

## Identification des chaînes de référence dans les systèmes automatiques : le cas des anaphores dans les *N de N*.

**Laurence Kister**

Université Nancy 2 - IUT A

LanDisCo

kister@iuta.univ-nancy2.fr

Les systèmes automatiques d'extraction de l'information, d'indexation, de génération de thesauri mais également de résumé tendent à se multiplier. Ces systèmes intègrent, bien entendu, des modules dits de traitement linguistique. Cependant, ce qui est désigné par *modules de traitement linguistique* correspond à des réalités très différentes. Il peut s'agir de modules qui permettent l'attribution d'étiquettes syntaxiques parfois complétés par des modules qui offrent la possibilité d'effectuer une description par l'attribution de traits sémantiques sous forme de codage. Ce qui est appelé analyse linguistique se limite alors souvent au repérage des chaînes de caractères, éventuellement à leur étiquetage syntaxique, puis à leur comptage avec ou sans pondération. Les modules les plus performants permettent de distinguer, en contexte, par une analyse distributionnelle la catégorie syntaxique d'un lexème : ils analysent, par exemple, le contexte d'apparition d'un lexème ambigu afin de déterminer s'il est en position de substantif ou de verbe. Si cette opération paraît assez facilement automatisable, il n'en est pas de même pour celle qui doit permettre de déterminer la valeur sémantique des entités linguistiques. Les modules dits sémantiques consistent dans la majorité des cas, en des systèmes de reformulation des requêtes fondés sur l'utilisation d'une sorte de thesaurus ou dictionnaire d'équivalence. Ils utilisent essentiellement les relations sémantiques de base (synonymie, hyperonymie, hyponymie...) afin d'établir une série de renvois vers des termes autorisés et d'exclure certains termes : il s'agit d'un moyen de réduire, en partie, le problème de la polysémie<sup>1</sup>. L'analyse des relations sémantiques et la mise en place de réseaux sémantiques se fait fréquemment par l'intermédiaire de dictionnaires qui précisent la forme canonique du lexème, son domaine sémantique... Certains donnent une définition sémantique au moyen de traits et associent ainsi les concepts et les mots susceptibles de figurer dans le texte. Les relations définies peuvent servir de fondement à l'analyse qui permet de mettre en place des réseaux sémantiques. Il

s'agit, alors, de modules, qui mettent en oeuvre des règles de reformulations, destinés à repérer les expressions lemmatisées et les idiomatismes, sans qu'il soit nécessaire d'utiliser des cotes, des booléens ou d'autres formes de codage pour montrer qu'il s'agit d'une expression. Il existe également des modules capables de réaliser une analyse morphologique c'est-à-dire de reconnaître automatiquement les signifiants quand ils apparaissent sous une forme dérivée ou une forme flexionnelle : cette reconnaissance peut être le fruit d'une véritable analyse automatique ou de la simple comparaison avec un dictionnaire qui liste l'ensemble des variantes. L'utilisation de ces systèmes montre que très souvent les modules se limitent au repérage des chaînes de caractères, à un traitement morpho-syntaxique afin d'obtenir la forme canonique des lexèmes et parfois les radicaux. Ce type de systèmes se contente lui aussi souvent de compter les entités reconnues, parfois en les pondérant, afin de déterminer les éventuels descripteurs. Un module statistique peut compléter l'amorce d'analyse linguistique.

Les systèmes de bases de connaissances sont, pour leur part, présentés comme des systèmes qui permettent de " produire une représentation organisée des concepts " accompagnée des rôles " qu'ils jouent et des liens qui existent entre eux " <sup>2</sup>. Ils travaillent généralement sur les titres et, de ce fait, n'analysent, dans de nombreux cas, que des constructions nominales. En effet, les titres qui ne comportent que des substantifs, des adjectifs et des mots grammaticaux tels les déterminants et les prépositions sont très fréquents. Une fois les unités linguistiques étiquetées, des règles de " compositions " permettent de reconnaître les séquences possibles ou autorisées. L'étiquetage apporte des informations en ce qui concerne les possibilités de composition des entités codées. Ces informations d'ordre syntaxique (place des unités en fonction de leur catégorie grammaticale, par exemple) se veulent de rendre compte des relations hiérarchiques entre les mots.

D'autres approches d'interprétation des textes et de représentation des textes ont été envisagées. On peut citer celles qui tentent de comprendre et de reproduire les réseaux de neurones ou encore la théorie de la sémantique des caractères<sup>3</sup> ou Readware (utilisée notamment par ConSearch de Management Information Technologies Inc). Les limites technologiques des systèmes disponibles sur le marché apparaissent de manière nette quand on précise que la plupart d'entre eux, notamment ceux qui s'attachent à l'indexation automatique, analysent le titre, plus rarement le résumé et extrêmement rarement le texte intégral.

Les informations que l'on peut obtenir sur les modules dits de traitement automatique des langues naturelles commercialisés sont présentées de manières succinctes. Il s'agit, en quelques sortes, de "boîtes noires" linguistiques, pour lesquelles, on sait, tout au plus, qu'elles intègrent des modules d'analyse morpho-syntaxique et/ou une ébauche d'analyse syntaxique et sémantique. On peut se demander si ce manque de transparence ne masque pas l'absence de véritable traitement automatique de la langue<sup>4</sup>.

## Un problème de coréférence : l'anaphore.

La nécessité d'inclure des modules d'analyse automatique qui prennent en compte les relations syntaxico-sémantiques est une problématique relativement ancienne comme le révèle, par exemple, les applications de SATO (Système d'Analyse des Textes par Ordinateur) développées essentiellement au Canada<sup>5</sup>. Les concepteurs de ce système qui a fait l'objet de nombreuses applications soulignent la nécessité de le compléter par des modules capables de procéder à "des analyses plus fines des lexies complexes"<sup>6</sup>. Ils précisent également que le logiciel présente l'avantage de pouvoir être couplé à d'autres logiciels, notamment statistiques ou documentaires.

La nécessité de modules de traitement automatique qui permettent de reconnaître un référent présent sous différentes formes apparaît comme une évidence. En effet, le thème de discours ou focus peut prendre la forme d'entités lexicales ou syntaxiques. Le changement de lexème et l'utilisation de mots grammaticaux pour désigner un même référent de la réalité extralinguistique sont des attitudes discursives courantes. L'identification de ces différentes entités comme des unités correspondant à un même concept, un même référent soulève de nouvelles limites des systèmes automatiques. La difficulté à reconnaître les chaînes de référence témoigne des limites des systèmes de traitement de la langue. Si on définit sommairement la *chaîne de référence* comme une série de lexèmes, de construction et d'expressions synonymes, plus ou moins complexes et de mots

grammaticaux (pronoms personnels sujets, relatifs...) qui désignent un même référent, on voit très vite émerger le problème de leur reconnaissance par la machine. Bien entendu, cette problématique est loin d'être nouvelle comme en témoigne la littérature abondante sur ce sujet. Ce phénomène de *chaîne de référence* est illustré par l'exemple<sup>7</sup> suivant :

(Essai général sur la préposition)<sub>1</sub>, (l'ouvrage de (J. Cervoni)<sub>5</sub>)<sub>1</sub> vise à combler [une attente]<sub>2</sub> : [celle (des)<sub>2</sub> nombreux grammairiens et linguistiques]<sub>3</sub> (qui)<sub>3</sub> pensent que les prépositions occupent une position privilégiée dans toute réflexion ayant pour objet (les fonctions syntaxiques, le phénomène de la transitivité, l'expression de la relation entre les langues, la charge sémantique des mots dits "grammaticaux", les rapports entre logique et langage, la portée de la grammaire des cas, l'articulation entre syntaxe et pragmatique)<sub>4</sub>. (Ces problèmes)<sub>4</sub> sont traités par (l'auteur)<sub>5</sub> dans un cadre théorique précis, issu de la psychomécanique du langage. Mais, (le schéma)<sub>6</sub> à partir (duquel)<sub>6</sub> (cette étude multidirectionnelle)<sub>1</sub> s'organise est (un élargissement du modèle guillaumien)<sub>1/6</sub>, en ce sens qu'(il)<sub>6</sub> intègre certains développements récents de la sémantique logique et les principaux acquis de la pragmatique. De (cette synthèse)<sub>1</sub>, découle une contribution nouvelle à l'élaboration d'une théorie de la préposition.

(Cervoni, Jean, 1991)

Ce résumé montre combien la reconnaissance d'une chaîne de référence peut s'avérer complexe. Il est aisé de comprendre que le traitement de ce genre de relations sémantiques par les systèmes du type de ceux que nous avons évoqué n'est pas suffisant. Aux relations sémantiques de type synonymie - que certains systèmes peuvent résoudre partiellement - s'ajoute celui de la reconnaissance de l'antécédent et de l'appariement de formes telles *celle de, duquel, qui, il...* Les travaux que nous évoquons ici n'ont pas la prétention de déboucher sur une proposition de traitement automatique de ce type de coréférence. Ils visent beaucoup plus simplement et pragmatiquement à sensibiliser le lecteur au problème complexe que pose le repérage et l'appariement des anaphoriques, notamment dans le cas des anaphores pronominales en *qui* qui recrutent un référent au niveau d'une structure en *dét. N1 de (dét.) N2 (le visage de la fillette, cette campagne de publicité, le directeur du journal, une pomme de terre...)*. Ils soulèvent également le problème que pose la non reconnaissance des pronoms et plus généralement des anaphores dans les systèmes actuels. Ils laissent aussi entrevoir les problèmes qui découleraient d'une tentative de prise en compte de ces relations par un système automatique.

## A propos de l'appariement de *qui* et d'un référent introduit au niveau d'un *dét. N1 de (dét.) N2*.

La reconnaissance des unités linguistiques qui désignent un même référent, et en particulier le cas des anaphores a été étudié par de nombreux linguistes et différents systèmes de reconnaissance automatiques sur lesquels nous ne revenons pas ici ont été envisagés. Le type particulier de reprise anaphorique qui nous intéresse est celui pour lequel le référent est introduit au niveau des *N de N*. Notre objectif est d'identifier des propriétés capables de prévoir ou du moins d'établir une échelle de préférence face aux possibilités d'anaphores. Pour ce type particulier de construction, le référent de l'anaphorique peut correspondre :

- au *N1* ;  
**La volonté de la jeune fille** s'oppose à **celle** de ses parents.
- au *N2* ;  
Ils agissent dans **la limite des pouvoirs qui** leur sont conférés.
- au *N1 de N2* ;  
Pour les assureurs, **le gel du bonus, qui** a entraîné des changements de comportement des assurés, est une aubaine.  
**L'affaire des bannis** prend tout son sens.  
**Elle/celle-ci/cet événement** ravive de vieilles blessures.
- au *N1* et au *N2* repris par deux anaphoriques différents ;  
**L'affaire des bannis** prend tout son sens.  
**Elle leur** rappelle d'affreux souvenirs.

Ces exemples illustrent quelques unes des combinaisons et des possibilités que nous avons observées. L'anaphore peut être nominale ou pronominale. Sa position grammaticale tout comme la position de l'expression qui introduit le référent peut varier : l'anaphorique et l'anaphorisé peuvent être en position sujet, objet direct, objet indirect, complément circonstanciel... La nature grammaticale et la position grammaticale des entités entre lesquelles s'établit la relation anaphorique ne sont pas suffisantes pour déterminer une éventuelle préférence pour l'un des constituants ou l'ensemble de la construction. L'argument qui consiste à dire que l'anaphorisant et l'anaphorisé occupent des positions grammaticales identiques n'est lui non plus pas valable. Si on prend l'exemple des reprises par *qui* en position sujet d'un référent introduit au niveau d'un *dét. N1 de (dét.) N2*, la position grammaticale de ce dernier peut varier. Elle peut être, par exemple, en position :

- sujet ;  
**La valeur du point, qui** sert à calculer les retraites, reste fixée à 23.03 francs.
- circonstanciel ;  
Chaque institution agit dans **la limite des attributions qui** lui sont conférées par le présent Traité.
- objet ;  
C'est **les pluies de printemps qui** sont bonnes pour les cultures.

Il est important de rappeler que, pour certains *dét. N1 de (dét.) N2*, on est appelé à s'intéresser au fonctionnement référentiel de la construction en *de* et au processus anaphorique à proprement parler. L'examen de la micro-structure qui correspond au *N de N* et la relation sémantique marquée par la préposition *de* suffisent, dans certains cas, à d'établir des préférences face aux possibilités de reprises<sup>8</sup>. Il s'agit, par exemple, des structures qui correspondent à des expressions lexicalisées pour lesquelles le *N2* n'est jamais précédé par un déterminant (*pomme de terre, lampe de poche...*) et de celles qui comportent un *N1* qui fait office de quantifieur du *N2* (*bande d'enfants, livre de fraises...*). Pour ces deux types particuliers de structure prépositionnelle, c'est respectivement l'ensemble du *dét. N1 de N2* et le *N2* qui sont préférés en cas de reprise anaphorique.

Ces deux exemples de fonctionnement lié à la structure et au sémantisme de la construction complexe attestent du rôle que jouent la présence et l'absence de déterminants devant les *N*. Les déterminants ainsi que la combinaison des deux déterminants du *dét. N1 de (dét.) N2* peuvent rendre l'un des constituants ou l'ensemble de la construction suffisamment saillant pour faire l'objet d'une reprise. Ils constituent aussi une marque de surface facilement repérable en traitement automatique. Ces deux facteurs sont à l'origine d'une étude sur corpus<sup>9</sup> qui vise à observer d'éventuelles régularités fonctionnelles en relation avec la distribution des déterminants. Cette analyse révèle que lorsque le *N2* suit directement la préposition on reprend, dans 89 % des cas, l'ensemble de la construction tandis que lorsque le *N2* est précédé par un déterminant la reprise sur l'ensemble de la construction ne se fait que dans 41 % des cas. Les observations sur corpus montrent aussi que 47.82 % des anaphores sont réalisées au moyen du relatif *qui*, 18.01 % grâce à des pronoms personnels sujets, 8.78 % se font par la structure *celui de*. Les autres formes anaphoriques (*où* ; préposition + *lequel*<sup>10</sup> ; pronom objet direct ; démonstratif + *N* ; *celui qui* ; pronom objet indirect ; défini + *N...*) n'atteignent pas 5 % du nombre total d'occurrences repéré.

Si on s'intéresse aux possibilités de reprise en fonction des combinaisons de déterminants, on observe que certaines d'entre elles sont peu ou pas

représentées. Nous n'avons trouvé aucune occurrence pour *dém. N1 de dém. N2*. Les occurrences de *ind. N1 de dém. N2*, *dém. N1 de ind. N2* et *déf. N1 de dém. N2* sont, quant à elles, peu nombreuses. Pour *ind. N1 de ø N2*, *déf. N1 de ø N2*, *déf. N1 de ind. N2*, *déf. N1 de déf. N2*, *dém. N1 de ø N2* et *dém. N1 de déf. N2* l'analyse montre une dépendance des paramètres<sup>11</sup> que constituent les deux déterminants. Pour ces six combinaisons, on observe la répartition suivante face aux possibilités de reprises<sup>12</sup> :

	Reprise du <i>dét. N1 de (dét.) N2</i>	Reprise du <i>N2</i>
<i>ind. N1 de ø N2</i>	89%	11%
<i>déf. N1 de ø N2</i>	84%	16%
<i>déf. N1 de ind. N2</i>	27%	73%
<i>déf. N1 de déf. N2</i>	45%	55%
<i>dém. N1 de ø N2</i>	89%	11%
<i>dém. N1 de déf. N2</i>	78%	22%

Ces résultats, qui ne prennent pas en compte la nature de l'anaphorique, semblent montrer que l'ensemble de la structure est préféré en cas de reprise anaphorique, excepté pour les *déf. N1 de ind. N2* et les *déf. N1 de déf. N2*. Ces résultats sont en accord avec ceux qui résultent des expérimentations menées en psycholinguistique<sup>13</sup>. Si on se focalise sur les seules anaphores en *qui* - sur lesquelles portent nos travaux actuels et auxquelles se sont, exclusivement, intéressés les psycholinguistes au cours des travaux que nous venons d'évoquer - on se rend compte qu'il est difficile de généraliser les résultats. En effet, les combinaisons pour lesquelles les résultats sont significatifs diffèrent. Il s'agit des *ind. N1 de ø N2*, des *déf. N1 de ø N2*, des *déf. N1 de ind. N2*, des *déf. N1 de déf. N2* et des *déf. N1 de dém. N2*. Pour ces cinq combinaisons on observe les préférences suivantes :

	Reprise du <i>dét. N1 de (dét.) N2</i>	Reprise du <i>N2</i>
<i>ind. N1 de ø N2</i>	91%	9%
<i>déf. N1 de ø N2</i>	86%	14%
<i>déf. N1 de ind. N2</i>	15%	85%
<i>déf. N1 de déf. N2</i>	40%	60%
<i>déf. N1 de dém. N2</i>	15%	85%

Le comptage effectué montre cette fois que l'anaphorique qui recrute son antécédent au niveau d'une construction complexe en *de* sélectionne plus souvent le *N2* que l'ensemble du *dét. N1 de (dét.) N2*. Ces résultats sont en contradiction avec ceux obtenus lors des manipulations psycholinguistiques évoquées ci-dessus. Ces divergences sont à l'origine d'une comparaison du matériel expérimental et des occurrences observées sur corpus : le matériel

expérimental ne comporte que des constructions pour lesquelles les deux *N* sont porteurs des traits *animé* et *humain* tandis que ce type de construction est quasi inexistant dans notre corpus. Le nombre de *dét. N1 de (dét.) N2* qui comportent deux *N* porteurs du trait *animé* est de l'ordre d'une vingtaine d'occurrences. Par ailleurs, les constructions utilisées pour les expérimentations psycholinguistiques comprennent toutes, exclusivement, deux déterminants définis. Un argument simpliste, en faveur des résultats obtenus en observant les occurrences du corpus, à savoir grand nombre de reprise du *N2*, revient à mettre en avant la proximité entre l'antécédent et le relatif. Cette régularité n'exclut pas la possibilité de reprendre l'ensemble de la construction plutôt que le seul *N2*. En effet, « si ce qui précède le relatif est un syntagme complexe on doit déterminer où se trouve effectivement l'antécédent »<sup>14</sup>. Les différents comptages effectués et la répartition des deux formes de saisie - l'ensemble de la construction ou le *N2* - confortent dans notre démarche qui consiste à identifier un plus grand nombre de facteurs intervenant dans le processus de reprise anaphorique pour établir une échelle de préférence.

L'idée de départ de nos travaux est de déterminer un ensemble de facteurs qui permette de pré-sélectionner le référent d'un anaphorique au niveau d'un *dét. N1 de (dét.) N2*. Les déterminants présentent l'avantage de constituer une marque de surface facilement repérable lors du traitement automatique de la langue. L'analyse des déterminants se révèle cependant insuffisante. Dans certains cas, même les résultats présentés comme significatifs ne permettent pas de trancher en faveur d'une saisie préférentielle du *dét. N1 de (dét.) N2* ou du *N2*. En effet, si on observe les résultats proposés ci-dessus, il paraît difficile de justifier une préférence en cas de reprise par *qui* d'un référent introduit au niveau d'un *déf. N1 de déf. N2*. On observe 40% d'anaphores sur l'ensemble de la structure prépositionnelle et 60 % sur le *N2*. Avant d'envisager la sélection d'un nombre plus important de facteurs à prendre en compte on a, bien entendu, envisagé un accroissement de notre corpus. L'accroissement du nombre d'occurrences n'a eu de répercussions notables ni sur les différentes combinaisons ni sur les possibilités de reprise. La proposition de reprise de l'ensemble de la structure et du *N2* par combinaison ainsi que celle du nombre d'occurrences par combinaison sont restées sensiblement identiques. Afin d'affiner les propositions de présélection envisagées et de déterminer de nouveaux paramètres à prendre en compte, on a entrepris de nouveaux comptages. De ces analyses en cours, il ressort une première série de facteurs susceptibles d'intervenir dans le processus de sélection du référent : traits sémantiques des *N* et position grammaticale du *dét. N1 de (dét.) N2*. Ces dernières ont été évoquées ci-dessus. Les *N* qui composent les constructions complexes sont, quant à

eux, de natures relativement diverses. Ils peuvent être *abstrait, animé, inanimé, quantifieur, collectifs...* (*les membres de la communauté, un panier de fraises, le visage de la fillette, une porte de la voiture, le capital de la banque, la plupart des participants, le bord du lac...*) mais aussi correspondre à un nom de lieu, une date (*la ville de Nancy, le traité de Varsovie, le traité du 15 juin...*). Ces exemples font apparaître un type de constructions que nos travaux laissent momentanément en attente : les *dét. N1 de (dét.) N2* qui comportent des noms propres ou des dates.

La diversité des formes anaphoriques, des constructions prépositionnelles en *de* capables d'introduire le référent, des positions grammaticales de l'anaphorisé et de l'anaphorisant, des traits sémantiques des deux *N*, des relations marquées par *de* et du fonctionnement référentiel de celles-ci rend la description linguistique des relations qu'on propose d'analyser et d'expliquer ardue.

## En guise de conclusion...

Les différentes possibilités de recrutement d'un référent au niveau d'un *dét. N1 de (dét.) N2* demandent qu'on repère les formes d'anaphores les plus courantes et qu'on analyse leurs régularités fonctionnelles avant d'envisager une automatisation du processus de reconnaissance du référent. Les observations sur corpus qui ont été entreprises, visent, dans un premier temps, à vérifier les prévisions envisageables et à préciser les paramètres qu'il est nécessaire de prendre en compte. Des divergences entre ce qui résulte des travaux d'observation sur corpus et des manipulations psycholinguistes<sup>15</sup> découlent nos travaux actuels. Ceux-ci visent vérifier l'importance qu'on a accordée aux déterminants et à leur combinaison. Pour cela, on envisage une expérimentation psycholinguistique destinée à vérifier les hypothèses formulées et la pertinence des premiers paramètres identifiés, à savoir, le rôle de la détermination et le type d'anaphoriques.

Dans un second temps, nos investigations actuelles s'attachent à la position grammaticale du *dét. N1 de (dét.) N2* et aux traits sémantiques des deux *N*. La réflexion que nous menons actuellement pour construire le matériel expérimental nous conduit, en raison du nombre et de la diversité des structures anaphoriques possibles à nous intéresser plus particulièrement aux reprises par le relatif *qui* d'un référent introduit au niveau d'une construction prépositionnelle en *de*.

## Références bibliographiques

- BACCINO, T., DE VICENZI M., JOB R., Cross  
Linguistic Studies of the late-closure Strategy :  
French and Italian, AMLaP'96, septembre 96,  
Torino. (à paraître)
- BALTAZART, Didier, KISTER, Laurence, Is it  
possible to predetermine a Referent included in a  
French *N de N* Structure ? , in *Corpus-Based and  
Computational Approches to Discourse Anaphora*,  
S.P. Botley and A.M. McEnnery eds., Amsterdam :  
John Benjamins Publishing Company. (à paraître)
- BERTRAND-GASTALDY, Suzanne, L'analyse du  
contenu textuel en vue de la construction de thesaurus  
et de l'indexation assistées par ordinateur ;  
applications possibles avec SATO. In *Documentation  
et bibliothèques*, avril-juin, 1992.
- CERVONI, Jean. *La préposition : étude sémantique  
et pragmatique*. Paris : Duculot, 1991. 309p.  
(Champs linguistiques).
- CORET, Annie, MENON, Bruno, SCHRIBLER,  
Daniele, TERRASSE, Christophe. Un système  
d'indexation structuré à l'Inist : bilan d'une étude  
préalable. In *Documentaliste-Sciences de  
l'information*, 1994, Vol 31, n°3, p 148-158.
- GINGRAS, Maurice. VIVIT : Un système d'analyse  
et de repérage de l'information textuelle pour la  
gestion des ressources humaines, In *Documentation et  
bibliothèques*, avril-juin, 1992
- GREVISSE, Maurice, *Le bon usage : grammaire  
française*, Paris : Duculot, 12<sup>ème</sup> édition, 1991.
- KISTER, Laurence, Accessibilité pronominale des *dét.  
N1 de (dét.) N2* : le rôle de la détermination, in  
*Linguisticae Investigationes*, Amsterdam : John  
Benjamins Publishing Company, 1995, XIX, 1,  
p.107-122.
- LACOTE GABRYSIK, Lylotte, KISTER,  
Laurence, Le traitement automatique des langues dans  
le processus d'indexation, in *L'indexation à l'heure  
d'internet*, Deuxième colloque du chapitre français de  
l'ISKO "International Society for Knowledge  
Organisation", 21/22 octobre, Lyon, 1999. (à  
paraître)
- LELOUP, Catherine, Moteurs d'indexation et de  
recherche : environnements client-serveur, Internet et  
Intranet, Paris : Eyrolle, 1998, 285p.
- ROCHON, Yves, SAGEE : un développement  
informatique adapté aux besoins en gestion de  
l'information de la Direction des évaluations  
environnementales. In *Documentation et  
bibliothèques*, avril-juin, 1992.
- PYNTE, Joël, The Time-course Attachment in  
French : Evidence from French, in *Syntax and  
Semantics*, 31, 1998, p227-245.
- PYNTE, Joël, FRENCK-MESTRE, C. Early-closure  
Attachment in French : a Replication, AMLaP'96,  
septembre 96, Torino, 1996.

ZAGAR, Daniel, *La lecture processus de base*, Habilitation à diriger les recherches, Université de Bourgogne, UFR Sciences Humaines, LEAD URA CNRS 1838, 1995.

ZAGAR, Daniel, RATIVEAU Sophie, Evidence of Early-closure Attachment of first-pass Reading Times in French, in *Quarterly Journal of Experimental Psychology*, 50A, 1997, p421-438.

---

<sup>1</sup> Lyette LACOTE GABRYSIK et Laurence KISTER (à paraître).

<sup>2</sup> Annie CORET, Daniele SCHIBLER et Christophe TERRASSE (1994).

<sup>3</sup> Catherine LELOUP (1998).

<sup>4</sup> Lyette LACOTE GABRYSIK et Laurence KISTER (à paraître).

<sup>5</sup> Maurice GINGRAS (1992) et Yves ROCHON (1992).

<sup>6</sup> Dans son article Suzanne BERTRAND-GASTALDY (1992) cite Aleth, Spirit, Fasit.

<sup>7</sup> Résumé présenté en quatrième de couverture : Jean CERVONI (1991).

<sup>8</sup> Laurence Kister (1995).

<sup>9</sup> Didier BALTAZART et Laurence KISTER (à paraître).

<sup>10</sup> Nous avons regroupé dans sous ce terme générique les formes duquel et auquel ainsi que leurs variantes.

<sup>11</sup> Le test du Chi 2 montre une dépendance des deux déterminants contenus dans le *dét. N1 de (dét.) N2*.

<sup>12</sup> Didier BALTAZART et Laurence KISTER (à paraître).

<sup>13</sup> Daniel Zagar (1994), Daniel Zagar et Sophie Rativeau (1997).

<sup>14</sup> Maurice Grevisse (1991), § 896d, p. 1353.

<sup>15</sup> T. Baccino, M. De Vicenzi et R. Job (1996), Joël Pynte et C. Frenck-Mestre (1996), Daniel Zagar et Sophie Rativeau (1997), Joël Pynte (1998).