



## Français parlé : données, représentations, questionnements théoriques

Turin, 16-17 juin 2022

Colloque international organisé par les Universités de Turin et Sorbonne Université, avec la participation du Do.Ri.F.-Università (Centro di documentazione e di Ricerca per la Didattica della Lingua Francese nell'Università Italiana) et de la CRL (Cellule de Recherche en Linguistique)

Après une époque pionnière, où l'étude de l'oral était confrontée à un déni de légitimité et à l'absence de cadres de référence précis et d'outils d'analyse – technologiques et conceptuels – adaptés, force est de constater que, malgré tout, le français parlé demeure encore assez peu représenté dans les ouvrages grammaticaux de référence. À l'heure où la *Grande Grammaire du Français* voit enfin le jour, il est donc légitime de revenir sur les représentations de l'oral, sur sa place et son rôle dans la théorie linguistique et dans la description du système unitaire de la langue française. En effet, si l'adoption d'une démarche empirique à partir de l'usage réel ne soulève pas d'objection pour les langues sans tradition grammaticale, pour lesquelles la norme, établie par les linguistes de terrain, correspond au fonctionnement « spontané » du système, il en est tout autrement pour des langues à la longue tradition grammaticale, comme le français. Pour ces langues, l'établissement de la description grammaticale est perpétuellement pris entre l'enclume et le marteau d'une double normativité : la norme interne au système lui-même et la norme externe, comme résultante de la stratification historique d'exigences émanant d'instances hétérogènes, institutionnelles, éducatives, esthétiques et sociétales. Les vœux pour une grammaire basée sur l'usage écrit et oral, sur le modèle de ce qu'ont fait Biber et ses collaborateurs pour l'anglais en 1999 (*cf.* Bybee 2006) butent donc tout à la fois contre trois obstacles majeurs, en plus du problème de légitimité dont il a déjà été question. Premièrement, les problèmes économiques et techniques liés à la constitution et à l'exploitation de grands corpus de référence, par rapport auxquels plusieurs projets essaient de donner une réponse ces dernières années (nous ne citerons ici, à titre d'exemple, que le projet Orféo, centralisant plusieurs corpus écrits et oraux, mais dont la taille est encore loin d'être comparable à celle des corpus anglophones BNC ou COCA). Deuxièmement, les problèmes épistémologiques liés à la possibilité d'appliquer le même cadre conceptuel et terminologique utilisé par les ouvrages courants ou de le remettre à plat suite à l'intégration des données orales. Le renouvellement du cadre d'analyse pose corollairement le problème de la création d'une terminologie « neutre ». En effet, l'abandon d'une terminologie héritée de la tradition, peu adaptée à la description de l'écrit lui-même, entraînerait néanmoins des problèmes de lisibilité pour un public non expert (comme cela s'est produit avec l'opus de Damourette et Pichon, qui intégrait déjà un très grand nombre de données orales), ainsi que des conflits entre approches linguistiques en concurrence. Troisièmement, le problème statistique de la représentativité des données orales et des conséquences sur la description qui en résulte. En effet, la fréquence de certaines structures est liée aux genres de parole de l'oral et aux dimensions de la variation linguistique. La dichotomie saussurienne langue/parole est présente ici de façon cruciale et pose la question de savoir si une telle description pourrait encore prétendre au statut de grammaire de la langue ou ne devrait pas plutôt être qualifiée de grammaire de la parole.

Sans exclure d'autres propositions concernant les multiples dimensions d'analyse de l'oral, les axes privilégiés par ce colloque visent donc prioritairement :

## 1) la représentation de l'oral :

- **Dans la transcription, qui effectue une sorte de pré-analyse implicite et qui conditionne la perception de phénomènes étudiés**
- **Dans les œuvres de fiction (œuvres littéraires, théâtre, cinéma, etc.)**
- **Dans le discours métalinguistique des non-experts (la « folk-linguistique »), en termes de connotations associées à la langue orale par ses utilisateurs**

Si la possibilité même d'une analyse de l'oral passe par sa transcription, c'est-à-dire sa représentation écrite, il est certain que les choix relatifs à la manière de transcrire et de représenter cet oral ont des conséquences sur les observables, dans la mesure où la transcription constitue déjà, par elle-même, une forme d'analyse, qui retient certains éléments et en laisse de côté d'autres. L'adoption de normes de transcription particulières rend souvent les corpus difficilement exploitables et la constitution de grands corpus informatisés rend le problème encore plus évident : difficultés liées à la normalisation orthographique, à la lemmatisation, à la multitranscription, à l'annotation morpho-syntaxique, prosodique et des modalités de phrase. Des représentations pluridimensionnelles combinant la transcription alphabétique et la disposition spatiale des segments transcrits permettent par exemple de faire ressortir certaines composantes propres à l'élaboration de l'oral : mise en grilles pour le croisement des axes syntagmatique et paradigmatique (Blanche-Benveniste et al. 1979) ; grilles prosodiques et mimogestuelles (par exemple Morel & Danon-Boileau, 2001, Bouvet & Morel, 2002) pour la multimodalité. Dans quelle mesure la transcription participe-t-elle de la construction de l'objet d'analyse ? Quelles sont les conséquences de ces choix de représentation et comment ces problèmes sont-ils gérés à l'heure actuelle ? Quelles solutions pour la normalisation et l'automatisation de ces tâches ?

Parallèlement, la représentation de l'oral se fait aussi de manière plus ou moins fantasmée, à travers la représentation fictionnelle du théâtre, du cinéma, de la chanson et de la littérature, qui oscille entre la mimésis et la re-création plus ou moins déformante, avec des phénomènes bien connus de stéréotypisation à partir d'un nombre réduit d'éléments dont la fréquence est artificiellement exagérée et dont le degré de représentativité est discutable. Quels sont les éléments récurrents de cette stylisation, contribuant à l'émergence d'un stéréotype du français parlé ?

Un troisième palier de représentation de l'oral concerne plutôt l'imaginaire linguistique et, plus largement, la « folk linguistique ». Il s'agit, dans ce cas, des connotations associées à la langue parlée, oscillant entre valeurs positives et négatives. D'une part, les valeurs positives de la spontanéité, de la fraîcheur voire de la langue verte, ou encore de l'éloquence qui, après une longue période où elle avait été ignorée, voire décriée, est aujourd'hui remise à l'honneur à la faveur des formations d'infocom et de concours d'éloquence plus ou moins médiatisés. D'autre part, les valeurs négatives pointant le caractère approximatif, voire fautif, attribué à la langue parlée en tant que telle, qu'on qualifie très souvent en lui associant les catégories sociolinguistiques (diaphasiques et diastratiques) des productions basilectales (langue familière, populaire, vulgaire...). Quel genre de productions orales étudie-t-on ? Quelle image du français parlé se dégage-t-il de ces études ?

## 2) Les conséquences, sur la description grammaticale courante, de l'intégration des données orales à l'ensemble des observables et l'extension du domaine de la syntaxe au-delà des rections de catégories

Les catégories d'analyse traditionnelles, telles que les parties du discours, ainsi qu'une tradition de description grammaticale centrée sur la parole écrite, sont de plus en plus souvent interrogées et mises à l'épreuve par l'analyse de l'oral, qui manifeste le caractère partiel de ces analyses, au niveau morphologique et syntaxique. Si la morphologie bénéficie en effet depuis longtemps des études sur les langues non écrites, la syntaxe, quant à elle, doit se confronter, d'une part, aux phénomènes marginaux, exclus de la description et, d'autre part, aux frontières du domaine de la rection de catégories. L'empan traditionnel de la phrase, notion reposant sur des routines communicatives basées sur la ponctuation et la typographie, est en effet remis en cause dès qu'on abandonne le médium écrit. Les notions de construction verbale (syntaxique), d'énoncé (pragmatique) et de structure thème-rhème (informationnelle), souvent utilisées dans les descriptions de l'oral, sont partiellement incommensurables par rapport aux notions de phrase et de discours et remettent en question le bien-fondé de la séparation entre domaine syntaxique et domaine sémantico-pragmatique. Les questions de cohésion et de

cohérence, à la faveur de phénomènes tels que la deixis, la dislocation ou la corrélation, ainsi que les questions d'incidence (adverbes de phrase et d'énoncé, par exemple, et, plus généralement, le statut des constituants non régis), interrogent profondément la frontière entre ce qui relève de la syntaxe réactionnelle et ce qui n'en relève pas. Les propositions et les recherches autour de la notion de macrosyntaxe montrent bien l'exigence d'intégrer certains phénomènes abondamment attestés en français parlé et traditionnellement rejetés hors du champ de la connexité syntaxique, dans la grammaire plutôt que dans la rhétorique, voire de remettre en cause le cloisonnement absolu entre langue et discours. Quels phénomènes de l'oral sont encore peu ou mal décrits ? Quels phénomènes pourraient recevoir un éclairage nouveau grâce à l'apport des données orales ? Où en est-on des recherches sur la macrosyntaxe et, plus largement, sur les frontières de la syntaxe ?

### **Calendrier de soumission**

**20 janvier** : diffusion de l'appel à communication

**15 mars** : date limite pour l'envoi des propositions de communication

**31 mars** : notification aux auteurs de la décision du comité scientifique

**16-17 juin** : colloque

Les propositions de communication (une page maximum, sans compter la bibliographie éventuelle) devront être anonymes et soumises à l'adresse suivante :

[francparle2022@gmail.com](mailto:francparle2022@gmail.com)

Les auteurs seront avertis au plus tard le 31 mars 2022 de la sélection opérée par le comité scientifique.

Une publication d'articles sélectionnés est prévue.

### **Déroulement du colloque : Turin, 16-17 juin 2022**

Ce colloque bénéficie du soutien du département LLC (Lingue e Letterature straniere e Culture moderne) de l'Université de Turin. Il constitue le premier volet du projet de recherche « Français parlé : descriptions, didactiques, traductions »

**Comité d'organisation** : Ruggero Druetta (Université de Turin), Maria Margherita Mattioda (Université de Turin), Patricia Kottelat (Université de Turin), Carole Conti (Université de Turin), Claire Martinot (Sorbonne Université), Sonia Gerolimich (Université de Udine), Isabelle Stabarin (Université de Trieste), Giulia D'Andrea (Université de Turin)

**Comité Scientifique** : en cours de constitution

### **Références bibliographiques citées dans le texte de l'appel**

ABEILLÉ, Anne, GODARD, Danielle, 2021, *La grande grammaire du français*, Paris, Actes Sud – Imprimerie Nationale Éditions

BIBER, Douglas, Stig JOHANSSON, Geoffrey LEECH, Susan CONRAD & Edward FINEGAN, 1999, *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited

BLANCHE-BENVENISTE Claire *et al.*, 1979. « Des grilles pour le français parlé », *Recherches sur le français parlé*, 2, pp. 163-206.

BOUVET, Danielle, MOREL, Mary-Annick, 2002, *Le ballet et la musique de la parole*, Paris-Gap, Ophrys.

BYBEE Joan, 2006, "From usage to grammar: the mind's response to repetition", *Language*, 82(4), 711–733.  
<http://www.jstor.org/stable/4490266>

DAMOURETTE, Jacques, PICHON, Edouard, 1930-1950, *Des mots à la pensée. Essai de Grammaire de la Langue Française*, Paris, Editions d'Artrey.

MOREL, M.-A. & L. DANON-BOILEAU, 1998, *Grammaire de l'intonation. L'exemple du français oral*, Paris : Ophrys.

### **Corpus, Banques de données et plateformes**

BNC: The British National Corpus, version 2 (BNC World). 2001. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>

CFPP2000 : Corpus de français parlé parisien des années 2000. <http://cfpp2000.univ-paris3.fr/>

CFPB : Corpus de Français Parlé à Bruxelles. <http://cfpp2000.univ-paris3.fr/cfpb.html>

CFPQ : Corpus de français parlé au Québec. <https://applis.flsh.usherbrooke.ca/cfpq/>

CLAPI Corpus de LAngue Parlée en Interaction. <http://clapi.ish-lyon.cnrs.fr/>

COCA: Corpus of Contemporary American English: <https://www.english-corpora.org/coca/>

Cocoon : COLLECTIONS de CORPUS ORAUX NUMÉRIQUES. <https://cocoon.huma-num.fr/exist/crdo>

C-PROM : Corpus libre de parole multigenre. <https://sites.google.com/site/corpusprom/>

Elicop : Étude linguistique de la communication parlée. [https://limo.libis.be/primos-explore/fulldisplay?docid=LIRIAS1797250&context=L&vid=Lirias&search\\_scope=Lirias&tab=default\\_tab&lang=en\\_US&fromSitemap=1](https://limo.libis.be/primos-explore/fulldisplay?docid=LIRIAS1797250&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US&fromSitemap=1)

ESLO: Enquêtes sociolinguistiques à Orléans. <http://eslo.huma-num.fr/index.php>

FONC : French Oral Narrative Corpus. <http://frenchoralnarrative.qub.ac.uk>

MPF (Multicultural Paris French) <https://www.ortolang.fr/market/corpora/mpf>

OFROM : Corpus oral de français de Suisse Romande. <http://www11.unine.ch>

Orfeo : <https://repository.ortolang.fr/api/content/cefc-orfeo/11/documentation/site-orfeo/index.html>

ORTOLANG: Outils et Ressources pour un Traitement Optimisé de la LANGue. <https://www.ortolang.fr/>

TCOF : Traitement de corpus oraux en français. <http://cnrtl.fr/corpus/tcof/>