sans engager celui-ci dans le contenu du discours (i.e. sans que cela implique qu'il est d'accord avec ce qui vient d'être dit). Ces structures sont repérées par la production d'un mot bref (moins de deux syllabes, par exemple *oui*, ou *ouais*, ou *hm*), ou d'un hochement de tête, ou d'un clignement ostensible des yeux, ou d'une modification de la posture.
- les **acquiescements forts**, quant à eux, marquent que l'énoncé produit est accepté et reconnu comme vrai par le récepteur (i.e. qu'il est d'accord avec ce qui vient d'être dit). Ceux-ci prennent la même forme que les précédents, à ceci près qu'ils combinent plusieurs des productions citées<sup>6</sup>.

L'intérêt de distinguer ces deux formes est que la première est le simple produit d'un feedback dialogique, alors que la deuxième constitue l'apport d'une information supplémentaire dans la définition de l'univers du discours présent. Leur analyse, leurs implications sont donc radicalement différentes, et il est impossible de les différencier uniquement en référence à la seule production linguistique. Cette distinction permet par la même occasion d'expliquer pourquoi nous nous intéressons ici uniquement aux structures brèves : en effet les structures plus longues, constituées notamment de plusieurs mots, correspondent le plus souvent à des acquiescements forts, et ne présentent pas l'ambiguïté d'interprétation que l'on vient d'évoquer, celle-ci étant levée par le lexique employé lui-même (comme on peut le voir dans l'exemple (2)). Nous reviendrons plus loin sur la différence impliquée par l'analyse des structures longues et brèves, et leurs conséquences sur le contenu de la grammaire.

- (2) Loc.A : ouais mais au-delà faut d'abord se poser la question initiale c'est-à-dire pourquoi pourquoi cette loi ce projet de loi je veux dire quel quel est le danger qu'il y ait des anglicismes  
Loc.B : *ouais c'est vrai*

Cette classification simple ne permet pas de considérer toutes les subtilités de l'acquiescement. Cependant elle permet de prendre en compte l'aspect dynamique et coopératif du dialogue. En outre, elle est théoriquement motivée par un ensemble de travaux qui considèrent les structures d'engagement, d'interaction, comme des éléments fondamentaux du discours, comme par exemple (Ginzburg & Sag, 2000).

---

<sup>6</sup>Il existe également d'autres formes d'acquiescements forts brefs, tels que la production synchrone d'un clignement de l'oeil et d'un claquement de langue ; cependant ceux-ci ne présentent pas d'ambiguïté possible avec les acquiescements faibles (on ne rencontre pas de production d'un claquement de langue seul comme signe d'acquiescement faible), et ils ne sont donc pas pertinents pour notre présente étude.

## 3.2 Formalisation

Représenter régulièrement une grammaire multimodale nécessite de répondre à deux questions : *comment représenter des informations issues de modes différents dans une même grammaire ?* d'une part, et *comment faire interagir ces modes ?* d'autre part. La réponse aux deux questions est unique dans notre proposition : elle est d'utiliser un formalisme capable de considérer toutes les modalités sur un plan comparable.

### 3.2.1 Choix du formalisme

A la première question nous avons répondu dans la section précédente : l'une des propriétés requises, essentielle, du modèle est sa capacité à **représenter efficacement** des informations de natures diverses. Par "efficacement", il faut entendre que le formalisme doit être en mesure d'exprimer tout type d'information nécessaire à l'analyse, en l'occurrence multimodale, et que la forme des représentations soit adaptée au traitement envisagé, en l'occurrence une analyse non-modulaire. Le formalisme des *Grammaires de Propriétés* (ci-après GP) offre une représentation flexible et tout à fait homogène des informations<sup>7</sup> qui nous permet d'exprimer sous une forme identique l'ensemble des indications pertinentes d'un énoncé multimodal. Cette flexibilité permet en outre de représenter des objets intrinsèquement multimodaux, toujours selon la même forme de base. L'un des effets de ceci sera que toutes les caractéristiques d'un objet sont placées sur le même plan, et l'interaction (entre objets, propriétés, niveaux, domaines, modalités, etc.) s'opère donc tout à fait naturellement.

L'intérêt des GP réside aussi dans le fait que toutes les informations sont exprimées en termes de contraintes (objets, relations et méthode d'analyse)<sup>8</sup>. Ceci nous permet d'obtenir en guise de résultat d'analyse, non pas une unique structure littéralement "grammaticale" (i.e. rigoureusement conforme à ce que la grammaire utilisée déclare), mais un ensemble d'informations concernant les relations entretenues par les unités et les groupes constituant notre entrée (la **caractérisation**). Ces informations pourront être partielles, superposées, sans que le déroulement de l'analyse n'en soit affecté. En ceci, l'intervention de plusieurs modes de communication pour la reconstitution du sens ne consitue pas un alourdissement du traitement à effectuer, puisque toutes les modalités sont traitées en même temps, comme un seul et même ensemble d'informations.

Enfin, le modèle des GP offre la possibilité de développer des propriétés multimodales de manière **partielle**. Pour en revenir à l'exemple des structures d'acquiescement longues, comme on l'a fait remarquer précédemment celles-ci ne présentent pas la même ambiguïté d'interprétation que les SBA, car elles sont désambiguïsées par leur seul contenu lexical ; dans ce cas de figure, et parce qu'alors la prise en compte de gestes serait redondante (et alourdirait inutilement le traitement), ceux-ci ne comportent pas, dans la partie de la grammaire les décrivant, de propriétés mentionnant des gestes. Ainsi l'on peut ajouter des propriétés gestuelles ou linguistico-gestuelles pour traiter un phénomène particulier (les structures brèves d'acquiescement, par exemple), sans avoir à modifier le reste en conséquence<sup>9</sup>.

<sup>7</sup>Cette propriété du modèle a été souvent mise en avant, comme par exemple dans (Blache, 2001).

<sup>8</sup>Nous supposons ici que le lecteur n'est pas étranger aux principes généraux du formalisme des GP. Dans le cas contraire, il pourra se référer par exemple à (Blache, 2001), ou concernant ce point précis du modèle, à (Vanrullen *et al.*, 2003).

<sup>9</sup>Dans l'absolu il est même possible d'écrire et d'utiliser une grammaire qui ne contiendrait *que* des propriétés traitant des SBA.

### 3.2.2 Application

La partie de grammaire que l'on a développée pour traiter la distinction entre SBA faibles et SBA fortes<sup>10</sup> est constituée des propriétés présentées dans le tableau 1 représentant, dans le formalisme des GP, les caractéristiques énoncées dans la section 3.1<sup>11</sup>.

EXIGENCE	(acquiescement fort) <b>verbal.type.acquiescement</b> $\Rightarrow$ <b>gestuel.type.acquiescement</b>
EXCLUSION	(acquiescement faible) <b>verbal.type.acquiescement</b> $\neq$ <b>gestuel.type.acquiescement</b>
DÉPENDANCE	(acquiescement fort) <b>verbal.type</b> $\approx$ <b>gestuel.type</b>
SIMULTANÉITÉ	(acquiescement fort) $\uparrow \{ \mathbf{verbal.type.acquiescement}, \mathbf{gestuel.type.acquiescement} \}$

TAB. 1 – Propriétés traitant des SBA dans la grammaire multimodale.

La propriété d'**exclusion** (cf. tableau 1) permet de spécifier que dans une SBA faible, la présence d'un élément verbal de type acquiescement exclut celle d'un geste de type acquiescement dans l'empant observé. Parallèlement à cette propriété, celle d'**exigence** indique que pour avoir une SBA forte, la présence d'un élément linguistique de type acquiescement exige celle d'un geste de type acquiescement dans l'empant observé. Dans le but de renforcer le lien entre les éléments constitutifs d'une SBA forte, on ajoute à cet ensemble grammatical une propriété de **dépendance** indiquant que dans une relation de SBA forte, la valeur du type des objets verbal et gestuel doivent concorder (en l'occurrence, cette valeur doit être "acquiescement"). Enfin, toujours en référence à la description de la section 3.1, on pose une relation de **simultanéité** entre l'objet verbal de type acquiescement, et l'objet gestuel de même type, dans les SBA fortes.

Une fois définies ces propriétés, encore faut-il pouvoir en décrire le mode de fonctionnement. Par exemple, il existe un grand nombre de façons d'évaluer si deux objets sont réalisés simultanément (i.e. *qu'est-ce que la simultanéité ?*), façons qui pourront mener à autant de résultats différents. La définition du mode de fonctionnement des propriétés est une part du développement formel d'une GP, et elle revient par conséquent au développeur de grammaire. Celui-ci dispose à cette fin d'un ensemble de caractéristiques définitoires de la **sémantique des propriétés**, qui lui permettent d'en décrire le fonctionnement, et dont on peut voir l'application pour notre cas dans le tableau 2<sup>12</sup> :

- **arité** exprime le nombre de membres possibles pour une propriété. Chaque membre peut lui-même être constitué d'un objet, ou bien d'un ensemble d'objets (liés par un "ou exclusif"), ou encore d'un ensemble de propriétés évaluées. Cette dernière façon de pointer permet de

<sup>10</sup>Partie de grammaire qui en soi, comme on vient de le voir, peut constituer formellement à elle seule une grammaire.

<sup>11</sup>Bien évidemment, la version complète formalisée de la grammaire a une forme beaucoup plus abstraite que ce que l'on peut voir dans les exemples de ce tableau. Cependant, pour des raisons évidentes de lisibilité, puisqu'il nous est bien impossible ici de faire figurer ne serait-ce qu'un extrait complet de grammaire, nous sommes dans l'obligation d'exprimer les objets désignés dans les propriétés par leur nom complet plutôt que par une (ou un ensemble de) représentation(s) formelle(s) correspondante(s).

<sup>12</sup>Il s'agit ici d'un exemple simplifié ; en effet dans les GP qui ont été développées voire implémentées, la sémantique des propriétés comporte d'autres caractéristiques, permettant par exemple de spécifier combien de fois une même propriété peut être satisfaite, si la satisfaction d'une propriété donnée à elle seule implique la construction d'un graphe entitaire ou non, etc. (cf. le travail en cours à ce propos dans (Vanrullen *et al.*, 2003)). C'est l'ensemble de ces caractéristiques qui confère aux GP leur flexibilité expressive, mais nous n'avons développé ici que quelques exemples jugés significatifs.

désigner un groupe par ses caractéristiques et non par son éventuelle étiquette, ce qui accentue la robustesse de la méthode ;

- **mobilité** permet de spécifier si l'ordre des membres dans la propriété est pertinent ou non. Par exemple, les membres d'une propriété de dépendance sont mobiles (puisqu'il s'agit uniquement de comparer deux valeurs de traits), alors que ceux d'une propriété d'exclusion ne le sont pas (puisque c'est le premier membre qui exclut la présence du deuxième et pas le contraire) ;
- **disponibilité** spécifie les conditions selon lesquelles une propriété sera évaluée ou non, étant donné un empant observé ; par exemple, si les deux membres d'une propriété de dépendance ne sont pas présents, il est tout à fait impossible de l'évaluer, et la propriété est par conséquent non disponible. Si les conditions indiquées sont vérifiées, alors la propriété sera évaluée (c'est-à-dire qu'on évaluera sa *satisfaisabilité*), dans le cas contraire on considérera qu'elle n'est pas pertinente pour l'empant donné ;
- **satisfaisabilité**, enfin, permet de signifier quelles conditions doivent être remplies pour que la propriété évaluée soit satisfaite. Si les conditions sont vérifiées, alors la propriété est satisfaite, dans le cas contraire elle est violée.

PROPRIÉTÉ	CARACTÉRISTIQUES	SÉMANTIQUE
<b>EXIGENCE</b> $a \Rightarrow b$	<i>arité</i> <i>mobilité</i> <i>disponibilité</i> <i>satisfaisabilité</i>	= 2 faux $\exists a$ $\exists b$
<b>EXCLUSION</b> $a \neq b$	<i>arité</i> <i>mobilité</i> <i>disponibilité</i> <i>satisfaisabilité</i>	= 2 faux $\exists a$ $\nexists b$
<b>DÉPENDANCE</b> $a.\text{trait} \approx b.\text{trait}$	<i>arité</i> <i>mobilité</i> <i>disponibilité</i> <i>satisfaisabilité</i>	= 2 vrai $\exists a \wedge \exists b$ $a.\text{trait.valeur} \equiv b.\text{trait.valeur}$
<b>SIMULTANÉITÉ</b> $\downarrow \{a, b, \dots, n\}$	<i>arité</i> <i>mobilité</i> <i>disponibilité</i> <i>satisfaisabilité</i>	$\geq 2$ vrai $\exists a \wedge \exists b \wedge \dots \wedge \exists n$ $\exists i (\max(a.\text{début}, b.\text{début}, \dots, n.\text{début}) \leq i \leq \min(a.\text{fin}, b.\text{fin}, \dots, n.\text{fin}))$

TAB. 2 – Sémantique des propriétés utilisées pour les traitement des SBA.

### 3.2.3 Analyses

A la lumière des propriétés définies plus haut, nous pouvons effectuer l'étape d'analyse de l'exemple (1). Dans cet exemple, l'objet verbal *ouais* et le hochement de tête cités satisfont la propriété de *dépendance* de SBA forte, ce qui constitue une première relation entre eux deux. Etant produits sur le même empant observé lors de l'analyse<sup>13</sup>, ils satisfont la propriété d'*exigence* concernant la SBA forte, et par là même violent la propriété d'*exclusion* de la SBA faible, créant deux liens supplémentaires. Enfin, conformément à la sémantique de la propriété de *simultanéité* donnée dans le tableau 2, on peut dire que le maximum des valeurs de début des objets (ici, 4, 91 qui marque le début du hochement) est strictement inférieur au minimum des valeurs de fin (ici, 5, 77 qui marque la fin du hochement), ce qui signifie que les deux éléments entretiennent également une relation de simultanéité. Cette caractérisation est illustrée dans la

<sup>13</sup>La taille de cet empant peut être, au choix, réglée par le développeur de grammaire, ou automatiquement calculée par le parseur.



figure 2, et met en évidence que les deux objets constituent à eux deux une SBA forte, en accord avec les 3 propriétés qu'ils satisfont et de celle qu'ils violent.

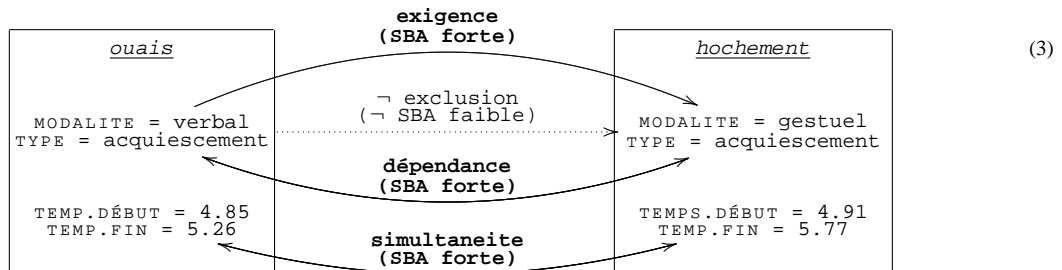


FIG. 2 – Caractérisation de l'exemple (1), SBA forte.

Dans le cas d'une SBA faible, comme dans l'exemple (4), la caractérisation aura la forme de la figure 3. On remarque dans cet exemple que les propriétés de *dépendance* et de *simultanéité* ne figurent pas dans la caractérisation ; la raison à cela est que leurs conditions de disponibilité n'ont pas été reconnues comme vraies lors de l'analyse, et que l'on n'a donc pas évalué leur satisfaisabilité.

- (4) Loc.A : on considère que c'est une horreur enfin c'est c'est une aberration  
 Loc.B : *ouais*  
 Loc.A : de faire une loi comme ça

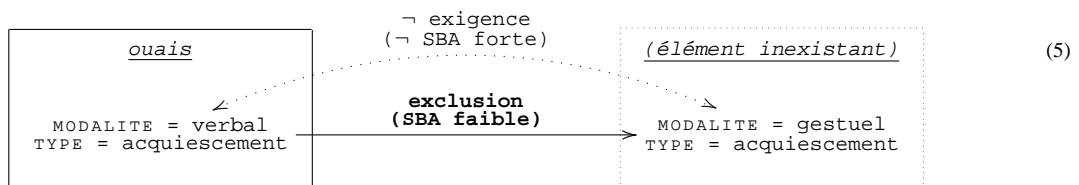


FIG. 3 – Caractérisation de l'exemple (4), SBA faible.

## 4 Conclusion

Dans cet article nous avons présenté une approche de l'analyse de discours multimodaux, basée sur une représentation rigoureuse et homogène des informations de natures diverses. Cette représentation permet d'utiliser une grammaire multimodale, au sein de laquelle est exprimée l'intégralité des objets, de leurs relations et de leur méthode d'analyse. Les éléments peuvent alors interagir simplement, et l'analyse génère un réseau de relations entretenues par les divers objets du message.

Le fragment de grammaire formalisé que nous présentons ici fait partie d'un projet plus large, dans lequel nous cherchons à intégrer à une grammaire linguistique classique (traitant principalement de syntaxe et de sémantique lexicale et prédicative) des informations multimodales,

informations qui nous l'espérons permettront d'obtenir de meilleurs résultats de parsing, que ce soit d'un point de vue quantitatif (dans le cadre d'un parsing non déterministe, il s'agit de réduire le nombre de résultats possibles) que qualitatif (c'est-à-dire de faire en sorte que les résultats les plus significatifs, les plus pertinents, apparaissent de façon plus évidente dans les résultats).

## Références

- BEAR J. & PRINCE P. (1990). Prosody, syntax and parsing. In *Proceedings of ACL-90*.
- BELLENGIER E. (2003). Un modèle multimodal basé sur les contraintes. In *Proceedings of Colloque Médiation et Ingénierie des connaissances*. (à paraître).
- BIRD S. & LIBERMAN M. (1999). *A formal framework for linguistic annotation*. Rapport interne MS-CIS-99-01, Dept of Computer and Information Science, University of Pennsylvania.
- BLACHE P. (2001). *Les grammaires de propriétés — des contraintes pour le traitement automatique des langues naturelles*. Paris : Hermès Sciences Publications.
- CROFT W. & CRUSE D. (2003). *Cognitive linguistics*. Cambridge University Press.
- GINZBURG J. & SAG I. (2000). *Interrogative investigations*. Stanford : CSLI.
- GODBERT E. (2002). Reformuler des expressions multimodales. In *Proceedings of TALN 2002*, p. 365–370.
- HIRST D. (1993). Detaching intonational phrases from syntactic structures. *Linguistic Inquiry*, **24**(4).
- KAMP H. & REYLE U. (1993). *From discourse to logic*. Dordrecht : Kluwer.
- KERBRAT-ORECCHIONI C. (1996). *La conversation*. Paris : Seuil.
- LANDRAGIN F. (1999). *Analyse de l'articulation entre parole et geste dans un corpus multimodal*. mémoire de dea, Université de Marne-La-Vallée.
- LANGACKER R. (1999). *Grammar and conceptualization*. Walter de Gruyter.
- MCNEILL D. (1992). *Hand and Mind*. Chicago and London : The University of Chicago Press.
- MULLER P. & PREVOT L. (2003). An empirical study of acknowledgement structures. In *Proceedings of DIABRUCK 2003*.
- PAVIA N. G. (2000). Hétérogénéité des corpus : vers un parseur robuste reconfigurable et adaptable. In *Proceedings of RECITAL 2000*.
- PINEDA L. & GARZA G. (2000). A model for multimodal reference resolution. *Computational Linguistics*, **26**(2).
- VANRULLEN T., GUÉNOT M.-L. & BELLENGIER E. (2003). Formal representation of property grammars. In *Proceedings of ESSLLI Student Session*.