SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

Impact des genres sur la composition des chaînes de référence : le cas des faits divers

Schnedecker, Catherine & Longo, Laurence

LiLPa (Linguistique, Langue, Parole), UR 1339, Université de Strasbourg
22 avenue René Descartes, 67084 Strasbourg Cedex
{cschnede, longo}@unistra.fr

1 Introduction

Cette étude porte sur la manière dont les expressions coréférentielles d'un texte, composant ce qu'on appelle les chaînes de référence (désormais CR, cf. Corblin, 1995, Charolles, 1988, Schnedecker, 1997) sont conditionnées par le genre de leur texte d'accueil. Des études antérieures ont, en effet, démontré l'étroite corrélation entre genre discursif et CR[1], soit en s'appuyant sur l'étude d'un genre précis, comme le portrait journalistique (Jenkins, 2002 et à sa suite Schnedecker, 2005), soit en comparant le comportement des CR dans des genres diversifiés (presse, romans, textes administratifs (Longo & Todirascu, 2010 ; Longo (à paraître)) ; nouvelles *vs* portraits journalistiques, Baumer, 2011[2]), quitte à limiter l'analyse à certaines catégories d'anaphores (hyperonymiques (Condamines, 2005), pronominales (Tutin *et al*., 2000)).

Il s'agit d'étayer cette thèse en prenant comme genre d'appui le fait divers (désormais FD). Pour étudier la composition des chaînes de référence, les FD offrent, en effet, plus d'un avantage. D'une part, il s'agit de textes courts, ce qui permet de rendre compte de l'intégralité des CR. D'autre part, ils permettent d'appréhender un genre de presse dont les caractéristiques thématiques et structurelles sont désormais bien établies (cf. Barthes, 1964 ; Auclair, 1970 ; Dubied, 2000), ce qui facilite théoriquement l'établissement de corrélations. Enfin, la concentration des phénomènes sur des formes brèves est de nature, peut-être, à faire ressortir des problèmes passés jusque là inaperçus.

Pour ce faire, nous avons établi un corpus de 46 textes (9838 mots/59097 caractères), indexés sous la rubrique des *Faits divers* du *Républicain Lorrain*, quotidien régional de la région lorraine (édition de Metz) et collectés durant l'été 2011 sur le site du journal. Notre objectif est de déterminer quelles sont les expressions référentielles mobilisées dans ce genre, quel est leur degré de solidarité, comment elles sont initiées et reliées les unes aux autres ; bref, en quoi elles répondent à un mode de composition précis.

Dans cette optique, nous commencerons par rappeler quelques résultats d'études contrastives antérieures qui ont démontré la corrélation entre genre et CR (cf. 2.), d'une part, ainsi que, d'autre part, les caractéristiques notoires des FD, susceptibles de peser sur la composition des CR (cf. 3.). Dans un troisième temps, nous procéderons à une étude quantitative du corpus construit pour cette étude, qui fera ressortir le matériau linguistique des CR des FD et, le cas échéant, permettra de dégager quelques modèles de chaînes (cf. 4.). Dans un quatrième et dernier temps (cf. 5.), nous procéderons à l'étude qualitative de notre corpus, en insistant particulièrement sur quelques spécificités référentielles des FD de nature à questionner les approches au long cours de la coréférence.

2 Impact du genre sur la composition des CR

Diverses études menées sur les portraits journalistiques (Jenkins, 2002 ; Schnedecker, 2005) et les textes expositifs (Goutsos, 1997) ont montré l'influence du genre textuel sur le matériau linguistique contenu dans les chaînes de référence. Ainsi, les genres textuels conditionneraient-ils l'introduction du référent (première mention) et les formes privilégiées de sa reprise. Pour vérifier la détermination des genres sur

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

Article available at http://www.shs-conferences.org or http://dx.doi.org/10.1051/shsconf/20120100061

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

les CR nous avons mené une étude comparative des CR sur un corpus mêlant cinq genres textuels (Longo & Todirascu, 2009)[3] : des articles politiques du *Monde* (2004), des éditoriaux du *Monde Diplomatique* (1980-1988), le début du roman *Les trois Mousquetaires* de Dumas (1884), des textes de lois européennes issues de l'*Acquis Communautaire* (Steinberger *et al.*, 2006), des rapports publics de *La Documentation Française* (2001).

En suivant la typologie de Schnedecker (2005), nous avons comparé les CR en retenant plusieurs critères :

- la longueur moyenne des CR : calculée sur le nombre moyen de maillons contenus dans une CR suivant le genre textuel ;
- la distance moyenne entre les antécédents : déterminée suivant le nombre de phrases séparant chacun des maillons d'une même CR ;
- la catégorie grammaticale privilégiée du premier maillon d'une CR suivant le genre textuel ;
- la catégorie grammaticale la plus fréquente des maillons des CR suivant le genre textuel ;
- la correspondance entre le thème phrastique (entendu ici comme le sujet de la phrase) et le premier maillon des CR suivant le genre textuel.

L'étude comparative des CR a révélé que la longueur des CR était de 3 maillons en moyenne pour les lois européennes (1) alors qu'elle était trois fois plus élevée en moyenne pour le roman (2).

(1) Au cas où la parité de la monnaie d'**un État membre** par rapport à l'unité de compte définie ci-dessus serait réduite, le montant de la quote-part de capital versé par **cet État** sera ajusté, proportionnellement à la modification intervenue dans la parité, moyennant un versement complémentaire effectué par **cet État** en faveur de l'Agence et limité au montant des avoirs effectivement détenus dans la monnaie de **cet État membre**. Le versement devra être effectué dans un délai de deux mois. *(Acquis Communautaire)*

(2) **M. de Tréville** était pour le moment de fort méchante humeur ; néanmoins **il** salua poliment le jeune homme, qui s'inclina jusqu'à terre, et **il** sourit en recevant son compliment, dont l'accent béarnais **lui** rappela à la fois **sa** jeunesse et **son** pays, double souvenir qui fait sourire l'homme à tous les âges. Mais, **se** rapprochant presque aussitôt de l'antichambre et faisant à d'Artagnan un signe de la main, comme pour lui demander la permission d'en finir avec les autres avant de commencer avec lui, **il** appela trois fois, en grossissant la voix à chaque fois, de sorte qu'**il** parcourut tous les tons intervallaires entre l'accent impératif et l'accent irrité. (Dumas, *Les Trois Mousquetaires*)

Concernant la distance moyenne entre les divers maillons, nous avons remarqué des disparités : pour certains genres comme le roman, les maillons d'une même chaîne sont situés dans la même phrase ou dans la phrase suivante (un nom propre suivi des pronoms réfléchis et personnels pour la plupart). Pour les rapports publics, les maillons sont plutôt éloignés les uns des autres (les maillons sont espacés de plus d'une phrase en moyenne) et on constate une forte densité de syntagmes nominaux pas repris par la suite.

Aussi, de même que Schnedecker (2005), nous avons constaté que la catégorie des noms propres (Npr) était privilégiée dans les premiers maillons des CR des textes journalistiques, alors qu'il s'agissait plutôt des syntagmes nominaux indéfinis pour les lois européennes, comme par exemple « *un Etat Membre », « une décision », « une loi »*.

Les expressions référentielles contenues dans les CR sont plutôt des pronoms et des possessifs pour le roman (respectivement 35.9% et 28.2%) et des syntagmes nominaux définis (33.4%) pour les éditoriaux (« *Le Centre de documentation d'études juridiques, économiques et sociales », « L'attentat d'Hébron », « Le comité d'orientation nationale »*).

Enfin, le premier maillon d'une CR, coïncide peu avec le thème phrastique (40%) pour les rapports publics mais plus souvent (80%) pour les textes journalistiques.

Bref, il ressort de cette étude que la composition des CR (nature du 1er maillon, distance inter-maillonnaire, catégorie grammaticale des maillons) est fortement tributaire du genre d'accueil des CR, même si, nous en avons bien conscience, il s'agit, pour l'instant de tendances qui demanderaient à être

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

confirmées à plus grande échelle. Ainsi, au terme de cette étude, nous avons dégagé des propriétés des CR issues d'un genre textuel particulier.

3 Les faits divers : quelques caractéristiques

3.1 Caractéristiques thématiques

Suivant la définition de Fragnon (2007 : 254), les FD « narrent des événements sociaux, eux-mêmes à la marge de l'espace social », dont ont été soulignés les écarts par rapport au déroulement quotidien des choses : ainsi Auclair a-t-il parlé de « rupture de l'univers réglé où l'homme trouve sa sécurité », Barthes (1981) de « déviations causales » : « en vertu de certains stéréotypes, on attend une cause, et c'est une autre qui apparaît ». Il est question aussi, dans la littérature, de dérogation à une norme ou d'extraordinaire[4]. (3) illustre la déviation causale dont parle Barthes : un moniteur d'auto-école est censé connaître la règlementation en matière de circulation routière et, théoriquement, devrait éviter les excès de vitesse. Or, non seulement il commet une infraction mais celle-ci est, qui plus est, d'importance :

(3) **Un automobiliste de 28 ans**, exerçant la profession de moniteur d'auto-école, a été contrôlé samedi à plus de 200 km/h au lieu des 130 km/h autorisés sur l'autoroute A16 à la hauteur de Pont-de-Metz (Somme), près d'Amiens. **Le conducteur**, contrôlé à 214 km/h, a eu **son** permis retiré et la voiture est partie à la fourrière.

3.2 Caractéristiques structurelles

3.2.1 Brièveté

Le FD se caractérise par sa brièveté[5]. Les textes de notre corpus comptent en moyenne 208 mots[6]. Structurellement,

> « <Ces> brèves sont composées d'un titre, d'une indication géographique et du texte proprement dit. […] Par souci de concision, les brèves répondent aux questions fondamentales (qui ? quoi ? quand ? où ? comment ?) : lieux et horaires de l'incident, sujet de l'acte délictueux, victimes, et, actions de l'anti-sujet policier. » (Fragnon, art. cit. : 256),

C'est ce que montre (4) où la mention du lieu figure généralement à gauche du titre ; elle est reprise dans le texte sous la forme d'une expansion du SN (une buraliste de Nancy) ou d'un circonstant localisateur.

(4) **Nancy.** Buraliste braquée
(p1) **UNE BURALISTE DE NANCY** a été agressée, hier à 6h, alors qu'ELLE ouvrait son commerce. (p2) << Deux hommes encagoulés et gantés ont surgi derrière ELLE alors qu'ELLE SE dirigeait vers la réserve de SON bar-tabac. (p3) Chacun portait une arme de poing. (p4) Ils ont bousculé LA GERANTE QUI est alors tombée à terre, avant de demander le coffre. (p5)Terrorisée, LA VICTIME a expliqué qu'il n'y en avait pas. (p6) Ils se sont fait remettre 4 000 € en chèques et espèces. (p7) Les deux agresseurs ont ensuite pris la fuite dans une direction inconnue, laissant LA BURALISTE sous le choc. (p8) ELLE a été soignée à l'hôpital central de Nancy pour une estafilade à l'épaule gauche, due à SA chute. >> (p9) L'affaire est confiée à **la sûreté départementale de Meurthe-et-Moselle.**
Les gendarmes et les agents de l'UTR de Bitche ont coupé la circulation sur la RD 35, un axe fréquenté qui relie Bitche à Deux-Ponts (Allemagne). >>

L'information qui constitue le FD apparaît dans les titres et/ou sous-titres, eux-mêmes repris et expansés dans la première phrase du texte ; elle est développée par le biais d'une narration constituant la partie

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

centrale du FD (entre chevrons dans l'exemple (4))[7]. La partie narrative est, le cas échéant, délimitée par des SN à caractère métadiscursifs comme *les faits*, p.e. dans (5) :

(5) L'auteur d'une fusillade survenue dans le cadre d'une opération de représailles, à la mi-juin, à Frouard, s'est constitué prisonnier cette semaine auprès de la gendarmerie. Il a été écroué en dépit de ses dénégations. Il reconnaissait avoir fait feu avec son fusil, mais pas avoir voulu tuer. Il est poursuivi néanmoins pour tentative d'homicide volontaire. Un mandat de recherche était lancé contre lui. **Les faits** s'étaient déroulés en deux temps, route de Liverdun. *Un passant avait été pris à partie et bastonné à coup de chaîne par deux frères turbulents, vivant dans une maison de MMH. Un peu plus tard, la victime avait envoyé des amis donner une leçon à ses agresseurs*. […]

3.2.2 Les individus dans les FD

L'une des particularités du FD tient à ce qu'ils racontent des événements touchant « des gens de tous les jours qui, à un moment de leur vie, ont vécu une situation exceptionnelle, souvent dramatique, et ont accompli un geste qui a sauvé des vies. » (Dubied, 2004 : 53). Dubied (*op. cit*. : 227) évoque également « des quidams, […] saisis dans leur vie de tous les jours ; des « Monsieur-tout-le-monde » (*ibid*.). De là viendrait que « les termes employés pour caractériser le sujet sont les plus neutres possibles » (Fragnon, art. cit. : 257). Nous aurons l'occasion d'en évaluer les répercussions sur les CR.

On peut d'ores et déjà souligner que cela transparaît dans la titraille. En effet, 67% des titres font référence à un individu humain dont les modes de désignation sont assez systématiques. Nous avons ainsi repéré 2 « patrons », schématisés sous **i.** et **ii.**, classés en fonction de la structure dominante : les SN expansés par un participe passé employé comme adjectif dominent dans 39% des cas (nous y incluons les cas où le titre se limite au participe passé) ; ils sont à égalité (20%) avec des phrases « complètes » instanciant le protagoniste principal, le plus souvent en fonction grammaticale de sujet. On observe, à part (cf. **iii.**), des SN, parfois déverbaux, dont les arguments sont laissés vides :

i. [Dét + N]+ppé : 39%
Accident mortel : conducteur poursuivi
Un homme tué par balles sur une aire de l'A31
Marseille. Garçonnet tué par un chauffard en fuite
Uckange. Poignardé à son domicile

ii. Phrase complète : 20%
Corse : un détenu violent s'évade d'un hôpital
Une octogénaire périt dans un incendie

iii. SN [déverbal] : 20%
Nancy. Mort suspecte dans un incendie
Imling. Le sauvetage de papy courage

3.3 Caractéristiques fonctionnelles

Certaines des visées du FD[8] ont un lien direct avec son « personnel » comme le nomme Dubied (1999) à la suite de Hamon. En effet, qu'il s'agisse de dénoncer les dysfonctionnements sociaux et d'en stigmatiser des groupes (Fragnon : 262-265) ou, au contraire, de décrire la condition humaine à des fins cathartiques ou d' « agrégation tribale » (expression de Maffesoli, citée *in* Dubied : 73-74), l'individu y est montré dans ses aspects les plus « stéréotypés », voire confiné dans des « rôles thématiques »[9] : le petit voyou, le criminel par accident, la victime âgée, etc. Bien entendu, cela aura des conséquences sur la dénomination des personnages et leur expression référentielle.

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

4 Les expressions référentielles dans les faits divers : les grandes tendances

4.1 Quelques chiffres

Sur un corpus de 9838 mots, nous avons recensé un total de 905 expressions référentielles référant à des personnes *i.e*. des entités individuelles (le motard sarrois en (6)) ou collectives (en italiques ou soulignées dans le texte), ce qui équivaut à une expression référentielle tous les 11 mots. Pour des textes d'une longueur moyenne de 208 mots, la densité est donc relativement importante. Cela signifie, en effet, que 19% environ des unités du texte sont des expressions référentielles renvoyant à des référents humains. Le texte (6) montre, en effet, qu'un référent au moins est instancié toutes les lignes :

(6) Schweyen **Un motard sarrois** tué dans une collision
Alors qu'**il** venait tout juste de passer la frontière avec *un groupe d'amis* et **0 se** dirigeait vers Bitche, **un motard sarrois** a trouvé la mort hier, à 10h, sur la RD 35 à hauteur de Schweyen.
Le pilote menait *le groupe de deux-roues*. **Il** aurait tenté de dépasser un camion dans une courbe à droite, pourtant marquée par une ligne blanche continue. Un autre poids lourd arrivait en face. La collision était inévitable. **Le pilote** et **sa** moto ont été projetés sur le bas-côté. **Joachim Platt**, 52 ans, est mort sur le coup.
Des pompiers français et allemands sont intervenus sur les lieux de l'accident. ***Ces derniers*** ont pris en charge *DEUX CHAUFFEURS ROUTIERS D'OUTRE-RHIN*, légèrement blessés et surtout choqués.
LES GENDARMES ET LES AGENTS DE L'UTR DE BITCHE ont coupé la circulation sur la RD 35, un axe fréquenté qui relie Bitche à Deux-Ponts (Allemagne).

Les 46 textes du corpus totalisent 89 chaînes de référence (désormais CR). Sont considérées comme telles les suites d'au moins 3 expressions coréférentielles[10] (p.e. le personnage du motard sarrois dans (6)). Nous pensons, en effet, (cf. Corblin, 1995 ; Schnedecker, 1997) qu'en deçà de ce nombre, la notion de CR serait inutile puisque celles de « coréférence » ou d'« anaphore » suffiraient alors amplement à rendre compte de la relation entre les « maillons ».

Plus de 60% de CR comportent un nombre situé entre 3 et 6 maillons, ce qui fait que, du point de vue de la longueur moyenne, elles sont conformes à ce qui a été constaté pour le genre journalistique (cf. *supra* 2).

Les CR se répartissent comme suit (cf. Tableau 1) dans les différents textes : la majorité des textes – presque 40% – sont centrés sur un personnage et un tiers sur 2. Par contraste, les textes restants, centrés sur plus de 2 personnages, ne constituent qu'un quart du corpus.

Nombre de textes	Nombre de chaînes de référence / texte	Pourcentage de textes concernés
18	1	39%
16	2	35%
9	3	19.5%
3	4	6.5%
46	--	100%

Tableau 1 – Répartition des chaînes de référence selon les textes

4.2 Catégories grammaticales des expressions référentielles

Les expressions référentielles se répartissent en différentes catégories grammaticales, que synthétise le Tableau 2 :

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

Catégories d'expressions référentielles	Npr	SNØ	SN indéfinis		SN poss.	SN définis		SN dém.	Pronoms					Dét. poss.
			nus	expansés		nus	expansés		dém.[11]	réfléchis	Ø	relatifs	Pers.	
Nombre d'occurrences	29	8	71	75	25	175	114	11	5	63	14	29	173	104
Pourcentages	3.2%	1%	8%	9%	2.5%	20%	12.5%	1.3%	0.5%	7%	1.5%	3%	19%	11.5%
		18%			35%			1.8%	11.5%				30.5%	

Tableau 2 – Répartition des catégories grammaticales des expressions référentielles

4.2.1 Un équilibre inattendu entre les expressions dites de haute et moyenne accessibilité référentielle

Suivant la terminologie d'Ariel (1990), les expressions dites de haute accessibilité référentielle comme le pronom personnel, auquel nous ajoutons le déterminant possessif du fait de son équivalence avec le pronom personnel, constituent 30% des expressions référentielles du corpus, ce qui ajouté aux expressions pronominales liées syntaxiquement (anaphore zéro, pronoms réfléchis et relatifs) aboutit à un pourcentage de 42%. De ce fait, les pronoms dominent nettement le corpus. Mais ils sont suivis de très près par les expressions dites de moyenne accessibilité référentielle dont le pourcentage s'élève à 36.3% et qui se répartissent entre les SN définis (35% dont 20% de SN « nus », 12.5% de SN expansés et 2.5% de possessifs) auxquels s'ajoutent 1.3% de démonstratifs.

A priori, ces chiffres démentent les pourcentages d'Ariel (*op. cit*.) dont le corpus fait apparaître 70% de pronoms personnels, ce qui les met en tête des emplois en tant que marques du topique, de la continuité référentielle. Cela étant, les pourcentages que nous obtenons s'expliquent. En effet, les textes de notre corpus sont pluri-référentiels, *i.e*. instancient au moins deux référents humains. Or, le nombre de référents – et, partant, les modalités de la reprise anaphorique – augmente les risques d'ambiguïtés pronominales. C'est ce que montre (8) – version remaniée de (7) – où la première apparition du pronom serait susceptible de brouiller l'interprétation coréférentielle[12] :

(7) Football : **le président d'un club** agressé par un ex-joueur
Les dirigeants du Sporting-Club de Moulins-lès-Metz pensaient en avoir terminé avec les problèmes extra-sportifs. Ils les ont violemment rattrapés le week-end dernier. A l'issue de l'entraînement de l'équipe seniors, **le président du club de football** a été agressé par *un ancien joueur*, exclu après avoir été suspendu par la fédération pour des menaces sur un arbitre. *L'auteur* a profité de l'obscurité du parking pour **le** frapper au visage, avec une batte de base-ball. […]

(8) A l'issue de l'entraînement de l'équipe seniors, **le président du club de football** a été agressé par *un ancien joueur*, exclu après avoir été suspendu par la fédération pour des menaces sur un arbitre. ***Il*** a profité de l'obscurité du parking pour **le** frapper au visage, avec une batte de base-ball.

De ce fait, le fort pourcentage d'expression de moyenne accessibilité référentielle pourrait, en première instance, prévenir les risques d'ambiguïtés[13].

4.2.2 Les expressions de faible accessibilité référentielle : domination des SN indéfinis

Les expressions de faible accessibilité référentielle s'élèvent à 21%. 3.2% d'entre elles sont constituées de noms propres[14]. Ceux-ci ne figurent que dans 17 (soit 37%) des 46 textes de notre corpus et rarement en première mention (dans 6 des 87 CR, soit 6.9% des cas). C'est le cas de (6) vu plus haut où le Npr apparaît dans le 7ème maillon[15]. En revanche, (9) illustre un cas de Npr en position de 1er maillon :

(9) Waltembourg Percuté et tué en marchant sur la RN4
Frédéric Jambois, 38 ans, d'Imling, est décédé, hier matin, sur la RN4. Percuté par une voiture, il marchait avec des amis sur la voie rapide. […]

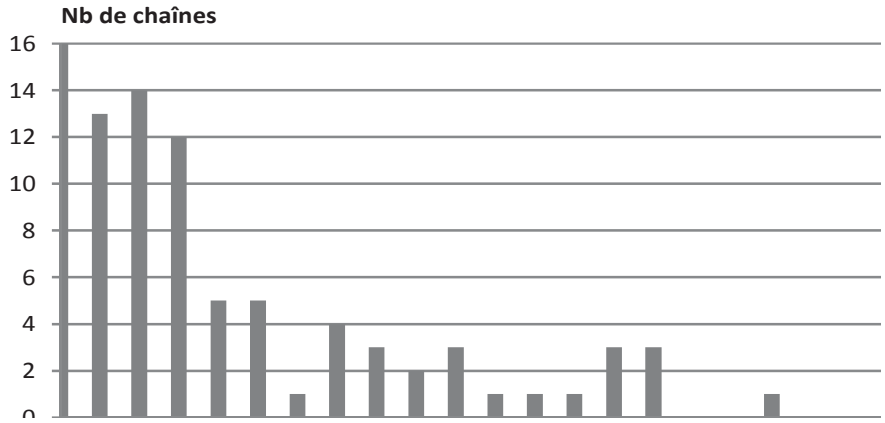
La majorité des expressions référentielles de faible accessibilité référentielle est constituée par les SN indéfinis dont le pourcentage s'élève à 18% et se répartit de manière équilibrée entre les SN « nus » (8%)

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

et expansés (9%)[16]. A quoi s'ajoutent 1% de SN sans déterminants qui figurent exclusivement dans les titres (cf. *supra* 3.2.2.).

4.3 Des patrons de CR ?

Il est difficile de dégager des patrons de chaînes dans les FD, tant est grande la disparité des CR, aussi bien du point de vue des expressions référentielles que de leur longueur. Néanmoins, il ressort de nos analyses que les CR des FD sont plutôt brèves, comptant en moyenne 3.42 maillons. Plus d'un tiers des CR en comptent 3 à 4 (cf. Graphique 1) ; un deuxième tiers (29%) 5 et 6. Cette longueur tient évidemment à la brièveté inhérente au genre du FD lui-même et se rapproche de celle des textes expositifs rappelée plus haut.



Graphique 1 – Longueur des CR dans les FD

Deuxièmement, le premier maillon est constitué à 45% de SN indéfinis, à 27% de SN définis et, pour les plus faibles pourcentages, à 8% de pronoms personnels ou de déterminants possessifs. Dans 84% des CR, le 1er maillon du titre coïncide avec la 1ère chaîne du texte, comme dans (10) et (11) :

(10) **Il** tente de **s'**immoler dans un terrain vague
Un jeune homme âgé de 29 ans, demeurant à Blainville-sur-l'Eau, en Meurthe-et-Moselle, a tenté de mettre fin à **ses** jours, samedi dernier, en **s'**immolant par le feu.

(11) Mirecourt. **Gendarme poignardé** : **sa** santé s'améliore
Les jours **du commandant de la communauté de brigade de Mirecourt** (Vosges) ne semblent plus en danger. Poignardé à la gorge par un individu, **le militaire** a été opéré, vendredi soir, au CHU de Nancy.

Troisièmement, en prenant en compte uniquement les 3 premiers maillons – donc le seuil-plancher des chaînes dans notre conception –, on a pu dégager 3 patrons correspondant à des débuts de CR configurées comme suit :

iv. SNdéf pro pro : 9%

(12) **Le conducteur de la voiture**, un homme de 83 ans, originaire d'Erstroff, est indemne. **Il** est rentré chez **lui** après l'accident.

v. SNindéf. pro pro : 9%

(13) **un blessé** et des dégâts
Vendredi après-midi, alors qu'**il** rentrait chez **lui** à pied, un habitant de Folschviller a été surpris par l'orage de grêle qui s'est soudainement abattu sur la région de Saint-Avold […].

vi. SNdéf. SNdéf. SNdéf. : 7%

(14) Alertés par des riverains, à 16h30, **les sapeurs-pompiers** ont déployé la grande échelle, ainsi que le fourgon pompe-tonne de Forbach. Mais à l'arrivée **des secours**, la locataire, Marie-Louise Pezzotta, avait déjà succombé, suite à l'inhalation de monoxyde de carbone. Âgée de 85 ans, la victime vivait seule, selon les premiers éléments de l'enquête. Selon la police de Forbach, si la cause exacte du sinistre reste encore à

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

déterminer, la thèse de l'accident est privilégiée. « Il est difficile de dire dans quelle pièce l'incendie s'est déclaré », indique l'adjudant-chef **des sapeurs-pompiers** Laurent Wetzl, qui a dirigé l'intervention.

Notons enfin que 18% des CR sont dépourvues de pronoms personnels. Nous y reviendrons (cf. *infra* 5.2)

4.4 Premier bilan

Cette analyse quantitative aura fait ressortir la spécificité des CR des FD à quatre points de vue :

- la pluri-référentialité des textes : 61% d'entre eux comprennent au moins 2 référents, ce qui les distingue d'autres genres de la presse comme, par exemple, les portraits ;
- le type d'expression référentielle dominante : ce sont les SN, tous déterminants confondus, qui dominent (55% des expressions référentielles) alors que les pronoms n'excèdent pas 42% ;
- le grand « perdant » des FD est le Npr qui n'atteint que 3.2% des cas et dans des emplois qui diffèrent de ce qui a été observé dans les romans ou les portraits ;
- la composition : les CR sont plutôt brèves (3.42 maillons en moyenne), plutôt hétérogènes, tant au plan des catégories grammaticales instanciées que des têtes lexicales des SN et assez difficiles à modéliser, puisque seuls 3 patrons ont émergé.

C'est toutefois l'analyse qualitative qui, comme on va le voir, est la plus riche d'enseignements.

5 Les expressions référentielles dans les FD : analyse qualitative

5.1 Diversité des noms d'humains

Au plan qualitatif, la première caractéristique notoire des CR de FD tient à la diversité des têtes lexicales des SN dont le Tableau 3 donne un aperçu[17].

Qui plus est, la diversité nominale est à la fois récurrente dans les FD du corpus, distribuée sur l'ensemble des référents instanciés dans un texte comme l'illustre (15)[18], et inhérente aux SN anaphoriques dont la tête lexicale varie d'une reprise à l'autre, comme le montrent, dans (15) ci-dessous, la CR du personnage principal de l'octogénaire et (4) (*supra*, 3.2).

(15) Forbach. *UNE OCTOGENAIRE* périt dans un incendie
Hier après-midi, un incendie s'est déclaré au quartier du Creutzberg à Forbach. Le feu a pris au premier étage d'une maison mitoyenne, au 6, avenue de Stiring-Wendel. Alertés par DES RIVERAINS, à 16h30, **LES SAPEURS-POMPIERS** ont déployé la grande échelle, ainsi que le fourgon pompe-tonne de Forbach. Mais à l'arrivée DES SECOURS, *LA LOCATAIRE*, Marie-Louise Pezzotta, avait déjà succombé, suite à l'inhalation de monoxyde de carbone. Âgée de 85 ans, ***LA VICTIME*** vivait seule, selon les premiers éléments de l'enquête.
Selon LA POLICE DE FORBACH, si la cause exacte du sinistre reste encore à déterminer, la thèse de l'accident est privilégiée.
« Il est difficile de dire dans quelle pièce l'incendie s'est déclaré », indique L'ADJUDANT-CHEF **DES SAPEURS-POMPIERS** LAURENT WETZL, qui a dirigé l'intervention.
En effet, les flammes ont entièrement ravagé l'appartement.
LES SAPEURS-POMPIERS ont longuement inspecté la toiture et les conduites de façon à empêcher tout risque de reprise du feu.

5.2 Des sous-catégories de N en nombre limité

Les types de N sont limités à 4 sous-catégories (voir Tableau 3) :

- les N « généraux » (*homme, femme*), auxquels s'ajoutent les N dits de phase, qui désignent l'individu par ses caractéristiques identitaires fondamentales (sexe et âge) ; ces N constituent un quart des emplois ;

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

© aux auteurs, publié par EDP Sciences, 2012

- les N de fonction/profession dont le poucentage s'élève à 21% mais dont une grande partie (19%), « incontournable » dans les FD, renvoie notamment aux forces de l'ordre, aux sauveteurs ou à la justice (désignée sous le nom de *parquet* (16 occurrences), cf. *supra*, ex. (6), (14)). Ce sont d'ailleurs ces N qui composent les chaînes les plus régulières, fondées sur la réitération de la tête lexicale, et donnent lieu au patron **vi.** ;
- la troisième catégorie, qui comprend 33% des SN, réunit 1) les N indiquant les relations entre le protagoniste principal du FD et ses proches (famille, voisinage mais aussi ses proches géographiquement, saisis dans leur ancrage spatial pourrait-on dire) ; à quoi s'ajoutent 2) les N de gentilés et, enfin, 3) les N saisissant les actants plus ou moins directs du « drame » décrit : acteurs (*agresseur, vengeur*) ou patients (*blessé, victime*) ainsi que les personnages plus extérieurs, simplement spectateurs (*témoin*) ;
- la dernière (1%) rassemble les rares N indiquant l'empathie du locuteur (cf. *supra* 3.2).

N Généraux (16%)	N de Phase (9%)	N Fonction/ Profession (21%)	N Relationnels (32%)			Gentilés (1%)	Axiologique Empathique (1%)
			relations sociales	ancrage spatial	ancrage situationnnel		
Personne (7) Individu (5) Homme (49) Femme (17)	Jeune / vieil homme (7) Personne âgée (4) Enfant (12) Petit garçon (3) Fillette (3) N en *-aire* (9) (*quinquagénaire*) Jeunes (4) Retraité (4)	Gendarmes (19) Gendarmerie (8) Enquêteurs (6) Police (24) Policier (13) Pompiers (24) Buraliste Automobiliste (8)	Mère (5) Fille (10) Fils (3) Compagnon (6) Voisin (18)	Habitant (12) Riverain (4) Occupant (3) Résident (2)	Agresseur (2) Vengeur Blessé (16) Victime (35) Suspect (6) Témoin (6) Détenu (2) ***Secours (23)*** ***Touriste*** Rescapé Conducteur (20) Gérante ***Auteur (4)***[19]	Meusien Mosellan Maizièrois Messin (3)	Bon samaritain Chauffard (3) Malheureux (2)

Tableau 3 – Classement (provisoire) des têtes lexicales des SN dans les FD

Là encore, les fonctions d'agrégation tribale ou d'identification/projection cathartique attribuées aux FD, rappelées plus haut, sont largement servies par le lexique qui sert, pour l'essentiel, à décliner l'identité des protagonistes au moyen des noms (sexe, âge, profession), communs à tous les humains et à les resituer dans un environnement ou dans des situations relativement ordinaires, banals. Autant d'éléments qui ramènent le personnel des FD à « monsieur tout le monde » et créent une certaine proximité avec le lecteur de FD.

5.3 La relative uniformité et objectivité des informations délivrées par les modifieurs

Près de 21% des expressions référentielles du corpus sont expansées, principalement par des compléments du nom (CDN) ou des adjectifs (éventuellement des participes passés employés comme adjectifs)[20], ainsi que le synthétise le Tableau 4. Ces informations concernent généralement le personnage principal du FD.

Type de modifieur	Complément du nom	Adjectif
Nombre	106	58
Pourcentage	56%	31%
Sous-catégories	Localisateur (32%)[21]	Localisation (20%)
	Age (12%)	Participe passé (31%)

Tableau 4 – Répartition et information des modifieurs

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

Les CDN dispensent, dans presque la moitié des cas, des informations relatives à la localisation géographique des référents (16)[22] – et ce, de manière systématique pour les forces de l'ordre, sauveteurs, etc. – ou à leur âge (17) ; quant aux adjectifs ou assimilés (cf. *supra* 3.3), ils informent principalement sur les faits (cf. *supra*, 3.2.2) ou la localisation (18) :

(16) d'un habitant **de Jezainville** / un habitant **de Folschviller** / au parquet **de Metz** / les enquêteurs de la compagnie **de Toul** et de la section de recherches **de Nancy**

(17) Un automobiliste **de 28 ans** / d'une mère **de 19 ans** / son petit-fils **d'un an** / un Mosellan **de 74 ans**

(18) Cette automobiliste **hollandaise** / Un quadragénaire **uckangeois** / un motard **sarrois** / un retraité **mosellan**

Les modifieurs sont assez peu utilisés pour « qualifier » physiquement ou moralement les référents. A cet égard, (19) et (20) constituent les 2 exceptions de notre corpus. Cette neutralité peut paraître paradoxale dans la mesure où les faits narrés prêtent facilement le flanc à l'émotion/indignation, etc. Mais, d'après la teneur des SN, il semble bien que les auteurs de FD sont, sans doute par devoir et par déontologie, astreints à une forme d'objectivité.

(19) Un détenu réputé **très violent**

(20) Un passant avait été pris à partie et bastonné à coup de chaîne par deux frères **turbulents**, vivant dans une maison de MMH.

5.4 Zoom sur …

Nous limiterons l'analyse, faute de place, aux N « relationnels » et aux N généraux d'humains, en nombre conséquent, dans notre corpus.

5.4.1 … les SN relationnels

Les N relationnels sont la sous-classe de N la plus nombreuse de notre corpus (32% des SN). Ils sont remarquables, lexicalement, par leur absence d'autonomie référentielle, que traduit, entre autres, la difficulté ressentie à leur emploi isolé (cf. Herslund, 1996, entre autres) :

(21) ***J'ai rencontré **un frère** (ex. de Herslund, art. cit., 36)

Aux N dénotant les relations sociales, bien repérés dans la littérature, telles que la parentèle (22) (26% des N relationnels) ou le voisinage (23), nous ajouterons d'autres N[23] renvoyant à d'autres formes de relations : spatiales (24) et (25) (14% des N relationnels) ou celles que nous appellerons, faute de mieux et provisoirement, situationnelles (60%) comme *témoin* ou *victime* dans (26) :

(22) **L'ex-compagnon d'une mère de 19 ans** retrouvée morte devant son domicile dans les Deux-Sèvres est suspecté de l'avoir étranglée.

(23) Il était environ 20h45 quand les secours ont été alertés par **des voisins**.

(24) Rapidement, la fumée envahit la cage d'escalier, empêchant **plusieurs habitants** de s'échapper

(25) Le choc a été tel qu'Ali Simsek, l'un des **deux occupants du véhicule**, a été éjecté.

(26) Selon **plusieurs témoins**, **la victime** s'était proposée pour le ramener à son domicile.

Ces dernières sous-catégories de N sont bien dépendantes comme le montre le test appliqué ci-dessus :

(27) *J'ai rencontré **un habitant/un occupant/un témoin/une victime**

Toutefois, à la différence des N de parenté, la dépendance des deux derniers types de N résulte de leur statut d'« agent » (agresseur) ou d'« acteur » (p.e. *conducteur/habitant*[24]), issu de la « transposition » selon Benveniste d'une construction verbale transitive (*agresser/conduire/habiter*) dont les arguments peuvent figurer dans les CDN (cf. *supra*, ex. (16) et (25)) ou se laisser récupérer *via* le contexte.

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

Ces types de N sont suscités par le genre du FD, dont les auteurs collectent les informations, pour une grande part, auprès des forces de l'ordre et au moyen d'enquêtes de voisinage, menées auprès des proches (familiaux ou locaux) des protagonistes. Ils montrent également que l'expression référentielle des FD est totalement focalisée sur le protagoniste et sa situation dramatique, qui servent de point d'ancrage aux anaphores (cf. *infra*).

Au plan théorique, la présence de ce genre de N dans les CR pose un double problème car elle montre que leur matériau a une origine diverse et que, par voie de conséquence, celui-ci a un empan, ou un potentiel d'utilisation sur la chaîne, variable.

Pour ce qui concerne l'origine, ou la source, on peut distinguer entre :

- celles qui sont alimentées par le cotexte propositionnel, comme, p.e. le SN défini *la victime* (cf. 15 *supra*) dont l'interprétation dépend du contenu d'une proposition antérieure (*avait succombé suite à l'inhalation de monoxyde de carbone*) et qui, selon la conception que l'on a de la source, peut être considéré comme anaphore associative (ancrée dans la situation et récupérée par les processus inférentiels décrits par Reichler-Béguelin, 1989)[25] ou comme une anaphore prédicative (puisque le SN fournit une information nouvelle). Dans cet ordre d'idées, il conviendrait d'ailleurs d'analyser[26] également des SN (bi-)nominaux appelés « discursifs » par Bartning (1996 : 30) :

  (28) Petit détail, **l'homme à l'échelle** est un retraité de 76 ans.

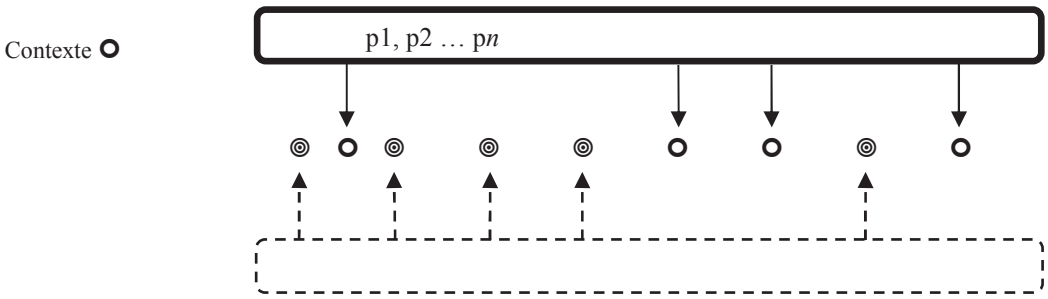
  (29) Après contrôles, **l'individu au scooter** est entendu à l'hôtel de police.

  (30) Cuisinier au restaurant Les Loges, installé à quelques centaines de mètres de l'accident, **le mis en cause** n'aurait jamais dû prendre le volant.

- celles qui proviennent du faisceau de propriétés intrinsèques au référent (p.e. les N généraux cf. (31) ci-dessous, de phase, etc.) ;

  (31) Malheureusement, il est trop tard pour **Pietro Lancellotti**, chauffeur-livreur de la société Autobar qui exploite des distributeurs automatiques de boissons et de sandwiches. **L'homme**, un Colmarien de 44 ans, père de deux enfants, est recroquevillé sur le siège conducteur de son camion frigorifique. Il a la tête à moitié arrachée. La vitre de sa portière a volé en éclats.

Le Schéma 1 synthétise la double nature des maillons dans les CR :



Propriétés intrinsèques du référent : ◎

Schéma 1 – Sources des maillons dans les CR

De là il résulte que les SN nominaux dont l'information dépend des propriétés intrinsèques du référent sont susceptibles d'être employés tout au long d'une CR : on dira alors qu'ils ont un empan large, voire maximal. Par contraste, les SN situationnels ou relationnels, fortement dépendants du cotexte, opèrent une saisie référentielle qui reste ponctuelle, ce qui fait qu'ils ont un empan « étroit ». C'est ce que traduisent notamment les modalités de leur reprise : dans (32), *la victime*, dont le référent est masculin, est repris par un pronom congruent en genre mais sur une brève distance[27], l'usage consistant apparemment à réinstancier assez rapidement le référent par un SN congruent en genre :

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

(32) **La victime** est tombée sur un chemin bétonné menant au hall d'entrée. **Elle** gisait sur le ventre quand les secours sont arrivés. Entourés de nombreux riverains, les pompiers ont tenté de réanimer **l'octogénaire** sur place avant de **le** transporter dans l'ambulance où **il** est décédé.

(33) Samedi, vers 19h, **la victime âgée de 29 ans et connue pour sa fragilité**, avait pris **son** vélo pour **se** rendre sur un terrain vague de Blainville. **L'homme s'**était aspergé d'essence avant d'y mettre le feu.

De là vient aussi que, certains emplois, comme en (34) où le référent désigné par *la victime* était manifestement connu par ses voisins antérieurement aux faits, doivent être considérés comme ce que Fauconnier (1984 : 47 et 49)[28] nomme des connexions entre « espaces-temps » :

(34) La police de Capellen a été rejointe par la police criminelle chargée par le parquet de rechercher des empreintes. La Protex a également envoyé son service d'aide psychologique afin d'assister *l'habitante de la rue qui connaissait **la victime** ainsi que ses voisins*, tous très choqués par ce geste désespéré.

Bref, il ressort de ces observations que la composition des CR est plus complexe qu'il n'y paraît compte tenu de l'origine variable de ses maillons. Cela ne semble guère avoir d'incidence dans une perspective d'annotation de corpus, où l'on « enregistre » les maillons les uns après les autres – encore faudrait-il le vérifier – mais, au plan cognitif, où l'on considère que le processus de traitement se fait de manière incrémentielle, en enrichissant le « fichier » d'un même référent au fur et à mesure du traitement des occurrences, il se pourrait que la disparité des maillons soit de nature à faciliter *vs* entraver le processus interprétatif. Ce qui serait, bien entendu, là encore à vérifier.

5.4.2 … les noms généraux d'humains

La seconde catégorie de N qui, en nombre, suit les relationnels est constituée par les N généraux d'humains, du type *homme, personne, individu,* pour lesquels le pourcentage s'élève à 16% des SN dont la tête est un nom commun. Leur originalité, dans les FD, est de faire office de 1er maillon dans 21% des cas[29], titraille exclue (cf. (36) et (37)), et, dans 40% des cas, d'anaphore. Ce point est important dans la mesure où, dans d'autres genres textuels, les SN à hyperonymes ont plutôt tendance à faire office de reprises anaphoriques comme l'illustrent (38-39) :

(35) Le 10 juillet dernier, **un homme** est mort après avoir été éjecté d'une voiture incontrôlable

(36) **Un homme** a été tué de plusieurs balles, dans la nuit de lundi à mardi.

(37) Quatre **personnes** sont décédées dans la nuit de samedi à dimanche dans un accident de la circulation dans le Cantal

(38) Le charme de **M. Giscard d'Estaing** opère. **Il** garde de la distance mais 0 sait séduire, avec humour**. L'homme** est aussi craint et **il** impose un secret total (*Le monde,* 07/06/2003)

(39) **Mon père** voulait devenir ingénieur comme lui. Qu'ai-je reçu en héritage de **cet homme** au destin tragiquement suspendu ? (Frantext, Flem, *Lettres d'amour*, 2006)

Ce mode de présentation initial, qui s'appuie sur les expressions référentielles spécialisées dans la référence indéfinie épistémique[30] selon Martin (2006 : 17), a pour effet d'anonymiser les référents et de contribuer à ce que les spécialistes du FD appellent la « platitude » des personnages (Dubied, *op. cit*. : 146) *i.e*. à l'absence d'attributs contradictoires et fixes.

6 Conclusion

De cette étude ressortent au moins trois résultats. Premièrement, nous avons fait valoir, sur la base d'une étude quantitative et qualitative des expressions référentielles, que les CR des FD sont originales au triple plan de :

- leur nombre : les FD sont, dans leur grande majorité (60% des cas), pluri-référentiels et instancient, qui plus est, des référents « obligés » comme les forces de l'ordre, de la sécurité et la magistrature ; corollairement, la densité des expressions référentielles dans les textes est assez forte ;

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

- la dimension : les CR sont, dans leur ensemble, brèves et comptent en moyenne 3.42 maillons ; cela tient évidemment à la brieveté même du genre ainsi que, dans une moindre mesure, à la pluri-référentialité ;
- la composition : les maillons des CR sont hétérogènes du point de vue de leur catégorie grammaticale qui se manifeste dans des proportions inattendues par rapport aux prédictions de certaines théories (p.e. l'accessibilité) puisque les formes de moyenne et haute accessibilité prévalent ; pour ce qui concerne les SN, leur tête lexicale est éminemment variable ; en outre, il est apparu que l'hétérogénéité des CR était renforcée, pour ainsi dire, par l'origine des maillons selon qu'ils s'ancrent dans la situation *vs* dans le faisceau de propriétés du référent : cela a des incidences sur leur empan long *vs* étroit selon leur nature ;

En cela, nos observations donnent des assises linguistiques non négligeables à celles des théoriciens du FD, qui, même lorsqu'ils examinent précisément, ainsi que le fait Dubied (*op. cit*. 227-245), le personnel du roman, en restent à des considérations sémiotiques très générales. Nos observations montrent également que les CR des FD se différencient aussi bien de genres non journalistiques (comme les textes de lois évoqués en 2) et de sous-genres journalistiques comme le portrait. La corrélation entre genre et CR se confirme donc, ce qui fait des CR un critère fort pour distinguer les genres et devrait inciter les approches des genres fondées sur l'exploration de gros corpus, à l'instar de celle de Biber, à systématiser la prise en compte de ce paramètre.

Au plan théorique, la distinction entre anaphores « situationnelles » *vs* anaphores fondées sur les propriétés du référent montre, par ailleurs, que la composition des CR, généralement schématisée de manière linéaire et « plate » par une série de maillons, n'est que le résultat d'expressions référentielles dont le point d'ancrage est éminemment variable avec des incidences sur le traitement cognitif qui mériteraient d'être explorées.

Enfin, à travers la typologie, encore sommaire, des N d'humains qui constituent, dans les FD, l'essentiel des expressions référentielles, nous avons montré que la distinction entre anaphores situationnelles *vs* fondées sur les propriétés du référent, nécessite une étude fine du lexique dénotant l'humain.

Références bibliographiques

Anscombre, J.-C. (2001). À propos de mécanismes sémantiques de formation de certains noms d'agent en français et en espagnol, *Langages, 143*, 28-48.

Anscombre, J.-C. (2003). L'agent ne fait pas le bonheur : agentivité et aspectualité dans certains noms d'agent en espagnol et en français. Thélème, *Revista Complutense de Estudios Franceses*, *11*, 11-27.

Ariel, M. (1990). *Accessing Noun Phrase antecedents*. London: Routledge.

Auclair, G. (1970). *Le mana quotidien : structures et fonctions de la chronique des faits divers*. Paris : Anthropos.

Barthes, R. (1964). Structure du fait divers. *Essais critiques*. Paris : Seuil, 188-197.

Barthes, R. (1981). *Essais critiques*. Paris : Seuil, (édition originale 1964).

Bartning, I. (1996). Eléments pour une typologie des SN complexes en de en français. *Langue française, 109,* 29-43.

Baumer, E. (2011). Noms propres et anaphores lexicales en anglais et en français : étude comparée des chaînes de référence. *Thèse de doctorat*, Paris Diderot, 378 p.

Benveniste, E. (1974). *Problèmes de linguistique générale*, t.2, chap. 8, 113-125.

Charolles, M. (1987). Contraintes pesant sur la configuration des chaînes de référence comportant un nom propre. *Cahiers du centre de recherches sémiologiques de Neuchâtel, 53*, 29-55.

Charolles, M. (1988). Les plans d'organisation textuelle : périodes, chaînes, portées et séquences. *Pratiques, 57*, 3-15.

Condamines, A. (2005). Anaphore nominale infidèle et hyperonymie : le rôle du genre textuel. *Revue de sémantique et pragmatique, 18*, 33-52.

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

Corblin, F. (1995). *Les formes de reprise dans le discours. Anaphores et chaînes de référence*. Rennes : Presses Universitaires de Rennes.

Cornish, F. (1986). *Anaphoric relations in English and French: a discourse perspective*. Londres : Croom Helm.

Deleu, C. (2005). Le monde selon le nouveau Détective : quand le fait divers renonce au réel. *Les Cahiers du journalisme n°14*– Printemps/ Eté 2005, 76-93.

Dubied, A. & M. Lits (1999). *Le fait divers*. Paris : PUF.

Dubied, A. (2000). Invasion péritextuelle et contaminations médiatiques. Le fait divers, une catégorie complexe ancrée dans le champ journalistique. *Semen, 13*, 51-66.

Dubied, A. (2004). *Les dits et les scènes du fait divers*. Genève : Droz.

Fauconnier, G. (1984). *Espaces mentaux*, Paris : Minuit.

Fragnon, J. (2007). Le fait divers dans la PQR : fenêtre ou miroir sur la violence ? *Les Cahiers du journalisme n°17* – Été 2007, 254-269.

Goutsos, D. (1997). Modeling Discourse Topic: sequential relations and strategies in expository text. *Advances in Discourse Processes, vol. LIX*. Norwood: Ablex Publishing Corporation.

Herslund, M. (1996). Partitivité et possession inaliénable. La relation d'appartenance. *Faits de Langue, 7*, 33-42.

Jenkins, C. (2002). Les procédés référentiels dans les portraits journalistiques. *XV Skandinaviske romanistkongress*. Oslo, 12-17 août 2002.

Longo, L. (à paraître). Un corpus pour optimiser l'identification automatique des chaînes de référence. *Cahiers de Praxématique, 51*.

Longo, L., Todirascu, A. (2010). Genre-based Reference Chains Identification for French. *Investigationes Linguisticae*, *Volume XXI*, 57-75.

Longo, L. Todirascu, A (2009). Une étude de corpus pour la détection automatique des thèmes, *actes des 6èmes journées de linguistique de corpus,* 10-12 septembre 2009, Lorient.

Lundquist, L. (2005). Noms, verbes et anaphores (in)fidèles. Pourquoi les Danois sont plus fidèles que les Français ? *Langue française, 145*, 73-91.

Luquet G. (1994). Remarques sur la structure des suffixes formateurs de noms d'agent et d'instruments en espagnol. *Recherches en linguistique hispanique*, *22*, 339-348.

Martin, R. (2006). « Définir l'indéfinition », in *Indéfini et prédication*, Corblin *et al.* (dir.). Paris : PUS, 11-24.

Reichler-Béguelin, M.-J. (1989). Anaphores, connecteurs, et processus inférentiels, in C. Rubattel (éd.). *Modèles du discours. Recherches actuelles en Suisse Romande*, Berne : P. Lang, 302-336.

Schnedecker, C. (1997). Noms propres et chaînes de référence. *Recherches Linguistiques*, *21*. Paris, Klincksieck.

Schnedecker, C. (2005). Les chaînes de référence dans les portraits journalistiques : éléments de description. *Travaux de linguistique* 2/2005, *51*. Duculot, 85-133.

Steinberger, R., Pouliquen, B., Widiger, A., Ignat, C., Erjavec, T., Tufis, D., Varga, D. (2006). The JRC-Acquis: A multilingual aligned parallel corpus with 20+languages. *Actes de LREC, Italie.*

Tutin, A., Clouzot, C., Antoniadis, G (2000). *Un corpus d'anaphores discursives pour les études en TAL*. Rapport non publié, http://w3.u-grenoble3.fr/tutin/Publis/rapport.pdf (consulté le 11/11/2011).

Ulland, H. (1993). *Les nominalisations agentives et instrumentales en français moderne*. Berne : P. Lang.

[1] L. Lundquist (2005) montre une corrélation entre formes de reprise et familles de langue.

[2] Qui étudie par ailleurs l'impact de la langue (cf. anglais vs français).

[3] Les textes sélectionnés sont de taille comparable pour les 5 genres (50 000 *tokens*, formes et signes de ponctuation, soit une quinzaine de pages par genre).

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

[4] Cf. la synthèse de Dubied (2004 : 84-85).

[5] Il existe aussi des FD longs composés d'articles judiciaires ou dont l'importance est telle qu'elle mérite une large couverture (cf. Fragnon : 256).

[6] Le plus court compte 52 mots et le plus long 579.

[7] On peut à l'instar de certains auteurs y voir la réalisation du fameux schéma quinaire de Propp, avec, dans le cas de (4), la force transformatrice (p2-3), l'action (p4-6), la force équilibrante (p7-8) et la situation finale (p9).

[8] Pour plus ample informé, nous renvoyons à Fragnon (art. cit.), Dubied (2004 : 84-85) et Deleu (2005 : 14-15).

[9] Au sens ici des théoriciens du récit (Propp, Greimas, etc.).

[10] Comme le montre l'annotation du texte (6), sont considérés comme « maillons » les SN nominaux ou pronominaux (pronoms personnels, réfléchis, etc.) explicites, les anaphores zéros et déterminants possessifs du fait qu'ils équivalent à « de SN ». Nous n'avons pas pris en compte ce que certains (cf. Cornish, 1986) nomment « indices », comme par exemple les accords des participes passés de « *Ces derniers ont pris en charge deux chauffeurs routiers d'outre-Rhin, légèrement blessé**s** et surtout choqué**s*** » qui sans référer à proprement parler réinstancient d'une certaine manière les référents dans le texte.

[11] Même si le regroupement des expressions démonstratives nominales et pronominales est peu orthodoxe, ce choix nous est paru pour cette étude pertinent compte tenu de la capacité commune des démonstratifs à opérer une forme de rupture avec le cotexte.

[12] En fait, comme le montre Schnedecker (1997 : 52-63) cette ambiguïté n'est, en l'occurrence, que potentielle.

[13] On verra plus bas que d'autres raisons, propres au genre, motivent ce type d'expressions référentielles.

[14] Qui constituent 17.5% de cette catégorie.

[15] Sur cette question, cf. Charolles (1987).

[16] Les SN indéfinis nus constituent 40% des expressions de faible accessibilité référentielle et les indéfinis expansés 42.5%, soit 82.5% du total.

[17] La classification est encore hasardeuse, mais le fait est que l'étude des noms d'humains reste à faire. Nous n'avons pas inclus les N à occurrence unique.

[18] Par contraste, p.e. avec les portraits journalistiques où la diversité des maillons se limite généralement à la CR du référent principal.

[19] Les N en italiques signalent un classement « faute de mieux ».

[20] Les relatives sont exploitées dans 13% des cas.

[21] Les pourcentages sont ici calculés sur le total de chacune des catégories.

[22] Avec un « jeu » entre têtes lexicales et modifieurs puisque les N délivrent, le cas échéant, ces informations : ***Une octogénaire*** *périt dans un incendie/Un jeune Maiziérois*.

[23] Encore peu pris en considération dans les classifications nominales et donc encore peu étudiés.

[24] Pour ces N d'agents, voir Benveniste (éd. 1974), Anscombre (2001), Luquet (1994) sur la différence de construction entre les N d'agent du type d'*enquêteur* (qui vient d'*enquêter*) et *conducteur* (qui provient du supin de *conducere* > conductus). Pour les N en *-ant*, voir Anscombre (2003), pour qui il s'agit de N d'acteurs, dotés d'une moindre agentivité que les N en *-eur*, et Ulland (1993).

[25] M.-J. Reichler-Béguelin (1989) a proposé une typologie de ce type d'anaphores.

[26] Ce que nous reportons, faute de place, à une étude ultérieure.

[27] Lorsque la victime renvoie à un référent féminin, la reprise pronominale peut opérer sur une longue distance. Quand le référent est masculin, il est assez rapidement réinstancié sour la forme d'un SN masculin.

[28] Qui donne d'ailleurs en guise d'illustration des extraits de FD.

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)

SHS Web of Conferences 1 (2012)
DOI 10.1051/shsconf/20120100061
© aux auteurs, publié par EDP Sciences, 2012

[29] Pourcentage calculé sur l'ensemble des noms d'humains généraux du corpus (78).

[30] « Dans l'indéfinition épistémique (dans l'indétermination), l'objet n'est pas déjà déterminé au moment de l'énonciation : soit que le locuteur estime qu'il ne l'est pas pour l'interlocuteur ; soit il est incapable lui-même de spécifier l'objet dont il s'agit parmi tous les objets de même nature […]. » (Martin, art. cit. : 19).

Article en accès libre placé sous licence Creative Commons Attribution 2.0 (http://creativecommons.org/licenses/by/2.0)