

ESPACE NUMÉRIQUE ET LITTÉRAL

Adriano ALLORA
Université de Turin

RESUMÉ

La présente recherche a pour but de montrer comment sont utilisés les déictiques spatiaux dans les newsgroups. A une étude linguistique attentive des variations diamésiques s'ajoute une analyse de la diaphasie et des degrés de formalité : sur les trois corpora analysés, le premier est constitué de textes de newsgroups censés être de registre haut, le deuxième, de textes de newsgroups non marqués par un style particulier, et le dernier, d'articles académiques. Les méthodes d'analyse, principalement quantitatives, se limitent à la fréquence d'occurrences de déictiques spatiaux et cooccurrences de déictiques spatiaux et autres parties du discours et/ou mots clé. Cette recherche met en évidence le rôle fondamental des démonstratifs dans la structuration du texte, en particulier dans les variétés hautes, et la spécialisation des deux types de démonstratifs : alors que le proximal est utilisé par les auteurs pour expliciter leurs propres textes, le distal renvoie aux textes d'autrui ou appartenant à d'autres segments du réseau (une spécialisation non présente dans l'usage des adverbes spatiaux proximaux et distaux).

ABSTRACT

This paper aims at showing how space indexicals are used in newsgroups. The linguistic inquiry regards as much medium constraints as style / formality degree constraints: of three corpora, the first one is built with newsgroups formal texts, the second one is built with newsgroups formality-unrelevant texts, the third one is built with academic articles. Analysis methods are merely quantitative: relative frequency of space indexicals and frequent co-occurrences of space indexicals and keywords or specific parts of speech.

From the corpora arose the basic role of demonstratives in structuring texts (more in formal texts than informal ones) and the specific way of use of two different kinds of demonstratives: proximal demonstratives are used to clarify / structure the writer's posts and articles, distal demonstratives are used for texts written by others or, indeed, texts belonging to other Net Mediated Communications (websites, chat lines, different newsgroups). This specific way in using space indexicals doesn't seem to occur with space deictic adverbs.

1. INTRODUCTION

Cette contribution¹ a pour but d'analyser la façon dont les usagers du réseau utilisent les déictiques spatiaux dans les textes des Newsgroups (que nous appellerons dorénavant NG), avec une double référence à deux axes de variation : diamésique d'un côté et diaphasique² de l'autre.

Le reste de cette intervention se concentrera sur certains aspects particuliers de la déixis spatiale ; suivront quelques annotations générales sur le concept d'espace en réseau et sur la déixis dans le réseau :

1) En ce qui concerne les démonstratifs, nous tenterons d'expliquer les raisons de leur recours aussi massif dans l'italien de registre haut ;

2) En ce qui concerne les adverbes spatiaux, nous mettrons à l'épreuve l'hypothèse d'un emploi différent des membres des couples italiens *lì* et *qui* (ponctuels) vs. *là* et *qua* (aréaux), déjà émise (cf. Allora 2002) pour les IRC (*Internet Relay Chat*).

2. PREMIERS RELEVÉS QUANTITATIFS : DÉIXIS ET NG

2.1. Les *corpora* examinés

La sélection des *corpora* étudiés répond à l'exigence de confronter des textes censément situés dans la même zone sur l'axe de variation diamésique, et des textes censément situés dans la même zone sur l'axe de variation diaphasique : NUNC-A et Athenaeum sont des recueils de textes de registre haut en langue italienne.

Dans le cadre de cette recherche, nous avons ainsi examiné :

- 1) le *corpus* NUNC-A, qui a été construit en sélectionnant dans les NUNC des hiérarchies en mesure de fournir une variété d'italien de registre haut (disciplines académiques telles que l'histoire, la géologie, la philosophie, le droit et les conversations relevant du domaine médical)³ ; une sélection ultérieure sur les posts a permis au *corpus* d'atteindre les 74.695.600 mots ;
- 2) le *corpus* Athenaeum composé d'articles de la revue du même nom, publiée par *l'Università degli Studi di Torino*, dont les textes scientifiques sont censés être de registre comparable à celui de NUNC-A, compte 306.927 mots ;

¹ Pour laquelle je souhaite remercier Manuel Barbera et Carla Marellò dont les précieux et pertinents commentaires ont permis d'éliminer la plupart des erreurs commises. La responsabilité des erreurs restantes est donc uniquement imputable à l'auteur.

² La diaphasie est, dans les limites du présent texte, démembrée ; les trois paramètres (domaine, réseau et interlocuteur) sont divisés d'une part en réseau – ce qui caractérise la variation diamésique – et d'autre part en champ et interlocuteur – que l'on attribuera par la suite au domaine de la variation diaphasique.

³ Cf. Onesti (2009 : 208), Cerruti (2009), Cerruti & Onesti (2010).

- 3) les *corpora* généralistes de langue italienne NUNC-IT 1 et NUNC-IT 2 – considérés ici comme un unique *corpus* et dorénavant indiqués sous le nom de NUNC-IT 1+2 – contiennent des hiérarchies de la plus grande variété possible et comportent 237.401.299 mots.

Athenaeum et NUNC-IT 1+2 sont diversement mais librement interrogeables à partir de leur page d'accueil respective.

2.2. Les déictiques examinés

Étant donné la quantité considérable de textes, et afin de réduire le plus possible le bruit dans les extractions, nous avons préféré adopter une politique qui favorise la précision (c'est-à-dire le rapport entre données utiles et données extraites) à celle qui, en *Information Retrieval*, est généralement appelée *recall* (le rapport entre données extraites et données disponibles dans le *corpus*) : la recherche a été effectuée sur les seuls déictiques inhérents, définis avec des critères plus restreints que la normale⁴ :

– pour les déictiques de personne, les pronoms personnels toniques et les pronoms et adjectifs possessifs des première, deuxième, quatrième et cinquième personnes, ainsi que les pronoms personnels clitiques des première et deuxième personnes ;

– pour les déictiques de temps, nous nous sommes limités à une série arbitraire de formes temporelles étiquetées par le *pos-tagger* (étiqueteur grammatical) en tant qu'adverbes : *ora, adesso, domani, ieri, oggi, stasera, stanotte e stamattina* [à présent, maintenant, demain, hier, aujourd'hui, ce soir, cette nuit et ce matin] ;

– pour les déictiques spatiaux, nous avons pris en considération les démonstratifs, le verbe de mouvement *venire* [venir] et les deux séries d'adverbes spatiaux (proximal et distal)⁵.

L'utilisation importante des démonstratifs dans des références anaphoriques et non pas déictiques pourrait plaider en faveur de leur exclusion, mais la spécificité de la déixis liminaire est au contraire un motif suffisant pour les inclure. Quant aux *verba movendi*, bien que le verbe de mouvement *andare* [aller] compte généralement au nombre des phénomènes liés à la

⁴ La *lectio maiorum* (cf. notamment Vanelli 1995) appelle déictiques inhérents les déictiques pouvant être utilisés de manière non déictique uniquement dans des références métalinguistiques (*qui* est déictique à moins qu'on ne l'utilise comme référent pour le mot *qui*, comme c'est le cas dans cette contribution) ; aux déictiques inhérents s'opposent les déictiques contextuels, les mots qui peuvent être déictiques, mais qui peuvent aussi ne pas l'être, comme les pronoms personnels des troisième et sixième personnes.

⁵ Le statut des démonstratifs comme topodéictiques nécessiterait de plus longues discussions et quelques éclaircissements sur lesquels nous préférons ne pas nous attarder ici pour des raisons d'espace et afin de conserver une argumentation linéaire.

déixis, il a été exclu ici en raison du grand nombre de cas où il n'est pas utilisé de manière déictique⁶.

2.3. Cadre théorique de référence

Lorsque, dans la présente étude, il sera fait référence aux caractéristiques du réseau de communication et à ses effets sur la variété de langue produite, le cadre théorique utilisé sera la Théorie de la Diamésie Générale (dorénavant appelée TDG. Cf. Allora 2009).

La TDG est le résultat des premières recherches sur les *corpora* distribués par corpora.unito.it. Née de la considération de l'inadéquation de l'opposition oral-écrit (*la Teoria della Diamesia Ristretta*⁷) pour décrire les textes des *Comunicazioni Mediate dalla Rete (Net Mediated Communication* ou *communication médiée par le réseau*), la TDG repose sur deux fondements :

- 1) le texte est avant tout une matière verbale dont la forme est aussi le résultat de contraintes imposées par le réseau ;
- 2) les spécificités du réseau, au plus haut niveau d'abstraction possible, doivent constituer l'élément de départ pour pouvoir décrire de manière scientifique le rapport entre texte et réseau, c'est-à-dire la variabilité diamésique.

À partir de ces bases, la TDG décrit la façon dont les caractéristiques particulières du réseau empêchent ou exigent certains types de manipulation de la matière verbale. Le présupposé est que, étant donné une signification X, les formes que peut prendre la matière verbale dans laquelle X s'actualise ne sont qu'un sous-ensemble de toutes celles théoriquement réalisables ; mais les options éliminées ne dépendent pas toutes du message, beaucoup dépendent en effet du réseau. Classifier les propriétés du réseau en tant que telles, et non pas à travers le filtre de la dichotomie oralité-écriture, permet donc d'isoler avec davantage de précision les effets du médium sur le message.

2.4. Un instantané des déictiques dans les NG

S'il était vrai qu'« un'immagine vale mille parole » [une image vaut mille mots], il suffirait de regarder l'histogramme de l'Illustration 1 pour comprendre intégralement l'utilisation des déictiques online.

⁶ Le verbe de mouvement a été pris en considération uniquement dans les extractions qui concernaient en général les déictiques, mais son utilisation ne sera pas traitée de manière approfondie.

⁷ Dont la référence standard est Koch & Oesterreicher (1985), mais la dichotomie oral-écrit est malheureusement invoquée également dans des textes par ailleurs admirables (et plus récents) comme Cristal (2006), malgré le fait que lorsqu'on tente de classifier et de définir la *NMC* (Net Mediated Communication) sur la base du paramètre oral-écrit, on obtient des résultats problématiques et peu clairs.

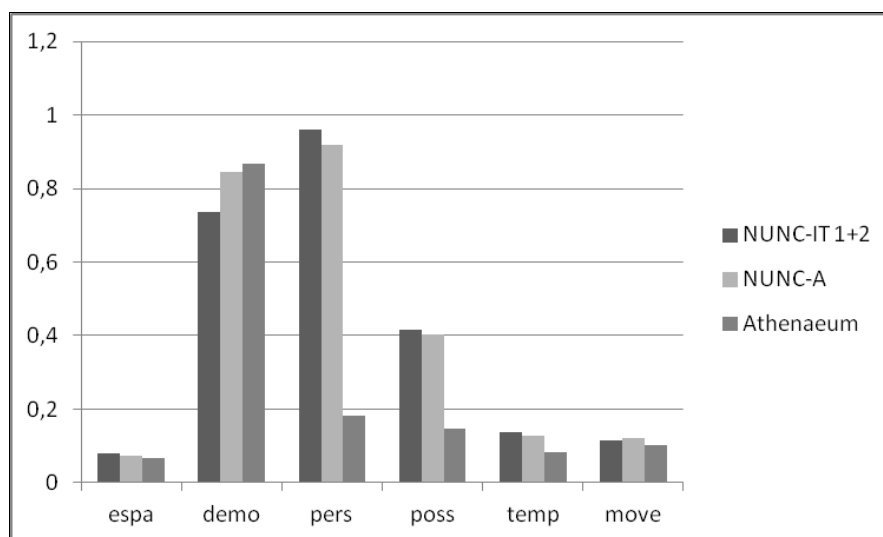


Illustration 1. – Distribution des déictiques (en pourcentages)

Dans ce graphique, en pourcentages sur le nombre total de tokens, sont indiquées les quantités d'indexicaux avec un grain légèrement plus fin que celui de la tripartition classique ; en particulier, les déictiques spatiaux ont été divisés en adverbes (espa), démonstratifs (demo) et verbes de mouvement (move) ; les déictiques de personne ont été divisés en personnels et possessifs et les déictiques de temps constituent un groupe à part.

Le Tableau 1, qui indique la distribution des déictiques, les types, les tokens, les sommes des indexicaux et la liste des types de déictiques sélectionnés, a une étendue encore plus vaste.

	NUNC-IT1+2	NUNC-A	Athenaeum
Token	237.401.299	74.695.600	306.927
Types	2.445.454	669.427	32.221
Indexicales	5.802.418	1.858.333	4.443
Espa	187.473	54.021	205
Demo	1.744.732	632.308	2.665
Pers	2.282.731	687.116	561
Poss	990.550	300.680	448
Temp	322.630	94.210	250
Move	274.302	89.998	314

Tableau 1. – Distribution des déictiques (quantités absolues)

Les données doivent être considérées en relation avec les deux aspects suivants : diamésie et diaphasie.

En ce qui concerne le premier point, les ressemblances entre NUNC-IT 1+2 et NUNC-A sont évidentes : les deux *corpora* sont caractérisés par la tendance à un plus grand recours à la déixis qui devient véritablement significative dans le cas des indexicaux de personne.

À la tendance générale qui confirme un plus grand recours aux déictiques dans les variétés de registre plus bas (NUNC-IT 1+2 avec plus de déictiques que les autres *corpora*) ainsi que dans les variétés caractérisées par un plus grand bruit⁸ (dans NUNC-IT 1+2 et NUNC-A plus que dans Athenaeum), s'opposent deux données : un plus grand recours en pourcentages aux verbes de mouvement déictiques dans NUNC-A et un plus grand recours aux démonstratifs dans les deux *corpora* de registre haut.

Cette dernière donnée, en particulier, pourrait se révéler intéressante notamment pour l'aspect diaphasique.

3. LA CONSTRUCTION DE L'ESPACE NUMÉRIQUE

3.1. L'espace métaphorique du réseau

Dès ses débuts, Internet a été décrit comme s'il s'agissait d'un lieu, d'un espace à la fois littéral et métaphorique. Littéral parce que le mot latin *spatium* signifie précisément vide, et le réseau est vide ; métaphorique parce que, étant donné l'étymologie du mot *métaphore*, il s'agit d'un espace au-delà de l'espace online ou d'un espace permettant de franchir les limites de la réalité online.

Dans cet espace, vide de manière différente de ce que peut l'être une pièce vide (qui a des murs, des portes, des fenêtres, un haut et un bas), dépourvu d'outils qui ne soient pas verbaux, comme l'ont souligné les premières recherches sérieuses sur le sujet (cf. Allora 2002), les usagers ont eu recours à leur propre langue pour rendre compte d'une telle spatialité : en décrivant le réseau, ils en ont créé une image dont on pouvait disposer.

Mais les descriptions ne pouvaient pas être dénotativement explicites, hormis dans certains cas pourvus de la structure informatique nécessaire pour conserver la mémoire de telles descriptions⁹. En d'autres termes, ces descriptions ne pouvaient pas « s'appuyer » sur des infrastructures technologiques et ne devaient impliquer aucune tradition orale nécessitant un travail

⁸ Dans le cadre de la théorie de la TDG.

⁹ MUD (*Multi User Dungeon* ou *Dimension*) à l'époque de la pure textualité, MMORPG (*Massive(ly) Multiplayer On line Role Playing Game*) aujourd'hui. Ces logiciels recréent, autrefois avec du texte pur et aujourd'hui graphiquement, des espaces dans lesquels les *personae* numériques des usagers, les *avatar*, agissent et interagissent. Pour cette raison, l'espace dans les MUD et dans les MMORPG est absolument indispensable.

explicite et conscient de mémorisation : aucun usager d'IRC n'aurait jamais écrit qu'un réseau était fait de telle manière, avec telles caractéristiques¹⁰.

Les descriptions passaient essentiellement dans cette zone interstitielle, entre grammaire et lexique, qui est occupée par les déictiques (cf. Allora 2005).

Les déictiques se sont révélés être un puissant outil sur ce plan également : comme vecteurs simples (classes fermées, mots de haute accessibilité, dont l'emploi correct dépend non pas de l'encyclopédie mais de la compétence linguistique) d'informations transparentes (l'opposition proximité-éloignement est la plus significative), ils garantissaient une transmissibilité facile indépendamment de l'infrastructure informatique en place.

3.2. L'espace enrichi

L'espace du réseau est qualitativement différent de l'espace online, et cette diversité est perçue par les usagers qui, bien qu'ayant recours aux déictiques pour décrire les deux espaces, les utilisent différemment. Comme pour d'autres aspects de l'analyse linguistique (cf. Allora 2009), le réseau a donc imposé un élargissement théorique.

La déixis enrichie par le réseau reconnaît trois nouvelles sous-catégorisations :

– déixis exomédiale : il s'agit de la déixis traditionnelle (et dans certains cas anaphore), qui fait référence à l'espace, aux participants à l'échange et au temps offline¹¹ :

- (1) Non so più a chi rivolgermi, qui a Sora dove abito stanno succedendo (da tre anni) cose incredibili !

[Je ne sais plus à qui m'adresser, ici à Sora où j'habite il se passe (depuis trois ans) des choses incroyables !].

– déixis endomédiale : il s'agit de la déixis qui fait référence à l'espace, aux participants et, dans une moindre mesure, au temps online :

- (2) oggi lunedì dicembre, scrivo qua a voi tutti, quello che purtroppo mi mi # ; successo.

[aujourd'hui lundi de décembre, j'écris ici à vous tous, ce qui malheureusement m'est m'est # ; arrivé.].

- (3) stasera provo a rileggermi il thread, se non lo capisco di nuovo posso chiedertelo ? Scusami ma sono arrivata qua in fondo soltanto ora

¹⁰ Même si le présent texte n'analyse pas l'utilisation des déictiques dans la perspective de la pragmatique néo-gricéenne, comme le fait par exemple Cornish (2008), le recours aux déictiques avec cette fonction peut être facilement expliqué en référence aux principes de minimisation des contenus (*I principle*) et minimisation de la forme (*M principle*).

¹¹ La seule modification apportée aux exemples est la suppression de l'espace avant les signes de ponctuation (qu'il n'était d'aucune utilité de diviser en *tokens* ici).

[ce soir j'essaie de relire le thread, si je le comprends toujours pas je peux te le demander ? Excuse-moi mais je suis arrivée en bas seulement maintenant].

– déixis liminaire : il s'agit de la déixis qui fait référence à des mouvements allant de l'espace offline à l'espace online et vice-versa ou aux participants à l'échange dans leur double apparence de personnes en chair et en os et *alter ego* numériques. Même les déixis textuelles entrent dans ce groupe, étant donné la nature à la fois purement textuelle et textuelle-in-quantum-numérique des textes online¹² :

- (4) Ad esempio , se uno venisse qui su FISA a dire che le Piramidi le hanno fatte gli alieni , noi tutti prima proveremmo a convincerlo del contrario

[Par exemple, si quelqu'un venait dire ici sur FISA que les Pyramides ont été construites par les aliens, on essaierait tous de le convaincre du contraire].

La nécessité d'une telle tripartition était apparue dans l'explication des différents usages de deux séries d'adverbes spatiaux dans les *Internet Relay Chat* (IRC) : *qui* et *là* (respectivement proximal et distal ponctuels) et *qua* et *là* (proximal et distal aréaux). L'opposition ponctuel-aréal, bien que dépassant la compétence métalinguistique des locuteurs dépourvus d'une préparation spécifique sur le sujet, est respectée dans la production linguistique ; par exemple, la locution *qua e là* est normalement utilisée pour désigner une zone qui n'est pas clairement définie, et remplacer le premier ou le second terme par ses équivalents ponctuels n'a généralement pas la même signification.

Dans les échanges en IRC, le terme ponctuel était principalement utilisé pour faire référence à des lieux ou des portions d'espace externes au réseau (déixis exomédiales), tandis que le terme aréal était utilisé pour des références à des lieux internes au réseau (réseau ou lieu de bavardage). Cette tendance a été interprétée comme un symptôme de la difficulté du locuteur à faire référence à un espace invisible et indivisible.

4. LES DÉICTIQUES SPATIAUX

4.1. Les démonstratifs

Les démonstratifs sont utilisés moins fréquemment dans les NG que dans les textes online de registre haut contenus dans Atheneum (cf. Tableau 2) :

¹² FISA est le nom du NewsGroup, le mouvement vers FISA est celui de la personne physique et de la personne numérique.

	NUNC-IT 1+2	NUNC-A	Athenaeum
Token	237.401.299	74.695.600	306.927
Démonstratifs	1.744.732	632.308	2.665
Pourcentage	0,73	0,85	0,87

Tableau 2. – Distribution des démonstratifs (quantités absolues)

Alors que le *corpus* NUNC-A distingue les démonstratifs utilisés comme déterminants de ceux utilisés comme pronoms¹³, les autres *corpora* (pour les raisons exposées dans Barbera 2002) ne le font pas, et il devient donc impossible d'établir une comparaison sur ce paramètre.

À seul titre de suggestion, on a essayé d'utiliser comme référence le rapport entre le pourcentage de pronoms sur le total des démonstratifs restitué par NUNC-A (45 %), et en supposant un rapport égal entre les dimensions étudiées dans les trois *corpora*, il en résulterait que dans NUNC-IT 1+2 les pronoms représentent 47 % et dans Athenaeum 39 %. Une telle estimation est partiellement confirmée par une sommaire extraction de tous les démonstratifs suivis d'un verbe (avec un critère plus étroit qui favorise la précision au détriment du *recall*) selon laquelle dans NUNC-IT 1+2 les pronoms ainsi définis représentent 0,054 %, dans NUNC-A 0,060 %, et dans Athenaeum 0,063 %.

Deux tendances générales apparaissent : d'une part, un recours plus fréquent aux pronoms démonstratifs dans les textes de NG, de l'autre un recours plus fréquent aux adjectifs démonstratifs dans les *corpora* de registre plus haut¹⁴.

On estime plus productive l'extraction des 50 premiers noms co-occurents avec un démonstratif.

Trois groupes de noms ont été identifiés : le premier groupe, étiqueté « actuel », est composé de noms qui font référence à l'actualité (tendanciellement des membres de syntagmes exodéictiques et anaphoriques) comme : *anni, articolo, cosa, mattina, mondo, guerra, giornata, gioco, film, vita, settore, situazione* [années, article, chose, matin, monde, guerre, journée, jeu, film, vie, secteur, situation].

Le deuxième groupe (« metadisc ») est composé de mots qui font référence au texte en tant qu'unité (tendanciellement des membres de syntagmes logodéictiques) comme : *affermazione, argomento, domanda, dati, frase, indagine, proposito, modo, quadro, vaccata* [affirmation, sujet, question, données, phrase, enquête, propos, mode, cadre, vacherie].

¹³ Dans NUNC-A on compte 119 522 *questo* pronom, 168 348 *quello* pronom, 240 170 *questo* déterminant, 104 268 *quello* déterminant.

¹⁴ On ne peut pas exclure que les extractions sur les pronoms aient été trop restrictives et que les locuteurs utilisent davantage, dans les *corpora* de registre haut, tous les pronoms démonstratifs et pas seulement les adjectifs démonstratifs.

Le troisième groupe (« metatext ») est constitué de mots qui font référence à la réalité médiale (tendanciellement des membres de syntagmes endodéictiques) comme : *newsgroup*, *NG*, *ng*, *post*, *messaggio*, *sito*, *link*, *thread*, *pagina* [*newsgroup*, *NG*, *ng*, *post*, *message*, *site*, *lien*, *thread*, *page*].

Les groupes sont distribués de la manière suivante (Tableau 3)¹⁵ :

Type	IT1	IT2	NUNC-A	Athenaeum
actuell	30	30	28	23
metatext	8	7	6	0
metadisc	12	13	16	27
	50	50	50	50

Tableau 3. – Mots co-occurents avec les démonstratifs
(50 premiers)

Les différences sont nettes : les textes traditionnels de registre haut contenus dans Athenaeum, plus denses¹⁶ d'un point de vue informatif, destinent une partie considérable de leurs démonstratifs à la structuration du texte (les mots du groupe metadisc, et dans une moindre mesure ceux du groupe metatext, ont une fonction métadiscursive et métatextuelle) ; les textes des NG qui ne sont pas caractérisés par un registre particulier comme ceux de NUNC-IT 1+2 font davantage référence au monde extratextuel ; les textes des NG de registre haut contenus dans NUNC-A se situent dans une position intermédiaire, avec plus de références métadiscursives (et moins de références métatextuelles) par rapport aux NUNC-IT 1 et NUNC-IT 2 et moins de références à l'actualité (mais plus de références métatextuelles) par rapport à Athenaeum.

Afin de déterminer une telle position intermédiaire, le réseau est peut-être plus influent que le registre (non seulement en raison de la présence des références métatextuelles telles qu'elles ont été définies dans la présente étude mais aussi et surtout en raison de la relative proximité avec les NUNC-IT 1+2). En d'autres termes, on peut sans doute dire que le registre influe sur la façon de définir NUNC-A par rapport aux autres *corpora* de NG, sans pour autant le « dénaturer » par rapport au médium.

Dans cette perspective, la TDG pourrait expliquer comment les paramètres de temps et d'accessibilité constituent une incitation à faire des références de type « actuell » et comment le paramètre d'espace conduit à

¹⁵ Travailler sur les 50 premières instances d'un certain type de cooccurrence nous empêche dans ce cas de traiter de façon unitaire et d'un point de vue quantitatif les *corpora* NUNC-IT 1 et NUNC-IT 2 ; par la suite nous ferons toutefois référence aux deux *corpora* ensemble, étant donné leur substantielle ressemblance.

¹⁶ Comme semblent le confirmer les relevés effectués dans Algozino (ics) relatifs au rapport entre mots grammaticaux et mots non grammaticaux.

des stratégies analogues de structuration de la matière verbale (aussi bien sur le plan « metadisc » que sur le plan « metatext »).

Notons que les six premiers rangs dans les listes de fréquence démonstratif + mots du groupe « metatext » restituent des déixis endomédiales ou liminaires et sont tous occupés par des formes de démonstratif proximal accompagnées de la sélection de mots indiquée dans le Tableau suivant :

NUNC-IT 1	NUNC-IT 2	NUNC-A
NG	NG	NG
Newsgroup	messaggio	Ng
Ng	newsgroup	Post
Messaggio	Ng	Messaggio
Post	post	Thread
Sito	Sito	Sito

Tableau 4. – Groupe metatext et démonstratif proximal

Avec différentes ordinations internes, ces six premières positions semblent définir les limites de la capacité à se référer à l'intérieur des NG : de l'« espace » dans lequel on y trouve – le *newsgroup*, quelle que soit la forme utilisée pour ce type de référence – à un « lieu » qui appartient encore à l'univers numérique de référence mais qui diffère du NG : le site.

Les références aux sites Web, les liens (déixis endomédiales, exemple (5)) et les références anaphoriques (déixis liminaires, exemple (6)) représentent la limite du « rapportable » online :

- (5) Non perdetevi l'altra sezione di questo sito che offre altre stupende foto di Zion

[Ne manquez pas l'autre section de ce site qui offre d'autres magnifiques photos de Zion].

- (6) Saiani sono un gruppo di tifosi di Bologna che ha un sito e all'interno di questo sito c'è un forum, frequentato da saiani e non

[Les Saiani sont un groupe de supporters de Bologne qui ont un site et à l'intérieur de ce site il y a un forum, fréquenté par des saiani, mais pas seulement].

Dans NUNC-IT 1+2, le septième rang est encore un proximal accompagné de *thread*, tandis que dans NUNC-A le septième rang est un distal co-occurent avec le mot *post*.

Le mot *sito* est le premier associé à un distal dans NUNC-IT 1+2 et le deuxième dans NUNC-A.

Si l'on veut tenter de dresser un premier bilan, on peut dire que l'emploi des démonstratifs est subordonné à cette norme de fonctionnement¹⁷ : les groupes démonstratif + nom, qui font référence non pas à des messages spécifiques mais à des canaux / situations communicatives ou lieux numériques (comme *NG*, *ng*, *newsgroup* et dans certains cas *thread*) sont habituellement déictiques tandis que les groupes démonstratif + nom qui font référence à des portions de texte (comme *post* ou *messaggio*) sont habituellement anaphoriques (cf. Conte 1999) ; si le démonstratif est un distal, le coréférent est dans un autre message, alors que si le démonstratif est un proximal, le coréférent est dans le même message.

4.2. Les adverbes spatiaux

Un cadre général de la distribution dans l'emploi des adverbes distaux et proximaux dans les NG révélera au moins deux points intéressants : alors que tous les *corpora* de référence présentent le proximal ponctuel comme l'adverbe le plus fréquent, Athenaeum révèle une fréquence insolite d'adverbes distaux aréaux ; les deux *corpora* de NG présentent des distributions semblables mais pas identiques – en particulier, les adverbes les plus fréquents dans le *corpora* de registre haut sont les deux adverbes ponctuels suivis du distal aréal, tandis que l'on retrouve les deux adverbes ponctuels suivis du proximal aréal dans le *corpus* non caractérisé par un registre haut :

	NUNC-IT 1+2	NUNC-A	Athenaeum
Qui	66,7	68,3	63,9
Qua	10,5	10,0	4,9
Lì	14,6	11,6	4,9
Là	8,2	10,1	26,3
	100	100	100

Tableau 5. – Distribution des adverbes spatiaux

Encore une fois, le registre semble donc avoir un poids, même s'il est mineur par rapport au réseau (les distances entre les couples de valeurs des *corpora* de NG sont toujours inférieures à celles entre les couples de valeurs des *corpora* de registre haut).

Pour approfondir la recherche, nous avons tenté d'extraire les cooccurrences de démonstratif + préposition + nom (sélectionner les seules prépositions nous a semblé la façon la plus rapide de comprendre à quoi elles se référaient et comment étaient utilisés les adverbes spatiaux).

La recherche s'est donc limitée à des échantillons de cent occurrences.

¹⁷ Sur la base de vérifications par échantillons. Une recherche plus détaillée serait nécessaire sur ce sujet.

Dans les listes de fréquence de prépositions co-occurentes on trouve aux rangs les plus bas les combinaisons indiquées dans le tableau suivant :

NUNC-IT 1	NUNC-IT 2	NUNC-A	Athenaeum
qui a	qui a	qui a	là di
qui in	qui di	qui in	qui di
qui per	qui in	là di	qui a
qui da	qui per	qui per	qui per
qui di	li a	qui da	qui con

Tableau 6. – Adverbes et prépositions (listes de fréquence)

On procède donc au cas par cas à la recherche d’occurrences capables de trahir les influences du médium ou du registre dans l’emploi des adverbes spatiaux.

Dans tous les *corpora*, le couple *là + di* se trouve à l’intérieur de la locution *al di là di* (qui a, dans la plupart des cas, une valeur métadiscursive, et plus rarement une valeur spatiale, les différences ne semblant pas être significatives par rapport aux axes de variation examinés).

Le couple *li di*, qui n’apparaît pas dans le tableau mais est utile pour une comparaison ponctuelle vs. aréale, est essentiellement utilisé pour les déixis exomédiales, comme c’est le cas dans les exemples (7) (NUNC-IT 1) et (8) (NUNC-A) :

- (7) si intravede la sua insegna, azzurra, su un edificio ormai pericolante, come pericolante è la casa cantoniera li di fianco, e tutto intorno sono scritte nere sul tipo « Viva il duce»

[*on entrevoit son enseigne, bleue, sur un édifice à présent croulant, croulant comme la maison d’à côté, et tout autour il y a des inscriptions noires du genre « Vive le Duce »*].

- (8) cos’è la tua normalità ? certo che non era una giornata ’ delle tante ’, quella li di genova,

[*c’est quoi ta normalité ? c’était sûrement pas une journée « comme les autres », celle de Gênes*].

Dans la plupart des cas, adverbe et préposition ne sont pas corrélés, certains adverbes ont une valeur empathique¹⁸ mais toujours en relation à des éléments étrangers au couple de mots extraits.

¹⁸ Cf. Lyons (1977 : 677) pour la première référence à la déixis empathique. Les adverbes ont une valeur de renfort par rapport aux démonstratifs qui les précèdent et respectent l’opposition étudiée par Glover (2000 : 918) entre (*empathic negotiable orientation* dans les proximales et *received* (et donc non négociable) *orientation* dans les distales.

Le couple *qui/qua per*, à de rares exceptions¹⁹ près, est utilisé pour introduire des subordinées ou des pronoms ; dans les subordinations on se trouve dans la plupart des cas face à des déixis endomédiales, avec les deux adverbes (exemples (9), (10) et (11) tirés de NUNC-IT 1+2, exemples (12), (13) et (14) tirés de NUNC-A) :

- (9) Vuoi litigare o sei qua per discutere di anticomunismo ?
[*Tu cherches la querelle ou t'es là pour discuter d'anticommunisme ?*].
- (10) Scrivo qua per proporvi la creazione di un newsgroup
[*J'écris ici pour vous proposer la création d'un newsgroup*]
- (11) beh raga io ero venuto qui per curiosare un po' da semi-lurker
[*Eh les gars j'étais venu ici pour fureter un peu comme une petit canaille*].
- (12) infatti ho postato qua per saperne di più.
[*en fait j'ai posté ici pour en savoir plus*].
- (13) Per i trolls, la soluzione si articola in due parti. Per prima cosa, dobbiamo renderci conto e capire che essi sono qui per farci del male.
[*Pour les trolls, la solution s'articule en deux parties. Premièrement, on doit se rendre compte et comprendre qu'ils sont là pour nous faire du mal*].
- (14) Questa non è storia, Arduino. Questa è la versione di 'Intimità', o di qualche altro giornale femminile. Non siamo qui per fare del romanticismo, ma per discutere di storia.
[*Ça c'est pas de l'histoire, Arduino. Ça c'est la version d' 'Intimità' ou de n'importe quel autre journal féminin. On n'est pas là pour donner dans le romantisme, mais pour discuter d'histoire*].

Cependant, les emplois logodéictiques des adverbes spatiaux ne manquent pas ((15) et (16) tirés de NUNC-A et (17) de NUNC-IT 1+2) :

- (15) magari si può partire da qua per ottenere l'equazione di Clapeyron ?
[*On peut peut-être partir de là pour obtenir l'équation de Clapeyron ?*].
- (16) Allora, se permetti, parto da qui per riprendere la discussione sui modelli di integrazione.
[*Alors, si tu permets, je pars de là pour reprendre la discussion sur les modèles d'intégration*].
- (17) io non sapevo proprio esistesse il manga di Iketeru Futari, lo ho letto qui per la prima volta
[*je ne savais pas du tout que le manga de Iketeru Futari existait, je l'ai lu ici pour la première fois*].

Naturellement, la logodéixis se prête à certaines ambigüités : il est difficile de savoir dans un cas comme celui de l'exemple (18) s'il s'agit d'une

¹⁹ Dans le cas de citations littérales de textes externes, par exemple.

déixis endomédiale ou liminaire (les textes originaux cités par le locuteur, les *farloccate*, sont dans un autre NG) :

- (18) leggo le farloccate di ***** MaMa ***** e le riporto qui per farvi fare due risate
 [je lis les âneries de ***** MaMa ***** et je les reporte ici pour vous faire rigoler un peu].

La déixis exomédiale est plus fréquente avec l'adverbe ponctuel, mais apparaît essentiellement dans des citations provenant d'autres médias (journaux, sites) et habituellement d'autres types de textes (non pas des posts mais des interviews). Nous avons donc décidé de ne pas la considérer comme importante dans le cadre de la présente étude.

Si l'on accepte la présence de la subordination comme symptôme d'une tendance à la référence endomédiale, il faut considérer que, avec l'adverbe aréal, le pourcentage de groupes adverbe spatial + *per* + verbe tourne autour de 65 %, tandis qu'avec le distal ce pourcentage chute à 45 %. Cette donnée semblerait contredire l'hypothèse selon laquelle les adverbes aréaux sont principalement utilisés pour les déixis endomédiales et les adverbes ponctuels pour les déixis exomédiales.

Le couple *qui/qua di* présente de la même manière une série d'usages prédominants : avec le proximal l'adverbe ponctuel est utilisé dans la locution *qui di seguito* (déixis liminaire) dans 40 % des cas dans NUNC-A et dans 60 % dans NUNC-IT 1+2. En dehors de cette locution, et quand la préposition n'introduit pas de déterminations spatiales ou temporelles (*di seguito, di fronte, di recente*) ou de modificateurs phrasaux (*di certo*), *qui di* est surtout employé pour les déixis endomédiales et liminaires dans des structures du type *qui di X (non) ce n'è (affatto / troppo / niente / proprio)*, où X représente des noms qui indiquent métaphoriquement richesse, opportunité, compétences ou des référents littéraires (*filosofi, donne, link*).

Le couple *qui/qua da* présente un profil plus varié et est difficile à systématiser ne serait-ce que dans ses grandes lignes. On peut toutefois dire que si elles sont précédées du verbe *essere* ou d'un *verbum dicendi* ces deux cooccurrences sont tendanciellement endomédiales – dans ce cas la préposition introduit une détermination temporelle – comme dans les exemples suivants (le (19) tiré de NUNC-IT 1+2 et le (20) de NUNC-A) :

- (19) non leggo e non posto messaggi qui da oltre un mese
 [je ne lis pas et je ne poste pas de messages depuis plus d'un mois].
 (20) Guarda che io discuto qua da anni
 [Figure-toi que ça fait des années que je discute ici].

Naturellement, les exceptions ne manquent pas. Le cas suivant en est un exemple intéressant :

- (21) > Vorrei riportare il sito presso un provider che seghi gli account.

[*Je voudrais rapporter le site chez un fournisseur d'accès qui coupe les comptes*]

- (22) *prego vieni qui da me*
[*stp viens ici chez moi*]

dans lequel la première réplique (21) est un texte cité d'un autre post et la seconde (22) la réponse. Le fait que le verbe vénitif ainsi que l'adverbe proximal ne se réfèrent pas au lieu/*provider* dans lequel se trouve le locuteur est curieux : l'adverbe, qui est ponctuel, a la fonction de contre-présupposition par rapport à un lieu différent préalablement mentionné au cours de l'échange²⁰.

Un rapide échantillonnage de *qui/qua a* et *qui/qua in*²¹ dans les *corpora* tirés de NG²² a enfin démenti l'hypothèse selon laquelle les adverbes aréaux sont de préférence utilisés pour les déixis endomédiales et les adverbes ponctuels pour les déixis exomédiales. Au contraire, dans tous les cas, les déixis exomédiales sont la majorité, comme le montre le Tableau 7 avec les moyennes des pourcentages sur 100 cas pris par échantillon parmi les types de déixis suivants :

	endom	esome	limin
<i>qui in</i>	16,30	27,17	11,96
<i>qui a</i>	4,35	30,43	1,09
<i>qua in</i>	14,13	44,57	5,43
<i>qua a</i>	2,17	30,43	1,09

Tableau 7. – Types de déixis (pourcentages)

5. CONCLUSIONS

Ce bref compte rendu sur l'utilisation de certains déictiques spatiaux a permis de jeter un éclairage nouveau sur un domaine plutôt obscur et généralement ignoré de l'analyse des déictiques online (les études sur la déixis

²⁰ Ce cas est intéressant dans la mesure où le vénitif peut être utilisé offline pour faire référence à un endroit dans lequel on ne se trouve pas au moment de l'énoncé mais dans lequel on se trouve habituellement (le médecin qui dit « *venga da me in studio* » même s'il ne se trouve pas « *in studio* » (dans son cabinet)), mais il en va différemment pour l'adverbe : aucun médecin ne dirait jamais « *venga qui da me in studio* » s'il ne se trouve pas dans son cabinet, alors que l'utilisateur ne se trouve pas dans le *provider* dans lequel il invite son interlocuteur. Ou peut-être que si, étant donné qu'online l'ubiquité est, avec une certaine approximation, possible.

²¹ Les prépositions prises en considération ici sont aussi les plus fréquemment co-occurentes avec le verbe *venire*.

²² Ce relevé n'aurait pas eu de sens dans le *corpus* Athenaeum, qui n'est pas sensible à la distinction entre déixis endomédiales et exomédiales.

online se réfèrent généralement non pas à la communication verbale mais à la communication gestuelle²³).

Nous avons notamment découvert que l'axe de variation diamésique joue un rôle tout aussi significatif que l'axe de variation diaphasique, qui est plus générale, sur l'emploi des démonstratifs : ces derniers sont utilisés plus fréquemment dans des registres plus hauts – pour structurer le discours – et, avec une référence spécifique à la variation diamésique, en cooccurrence avec des termes se rapportant à des segments de la réalité numérique. Les démonstratifs proximaux sont employés par les auteurs pour faire référence à des parties de leur propre texte, tandis que les démonstratifs distaux sont utilisés pour renvoyer à des parties de textes d'autrui ou d'autres secteurs du réseau (sites, autres *threads* ou NG).

Enfin, en ce qui concerne les adverbes de lieu, nous ne sommes pas parvenus à confirmer l'hypothèse selon laquelle les adverbes aréaux sont plus fréquents dans les références à la réalité numérique digitale et les adverbes ponctuels dans les références à la réalité concrète (cf. Allora 2002).

BIBLIOGRAPHIE

- ALGOZINO Elisa (sous presse). Lessico e variazione di registro : un confronto tra i corpora NUNC, LIP e Athenaeum. In : M. Cerruti, E. Corino, C. Onesti (eds), *Formale e informale. La variazione di registro nella comunicazione elettronica*. Roma: Carocci.
- ALLORA Adriano (2002). La deissi in IRC. *Linguistica e Filologia* 15, 61-87.
- ALLORA Adriano (2005). *La deissi on-line*. Ph.D. Thesis.
- ALLORA Adriano (2009). Variazione diamesica generale nelle Comunicazioni Mediate dalla Rete. *Rassegna Italiana di Linguistica Applicata* 3, 147-170.
- ALLORA Adriano (sous presse). Annotazioni sulla sintassi dell'italiano di registro alto nei NewsGroups. In : M. Cerruti, E. Corino, C. Onesti (eds), *Formale e informale. La variazione di registro nella comunicazione elettronica*. Roma: Carocci.
- BARBERA Manuel (2002). Pronomi e determinanti nell'annotazione dell'italiano antico. La POS "PD" del Corpus Taurinense. In : R. Bauer, H. Goebel (eds), *Parallela IX. Testo – variazione – informatica / Text – Variation – Informatik*. Atti del IX Incontro italo-austriaco dei linguisti (Salisburgo, 1-4 novembre 2000) / Akten des IX Österreichisch-italienischen Linguistentreffens (Salzburg, 1.-4. November 2000). Wilhelmsfeld : Gottfried Egert, "Pro Lingua" 35. 35-52.
- BARBERA Manuel, CORINO Elisa, ONESTI Cristina (eds) (2007). *Corpora e linguistica in rete*. Perugia : Guerra Edizioni.

²³ Voir notamment Lonchamp (ics), Lester, Voerman, Towns, Callaway (1999), Hill, Hollan (1991).

- BAUER Roland, GOEBL Hans (eds) (2002). *Parallela IX. Testo – variazione – informatica / Text – Variation – Informatik*. Atti del IX Incontro italo-austriaco dei linguisti (Salisburgo, 1-4 novembre 2000) / Akten des IX Österreichisch-italienischen Linguistentreffens (Salzburg, 1.-4. November 2000). Wilhelmsfeld : Gottfried Egert, “Pro Lingua” 35.
- CERRUTI Massimo (2009). Premesse per uno studio della variazione di registro in italiano. *Rivista Italiana di Dialettologia* 33, 267-282.
- CERRUTI Massimo, CORINO Elisa, ONESTI Cristina (eds) (sous presse). *Formale e informale. La variazione di registro nella comunicazione elettronica*. Roma : Carocci.
- CERRUTI Massimo, ONESTI Cristina (2010). Formal varieties in an Italian News-Groups corpus. Contribution présentée lors de la Fifth Inter-Varietal Applied Corpus Studies (IVACS) group International Conference, 2010, June 18-19, Edinburgh.
- CONTE Maria-Elisabeth (1999/88). *Condizioni di coerenza*. Alessandria: Edizioni dell’Orso, 1999. Nouvelle édition, avec l’ajout de deux essais, édité par Bice Mortara Garavelli, di Maria-Elisabeth Conte, *Condizioni di coerenza. Ricerche di linguistica testuale*. Firenze: La Nuova Italia Editrice, 1988 “Pubblicazioni della Facoltà di Lettere e filosofia dell’Università di Pavia” 46.
- CORNISH Francis (2008). How indexicals function in texts : Discourse, text and one neo-Gricean account of indexical reference. *Journal of Pragmatics* 40, 997-1018.
- CRISTAL, David (2006). *Language and the Internet*. Cambridge: Cambridge University Press.
- HILL William C., HOLLAN James D. (1991). Deixis and the future of visualization excellence. In: *Proceedings of the 2nd conference on Visualization ’91*. Los Alamitos : IEEE Computer Society Press, 314-320.
- GLOVER Karen (2000). Proximal and distal deixis in Negotiation Talk. *Journal of Pragmatics* 32, 519-926.
- KOCH Peter, OESTERREICHER Wulf (1985). Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36, 15-43.
- LESTER James, VOERMAN Jennifer, TOWNS Stuart, CALLAWAY Charles (1999). Deictic Believability: Coordinated Gestures, Locomotion and Speech in Lifelike Pedagogical Agent. *Applied Artificial Intelligence* 13, 383-414.
- LONCHAMP Jacques (sous presse). Deixis in Synchronous CSCL Systems. In: *Proceedings: 3rd International Conference on Computer Supported Education (CSEDU’11)*. Noordwijkerhout, Mai 2011.
- LYONS John (1977). *Semantics*. Cambridge: Cambridge University Press.
- ONESTI Cristina (2010). Varietà formali in rete. Particolarità linguistiche e ricadute glottodidattiche. *Italiano LinguaDue* 1, 206-219.

Corpora et sites de référence

Athenaeum	http://www.bmanuel.org/projects/at-HOME.html
bmanuel.org	http://www.bmanuel.org
corpora.unito.it	http://www.corpora.unito.it/
NUNC	http://www.bmanuel.org/projects/ng-HOME.html
NUNC-A	http://www.progettovalere.org