

Outils Linguistiques

Jean-Yves Antoine

LI - Université Rabelais de Tours

Jean-Yves.Antoine@univ-tours.fr

Mastere L&R – U. Tours



Outils linguistique — Mastere L&R — © J.Y. Antoine — 1

Outils Linguistiques

LANGAGE PARLE

Mastere L&R – U. Tours



Outils linguistique — Mastere L&R — © J.Y. Antoine — 2

Qu'est-ce que le langage parlé

- **Le point de vue du grammairien**
 - compétence / performance
 - l'oral : un langage fautif, non digne d'intérêt
- **Le point de vue du linguiste (de corpus)**
 - oral : témoin direct de la production langagière
 - oral : indice de l'évolution diachronique du langage
 - bref, un objet d'étude à part entière...
- **Ecrit / Oral : quelles différences réelles ?**

Outils linguistique — Mastere L&R — © J.Y. Antoine — 3

Texte libre

Loi internationale : des règles à inventer. Depuis la fin de la guerre froide, qui paralysait le système de l'ONU, la communauté internationale paraît à la recherche des règles qui doivent présider aux interventions répressives, qu'il s'agisse de sanctions économiques, d'intervention militaires, ou encore de sanctions pénales. Ces interventions contrevenaient à un principe de base énoncé par la Charte des Nations Unies : celui de la souveraineté des États et de la non-ingérence dans les affaires des autres.

[Le Monde Dossiers et Documents, 2002]

1 - Sur la base :

- branchez le cordon ligne, puis le bloc d'alimentation secteur (dans une prise de courant facilement accessible) comme décrit ci-contre. Le voyant secteur s'allume.

2 - Sur le mobile :

- ouvrez la trappe d'accès à la batterie en pinçant de chaque côté (1) puis en soulevant.

[Documentation Matra Communication Sensoria 220]

Corps parfumés, tièdes, fermes. Tous baisés, donnés : dans les prés profonds de l'été, herbes couchées enchevêtrées, dans les couloirs suintant de maisons de pauvres, sur des divans, des lits qui craquent. Jack, mon amour ! Chérie ! Embrasse-moi, Reggy ! Mon petit ! Mon amour ! Le cœur en branle il poussa la porte du restaurant Burton. L'odeur le saisit à la gorge : sauces de viandes pénétrantes, lavasses de légumes verts. Le repas des fauves. Des hommes, des hommes, des hommes.

[James Joyce, Ulysse, traduction Auguste Morel, Gallimard, 1957]

un foie, deux reins, trois raisons de boire Contrex

[Message publicitaire]

Outils linguistique — Mastere L&R — © J.Y. Antoine — 4

Parole « brute »

- **Monologue** [Blanche-Benveniste, 1997]

et donc il y avait également sans doute et qui va avec tout ce que je viens de dire avant il y avait aussi un homme mythique complètement mythique et c'est la / première et là en plus c'est intéressant parce que c'est la première image que j'ai eue de la Tchécoslovaquie lorsque je suis arrivé à la frontière je me rappelle il y avait une brume c'était pratiquement en fin de journée / à la frontière entre l'Allemagne et la Tchécoslovaquie dans les euh dans les guérites des douaniers on imaginait à travers la brume on imaginait encore des des bunkers des de tout ce qui avait été supprimé très rapidement mais sans enlever toutes les infrastructures / c'est à dire on sentait encore le poids architectural de cette frontière qui avait été fermée pendant tant d'années / et là dans cette dans ces guérites de douaniers / il y avait partout étaient placardées donc des photos de Vaclav Havel et ça c'était un un point important parce qu'on voyait dès le départ dès l'arrivée sur le territoire tchécoslovaque donc on voyait l'image de cet homme

Outils linguistique — Mastere L&R — © J.Y. Antoine — 5

Dialogue oral

- **Dialogue sans chevauchement** [Blanche-Benveniste, 1997]

- (1) *c'est certain je veux dire je je c'est vrai que*
- (2) *je sais pas non pour le moment je je me sens concerné parce que bon c'est vrai euh mais bon euh*
- (1) *non c'est parce que tu es déçu par ce qu'on t'offre en face mais euh je crois qu'il faut il faut faire cet effort je veux dire moi je te dis j'ai assisté à plein de réunions bon j'ai été très déçue effectivement mais je crois qu'il faut faire cet effort justement d'aller écouter les gens les gens de tout bord tu sais*

Outils linguistique — Mastere L&R — © J.Y. Antoine — 6

Dialogue oral

• Dialogue avec chevauchement (corpus OTG [Nicolas *et al.*, 2002])

- <01> c : j'aimerais avoir les programmes de la DEL # s'il vous plaît
<02> h : de la DEL je crois qu'on les reçoit pas
<03> c : ah ouais
<04> h : oui en fait le problème c'est qu'ils nous les amènent pas à l'accueil
<05> c : hum
h : par
<06> h : moment # c'est irrégulier
<07> c : et là vous l'avez pas
<08> h : non non non je veux dire ils nous l'amènent pas là hein
<09> c : d'accord
h : ils
<10> h : le déposent quand ils les amènent ils les déposent discrètement là sur "servez-vous" mais on n'est pas au courant de leur
<11> c : d'accord ben apparemment il y a
h : régularité je dirais # c'est
<012> c : rien
<013> h : apparemment il y a rien alors c'est qu'ils nous les ont pas

Outils linguistique — Mastere L&R — © J.Y. Antoine — 7

Dialogue oral

• Dialogue avec chevauchement (suite)

- <014> c : d'accord sinon
h : eus
<015> c : vous savez pas où on peut les trouver
<016> h : à DEL j'ai peut être leurs coordonnées # 3
<017> c : c'est ouais les coordonnées je les ai c'est bon c'est
<018> h : ah
c : pour
<019> c : savoir s'il y avait des présentoirs particuliers donc e on m'a dit bibliothèque j'irai voir
<020> h : oui c'est vrai les bibliothèques aussi
<021> c : d'accord
h : et c'est
<022> h : toujours pareil hein si on leur dit pourtant on leur dit # passez là vous en mettez là mais vous en passez là comme ça
<023> c : ils ne le font pas
h : on est au courant quand
<024> h : vous les non et [pf] on les voit même pas passer
<025> c : ok et sinon e si vous aviez de de la documentation sur les musées de Grenoble

Outils linguistique — Mastere L&R — © J.Y. Antoine — 8

Dialogue oral

• Dialogue avec chevauchement (suite)

- <026> h : bien sûr
c : tout ce
- <027> c : qui est # de Grenoble et de l'Isère si c'est possible
- <028> h : vous les aurez tous là
- <029> c : super # et est ce que je pourrais aussi vous demander cette brochure là e
- <030> h : laquelle
c : la bleue
- <031> c : les musées
- <032> h : départementaux
c : départementaux
- <033> c : voilà
- <034> h : bien sûr mais là c'est "visite guidée en Isère"
- <035> c : ouais aussi ouais # merci beaucoup
- <036> h : je vous en prie
- <037> c : au revoir
h : au revoir

Outils linguistique — Mastere L&R — © J.Y. Antoine — 9

Genres

• Genre ou registre (*register*) [Halliday 1985 ; Biber 1986, 1988]

- il n'existe pas un seul type d'écrit ou un seul type d'oral, de même qu'il y a parfois moins d'écart entre deux productions orale et écrite qu'entre deux oraux (ou deux écrits) différents
- les productions linguistiques relèvent d'un continuum de genres (ou registres)

• Principaux registres de langues [Biber *et al.* 1999]

- **deux modalités** écrit ou oral
- **quatre genres** dialogue interactif, fiction, langage journalistique, prose académique (typologie de base pouvant être affinée)

• Registre et situation discursive [Gadet 1999]

- le registre ne se limite pas uniquement à la situation discursive : une personne peut changer sans peine de registre en cours de discours (variabilité diaphasique)

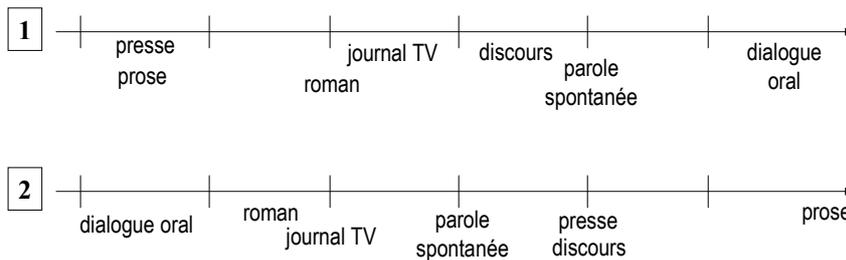
Outils linguistique — Mastere L&R — © J.Y. Antoine — 10

Genre et registre

• Exemple d'analyse de corpus

[Biber 1986]

- Analyse factorielle discriminante suivant 41 indices, sur 545 textes différents.
- Facteur 1 information exprimée de manière explicite vs. implicite
- Facteur 2 style élaboré et abstrait vs. informel et concret



Outils linguistique — Mastere L&R — © J.Y. Antoine — 11

Langage parlé

• Hétérogénéité du langage parlé

- **plusieurs registres** : dialogue interactif, interview, monologue, discours ou présentation orale, journal TV, théâtre...
- **facteur discriminant important** : parole spontanée vs. oral préparé

• Ce que n'est pas le langage parlé

- un langage fautif par rapport à l'écrit
- un langage moins complexe que l'écrit

[Poole, Field 1976; Chafe 1982] longueur supérieure des énoncés écrits
[Chafe 1982] plus de subordonnées à l'écrit

MAIS

[Poole & Field 1976] autant de subordonnées à l'oral qu'à l'écrit
[Halliday 1985] plus grande complexité de l'oral : incises, parenthèses

Outils linguistique — Mastere L&R — © J.Y. Antoine — 12

Parole spontanée

- **Situations discursives**

dialogue, interview, monologue

- **Dialogue oral**

- conversation en face à face ou téléphonique
- **parole spontanée** : registre très spontané car situation interactive
- deux sources de spontanéité

planification en direct
dialogue

traces du processus de production
improvisation sous la pression de l'interlocuteur

Outils linguistique — Mastere L&R — © J.Y. Antoine — 13

Parole spontanée : procédés

- **Procédés spécifiques de l'oral spontané**

- traces du processus de production en direct de l'énoncé
- reflètent une recherche de dénomination ou d'adaptation en direct à la situation (réaction de l'interlocuteur ou autre)
- perturbent la régularité syntaxique de l'énoncé : problèmes en TAL

reconnaissance de parole 5% de taux d'erreur de mots en parole lue (*broadcast news*)
30% au mieux en parole spontanée (*switchboard*)

- **Hésitations**

- pauses remplies (*eah, mmh, ...*) vs. pauses vides
- problème pour la reconnaissance de parole

Outils linguistique — Mastere L&R — © J.Y. Antoine — 14

Reprises ou réparations (*repairs*)

• Répétitions

- **répétition simple** arrêt syntagmatique du à la recherche de dénomination
*ce qui était fantastique dans ce **dans ce** camping*
- **enrichissement lexical** : reprise précisant la dénomination en cours d'élaboration
*c'est le voisin **le voisin de gauche Mr Pichu** qui demande si Pierre est là
j'ai eu mon mon brevet **le diplôme de fin de collège là le brevet des écoles**
cà fait un comment on dit je sais plus **un chose là une écharpe un col roulé***

• Auto-corrrections

- rattrapage de dénomination jugée erronée, imprécise ou hors des attentes dialogiques
- auto-corrrections marquées (*non, enfin, ...*) ou pas : *reparandum*
*c'est le Papi de Lourmarin qui me l'a donné non **qui me l'a offert aussi**
dans ce **dans ce camping dans ce dans cet** hôtel*

Outils linguistique — Mastere L&R — © J.Y. Antoine — 15

Réparations et syntagme

• Arrêt syntagmatique

- le plus souvent, la reprise se traduit par une pause sur l'axe syntagmatique
- la réparation est réalisée à partir du début du syntagme courant
*je cherche un petit ensemble un petit ensemble pour ma fille de 15 ans
* je cherche un petit ensemble petit ensemble pour ma fille de 15 ans
* je cherche un petit ensemble ensemble pour ma fille de 15 ans*

• Anticipation

- la reprise peut concerner des éléments tenus en réserve et prononcés trop tôt
ils faisaient / ils disaient à ma mère de faire attention

• Retour en arrière

- le locuteur peut cependant revenir en arrière sur l'axe syntagmatique pour se corriger
*quand on remonte dans une voiture avec des g- des gens qu'on ne connaît pas ou même
qu'on connaît d'ailleurs et euh quand c'est pas nous qui prenons le volant donc*

Outils linguistique — Mastere L&R — © J.Y. Antoine — 16

Analyse en grille

- **GARS - DELIC** [Blanche-Benveniste 1991]

représentation des énoncés en plaçant les éléments paradigmatiques à la verticale

- **Exemple**

quand on remonte dans une voiture avec des g- des gens qu'on ne connaît pas ou même qu'on connaît d'ailleurs et euh quand c'est pas nous qui prenons le volant donc

- **Analyse en grille**

*quand on remonte dans une voiture avec des g-
des gens qu'on ne connaît pas
ou même qu'on connaît
quand c'est pas nous qui prenons le volant donc et euh*

Outils linguistique — Mastere L&R — © J.Y. Antoine — 17

Réparations

- **Étude de corpus** [Lettelier *et al.*, 1999]

- **Dialogue oral homme-homme** réel ou simulé
- **corpus** : Air France (réservation aérienne), Murol (renseignement touristique), Levelt (conception de dessin architectural)
- **observation** : taux moyen de tours de parole présentant une répétition ou une auto-correction (écart-type calculé par dialogue)

	Air France	Murol	Levelt
Freq ^{ce} moy.	9,3 %	34,8 %	13,7 %
écart-type	6,6 %	3,3 %	10,2 %

- **Influence du degré d'interactivité du dialogue**

Outils linguistique — Mastere L&R — © J.Y. Antoine — 18

Autres phénomènes

- **Procédés non spécifiques à l'oral**
 - se retrouvent dans tous les registres oraux ou écrits
 - fréquence plus importante en parole spontanée et plus encore en dialogue oral du fait de la pression interactive
- **Ellipses, parataxes et autres raccourcis d'expression**
- **Incises et autres parenthèses**
- **Extractions et autres modifications de l'ordre linéaire**
 - motivation : mise en relief d'un élément dans l'énoncé
 - variabilité forte (structure syntaxique discontinue) ou faible
 - plusieurs procédés

Outils linguistique — Mastère L&R — © J.Y. Antoine — 19

Extractions

- **Extraction simple**
 - simple mouvement d'un élément de l'énoncé
 - **antéposition** *dans l'armoire du haut tu trouveras des verres à pied*
 - **postposition** *elle a demandé à Pierre tout de suite son adresse*
- **Dislocation (double-marquage)**
 - rappel (généralement par un pronom) de l'élément régi déplacé
 - **antéposition** *les moules j'ai jamais pu en manger*
le facteur il passe jamais sans me saluer
 - **postposition** *j'ai jamais pu en manger de moules*
je préfère la jaune de robe
- **Structure à présentatif**
 - **clivées (cleft)** *c'est Jean qui sonne à la porte / il y a Jean qui sonne*
 - pseudo-clivée *celui qui sonne c'est Jean / ce que tu veux c'est tout à la fois*

Outils linguistique — Mastère L&R — © J.Y. Antoine — 20

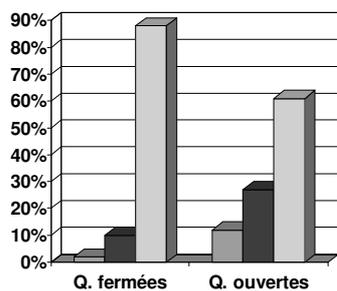
Extraction et interrogation



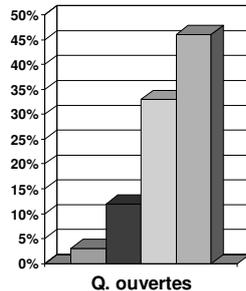
- **Interrogation**

- par inversion viens-tu ? inversion
- par est-ce-que est-ce que tu viens ? SVO respecté
- par intonation tu viens ? SVO respecté

- **Observation sur corpus oral**



[Gadet 1989]



[Valdman 1982]

Outils linguistique — Mastere L&R — © J.Y. Antoine — 21

Extraction: quelle spécificité de l'oral

- **Ordre linéaire** ordre SVO privilégié en français [Gadet 1989]

- **Variabilité idiomatique**

Le chien regarde le chat
The dog sees the cat

mais...

Sobaka vidit koshku
Sobaka koshku vidit
Vidit sobaka koshku
Vidit koshku sobaka
Koshku vidit sobaka
Koshku sobaka vidit



- a : morphème du nominatif
 - u : morphème accusatif

Outils linguistique — Mastere L&R — © J.Y. Antoine — 22

Langages à ordre variable

- Morphologie \Leftrightarrow Variabilité linéaire
- Variabilité de l'ordre linéaire [Covington 1990]

– aucune ou très peu



– variabilité limitée



– forte variabilité



– variabilité extrême

Walbiri (aborigènes d'Australie)

Outils linguistique — Mastere L&R — © J.Y. Antoine — 23

Ellipses

- Raccourci qui peut également concerner l'absence d'un élément au sein d'un syntagme, contrairement aux énoncés incomplets

<i>je prendrai le premier [train] de l'après-midi</i>	contexte discursif
<i>je cherche un [hôtel] deux étoiles au centre-ville</i>	univers du discours
<i>j'ai deux [réservations en] classes affaires sur Agadir</i>	univers du discours
<i>et pareil pour le retour sur Paris</i>	contexte discursif
<i>tu as gagné une fois moi [j'ai gagné] trois fois</i>	contexte discursif
<i>Paul est déjà allé à Paris, moi [j'y suis allé] aussi / jamais</i>	contexte discursif

- Se retrouve dans tous les genres et pas uniquement à l'oral

*je t'aimais inconstant, qu'aurais-je fait [si tu avais été] **fidèle** !* [Racine]

- Ellipses marquées (*pareil, idem, aussi, jamais, non ...*) vs. non marquées

Outils linguistique — Mastere L&R — © J.Y. Antoine — 24

Commentaires : incise et apposition

• Incise

- partie de l'énoncé, présente à quelque endroit que ce soit, qui ne relève pas directement du message principalement véhiculé par l'énoncé
- très fréquent à l'oral spontané, mais non limité à ce registre

- **digressions ou marques d'hésitations**

*bonjour **qu'est-ce que je veux dire / e** / vous avez des vols pour Agadir (AF II.4.1)*

*auriez-vous deux places Paris Nice sur Air France **ben attendez j'ai même pas marqué le numéro de vol le 23 janvier déjà le temps que je cherche le numéro j'ai tout noté sauf les numéros de vol ...** (AF I-30)*

- **commentaires pouvant aider à préciser l'intention du locuteur**

*quel est le prix d'un voyage **en avion évidemment** de Paris à Bruxelles (AF I.9)*

***ça se visite pas le château de Vizille** donnez-moi **e Longpra c'est fait Virieu c'est fait** qu'est ce qui a comme autres châteaux (OTG)*

Outils linguistique — Mastere L&R — © J.Y. Antoine — 27

Commentaires : incise et apposition

• Apposition [Blasco & Caddeo 2001]

- juxtaposition d'un syntagme ou d'une proposition à un nom ou pronom qu'il qualifie

***Paris, capitale de la France,** accueille chaque année plusieurs millions de visiteurs*

- apparaît fréquemment avec une dislocation avec double-marquage pronominal

***moi, le second de ce navire,** j'ai été traité comme un chien ! [Verne, Révoltés de la Bounty]*

*et **elle, Mme Arnoux,** comment **la** revoir maintenant ? [Flaubert, L'Éducation Sentimentale]*

- se retrouve dans tous les registres oraux et écrits

*certains ont été jusqu' à **lui** reprocher une fictive collaboration avec la Stasi - à **lui, le dissident du régime des Ulbricht et Honecker** ! [Le Monde Diplomatique, oct. 1995]*

*pourquoi **lui Français** il aurait pas besoin d'un dico [oral adulte, corpus Caddeo]*

Outils linguistique — Mastere L&R — © J.Y. Antoine — 28

Bibliographie

- **Articles ou ouvrages d'entrée**

- **Biber D.** (1986) Spoken and written textual dimensions in English : resolving the contradictory findings, *Language*, 62(2), 384-414.
- **Blanche-Benveniste C.** (1997) *Approches de la langue parlée en français*. Collection L'essentiel français. Ophrys, Paris, France.
- **Gadet F.** (1992) *Le français populaire*, PUF, Paris, France.

- **Articles ou ouvrages de référence**

- **Biber D.** (1988) *Variations across speech and writing*. Cambridge University Press.
- **Biber D. et alii** (1999) *A grammar of spoken and written English*. Longman, London, UK.
- **Blanche-Benveniste C., Bilger M., Rouget C., van den Eynde K.** (1991) *Le français parlé : études grammaticales*. CNRS Editions, Paris, France.
- **Gadet F.** (1989) *Le français ordinaire*. Armand Colin, Paris, France.
- **Halliday M. A. K.** (1985) *Written and spoken language*. Oxford University Press. Oxford.

Outils linguistique — Mastere L&R — © J.Y. Antoine — 29

Bibliographie (2)

- **Travaux cités**

- **Abeillé A., Blache P.** (2000) Grammaires et analyseurs syntaxiques. In Pierrel J.-M. (Ed.), *Ingénierie des Langues*. Hermès, Paris, France. 51-76.
- **Antoine J.-Y., Goulian J.** (2001) Etude des phénomènes d'extraction en français parlé sur deux corpus de dialogue finalisé. *Traitement Automatique des Langues*. Hermès, Paris, France. 42(2). 413-440.
- **Ashby W.** (1981) The loss of the negative particle *ne* in French, a syntactic change in progress. *Language*, 57.
- **Blasco-Dubbelco M., Caddeo S.** (2001) Apposition et dislocation : la séquence pronom + lexique + clitique. *RSFP, Recherches sur le français parlé*, 16, 125-149.
- **Chafe W.L.** (1982) Integration and involvement in speaking, writing, and oral literature. *Language*. 58, 1-21.
- **Covington M. A.** (1990) A dependency parser for variable-word-order languages. *Research report AI-1990-01*, U. of Georgia, Athens, Georgia, USA.

Outils linguistique — Mastere L&R — © J.Y. Antoine — 30

Bibliographie (3)

- **Travaux cités**

- **Gadet F.** (1999) La variation diaphasique en syntaxe. In Barbiéris J.-M. (Ed.) *Le français parlé: variation et discours. Praxiling.* U. Montpellier III, Montpellier, France. 211-228.
- **Letellier-Zarshenas S., Nicolas P., Goulian J., Antoine J.-Y.** (1999) Inattendus structurels et communication orale finalisée : influence de la tâche et du contexte interactif. Actes *Journées Internationales de Linguistique Appliquée, JILA '1999*, Nice, France.
- **Nicolas P., Letellier S., Schadle I., Antoine J.-Y., Caelen J..** (2002) Towards a large corpus of spoken dialogue in French that will be freely available : the « Parole Publique » project and its first realisations. Actes *3rd International Conference on Language Resources and Evaluation, LREC '2002*. Las Palmas de Gran Canaria, Espagne.
- **Poole M.E., Field T.W.** (1976) A comparison of oral and written code elaboration. *Language and Speech.* 19, 305-311.
- **Valdman A.** (1982) Français standard et français populaire : sociolectes ou fictions ?, *The French review*, 2.