

# De la résolution des anaphores

**Dominique Laurent**  
**Synapse Développement**

Ce texte est destiné à regrouper les résultats de Cordial sur la résolution des anaphores. En préambule, rappelons que ce travail visait à retrouver le plus possible de références pour les anaphores pronominales et possessives (adjectifs possessifs) ainsi que pour les noms génériques comme "ce dernier", "le premier", "la seconde", "un seul", "un autre", etc. Ce référencement sert en particulier pour le moteur de questions-réponses.

## 1. Méthodologie

Résoudre les anaphores, c'est-à-dire retrouver le ou les référents d'un pronom, d'un adjectif possessif, d'un adverbe générique ("beaucoup"), d'un nom générique ("ce dernier", "le premier",...), suppose que l'on dispose d'hypothèses théoriques et d'un corpus d'évaluation et de test. Les hypothèses théoriques avaient été formulées par Joaquin à partir des textes fournis en bibliographie, à partir d'une étude cas par cas des différents types d'anaphores et, enfin, à partir de nos bases lexicales, grammaticales et sémantiques, les algorithmes de résolution ne pouvant être construits qu'à partir de ces ressources.

Le corpus d'évaluation utilisé est un corpus fourni par Agnès Tutin de l'université Stendhal de Grenoble. Il compte un peu plus de 100 000 mots et étiquette 2510 anaphores. Ces anaphores sont essentiellement pronominales mais on y trouve aussi des adjectifs possessifs, quelques nombres et des adverbes. N'ont pas été annotés les pronoms relatifs, la plupart des adverbes anaphoriques et les nominalisations.

Afin d'utiliser ce corpus au format XML, il a fallu créer des moulinettes de conversion dans un format Ascii simplifié. A partir de ce texte Ascii, ont été constitués plusieurs corpus distincts pour la résolution des différents types d'anaphores. Un tableau Excel regroupant les anaphores et leurs caractéristiques a par ailleurs été constitué .

Ce tableau Excel fournissait les données suivantes :

- intitulé de l'anaphore (pour les reprises du sujet, le trait d'union précédant l'anaphore a été repris en fin d'anaphore et entre parenthèses afin de simplifier les tris ultérieurs.
- numéro de l'anaphore dans le texte (ce numéro peut aller de e0 à h981).
- fonction grammaticale de l'anaphore.
- numéro de proposition de l'anaphore dans la phrase.
- intitulé du référent (dans le corpus étiqueté le début du référent était étiqueté, donc en général un article, ici le référent fourni est le nom, parfois les noms; très rarement le référent peut être un verbe)
- distance en phrases. 0 signifie que l'anaphore et le référent sont dans la même phrase. -1 que le référent est dans la phrase précédente, +1 qu'il est dans la phrase suivante, etc.
- distance en mots. Est indiqué ici le nombre de mots séparant l'anaphore du référent. Ce compte peut être nul, par exemple si l'anaphore suit le référent (exemple : "un substrat lui-même monocristallin).
- distance en noms. Est indiqué ici le nombre de noms, donc de référents possibles, trouvés entre l'anaphore et le référent.
- fonction grammaticale du référent.
- numéro de proposition du référent dans la phrase (cette donnée n'est vraiment utile que lorsque l'anaphore et le référent sont dans la même phrase).

Nous disposons d'un autre corpus pour test, gracieusement fourni par Catherine Clouzot (qui travaille avec Agnès Tutin). Malheureusement le format d'étiquetage de ce second corpus est très différent du premier format et pour l'exploiter, il serait nécessaire de créer un autre programme de conversion.

L'étiquetage du corpus utilisé ici, une fois simplifié, fournit une étiquette du type "<"e108"> qui est placée au début du groupe référent. L'anaphore ayant ce référent disposera d'une étiquette <="e108"=> placée avant l'anaphore. Les tests effectués ont consisté à vérifier que le groupe nominal renvoyé en référence comportait bien l'étiquette placée devant l'anaphore. Cette méthode ne garantit pas que le référent trouvé est toujours strictement identique au référent étiqueté. Ainsi, lorsque le référent est "l'injustice de l'exclusion", si nos algorithmes trouvent comme référent "l'injustice", la réponse sera considérée comme bonne. A contrario, si le référent étiqueté était "Chacune des familles" mais que le programme trouve "familles" (parce que l'anaphore est au pluriel), alors la réponse sera considérée comme fautive. L'étiquetage a été généralement fait pour que le référent choisi soit le plus récent lorsqu'il y a avait plusieurs référents identiques successifs dans le texte. Les algorithmes ont donc utilisé ce principe mais, malheureusement, quelques cas d'espèce ont montré que l'étiquetage n'avait pas toujours renvoyé aux référents les plus récents, ce qui amène quelques fausses erreurs apparentes de référencement.

Afin de valider les résultats du corpus d'évaluation décrits ci-dessous, il faudrait, sans doute à partir d'un texte de difficulté moyenne, en tous cas moins complexe que le corpus d'évaluation, étiqueter automatiquement les anaphores puis vérifier manuellement l'ensemble des étiquettes pour obtenir des pourcentages fiables.

## 2. Différences entre les anaphores

Avant de détailler les anaphores, il faut dire deux mots du pronom impersonnel "il". Dans le corpus de travail, on compte 401 usages impersonnels du pronom "il", soit plus que le nombre d'emplois personnels (296). La distinction de l'emploi impersonnel ou personnel est donc une étape préliminaire importante pour la résolution de l'anaphore. Cette phase a fait l'objet d'un effort particulier car le repérage de l'impersonnel est en lui-même intéressant pour l'analyse globale du discours. Sur le corpus du travail, 397 impersonnels sont correctement identifiés (et 4 sont mal identifiés) tandis que 2 formes personnelles sont étiquetées comme impersonnelles. On peut donc estimer que cette différenciation s'effectue très correctement. Une validation sur un nouveau corpus est toutefois nécessaire.

Le tableau ci-dessous fournit les différents types d'anaphores par catégorie grammaticale. Si les pronoms personnels représentent plus de 50 % de l'ensemble des anaphores, le nombre des anaphores possessives (29,4 %) est tout à fait remarquable, d'autant plus que les études universitaires sur la référence ont largement sous-estimé ce type d'anaphores.

	Nombre d'occurrences	%
pourcentages + nombres	54	2,1 %
adverbes (beaucoup)	5	0,2 %
noms (ce dernier, le premier, un seul...)	47	1,9 %
pronoms personnels sujets	841	33,5 %
pronoms personnels COD	220	8,8 %

pronoms personnels COI	320	12,7 %
pronoms démonstratifs	195	7,8 %
pronoms indéfinis	91	3,6 %
adjectifs possessifs	737	29,4 %
TOTAUX	2510	100,0 %

Le tableau ci-dessous fournit le positionnement global des anaphores par rapport aux référents. On note que plus d'une anaphore sur 10 trouve son référent dans une phrase qui n'est ni la phrase courante ni la précédente et qu'environ deux anaphores sur trois ont leur référent dans la phrase courante.

	nombre	%	nombre cumulé	% cumulé
référent 25 phrases avant l'anaphore	1	0,04 %	1	0,04 %
référent 18 phrases avant l'anaphore	1	0,04 %	2	0,08 %
référent 17 phrases avant l'anaphore	1	0,04 %	3	0,11 %
référent 16 phrases avant l'anaphore	2	0,08 %	5	0,20 %
référent 15 phrases avant l'anaphore	1	0,04 %	6	0,24 %
référent 14 phrases avant l'anaphore	1	0,04 %	7	0,28 %
référent 13 phrases avant l'anaphore	1	0,04 %	8	0,32 %
référent 12 phrases avant l'anaphore	4	0,16 %	12	0,48 %
référent 11 phrases avant l'anaphore	2	0,08 %	14	0,56 %
référent 10 phrases avant l'anaphore	3	0,11 %	17	0,68 %
référent 9 phrases avant l'anaphore	2	0,08 %	19	0,76 %
référent 8 phrases avant l'anaphore	4	0,16 %	23	0,92 %
référent 7 phrases avant l'anaphore	6	0,24 %	29	1,16 %
référent 6 phrases avant l'anaphore	11	0,44 %	40	1,59 %
référent 5 phrases avant l'anaphore	13	0,52 %	53	2,11 %
référent 4 phrases avant l'anaphore	27	1,08 %	80	3,19 %
référent 3 phrases avant l'anaphore	51	2,03 %	131	5,22 %
référent 2 phrases avant l'anaphore	129	5,14 %	260	10,36 %
référent phrase précédant l'anaphore	563	22,43 %	824	32,83 %
référent même phrase que l'anaphore	1679	66,89 %	2503	99,72 %
référent phrase suivant l'anaphore	7	0,28 %	2510	100,0 %

Ce tableau reflète un comportement global. Or la distance entre l'anaphore et le référent est liée au type de l'anaphore. Le tableau suivant fournit les pourcentages pour chacun des différents types d'anaphores :

	25 à 3 phrases avant	2 phrases avant	phrase avant	phrase courante
personnels sujets	70 (8,2 %)	66 (7,6 %)	267 (30,9 %)	460 (53,3 %)
personnels COD	6 (2,8 %)	9 (4,2 %)	40 (18,6 %)	160 (74,4 %)
personnels COI	19 (5,9 %)	10 (3,1 %)	87 (27,2 %)	204 (63,7 %)
démonstratifs	0 (0,0 %)	3 (1,6 %)	28 (14,4 %)	163 (84,0 %)
indéfinis	7 (6,3 %)	5 (4,5 %)	21 (18,7 %)	79 (70,5 %)
adjectifs possessifs	29 (4,0 %)	32 (4,3 %)	103 (14,0 %)	573 (77,7 %)
Totaux	131 (5,4 %)	125 (5,1 %)	547 (22,4 %)	1639 (67,1 %)

Ce tableau ne porte que sur 2442 anaphores, soit 68 de moins que le total de départ, car les anaphores nombres et pourcentages n'ont pas été étudiées et, par ailleurs, quelques anaphores douteuses n'ont pas été comptabilisées ici. On notera le comportement très atypique des anaphores démonstratives qui ont leur référent essentiellement dans la phrase courante (en fait le référent d'une anaphore démonstrative est dans deux cas sur trois le groupe nominal la précédant). On peut également noter que les pronoms personnels sujets (il, ils, elle, elles) sont ceux qui ont le référent le plus éloigné. En fait, si l'on considère les reprises de référents, c'est-à-dire les cas où "il" ou "elle" renvoie vers une autre anaphore "il" ou "elle" qui elle-même ou la précédente (ou celle qui précède) a le référent nominal de départ, il est probable que le référent d'un pronom personnel sujet n'est pas plus éloigné que celui d'une anaphore COD ou COI.

La cataphore (anaphore pour laquelle le référent est placé après, comme dans "son heure de gloire, Matthieu l'a connu...") est finalement d'usage assez restreint : 81 cas dans le corpus de travail, soit 3,2 %. Le tableau ci-dessous fournit les différents cas. Il y apparaît clairement que si la cataphore est inimaginable pour une anaphore démonstrative, elle est par contre d'usage réel pour les personnels COD, prenant d'ailleurs parfois la forme d'une cataphore en phrase suivante comme dans : " Mais dans tous les cas, ainsi que le note Landré (1991) : " Leur théorisation (celle de Sperber et Wilson) fait jouer un rôle central à ces connecteurs qui permettent des raisonnements syllogistiques, étant donné que, pour eux, les raisonnements mis en œuvre sont essentiellement de nature déductive. ".

	cataphore	anaphore	total	pourcentage
personnels + noms sujets	14	849	863	1,6 %
personnels COD	20	200	220	9,1 %
personnels COI	6	314	320	1,9 %
démonstratifs	0	195	195	0,0 %
indéfinis + noms <> sujets	3	109	112	2,7 %
adjectifs possessifs	38	699	737	5,2 %
Totaux	81	2366	2447	3,3 %

### 3. Étude comparative des résultats de résolution

Voici les résultats globaux pour la résolution des anaphores. On notera que les totaux portent sur des montants légèrement inférieurs aux chiffres précédemment fournis, certains

pronoms ayant été mal identifiés à la désambiguïstation (par exemple confusion entre le pronom personnel et l'adverbe "y").

	résolus	non résolus	totaux	% résolus
personnels sujets	716	147	863	83,0 %
personnels COD	179	40	219	80,4 %
personnels COI	242	64	306	79,1 %
démonstratifs	182	12	194	93,8 %
indéfinis	87	25	112	77,7 %
totaux pronominaux	1406	288	1694	<b>83,0 %</b>
adjectifs possessifs	500	222	722	69,0 %
Totaux	1906	510	2416	<b>78,9 %</b>

Le taux global de résolution est donc pour les anaphores pronominales de 83 %. Pour les anaphores pronominales et les anaphores possessives, il est de 78,9 %.

Ces pourcentages couvrent des réalités bien différentes. Voici par exemple les types d'anaphores pour lesquelles les taux de résolution sont les plus faibles :

"en" à référent phrastique (7/14) .....	50,0 %
"y" à référent phrastique (5/9) .....	55,6 %
anaphore nominale (ce dernier, le précédent...) (33/51).....	64,7 %

Pour les anaphores COD (l', le, la, les), les différences selon les types de référents sont également importantes :

	nombre total	bien résolues	pourcentage
référent adjectival	21	20	95,2 %
référent phrastique	60	53	88,3 %
référent nominal	138	103	74,6 %
totaux	219	176	80,4 %

Si l'on essaye de comparer les résultats selon les types d'anaphores et surtout de référents, on peut établir l'ensemble de règles suivants :

- plus la distance entre le référent et l'anaphore est importante, plus difficile est la résolution;
- les anaphores typées en genre et en nombre (il, lui...) sont plus faciles à résoudre que celles uniquement typées en genre ou en nombre (son, sa, leur...), elles-mêmes plus faciles à résoudre que celles non typées comme "en" ou "y".
- les anaphores possessives, au-delà de l'indétermination en genre, sont plus difficiles à résoudre car elles nécessitent des connaissances pragmatiques, sur la possession, l'attribut, plus importantes que les autres types d'anaphores.
- les anaphores démonstratives, référant fréquemment au groupe nominal précédent, sont les anaphores les plus faciles à résoudre.

La méthode brutale consistant à prendre pour référent le groupe nominal placé avant l'anaphore et s'accordant en genre et en nombre donne les résultats suivants :

	% groupe précédent	% Synapse
personnels sujets	(257/863) 29,8 %	(716/863) 83,0 %

personnels COD	(78/219) 35,6 %	(179/219) 80,4 %
personnels COI	(86/306) 28,1 %	(242/306) 79,1 %
démonstratifs	(161/194) 83,0 %	(182/194) 93,8 %
indéfinis	(41/112) 35,7 %	(87/112) 77,7 %
totaux pronominaux	(623/1694) <b>36,8 %</b>	<b>83,0 %</b>
adjectifs possessifs	(234/722) 32,4 %	(500/722) 69,0 %
Totaux	(857/2416) <b>35,5 %</b>	<b>78,9 %</b>

On constate qu'existe un lien évident entre la proximité du groupe nominal et l'anaphore, les anaphores démonstratives en étant la meilleure des illustrations. On constate aussi que la méthode brutale consistant à rechercher uniquement un accord en genre et en nombre ne permet d'obtenir que des résultats médiocres (environ une anaphore sur trois).

Comment est-on parvenu d'un pourcentage de 35,5 % à 78,9 %, soit plus du double ? Essentiellement grâce aux données sémantiques mais également grâce à quelques astuces. Voici une liste des traitements qui ont permis d'obtenir ce taux final de résolution :

- prendre en compte le type sémantique du référent potentiel et celui de l'anaphore lorsque celle-ci était sujet, COD ou COI. Dans ces cas-là, nous disposons en effet, grâce aux données sur le verbe, d'informations sémantiques sur le sujet, le COD ou le COI;
- traiter différemment selon les types d'anaphores. On ne peut pas traiter de la même façon des pronoms personnels à référent phrastique et des adjectifs possessifs;
- donner un poids plus élevé aux référents sujets par rapport aux référents attributs, aux référents COD, aux référents COI, aux autres types grammaticaux. Ces poids ne sont pas les mêmes selon que l'on traite des anaphores sujets, COD, COI, etc.;
- traiter spécifiquement les relatives pour rechercher l'antécédent de la relative, qui tantôt constituera un référent potentiel puissant, tantôt ne pourra jamais être le référent;
- donner un poids plus important aux référents proches, en dosant différemment ce poids selon les types d'anaphores;
- gérer spécifiquement les cataphores possibles, en particulier pour les pronoms personnels COD (le, l', la, les);
- prendre en compte le type de déterminant introduisant le groupe nominal;
- pour les adjectifs possessifs, prendre en compte une donnée supplémentaire: le groupe sémantique d'appartenance lors d'un rapport de possession ou de méronymie;
- enfin, pour les adjectifs possessifs, la classe ontologique du nom précédé de l'anaphore et la classe ontologique du référent potentiel sont comparées et une proximité ou une équivalence améliore la probabilité du référent. Ce paramètre a toutefois peu d'impact sur le résultat global.

Pour améliorer encore la résolution des anaphores, il me semble nécessaire d'aller beaucoup plus loin dans l'appel à des données contextuelles et sémantiques. Prenons un exemple simplifié : "L'équipe du professeur Wightman a vu la cellule communiquer avec ses voisins au moyen d'une ultramicroélectrode et du signal électrique qu'elle délivre". Pour trouver que le sujet de "délivre" n'est pas "équipe" ou "cellule" mais bien "ultramicroélectrode", il faut disposer de données sémantiques importantes, permettant de relier "délivrer un signal" à "ultramicroélectrode". Cela suppose déjà de déduire de "ultramicroélectrode", inconnu des dictionnaires, le mot "électrode", puis de lier "électrode" et "délivrer un signal"...

Resteront toutefois quelques pierres d'achoppement comme :

- les référents qui ne correspondent pas en genre ou en nombre à l'anaphore (comme dans : " La diversification du gibier pour la chasse et la gestion de la population de **chevreuils** en France (environ 600 000 têtes) nécessite qu'on **le** réintroduise dans les régions d'où **il** a disparu".
- certains cas de cataphores qui, s'ils sont traités, risquent de provoquer plus d'erreurs que d'améliorations (ainsi, pour les personnels sujets, je n'ai pas pris en compte la possibilité d'une cataphore sauf pour des cas bien précis comme une proposition débutant par "bien que").
- certains référents pronominaux. Il arrive que des anaphores pronominales indéfinis ou personnelles renvoient à des pronoms qui constituent en fait des anaphores sans référent. Ainsi du référent "ceux" dans "Ceux qui le peuvent s'exilent, les autres meurent." Où "d'autres" renvoie à "ceux" qui a pour référent l'ensemble de la population du pays dont il est question dans l'article.
- certains cas de référents multiples, en particulier lorsque plusieurs parties de propositions constituent le référent phrastique d'un pronom personnel.

L'ensemble de ces cas particuliers représente toutefois moins de 5% des anaphores. On voit donc que la résolution de l'anaphore conserve encore une marge d'amélioration non négligeable, même si cette amélioration suppose des ressources linguistiques nettement plus importantes que celles dont nous disposons actuellement, du moins sous une forme accessible par les routines informatiques.

## Conclusion

La véritable validation des résultats aura lieu lorsque j'aurai réalisé une routine d'étiquetage référentiel automatique, laquelle nous permettra, après validation manuelle, de vérifier la proportion d'anaphores résolues. En attendant, on peut penser que le fait d'avoir utilisé le même corpus pour mettre en place et tester les algorithmes, et pour calculer les taux de résolution, induit un biais certain.

Il faut cependant prendre en compte la complexité des textes du corpus, qui pourrait grever les résultats. Il n'est donc pas absurde d'espérer que sur des textes plus simples, plus classiques, les résultats seront assez similaires.

## **Bibliographie**

### **Centering in Discourse**

Marilyn A. Walker, Aravind K. Joshi & Ellen F. Prince, Oxford University Press, 1998

### **L'anaphore et les pronoms**

Anne Zribi-Hertz, Presses Universitaires du Septentrion, Lille, 280 p.

### **Les formes de reprise dans le discours, Anaphores et chaînes de référence**

Francis Corblin, Voll. Langue/discours. Presses Universitaires de Rennes

voir <http://www.linguist.jussieu.fr/~corblin/reprise.pur.html>

## **Bibliographie :**

### **Sources for pronominal anaphor resolution**

Ruslan Mitkov, 3 p. [<http://www.lb.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/nlp/doc/bib/anaphora/anaphora/txt>]

### **A Discourse Copying Algorithm for Ellipsis and Anaphora Resolution**

Andrew Kehler, 10 p. [eac193.ps]

### **A Flexible Architecture for Reference Resolution**

Donna K. Byron & Joel R. Tetreault, 4 p. [testbed.ps]

### **An Algorithm for Pronominal Anaphora Resolution**

Shalom Lappin, Herbert J. Leass, 26 p. [lappin94algorithm.ps]

### **Analysis of Syntax-Based Pronoun Resolution Methods**

Joel R. Tetreault, 4 p. [tetreault.ps]

### **Anaphora for Everyone: Pronominal Anaphora Resolution without a Parser**

Christopher Kennedy, Branimir Boguraev, 6 p. [kb-coling96.ps]

### **Anaphora in a Wider Context: Tracking Discourse Referents**

Harry Sauers, 9 p. [<http://www.engr.uconn.edu/~harry/NLP/anaphor.html>]

### **Anaphora Resolution: Recent developments, future Directions**

Ruslan Mitkov, Univ. of Wolverhampton, 8 p.  
[[http://www.wlv.ac.uk/~le1825/coling98\\_tutorial/long.htm](http://www.wlv.ac.uk/~le1825/coling98_tutorial/long.htm)]

### **Anaphoric Dependencies in Ellipsis**

Andrew Kehler, Stuart Shieber, 16 p. [ellipsis-prag-squib.ps]

### **Applying Genetic Algorithms to Pronoun Resolution**

Donna K. Byron and James F. Allen, 1 p. [aai99.ps]

### **A Preliminary Model of Centering in Dialog**

D. Byron and A. Stent, 3 p. [98.Byron-Stent.ACLCentering.ps]

### **A statistical Approach to Anaphora Resolution**

Niyu Ge, John Hale & Eugene Charniak, 10 p. [jtwWVLC98.ps]

**A Tableau calculus for Pronoun Resolution**

Christof Monz &amp; Maarten de Rijke, 16 p. [tableaux99.ps]

**Automatic Resolution of Anaphora in English**

Michel Denber, Eastman Kodak, 68 p. [denber.ps]

**Centering: A FrameWork for Modelling the Local Coherence of Discourse**

Barbara J. Grosz, Aravind K. Joshi, Scott Weinstein, 29 p. [centering.ps]

**Centering, Anaphora Resolution, and Discourse Structure**

Marilyn A. Walker, 35 p. [cent-alg13.ps]

**Centering, Global Focus, and Right Dislocation**

Barbara Grosz, Yael Ziv, 14 p. [GrossZiv96.ps]

**Centering in Naturally-Occurring Discourse: An Overview**

Marilyn A. Walker, Aravind K. Joshi, Ellen F. Prince, 26 p. [intro6.ps]

**Centering in-the-Large: Computing Referential Discourse Segments**

Breck Baldwin, 13 p. [coGNIAC.ps]

**CogNIAC : a High Precision Pronoun Référence Engine**

Udo Hahn &amp; Michael Strube, 8 p. [acl97-text.ps]

**Common Topics and coherent Situations: Interpreting Ellipsis in the context of Discourse Inference**

Andrew Kehler, 8 p. [acl94.ps]

**Control issues in Anaphor Resolution**David M. Carter, SRI Int. Cambridge Research Center, 14 p.  
[http://www.cam/sri.com/tr/crc016/paper/paper.html]**Cooperation between Pronoun and reference Resolution for Unrestricted Texts**

Andrei Popescu-Belis &amp; Isabelle Robba, 6 p. [andrei-acl-97.ps]

**Coreference as the Foundations for Link Analysis over Free Text Databases**

Breck Baldwin, Amit Bagga, 6 p. [paper.ps]

**Coreference Resolution in a Multilingual Information Extraction System**

Salima Azzam, Kevin Humphreys &amp; Robert Gaizauskas [azzam-granada-98.ps]

**Cue Phrase Classification Using Machine Learning**

Diane J. Litman, 42 p. [litman96a.ps]

**Current Theories of Centering for Pronoun Interpretation: A Critical Evaluation**

Andrew Kehler, 8 p. [centering-squib.ps]

**Des homographes aux références pronominales**Didier Baltazart et Philippe Trigano, 15 p.  
[http://www.hds.utc.fr/~ptrigano/baltazart.htm]

**Dynamic deduction for natural Language**

Susanna Kuschert, 14 p. [ded-long.ps]

**Évaluation numérique de la résolution de la référence - Critiques et propositions**

A. Popescu-Belis, Limsi, 30 p. [apb-tal.ps]

**Evaluation-Driven Design of a Robust Référence Resolution System**

A. Popescu-Belis, 21 p. [envoi de l'auteur]

**Evaluation of Coreference Rules on Complex Narrative Texts**

A. Popescu-Belis, Isabelle Robba, 9 p. [daarc2.ps]

**Event Coreference for Information Extraction**

Kevin Humphreys, Robert Gaizauskas &amp; Saliha Azzam, 7 p. [kevin-acl-97.ps]

**Extending a Simple Coreference Algorithm with a Focusing Mechanism**

Saliha Azzam, Kevin Humphreys, Robert Guizauskas, 13 p. [daarc98.ps]

**Factors in anaphora resolution: they are not the only things that matter**

Ruslan Mitkov, 7 p. [acl97reg.doc]

**Focusing for Pronoun Resolution in Discourse: An Implementation**

Varol Akman and Ebru Ersan, 16 p. [focus.ps]

**Functional Centering**

Michael Strube &amp; Udo Hahn, 8 p. [acl96.ps]

**Getting Useful Gender Statistics from English Text**

John Hale, 8 p. [jth\_thesis.ps]

**How Corpora with Annotated Coreference Links Improve Reference Resolution**

A. Popescu-Belis, Limsi, 5 p. [popescu-belis.ps]

**How To Fit Generalized Quantifiers Into Terminologica Logics**

J. Joachim Quantz, 5 p. [ecai92.ps]

**Incremental Anaphora Resolution in a Chart-based Semantics Construction Framework using l-DRT**

Ingrid Fisher, Bernd Geistert, Günther Görz, 10 p. [daarc96.ps]

**Incremental Centering and Center Ambiguity**

Udo Hahn &amp; Michael Strube, 6 p. [cogsci96-center.ps]

**Indefinite associative Anaphora in French**

Michel Cosse, Paris 7, 3 p.

[http://www.comp.lancs.ac.uk/computing/users/spb/F.html]

**Intrasentential Constraints on Intersentential Anaphora in Centering Theory**

Andrew Kehler, 7 p. [command.ps]

**Long Distance Pronominalisation and Global Focus**

Janet Hitzeman and Massimo Poesio, 7 p. [ACL98.ps]

**Multilingual robust anaphora resolution**

Ruslan Mitkov, Lamia Belguith, Malgorzata Stys, 10 p. [emnlp-3.doc]

**Never Look Back : An Alternative to Centering**

Michael Strube, 7 p. [colacl98.ps]

**Noun Phrase Coreference as Clustering**

Claire Cardie and Kiri Wagstaff, 8 p. [emnlp-99.ps]

**On the proper treatment of context in NL**

D.J.N. Van Eijck, 16 p. [INS-R0018.ps]

**Probabilistic Coreference in Information Extraction**

Andrew Kehler, 11 p. [emnlp97.ps]

**Processing Complex Sentences in the Centering Framework**

Michael Strube, 3 p. [acl96-student.ps]

**Pronoun resolution made simple: a robust, knowledge-poor approach in action**

Ruslan Mitkov &amp; Lamia Belguith, 10 p. [Actes TALN 1999, Carghèse]

**Pronoun Resolution without Pronouns: Same Consequences of Memory Based Text Processing**Gail McKoon, Richard J. Gerrig, and Steven B. Greene, 18 p.  
[<http://www.psych.nwu.edu/folks/mckoon/ruppaper.html>]**Quantification et anaphore : entité anaphorique complexe (méronymique, processuelle, situationnelle)**

Christiane Panissod, 6 p. [Actes TALN 1999, Carghèse]

**Quantitative Evaluation of Coreference Algorithms in an Information Extraction System**

Robert Gaizauskas &amp; Kevin Humphreys, 17 p. [kevin\_daarc\_96.ps]

**Recognizing Referential Links: An Information Extraction Perspective**

Megumi Kameyama, 8 p. [megumi-acl-97.ps]

**Reference Resolution beyond Conference: a Conceptual frame and its Application**

A. Popescu-Belis, Isabelle Robba, Gérard Sabah, 7 p. [coling-acl98.ps]

**Résolution de l'anaphore pronominale et de la référence**A. Popescu-Belis, Limsi, 2 p.  
[<http://www.limsi.fr/RS97FF/CHM97FF/LC97FF/LC1.html>]**Resolving Anaphors in Embedded Sentences**

Saliha Azzam, 7 p. [coref-acl96.ps]

**Resolving Demonstrative Anaphora in the TRAINS93 Corpus**

Donna K. Byron and James F. Allen, 3 p. [daarc.ps]

**Resolving Pronominal Reference to Abstract Entities**

Donna K. Byron, 68 p. [tr714.ps]

**Resolving Pronoun References**

Jerry R. Hobbs, 1977, 13p. [papier]

**Resolving Pronoun References**

David M. Höppner, 3p. [rpr.ps]

**Robust Pronoun resolution with limited knowledge**

Ruslan Mitkov, 7 p. [coling98.doc]

**Robust reference resolution with limited knowledge: high precision genre-specific approach for English and Polish**

Ruslan Mitkov, Malgorzata Stys, 9 p. [ranlp97.doc]

**Rule-based reference resolution for unrestricted text using part-of-speech tagging and noun phrase parsing.**

Sandra Williams, Mark Harvey and Keith Preston, BT Labs.

[<http://www.mri.mq.edu.au/~swilliam/publications/daarc1.html>]

**Situated Processing of Pronominal Anaphora**

Erkan Tin & Varol Akman, [tin-Akman.ps]

**Supporting anaphor resolution in dialogues with a corpus-based probabilistic model**

Marco Rocha, 8 p. [marco-acl-97.ps]

**Taking time to Structure Discourse: Pronoun Generation Beyond Accessibility**

Kathleen F. McCoy, Michael Strube, 6 p. [cogsci99.ps]

**The effect of Establishing Coherence in Ellipsis and anaphora Resolution**

Andrew Kehler, 8 p. [acl93.ps]

**The effect of thematic roles on pronoun use and frequency of reference continuation**

Jennifer E. Arnold, 28 p. [throles.ps]

**Toward a Model of the Interaction of Centering with Global Discourse Structure**

Marilyn A. Walker, 23 p. [cent-cache.ps]

**Un atelier de traitement de la référence**

A. Popescu-Belis, Isabelle Robba, 3 p. [Actes TALN 1999, Carghèse]

**Underspecification in Anaphoric Reference**

Massimo Poesio, Uwe Reyle, 12 p. [tilburg.ps]

**Using Discourse Analysis and Automatic Text Generation to Study Discourse Cue Usage**

Megan Moser, Johanna D. Moore, 7 p. [aaai-spr-symp95.ps]