LiDiFra
Linguistique, Didactique, Francophonie

DESCILAC
Département des Sciences du Langage et de la Communication

UNIVERSITÉ DE ROUEN

JOURNEE D'ETUDES

Ouverte aux étudiants de Master et de Doctorat

Vendredi 13 février 2009 - Université de Rouen

LES DONNEES EN LANGUE : DESCRIPTION, MODELISATION ET APPRENTISSAGE

Qu'il s'agisse de décrire et de comprendre le fonctionnement des langues ou d'en étudier l'acquisition et l'apprentissage, le chercheur, comme l'enseignant, est amené à se positionner vis-à-vis des données linguistiques : *quelles doivent être ces données ? Comment les recueillir ? Comment les modéliser ? Comment les utiliser avec des apprenants ?*

Jusque dans les années 80, la linguistique était marquée par une relative stabilité dans son rapport aux données langagières, quelle que soit l'école dominante (structuralisme, énonciativisme, générativisme…). Depuis cette époque, le développement d'outils informatisés plus puissants, ainsi que l'influence grandissante des démarches expérimentales, notamment à travers la psycholinguistique, a conduit à un tournant conceptuel et méthodologique sur cette question. Les approches formelles classiques ont ainsi cédé du terrain à des approches dans lesquelles l'induction, la recherche d'assises substantielles (acoustiques, neurophysiologiques, etc.) à la modélisation des phénomènes langagiers, ainsi que le traitement automatisé et statistique de larges corpus figurent au premier plan des préoccupations des chercheurs.

Quel a été l'impact de ce tournant sur les pratiques effectives des linguistes et des enseignants de langue ? Ce renouvellement méthodologique est-il réellement pertinent pour tous les domaines des Sciences du Langage ? Si, pour certains, cette évolution représente une véritable rupture méthodologique et conceptuelle, qui devrait affecter à (court) terme l'ensemble de la discipline et de ses pratiques, pour d'autres, selon les domaines et les objets d'étude considérés, ces approches récentes semblent sinon totalement inadéquates, du moins prématurées.

L'objet de cette journée d'étude est d'offrir ainsi plusieurs éclairages sur cette question très générale du rapport aux données en langue, tant d'un point de vue théorique que méthodologique, et cela : en (neuro)psycholinguistique (*Jean-Luc Nespoulous*), en phonologie (*Jacques Durand*), en acquisition du langage (*Sophie Wauquier*), en apprentissage des langues étrangères (*Rémy Porquier*), et en sémantique (*Laurent Gosselin*).

PROGRAMME (chaque intervention durera 40 minutes, suivies de 15 minutes de discussion)

9h15-9h30 : Ouverture de la journée
9h30-10h25 : Jean-Luc NESPOULOUS (Université Toulouse II) - *La variabilité inter-tâches dans la performance linguistique. Apport de l'aphasiologie*
10h25-10h35 : Pause-café
10h35-11h30 : Jacques DURAND (Université Toulouse II) - *Données, modélisation et cognition : le cas de la liaison en français*
11h30-12h25 : Sophie WAUQUIER (Université Paris VIII) - *Acquisition phonologique et données de français. Enjeux théoriques*
12h25-14h00 : Pause-déjeuner
14h00-14h55 : Rémy PORQUIER (Université Paris X) - *Corpus et données comme sources et comme observables dans l'appropriation et l'enseignement des langues*
14h55-15h05 : Pause-café
15h05-16h00 : Laurent GOSSELIN (Université de Rouen) - *Le statut des données en sémantique grammaticale : négation et modalité*
16h00-16h15 : Clôture de la journée

Date et lieu : vendredi 13 février 2009 - Maison de l'Université, salle divisible sud, Université de Rouen
Comité d'organisation : Sylvain Detey & David Le Gac ({sylvain.detey & david.legac}@univ-rouen.fr)
Cristina Alexandru, Cécilia Hutter, Mito Katano – E.A. 4305 *LiDiFra*, DESCILAC
Public concerné : cette journée intéressera fortement les étudiants de Master et de Doctorat, non seulement en SDL mais également en lettres, langues, psychologie et autres secteurs des SHS. Ils sont donc invités à y participer.

Jean-Luc NESPOULOUS

Professeur, Département des Sciences du Langage, Université Toulouse II. Membre senior de l'Institut Universitaire de France. Directeur de l'Unité de Recherche Interdisciplinaire OCTOGONE, E.A. 4156. Institut des Sciences du Cerveau de Toulouse IFR 96.

La variabilité inter-tâches dans la performance linguistique. Apport de l'aphasiologie

Au début de l'aphasiologie, et y compris chez les premiers linguistes qui se sont intéressés aux performances linguistiques des patients aphasiques (cf. R. Jakobson), les dysfonctionnements langagiers consécutifs à des lésions cérébrales focales ont été, le plus souvent, considérés comme stables, traduisant une « perte » de compétence. Plus récemment cependant, le recours à des performances linguistiques recueillies dans des situations ou dans des tâches expérimentales de plus en plus variées a clairement montré que la qualité des performances tendait à varier grandement d'une situation à une autre, d'une tâche à une autre. La prise en considération d'une telle hétérogénéité symptomatologique si elle discrédite, d'une part, l'hypothèse d'une perte de compétence linguistique chez ses sujets, est, d'autre part, pleine d'enseignements pour les linguistes de corpus qui auraient dès lors grand tort à ne pas prendre en compte les caractéristiques psycholinguistiques des situations/tâches auxquelles ils soumettent leurs « informateurs » lors du recueil de données.

- Avec Dordain, M., Perron, C., Ska, B., Bub. D., Caplan, D., Melher, J. & Lecours, A.R. Agrammatism in sentence production without comprehension deficits: reduced availability of syntactic structures and/or of grammatical morphemes. A case study. *Brain and Language*. 33 : 273-295, 1988.
- Avec Dordain, M. Dissociations procédurales chez l'aphasique et variabilité inter-tâches chez l'apprenant de langue seconde. *Revue de Phonétique Appliquée*, N° spécial sur "Psychologie et Didactique". 112/113 : 263-278, 1994.
- Universal vs. language-specific constraints in agrammatic aphasia. In C. Fuchs & S. Robert (Eds.) *Language diversity and cognitive representations*. Amsterdam : John Benjamins : 195-207, 1999.

Jacques DURAND

Professeur, Département des Etudes du Monde Anglophone & Directeur de l'UMR 5263, *Cognition, Langues, Langage, Ergonomie* (CLLE), CNRS & Université Toulouse II. Membre senior de l'Institut Universitaire de France.

Données, modélisation et cognition : le cas de la liaison en français

De par le passé, la liaison en français a trop souvent été appréhendée à partir d'exemples construits, artificiels ou normatifs. Nous avons désormais à notre disposition un grand corpus codé pour la liaison, en l'occurrence le corpus PFC (*Phonologie du Français Contemporain*), avec plus de 28 000 codages. Comment interpréter ces résultats ? Comment les modéliser ? Et comment les intégrer à une approche cognitive ?

- Mapping French Pronunciation. The PFC project. In J.-P. Montreuil & C. Nishida (Eds.) *New Perspectives on Romance Linguistics*. Vol. 2 : *Phonetics, Phonology and Dialectology*. Amsterdam : John Benjamins : 65-82, 2006.
- Avec Lyche, C. French liaison in the light of corpus data. *Journal of French Language Studies*. 18/1 : 33-66, 2008.
- On the scope of linguistics : data, intuitions, corpora. In Y. Kawaguchi, M. Minegishi & J. Durand (Eds.) *Corpus and Variation in Linguistic Description and Language Education*. Amsterdam : John Benjamins. sous presse.

Sophie WAUQUIER

Professeur, UFR de Sciences du Langage & Directrice de l'UMR 7023, *Structures Formelles du Langage* (SFL), CNRS & Université Paris VIII.

Acquisition phonologique et données en français. Enjeux théoriques

Les enjeux théoriques que pose l'acquisition de la phonologie sur la base de données francophones ont longtemps été négligés ou exclusivement envisagés à l'aune des hypothèses faites pour l'anglais et les langues germaniques. Mais ce champ de la phonologie du français a connu un essor important depuis les années 1990. Des corpus longitudinaux et transversaux sont actuellement disponibles ou en passe de l'être, notamment grâce à l'enrichissement de la base de données « Childes » et la mise en correspondance de celle-ci avec la base « Phon », élaborée par Yvan Rose.

Les problèmes spécifiques que pose le français (absence d'accent lexical et présence d'un « accent de groupe », liaison, présence du schwa, structure rythmique et syllabation) sont donc maintenant susceptibles d'être étudiés et approfondis à partir d'une base empirique beaucoup plus riche.

Nous ferons le point sur les données, et nous nous attarderons, sur la base des résultats actuels et des perspectives de recherche auxquelles invitent ces nouvelles données, sur le point suivant : comment le recours à une modélisation hautement abstraite des représentations phonologiques permet de renseigner adéquatement la compréhension que nous avons des mécanismes d'apprentissage.

- Acquisition et développement phonologiques. In N. Nguyen, S. Wauquier-Gravelines & J. Durand (Dir.) *Phonologie et phonétique, forme et substance*. Paris : Hermès : 325-345, 2005.
- Du son au sens, acquérir ou apprendre la phonologie, *Recherches Linguistiques de Vincennes*, 35 : 5-31, Presses Universitaires de Vincennes, Université Paris VIII, 2006.
- Avec Rose, Y. Acquisition of speech in French. In S. Mc Leod (Ed.), *International guide of Speech Acquisition*, Thomson Delmar Learning : 364-386, 2007.

Rémy PORQUIER

Professeur honoraire, Université Paris X.

Corpus et données comme sources et comme observables dans l'appropriation et l'enseignement des langues.

Les questions concernant la nature des données, le recueil des données, leur modélisation et leur exploitation s'abordent et s'imbriquent de façon spécifique lorsqu'il s'agit de l'acquisition/apprentissage de langues non maternelles, et d'interactions en rapport avec l'appropriation et l'enseignement de langues non maternelles. Les données peuvent être alors de divers types, et pour des visées diverses : soit des corpus de langue, spécialisés ou non, en amont, pour élaborer ou adapter des descriptions à finalité didactique; soit des corpus de productions d'apprenants et/ou de locuteurs non natifs, recueillis en milieu guidé ou non guidé, y compris dans des protocoles de recherche semi-expérimentaux ; soit des corpus d'interactions exolingues dans des contextes divers.
Au-delà des questions méthodologiques concernant le recueil et le traitement de telles données, on peut constater leur relative complémentarité. Ainsi, les descriptions sur lesquelles s'appuie un enseignement vont déterminer en partie l'analyse des productions d'apprenants. Ainsi, les études sur les productions d'apprenants et les interactions sont de nature à informer des descriptions linguistiques, et plus particulièrement celles traitées et investies dans les méthodes et les cours de langue. Cela pose le problème de la circularité et de l'articulation entre les modèles de description et d'analyse adoptés dans les deux cas, et donc des choix opérés au préalable pour les recueils de corpus.

- L'analyse des erreurs. Problèmes et perspectives. *Etudes de Linguistique Appliquée*, 25, 1977
- Avec Besse H., *Grammaires et didactique des langues*. Paris, Hatier-Crédif, 1984, 1993, 2005
- Avec Py B. *Apprentissage d'une langue étrangère : contextes et discours*. Didier, 2004.

Laurent GOSSELIN

Professeur, Département des Sciences du Langage et de la Communication, E.A. 4305 *Linguistique, Didactique, Francophonie* (LiDiFra), Université de Rouen.

Le statut des données en sémantique grammaticale : négation et modalité.

Jusqu'à présent, les études de sémantique grammaticale n'ont été que marginalement concernées par l'utilisation de corpus informatisés. On examinera les raisons de cette situation en interrogeant le statut des « données » en sémantique. On montrera, en prenant pour exemple les rapports complexes entre négation et modalité, que loin de pouvoir s'appuyer sur des données sémantiques évidentes (repérables automatiquement), la tâche du sémanticien est d'abord de les « déconstruire » afin d'identifier des éléments très abstraits qui pourront ensuite servir d'entrées pour des analyseurs sémantiques.

- *Sémantique de la temporalité en français. Un modèle calculatoire et cognitif du temps et de l'aspect*. Louvain-la-Neuve : Duculot, 1996
- *Temporalité et modalité*. Bruxelles : Duculot-De Boeck, 2005
- *Les modalités en français. La validation des représentations.* Rodopi, Amsterdam / New York, sous presse.