



HAL
open science

Petits et grands corpus en analyse linguistique des gestes

Camille Debras

► **To cite this version:**

Camille Debras. Petits et grands corpus en analyse linguistique des gestes. Corpus, 2018, Les petits corpus, 18, 10.4000/corpus.3287 . hal-03110749

HAL Id: hal-03110749

<https://hal.parisnanterre.fr/hal-03110749>

Submitted on 14 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Petits et grands corpus en analyse linguistique des gestes

Small and big corpora in the linguistic analysis of gesture

Camille Debras

1. “Big is beautiful”? Size matters

- 1 Si l'on devait définir la linguistique comme champ disciplinaire aujourd'hui, ce serait moins par ses cadres théoriques que par ses cadres méthodologiques et ses objets. L'époque où la théorie était reine est, en quelque sorte, révolue : la linguistique est entrée dans l'ère de la méthode. Ce qui norme aujourd'hui la discipline, c'est-à-dire ce qui fournit des évaluations positives en vue de la publication dans les revues les plus prestigieuses, ce qui permet de dérocher des financements de recherche sur appel à projet, ce sont majoritairement les études basées sur grands corpus écrits, objets à partir desquels on peut produire des statistiques inférentielles (aux résultats statistiquement généralisables) sur la base d'annotations automatisées. La méthodologie est devenue centrale : la recherche de certains de mes collègues sémanticiens, par exemple, s'appuie sur des méthodes d'analyse statistique complexe (analyse multinomiale et régression logistique) qui, loin de prendre une théorie pour point de départ, soumettent diverses théories aux tests statistiques pour identifier celle qui rend compte de l'organisation des données de la manière la plus juste et la plus cohérente. Par comparaison à ces pratiques dominantes sur corpus écrits, l'étude de l'oral et les analyses qualitatives tendent à être minorées.
- 2 Les normes contemporaines de notre champ disciplinaire, qui permettent de produire des connaissances nouvelles sur les langues et le langage, ne sont pas apparues *ex nihilo* : elles s'inscrivent dans une histoire de la discipline. La linguistique des grands corpus s'inscrit dans un mouvement dialectique, visant à remédier aux faiblesses des méthodes de *l'armchair linguist*, qui, depuis son fauteuil, produit sa théorie de la langue de manière totalement introspective à partir de quelques énoncés inventés pour l'occasion, prenant tout de même la peine de confirmer ses intuitions auprès de quelque locuteur natif qui

serait passé par là pour prendre le thé. Par contraste, les grands corpus de l'*empirical turn* des années 90, notamment de langue anglaise (CoCA, BNC), apparaissent plus robustes et plus fiables car ils permettent de fonder l'analyse sur un plus grand nombre d'occurrences attestées d'un phénomène, qui ne dépendent pas de la subjectivité, devenue suspecte, d'un seul linguiste inventeur d'exemples dans son fauteuil. Comme l'annonce fièrement le slogan du corpus BYU, en haut de sa page d'accueil¹, c'est désormais la quantité qui est gage de qualité : « *corpora, size, queries = better resources, more insight* ».

- 3 Le changement d'échelle des corpus implique la nécessité d'un changement de paradigme méthodologique : l'*empirical turn*, temps de l'agrandissement des corpus pour mettre les théories à l'épreuve, engendre le *quantitative turn* (voir Janda (2013) pour la linguistique cognitive) fondé sur les méthodes statistiques. L'apport interdisciplinaire des méthodes statistiques venues des sciences exactes a pour ainsi dire mis au rencart l'intuition et la réflexion du linguiste dans son fauteuil. L'artisan a laissé la place à l'ingénieur, à mesure que s'est développé un idéal épistémologique, peut-être vaguement positiviste, de produire des résultats plus objectifs, car (en théorie) répliquables, à partir d'hypothèses réfutables, et valorisés comme tels par les outils de l'évaluation internationale de la recherche qui s'homogénéisent ainsi en s'alignant sur celles des sciences dites exactes. Pour certains de mes collègues, nulle hésitation sur l'appartenance, dans les faits pourtant fluctuante, de la linguistique aux humanités ou aux sciences : « *we are social scientists* », me disait tout récemment l'un d'entre eux. Et l'interprétation n'est que littérature.
- 4 La norme actuelle des études quantitatives sur grands corpus écrits n'a pas pour seul parent la quête d'objectivité dans l'évaluation et la production de la recherche. Elle provient aussi des formes contemporaines du développement technologique. Le processus de grammatisation identifié par Auroux (1995) pour la linguistique du monde moderne dès le XVII^e siècle a son pendant à l'époque contemporaine. L'évolution de la discipline dépend des techniques et des ressources dont elle s'empare et qu'elle développe. L'essor de l'informatique et d'internet (archivage et partage de grandes quantités de données) et les progrès des ordinateurs en puissance de calcul (outils statistiques), sont autant de facteurs qui ont déterminé l'évolution des usages, entraînant l'orientation de la linguistique vers l'étude des grands corpus.
- 5 Cependant, ces évolutions méthodologiques ont été constamment accompagnées de mises en garde contre le mirage du tout électronique, dès les débuts de la linguistique de corpus dans les années 90. Svartvik (1992) énumère les possibles périls d'une linguistique dont les corpus de grande taille constitueraient le critère unique :

However, there may also be potential hazards embedded in heavy dependence on corpus data alone. One danger is the convenient replacement of laborious hands-on analysis by rapid, automatic processing: yet in many areas of linguistic study, careful manual analysis cannot be dispensed with. Another trap, easy enough to fall into in this electronic age of ours, is the delusion that corpus size ("big is beautiful") is more important than, or can actually provide, corpus adequacy. The greatest risk of all, however, is the distance that may arise between the end user of a standard corpus and the primary textual material - a danger that is particularly imminent in the case of impromptu speech which has been recorded and orthographically transcribed by others than the user and where the actual audio-recording is not readily available or properly consulted. A corpus text may then be treated as a kind of canon and context-free object. (Svartvik 1992 : 10)

- 6 Comme l'explique Svartvik, l'exploration automatisée de grands corpus, ne peut pas, et ne doit pas se passer d'un minutieux travail d'analyse qualitative des occurrences en contexte. Même de taille immense, un corpus demeure un objet contextualisé, qui doit être en adéquation avec les hypothèses du linguiste autant qu'avec le type d'usage de la langue qu'il souhaite étudier. Dans le même recueil que Svartvik, Fillmore parodie l'incommunicabilité naissante entre *armchair linguist* et linguiste travaillant sur grand corpus, jusqu'à l'incompréhension mutuelle :

Armchair linguistics does not have a good name in some linguistics circles. A caricature of the armchair linguist is something like this. He sits in a deep soft comfortable armchair, with his eyes closed and his hands clasped behind his head. Once in a while he opens his eyes, sits up abruptly shouting, "Wow, what a neat fact!", grabs his pencil, and writes something down. Then he paces around for a few hours in the excitement of having come still closer to knowing what language is really like. (There isn't anybody exactly like this, but there are some approximations.)

Corpus linguistics does not have a good name in some linguistics circles. A caricature of the corpus linguist is something like this. He has all of the primary facts that he needs, in the form of a corpus of approximately one zillion running words, and he sees his job as that of deriving secondary facts from his primary facts. At the moment he is busy determining the relative frequencies of the eleven parts of speech as the first word of a sentence versus as the second word of a sentence. (There isn't anybody exactly like this, but there are some approximations.)

These two don't speak to each other very often, but when they do, the corpus linguist says to the armchair linguist, "Why should I think that what you tell me is true?", and the armchair linguist says to the corpus linguist, "Why should I think that what you tell me is interesting?" (Fillmore 1992 : 35)

- 7 Pour caricaturale qu'elle soit, cette représentation d'une opposition catégorique dressant analyse quantitative contre analyse qualitative, statistiques contre interprétation, objectivité contre subjectivité, continue, dans une large mesure, de structurer les rapports entre linguistes issus de différents « cercles ». Le but de cet article est, tout au contraire, de faire apparaître les complémentarités et les continuités entre petits et grand corpus, entre analyse qualitative et quantitative, en somme, de remettre en question la dichotomie grand corpus/analyse quantitative *versus* petit corpus/analyse qualitative, et ce à l'appui d'outils et de méthodes dédiées à l'analyse de corpus relativement nouveaux : les corpus oraux multimodaux. L'article interroge en particulier la question des corpus en analyse multimodale des corpus oraux d'interactions filmés, dans le cadre d'une linguistique qui intègre l'étude des gestes à celle du discours pour rendre compte des pratiques langagières dans leurs multiples dimensions. Ce type d'approche est traditionnellement associé à des petits corpus supports d'analyses qualitatives. Cependant, à l'appui d'un tour d'horizon de recherches contemporaines en analyse multimodale, il s'agit de montrer que l'étude linguistique des gestes au cours des interactions orales se prête à une large variété d'approches complexes et innovantes au plan méthodologique, qui remettent en question la dichotomie entre analyses qualitatives et analyses quantitatives, entre petits et grands corpus. Des corpus de taille variée permettent en effet de produire des résultats de diverses natures, pas moins « utiles » les uns que les autres, et complémentaires de surcroît.

2. À la découverte des fonctions linguistiques des gestes

- 8 La langue parlée est, de longue date, l'un des parents pauvres de l'analyse linguistique. Par contraste, l'étude de la langue écrite, objet plus stable, moins éphémère, plus facile à archiver et à exploiter de manière automatisée, lui est largement préférée, à tel point que, par un subtil glissement digne du plus élémentaire des syllogismes, pour nombre de linguistes qui étudient la langue écrite, étudier la langue c'est étudier l'écrit. Ce biais en faveur des corpus écrits (*written language bias* pour Linell (2005)), conçus comme étant la langue même, tend à reléguer les pratiques orales de la langue, jugées trop instables et trop hétérogènes, hors du champ d'étude de la linguistique. Elles sont, au mieux, jetées dans le grand sac fourre-tout de la « pragmatique » (*the ragbag of performance*, pour reprendre l'expression maintes fois attribuée à Chomsky).
- 9 Cependant, les avancées récentes des nouvelles technologies, celles-là même qui ont donné naissance à la linguistique des grands corpus, ont également déclenché un regain d'intérêt pour l'étude linguistique de l'oral, notamment sous la forme de corpus oraux vidéos. L'omniprésence de l'image et de la vidéo dans le monde contemporain ont, à leur tour, suscité l'intérêt croissant des linguistes pour les corpus oraux filmés, et plus particulièrement pour des phénomènes jusqu'alors ignorés car cantonnés à la dimension paralinguistique du langage : la gestualité, ainsi que l'étroite interdépendance formelle, sémantique et séquentielle qu'entretiennent les formes mimo-posturo-gestuelles avec les énoncés du discours.
- 10 Mais qu'entend-on exactement par « geste » ? Un geste, c'est avant tout une forme observable de nature kinésique : un mouvement du corps. Le geste peut être réalisé par différentes parties du corps, ou articulateurs. La conformation des articulateurs détermine bien sûr les formes de gestes possibles. Les sourcils peuvent par exemple être haussés ou froncés à des degrés variables. La tête peut bouger selon quatre axes (et les combiner de multiples façons) : sagittal (hochement de tête), latéral (tête penchée sur la gauche ou la droite), horizontal (comme dans le secouement de tête), mais aussi bouger d'avant en arrière. Les épaules peuvent être haussées, ensemble ou indépendamment l'une de l'autre. Les mains sont les articulateurs les plus complexes, et peut-être pour cette raison, les plus étudiés. L'analyse des gestes manuels repose principalement sur quatre paramètres, hérités des conventions de description des langues des signes : l'orientation de la paume, la forme de la main, sa localisation dans l'espace gestuel, et le mouvement qu'elle accomplit.
- 11 Comme l'explique Kendon, la plupart des gestes sont de nature « excursive » (*excursionary*, Kendon 2002 : 149), c'est-à-dire que l'articulateur a tendance à retourner au repos une fois le geste terminé, parfois quasiment à la place qu'il occupait avant le geste (ou *home position*, selon Sacks & Schegloff 2002). L'excursion du geste s'organise en phases (Kendon 1980, 2004 : 108-126) : le geste commence par une préparation (mise en place des articulateurs) jusqu'à sa réalisation (l'« acmé » du geste pour Schegloff 1984, plus communément dénommée *stroke* suivant Kendon 2004 : 115) ; il peut s'ensuivre une tenue du geste (*hold*, voir Kendon 2004 : 123) qui maintient l'articulateur dans un certain état de tension, avant le relâchement de l'articulateur dans une phase de rétraction. L'excursion est probablement moins typique des mouvements du buste, qui peuvent parfois se confondre avec des changements de posture. Dans ce cas, c'est plutôt le moment du

changement d'orientation qui importe (par exemple : buste rejeté en arrière au moment d'un désaccord) que la tenue du geste ou son retour à une place initiale.

- 12 Le moindre mouvement du corps constitue-t-il un geste ? On serait tenté en première approche de définir le geste comme un mouvement du corps qui porte une intention de communication. Cela ne serait pourtant pas juste, et cela pour deux raisons : on fait nombre de gestes en parlant sans s'en rendre compte (ce qui, d'ailleurs, n'empêche pas nos interlocuteurs de les interpréter), et il est souvent impossible pour le chercheur de déterminer dans quelle mesure un geste est produit de manière consciente voire délibérée², pour chaque geste d'un corpus filmé, si le participant avait l'intention délibérée de produire le geste en question. Comme en témoigne son organisation en phases, le geste est, comme la parole, temporel par nature. Au cours de l'interaction, parole et gestes des participants sont autant d'actions (Schegloff 2007 : xiv) qui se déploient, en harmonie ou en contrepoint, pour écrire une partition à plusieurs voix où le rôle et le « *timing* » de chaque action sont intimement liés aux actions qui la précèdent, la suivent, ou l'accompagnent (Sacks *et al.* 1974).
- 13 Voix et geste peuvent fonctionner de manière synchrone, comme lorsque la saillance prosodique semble accomplie à la fois par la voix et les gestes (Ferré 2014) ; le geste peut aussi anticiper un segment verbal auquel il est lié au plan référentiel (*lexical affiliate* pour Schegloff 1984). Les gestes co-verbaux (qui accompagnent la parole) contribuent pleinement à la construction du sens d'un énoncé (Kendon 1980, 2004 : 5), parfois au point que, pour gloser le sens d'un énoncé associant la parole au geste, il est nécessaire de reformuler l'énoncé dans son ensemble (Kendon 2002 : 147-148). Cependant, un geste n'a pas besoin d'être combiné à la parole pour contribuer à l'interaction, et peut constituer une réponse en soi (voir par exemple Stivers 2008 sur les hochements de tête, ou les *gestural responsive actions* identifiés par Berger & Rae 2012). En d'autres termes, un geste, c'est un mouvement qui n'est pas là par hasard, mais qui coïncide d'une manière ou d'une autre avec les autres dimensions de l'interaction : « *This co-occurrence is an organized achievement and not "mere coincidence," it is a product of an organized effort and not a byproduct of some otherwise-focused organization* » (Schegloff 1984 : 273). On peut donc définir le geste comme une forme kinésique (un mouvement du corps) qui devient signifiante en tant qu'elle s'inscrit dans la séquentialité de l'interaction (Schegloff 1988³), c'est-à-dire la succession coordonnée des actions constitutives de l'interaction et par l'intégration simultanée et séquentielle desquelles se construit le sens.
- 14 Ainsi, certains mouvements du corps sont traditionnellement écartés de l'analyse gestuelle : il s'agit des « gestes d'auto-contact » ou « adaptateurs »⁴ (*self-adaptors*) identifiés par Ekman et Friesen (1969). Cette catégorie rassemble tous les gestes orientés vers l'aise du locuteur (identiques aux « *self grooms* » des animaux identifiés par les éthologues, pour Schegloff (1984 : 271)) : se gratter la tête, repositionner ses lunettes, étouffer un bâillement, par exemple. Le changement de posture, « compris comme une modification de la posture de repos » (Ferré 2011 : 3) sert à relâcher les tensions du maintien d'une posture, et entre donc habituellement dans la catégorie des adaptateurs. La limite entre gestes et adaptateurs est parfois difficile à tracer : un participant se gratte l'arrière de la tête ou se tient le menton n'est-il pas justement en train de faire part de son incertitude ? C'est l'inscription du mouvement dans le contexte simultané et séquentiel des actions des participants qui permet en dernière instance de déterminer s'il fonctionne comme geste ou non.

- 15 L'étude du geste étant un champ fondamentalement interdisciplinaire (incluant approches formelle, cognitive, psycholinguistique, sémiotique, pragmatique, anthropologie linguistique ou analyse conversationnelle ; voir Stam et Ishino (2011) ; voir Kendon (2004) pour une mise en perspective historique et théorique), on se bornera à présenter brièvement quelques-unes des principales approches linguistiques de la gestualité en interaction.
- 16 Les approches psycholinguistiques ont montré que la situation de dialogue est une variable qui influence la forme (dont l'ampleur et l'orientation, voir par exemple Özyürek (2002) sur les gestes représentationnels) mais aussi la fréquence des gestes. La linguistique cognitive s'intéresse principalement aux correspondances entre espaces mentaux et espace gestuel (Sweetser 2012), à la schématicité (Cienki 2005, Ladewig 2011) et à la métaphoricité du geste (Cienki & Müller 2008). Les gestes métaphoriques, qui figurent des idées abstraites sont selon Ferré « plus souvent liés à l'information grammaticale qu'à l'information lexicale du message » (Ferré 2011 : 6). Plus récemment, c'est par le biais des grammaires de constructions que la linguistique cognitive s'intéresse aux gestes, sous l'angle de l'analyse des constructions multimodales (Steen & Turner 2013, Hinnell 2014, Zima 2014, Wu & Cienki 2014, Brône & Sambre 2015, Schoonjans *et al.* 2015 *inter alia*).
- 17 Certains gestes manuels, dits interactifs, sont spécifiques à la situation d'interaction (*interactive gestures* chez Bavelas *et al.* 1992 et 1995) et servent à la gestion du dialogue. Ils ont généralement la forme d'une paume ouverte aux doigts légèrement repliés, orientée en direction de l'interlocuteur. Ce type de gestes dits performatifs (Müller & Haferland 1997) ou pragmatiques (Streeck 1994), qui sont dérivés d'actions concrètes, en l'occurrence *donner* et *recevoir* pour la paume ouverte tendue, servent généralement à exprimer des actes de parole (Kendon 1995, Müller 2004). Le pointage peut également servir à la gestion de la parole en contexte d'interactions multipartites (Mondada 2007).
- 18 Au-delà des manifestations d'accord et de désaccord conventionnelles dans de nombreuses cultures occidentales, les mouvements de tête revêtent de nombreuses fonctions sémantiques : un balayage horizontal signifie l'inclusion, une oscillation latérale est synonyme d'incertitude (McClave 2000), tandis qu'un secouement de tête (*head shake*) peut servir à intensifier une évaluation (M.H. Goodwin 1980, Kendon 2002). Une orientation nouvelle de la tête peut également marquer différents énonciateurs en accompagnant une citation rapportée (Maury-Rouan, 2011). La tête sert à des mouvements déictiques de repérage concret ou abstrait dans l'espace, et permet de structurer la mention de listes ou d'alternatives par des orientations alternées (McClave, 2000). L'utilisation d'un hochement unique (*single nod*) synchronisé avec l'accent tonique d'un mot sert aussi à exprimer l'emphase (C. Goodwin et M.H. Goodwin 1987, 1992).
- 19 À la suite de Birdwhistell (1970), qui envisageait la possibilité de fonctions « linguistiques » des expressions du visage, Chovil (1991) et Bavelas (Bavelas & Chovil 1997, 2000) rappellent que l'abondante littérature sur l'expression faciale comme *extériorisation* d'un état psychologique interne (réaction émotionnelle) tend à nous faire oublier le rôle *communicationnel* que les expressions faciales peuvent jouer en interaction. Chovil (1991) identifie trois grandes fonctions pour les expressions du visage en conversation : une fonction linguistique-syntaxique (emphase, marquage des questions ou des étapes de l'organisation du discours, généralement réalisée par un mouvement des sourcils), une fonction sémantique (attitude, émotion), et une fonction de réponse minimale (*backchannels* et *listener commentaries*, commentaires silencieux de

l'interlocuteur pour exprimer une réaction sans pour autant interrompre le discours du locuteur, qui peuvent également être assurés par des hochements de tête, voir Stivers, 2008).

3. Approche multimodale des interactions orales : l'analyse qualitative en héritage

- 20 En s'emparant d'un objet d'étude comme le geste, auparavant laissé au soin des sémiologues, nombre de linguistes s'intéressent aujourd'hui à construire les cadres théoriques et méthodologiques d'une analyse multimodale des pratiques orales de la langue, c'est-à-dire une analyse qui élargirait la définition du « linguistique » pour prendre en compte des formes issues de multiples *modes* de communication (discours, voix, gestes ; voir Norris (2004), Jewitt (2009), O'Halloran & Smith (2011)) et de différentes *modalités* d'interaction (audio-orale, gestuo-visuelle ; voir Boutet *et al.* 2010 *inter alia*). L'analyse multimodale des corpus interactionnels filmés propose de traiter à statut égal les formes gestuelles aux côtés des formes verbales, par exemple dans le domaine de l'acquisition du langage (Beaupoil-Hourdel 2015, Morgenstern & Chang 2015, Morgenstern *et al.* 2016 *inter alia*). Ces travaux s'appuient notamment sur une première génération de recherches qui se sont d'abord employées à poser les cadres d'une analyse linguistique des formes gestuelles (Müller *et al.* 2013a). À rebours des approches psycholinguistiques qui caractérisent les gestes par leur fonction vis-à-vis de la parole (dans la lignée des travaux de McNeill (1992)), les travaux d'Adam Kendon et de Cornelia Müller ont contribué à poser les bases d'une approche fondée sur les formes gestuelles (*form-based approach* pour Müller *et al.* 2013b, voir aussi : Kendon 2004), sans les considérer comme secondaires par rapport à la parole. Une telle approche postule que certains gestes de la parole sont récurrents dans une langue (*recurrent gestures* d'après Ladewig 2014a, Bressemer & Müller 2014a), c'est-à-dire que à travers l'ensemble de leurs différents contextes d'usage (*context of use*, Kendon 2004 : 84) ils présentent un thème sémantique commun (*common semantic theme*, (Kendon 2004 : 227 ; Müller 2004 ; Müller *et al.* 2013b)). À l'appui de l'étude qualitative d'un ensemble aussi varié que possible d'occurrences de la forme gestuelle en contexte, le plus petit dénominateur sémantique commun de toutes les occurrences du geste étudié est ainsi identifié comme le thème sémantique que cette forme exprime. Ce type de méthode inspire des études permettant de caractériser le thème sémantique de nombreuses formes gestuelles : le secouement de tête (Kendon (2002) sur le *head shake*), le geste de la pince (entre le pouce et l'index, ou *precision grip*, Kendon 2004 : 225-247), les gestes manuels de paume ouverte tendue vers l'interlocuteur (*palm-up open hand*, Müller 2004, voir également Beaupoil & Debras (2017) pour une perspective développementale) ou de balayage (Bressemer & Müller 2014), le geste cyclique de la main (Ladewig 2011), le geste de la paume présentée verticalement à l'interlocuteur (*vertical palm* ou *open-hand prone*, Kendon 2004 : 251-255, Harrison 2015 ou *holding away gesture*, Bressemer *et al.* 2015 ; voir également Calbris 2011 sur les gestes manuels de la pince, de la paume ouverte tendue, de balayage, le geste cyclique et la paume verticale), la gestion des attentes exprimée par les haussements de sourcils (Debras 2013), les formes de la posture composite de désengagement (Debras 2017).
- 21 L'attention portée à la séquentialité des formes étudiées et à leur contexte d'usage inscrit pour une part l'analyse multimodale dans la lignée de l'analyse conversationnelle. Cette influence est mutuelle, car un nombre grandissant de chercheurs en analyse

conversationnelle incluent la dimension kinésique dans leurs analyses (pour une synthèse, voir Mortensen (2012)). Issue du courant sociologique de l'ethnométhodologie (Garfinkel 1984), l'analyse conversationnelle (Sacks *et al.* 1974, Sacks 1992) constitue l'une des approches les plus structurées de l'étude de l'oral au plan institutionnel tant que méthodologique. Basée sur l'analyse qualitative séquentielle minutieuse d'interactions orales à l'appui de conventions de transcription très précises (Jefferson 2004) pour mettre au jour les fonctions de certaines formes à travers divers contextes, elle s'appuie généralement sur des collections d'exemples choisis qui mettent au jour la variété fonctionnelle d'une forme. Un reproche majeur fait à ce type d'approche, notamment par les tenants des méthodes quantitatives, est celui du « *cherry-picking* » : si le chercheur est complètement libre de choisir les exemples qui confortent sa démonstration, alors il peut aisément laisser de côté des cas plus problématiques, voire qui pourraient remettre en question la ligne d'analyse choisie. Les études qualitatives qui permettent de caractériser la variété de fonctions d'une forme à travers divers contextes constituent cependant un premier apport considérable pour la caractérisation des systèmes langagiers multimodaux à l'oral. Il appartient désormais à la communauté des chercheurs en analyse multimodale des interactions orales de développer des complémentarités entre analyse qualitative et méthodes quantitatives, et entre petits et grands corpus.

4. Les corpus pour l'analyse des gestes en interaction : un enjeu de taille

- 22 Aujourd'hui, la vidéo est partout (internet, smartphones, télévision, réseaux sociaux), et la relative démocratisation des technologies numériques facilite la construction, l'analyse et le partage de ce nouveau type de corpus (Mittelberg 2007). Pour pouvoir étudier les gestes utilisés par les locuteurs et l'intégration de ces formes visuelles (ou modalité kinésique) avec le discours et la prosodie (ou modalités verbale et vocale) dans la construction du sens, un travail préalable de transcription-annotation conséquent est nécessaire. Pour être analysé d'un point de vue multimodal, tout enregistrement doit faire l'objet d'un important travail de préparation, incluant la transcription des énoncés prononcés, et l'annotations des formes gestuelles et/ou vocales étudiées, qui sont « alignées » sur le temps de l'enregistrement vidéo à l'aide de logiciels dédiés tels que CLAN⁵, PRAAT, ANVIL ou ELAN⁶. Cela signifie par exemple que la transcription d'un énoncé prononcé pendant une durée définie de l'enregistrement vidéo comporte non seulement l'information du contenu de ce qui a été dit, mais aussi l'information du moment et de la durée de cet énoncé dans l'enregistrement. Concrètement, cela permet, dans les logiciels dédiés, de faire défiler la transcription et les annotations pendant le visionnage de la vidéo, qui peuvent par exemple apparaître sous la forme d'un script ou de sous titres qui défilent au rythme du film. Les logiciels permettent dans un second temps d'effectuer des requêtes automatiques sur des formes issues des transcriptions-annotations, pour retrouver chaque occurrence de la forme étudiée (mot, geste) dans son contexte d'usage (*context of use*, Kendon 2004 : 84) multimodal.
- 23 Si les outils informatiques permettent aujourd'hui la reconnaissance automatique des formes textuelles et l'automatisation des requêtes sur les corpus écrits, la technologie dont les chercheurs disposent pour le repérage des formes verbales (discours parlé) et kinésiques (gestes) dans les corpus oraux est pour l'heure bien moins avancée. La plupart du travail de transcription des énoncés s'effectue donc à la main. En effet, les logiciels de

reconnaissance vocale, qui permettraient de dégrossir le travail, sont peu adaptables aux chevauchements de parole et aux variations de voix typiques des corpus d'interaction. L'analyse automatique de la voix est l'aspect le plus avancé : un logiciel tel que PRAAT⁷ permet un grand nombre d'analyses automatiques des formes sonores de la voix parlée ; le logiciel SPPAS⁸, permet de gagner beaucoup de temps sur l'alignement des annotations avec le son, notamment en anglais et en français et avec de bons résultats sur les corpus de conversations. L'outil d'alignement EasyAlign⁹, utilisable dans PRAAT, fonctionne lui aussi bien pour le français. Malgré l'aide fournie par ces outils, le travail manuel de correction et de transcriptions reste conséquent. Concernant la dimension kinésique, l'avancée rapide des technologies de capture de mouvement pourrait permettre, dans un avenir que l'on espère proche, l'automatisation au moins partielle de l'annotation de formes gestuelles (Priesters et Mittelberg 2013, Guez *et al.* 2013), qui, pour l'heure, n'en est qu'à ses balbutiements. L'étiquetage des formes gestuelles est un travail de fourmi que les chercheurs réalisent pour l'instant à la main, geste par geste, dans des logiciels d'annotation dédiés comme ELAN¹⁰ ou ANVIL¹¹. On comprend ainsi plus aisément la taille limitée des corpus disponibles, dont la préparation représente déjà des heures de travail.

24 À ce jour, constituer des corpus oraux d'interactions filmées est donc une entreprise longue, fastidieuse, et coûteuse lorsqu'il s'agit d'employer des petites mains pour transcrire les énoncés et annoter les phénomènes observés (gestes, actions, éléments de contexte). Prenons pour exemple le travail de transcription réalisé au sein de l'ANR CoLaJE¹² (2008-2012), qui avait pour objectif de reconstituer l'émergence et le développement de la communication langagière chez le jeune enfant en adoptant une approche pluridisciplinaire et multimodale. Pour chaque heure d'enregistrement, la transcription-annotation des multiples niveaux d'analyse (phonologie, prosodie, morpho-syntaxe, colocation et éléments posturo-mimo-gestuels) s'élevait à 20 heures de travail, en moyenne. En analyse multimodale des corpus oraux vidéos, la nécessité de travailler en équipe au sein de projets collectifs, et dans la mesure du possible, financés, n'est ainsi pas à démontrer. Si créer un corpus oral multimodal est un travail long est difficile, il faut néanmoins préciser que transcrire un corpus à la main permet d'en acquérir une connaissance extrêmement précise. C'est au prix des heures passées à observer minutieusement de nombreux gestes dans leurs contextes d'usage que se construit l'expertise d'un chercheur en gestualité, et que s'aiguise la justesse de ses analyses.

25 Construire l'équivalent des COCA et autres BNC pour des corpus multimodaux filmés, qui seraient librement consultables en ligne et permettraient des requêtes automatiques sur les transcriptions et les annotations gestuelles, reste un objectif encore difficilement atteignable à ce jour. Même si cela était possible, les participants filmés ont, bien entendu, à tout moment le droit de refuser que des enregistrements où ils apparaissent soient librement consultables en ligne ; il revient d'ailleurs au chercheur de vérifier que le droit à l'image de chaque participant soit respecté comme celui-ci le souhaite (Baude 2006). Dans l'état actuel de la recherche, ce sont bien souvent des contraintes très concrètes de temps et de faisabilité qui déterminent la taille des corpus oraux multimodaux existants, et le détail des annotations dont ils font l'objet. Un chercheur qui conduirait un projet de recherche seul ne peut, pour des raisons pratiques, construire qu'un corpus de relativement petite taille. À titre d'exemple, un corpus de thèse comme le mien (Debras 2013) rassemble 16 locuteurs, environ 2 heures d'interactions orales en anglais britannique, et comporte une à deux centaines d'occurrences de chaque phénomène gestuel observé (220 haussements de sourcils, 102 occurrences de la posture

de désengagement ou *shrug* (voir aussi Debras 2017)). Les premières études sur les gestes récurrents (*recurrent gestures*) en allemand, portées par Cornelia Müller et ses collègues dès 2006 étaient fondées sur un peu plus de 219 occurrences pour l'étude de 7 gestes¹³ ; l'étude du geste cyclique en allemand proposée par Silvia Ladewig (2011) ou celle de Jana Bressemer sur la paume verticale tendue vers l'interlocuteur en Savosavo (langue de Papouie) (Bressemer *et al.* 2015) reposent chacune respectivement sur 56 occurrences de la forme observée. Un travail mené en équipe permet de créer des corpus plus grands, aux annotations plus riches, comme le montre l'exemple du *Corpus of Interactional Data* (CID, Bertrand *et al.* 2008) développé au Laboratoire Parole et Langage (Aix-Marseille Universités), qui rassemble 8 heures d'interactions orales en français annotées à plusieurs niveaux (phonétique, prosodie, morphologie, syntaxe, discours et mimo-gestualité). Ces quelques exemples montrent que ce domaine d'étude naissant s'organise au départ de manière plutôt artisanale, plus en collections de petits corpus qu'en grands corpus.

26 Une première solution pour l'agrandissement des corpus oraux vidéos est la création de bases de données nationales et internationales, fruit de la coordination des chercheurs à travers le monde, et qui permet autant le partage que l'archivage des données collectées. Grâce à des bases de données comme la *Talk Bank*¹⁴, les linguistes du monde entier peuvent mettre des corpus oraux filmés à la disposition des communautés des chercheurs. Les spécialistes d'acquisition du langage ont été précurseurs dans la mutualisation des corpus oraux. Brian MacWhinney et Catherine Snow ont commencé à développer la base de données CHILDES¹⁵ dédiée à l'acquisition du langage dès 1984, et c'est plus récemment, en 2002, que Brian MacWhinney a créé la *TalkBank* pour l'étude plus générale de la communication orale humaine. CHILDES constitue aujourd'hui la composante « acquisition du langage » de la *TalkBank*. Les initiatives de valorisation des corpus oraux se multiplient également au niveau européen grâce à l'infrastructure de recherche CLARIN¹⁶, ainsi qu'en France, notamment dans le cadre de la TGIR Huma-Num¹⁷. Celle-ci vise à faciliter le tournant numérique de la recherche en sciences humaines et sociales, en favorisant le traitement, la conservation, l'accès et l'interopérabilité des données de la recherche. Des équipements d'excellence validés dans le cadre des investissements d'avenir comme la plateforme ORTOLANG¹⁸ ou des outils d'accompagnement et de stockage comme la plateforme COCOON¹⁹ rattachée à Huma-Num accompagnent les chercheurs travaillant sur corpus oraux vidéos. Nombre de ces chercheurs comptent aujourd'hui parmi les membres du Consortium Multi-Com²⁰-CORLI²¹ (anciennement IRCOM²²) de la TGIR Huma-Num, où ils se concertent pour développer collectivement des pratiques communes d'enregistrement, de transcription et de stockage des corpus oraux au niveau national. Des pratiques harmonisées de transcription et d'annotation pourraient, à l'avenir, favoriser une plus grande cohérence dans les futures collectes de corpus, ce qui permettrait de constituer de grands corpus relativement homogènes en rassemblant plusieurs petits.

27 La construction de grands corpus oraux multimodaux reste une gageure. Un des plus larges corpus d'interactions orales disponibles en ligne, le *Santa Barbara Corpus of Spoken American English*²³ développé par John Du Bois et ses collègues (Du Bois *et al.* 2000-2005) à l'Université de Santa Barbara en Californie (États-Unis), compte 249 000 mots ; il est fondé sur des enregistrements sonores, la dimension visuelle de l'interaction n'a donc malheureusement pas pu être conservée. Les grands corpus oraux partagés en ligne tendent cependant à se développer. Pour le français parlé, la base CLAPI²⁴, développée au laboratoire ICAR²⁵ à l'université de Lyon rassemble 63h de corpus d'interactions orales

dont une grande partie est fondée sur des enregistrements filmés et propose un outil de requêtes automatiques en ligne pour interroger les transcriptions de ces corpus. À ce jour, le plus ambitieux projets de création de corpus oraux multimodaux est le *Red Hen Lab*²⁶ porté par Mark Turner²⁷ et Francis Steen²⁸ (Steen & Turner 2013). Il s'agit d'un immense corpus vidéo de plus de 200 000 heures, stocké à UCLA en Californie, sous la forme du *UCLA Library Broadcast NewsScape*²⁹. Il est fondé sur l'extraction automatique d'émissions d'information sur des chaînes de télévision du monde entier (États-Unis, pays Scandinave, Espagne notamment). Lorsque des transcriptions télétextes sont disponibles, celles-ci sont automatiquement alignées sur les enregistrements collectés. Le dispositif de captation automatique de ces chaînes d'information permet d'enregistrer plus de 100 heures de corpus chaque jour, et le volume de textes transcrits et alignés atteignait le milliard de mots en 2013. Steen et Turner ont depuis créé un réseau de chercheurs qui collectent du corpus vidéo télévisé dans leur propre pays, le *Distributed Little Red Hen Lab*, pour enrichir ce corpus multimodal géant de composantes de diverses langues, qui propose des perspectives intéressantes en termes de reconnaissance automatique des gestes. Ce projet international mobilise d'impressionnants moyens techniques ; et même si les interactions qui se déroulent sur les plateaux télévisés constituent un genre d'interactions orales fortement déterminées par des codes sociaux et langagiers spécifiques et distincts de ceux de la conversation ordinaire, l'immense quantité de données disponibles représente une ressource riche de potentialités et de recherches nouvelles en analyse multimodale des corpus oraux filmés.

5. Petits et grands corpus : complémentarités et continuités

- ²⁸ Les standards internationaux de l'évaluation (revues de haut rang, projets financés) de la recherche tendent aujourd'hui à valoriser l'analyse quantitative, basée sur l'analyse automatique et statistique de grands corpus, aux dépens de l'analyse qualitative, fondée sur des corpus plus petits³⁰. Une des raisons de cette tendance est probablement que les résultats produits par le biais de méthodes quantitatives sont en théorie répliquables et donc réfutables (au sens de Popper (1972)), et semblent ainsi plus objectifs, plus indépendants de la subjectivité du chercheur, donc plus facilement évaluables. À l'opposé, les résultats de recherches fondées sur des méthodes qualitatives peuvent sembler à certains trop dépendants de l'interprétation du chercheur pour être tout à fait fiables – indépendamment du degré d'expérience, de connaissance ou d'expertise dudit chercheur vis-à-vis de son objet de recherche. Cette dichotomie entre approches quantitatives et qualitatives, née de la mutation des outils et des pratiques d'évaluation de la recherche, tend à faire oublier que ces deux types d'approche ne sont en fait pas comparables. En effet, leurs objectifs de recherche sont non seulement différents, mais tout à fait complémentaires. Comme le rappelle Maxwell (2005) tandis que l'analyse quantitative vise la généralisabilité et la reproductibilité des *résultats* obtenus, l'analyse qualitative propose la généralisabilité des *processus* observés. L'analyse qualitative permet elle aussi de produire des savoirs transférables et réutilisables, qui prennent plutôt la forme de théories que celle de résultats chiffrés. Elles sont également complémentaires dans la mesure où les points forts de l'une correspondent aux points faibles de l'autre. Les méthodes qualitatives visent à respecter une plus grande validité écologique, c'est-à-dire qu'elles tentent de s'approcher au plus près de la complexité des pratiques humaines en

contexte. Elles ne peuvent ainsi viser en même temps la reproductibilité exacte de leurs résultats, propres à des contextes bien spécifiques. En revanche, les méthodes quantitatives, qui visent avant tout la reproductibilité des résultats, sont forcées de s'appuyer sur des modèles simplifiés des processus observés, et ce faisant, perdent en validité écologique. Enfin, la nécessité de neutraliser toutes les variables autres que celle étudiée dans les méthodes quantitatives a pour effet que ce type d'analyse se concentre sur des questions de recherches très précises, mais donc très étroites et qui ont tendance à isoler des variables d'une manière qui fractionne artificiellement le réel, parfois à outrance. Les méthodes qualitatives en revanche, qui visent à identifier des processus à l'œuvre dans un contexte précis mais qui soient transposables à d'autres contextes, permet d'envisager des questions de recherches plus larges et plus complexes.

- 29 De plus, l'idéal d'objectivité des méthodes quantitatives a, dans les faits, du mal à tenir ses promesses. C'est ce qu'a montré le *Reproducibility Project* mené par le psychologue B. Nosek et ses collègues (2015) au *Center for Open Science* de l'Université de Virginie (États-Unis). Dans une méta-étude visant à tester la fiabilité des méthodes quantitatives en psychologie, Nosek et ses collègues ont reproduit 100 études en suivant le protocole défini dans la publication de chaque étude, pour tester la répétabilité et donc la fiabilité des résultats obtenus. Sur les 100 études répétées, seulement 36 ont donné des résultats similaires aux résultats publiés dans l'étude initiale : ainsi, en pratique, presque deux publications sur trois en psychologie ne remplissent pas le critère de reproductibilité qui fonde implicitement leur légitimité scientifique. Cette étude met en lumière les limites des méthodes quantitatives en psychologie, domaine voisin de la linguistique, et pourtant souvent brandie comme modèles méthodologiques par les linguistes partisans du tout quantitatif. À ceux-là, on pourra également opposer l'argument que le fait de tester une hypothèse ne constitue pas la seule forme de résultat de recherche possible. Les méthodes qualitatives proposent des conclusions dont la valeur est d'une nature différente, sous la forme de ce que l'historien des sciences William Whewell nomme la « consilience » (1840), c'est-à-dire le fait que des sources ou des personnes différentes convergent autour d'une même interprétation des données, renforce la validité des interprétations proposées. Les approches qualitatives proposent ainsi de repérer des faisceaux convergents vers l'hypothèse ou l'interprétation la plus probante, et ainsi l'accord des subjectivités des chercheurs spécialistes d'un domaine de recherche peut être considéré comme une source fiable de connaissance.
- 30 Parce qu'elle constitue un domaine encore relativement récent et peu balisé, l'étude linguistique du geste est l'occasion de proposer des méthodes originales qui associent plusieurs approches, réinterrogeant notamment la dichotomie entre analyses qualitatives et quantitatives, entre petits et grands corpus. Nombre de travaux récents en analyse gestuelle encouragent en effet à remettre en question la dichotomie traditionnelle entre analyses qualitatives et quantitatives, de deux manières au moins. D'abord en proposant des études qui intègrent les deux approches : Stivers et ses collègues montrent l'intérêt d'adosser les méthodes de l'analyse conversationnelle à des méthodes quantitatives (codage, statistiques ; voir Stivers & Enfield 2010, Stivers 2015). Pour leur part, Cienki *et al* . (2011) proposent d'associer approches qualitatives et quantitatives : ils formulent dans un premier temps des hypothèses précises et documentées à l'appui d'analyses qualitatives des formes observées, qui sont dans un second temps mises à l'épreuve grâce à un test de perception psycholinguistique. En associant analyses qualitatives multimodales et méthodes statistiques sur corpus écrits, Debras et L'Hôte (2015, L'Hôte &

Debras 2016) proposent une méthode qui met en évidence l'observabilité de certains phénomènes discursifs à différentes échelles : à l'échelle micro (analyses qualitatives du discours multimodal en contexte) et macro (analyses statistiques des discours étudiés).

- 31 Par ailleurs, l'analyse qualitative systématique de certains traits d'un corpus multimodal, ou codage (Morgenstern & Chang 2014, Beaupoil-Hourdel 2015, Morgenstern *et al.* 2016, Stivers 2015, Debras 2017, *inter alia*) constitue une méthode intermédiaire intéressante entre analyses qualitatives et quantitatives. Le codage systématique des traits et fonctions de toutes les occurrences d'une forme dans un corpus donné permet notamment d'éviter le « *cherry picking* » souvent reprochés aux analyses qualitatives, c'est-à-dire la cueillette aux exemples choisis qui illustrent parfaitement le propos et peuvent ainsi proposer une image distordue des données. Il permet également de ne pas laisser de côté les cas problématiques, qui sont les plus intéressants car ils permettent de remettre en question les étiquettes de catégorisation proposées, et de rassembler une collection d'exemples plus large pour renforcer l'argumentaire et donner du poids aux propositions de généralisation. Un corpus vidéo d'interactions orales de quelques heures, à première vue « petit », peut contenir une grande densité d'occurrences selon la forme observée (plusieurs centaines d'occurrences), fournissant ainsi un type d'étude intermédiaire entre analyses qualitatives et statistiques, qui invite à explorer les fréquences brutes observées sur des corpus plus grands pour confirmer et/ou enrichir l'analyse. Enfin, les évolutions récentes des méthodes statistiques, notamment en analyse multivariée et régression logistique (voir Glynn & Robinson 2014, Krawczak 2015), rendent les méthodes statistiques de plus en plus compatibles avec les petits corpus, en proposant de prendre en compte des paramètres plus nombreux dans des corpus plus petits mais plus complexes. Elles offrent des perspectives particulièrement riches pour le développement des études en analyse multimodale (Debras et Krawczak 2016).

Conclusion

- 32 Parce qu'elle est un domaine de recherche relativement récent, l'étude linguistique des gestes permet de questionner des présupposés qui font autorité dans certains cercles en linguistique. Le domaine naissant de l'étude linguistique des gestes fournit l'occasion de rappeler les apports indispensables des petits corpus et de leur analyse qualitative, pourtant parfois mis au rencart par des tenants de la linguistique de corpus qui ne jureraient que par les méthodes quantitatives sur grands corpus. Poser les bases méthodologiques et théoriques d'une analyse multimodale des corpus oraux nécessite de passer par des étapes d'analyses qui ont été parcourues il y a longtemps pour l'analyse des corpus écrits. L'analyse qualitative des formes kinésiques dans des petits corpus d'interactions filmées fait partie de ces étapes fondatrices incontournables, pour commencer par cerner et définir précisément l'objet d'étude d'une analyse linguistique des gestes.
- 33 Par comparaison avec les possibilités offertes par les outils de recherches automatiques sur corpus écrits, l'automatisation d'annotations gestuelles et d'analyses multimodales pour les corpus oraux vidéos n'en est, certes, qu'à ses balbutiements. Cependant, la constitution de grands corpus multimodaux a déjà commencé (notamment à l'initiative du *Red Hen Lab*), et grâce aux multiples initiatives de coordination nationales, européennes et internationales (Consortium Multicom CORLI, TGIR Huma-Num, réseau CLARIN, *TalkBank*) dans la constitution de corpus oraux filmés, de grands corpus pourront

également être construits dans un avenir proche, à partir de collections de corpus plus petits.

- 34 Plus largement, l'analyse multimodale est l'occasion de réinterroger les limites du linguistique, en proposant d'intégrer d'autres ressources sémiotiques que la parole dans l'analyse d'une langue. Elle invite à remettre en question la dichotomie entre analyses qualitatives sur petits corpus et analyses quantitatives sur grands corpus. Loin de s'opposer, analyses qualitatives et quantitatives, petits et grand corpus, sont à maints égards complémentaires, notamment parce qu'ils permettent de répondre à des questions de recherche différentes (larges ou étroites) et de proposer des savoirs d'origines différentes (généralisabilité des processus observés ou généralisabilité des résultats), dont la légitimité est elle aussi de nature différente (consilience des experts ou réfutabilité des résultats). Qu'il s'agisse d'approches qualitatives ou quantitatives, c'est le travail d'interprétation des données et des processus observés, qu'on le nomme théorisation ou discussion, qui constitue le véritable apport d'une recherche. Comme le conclut Fillmore (1992), la linguistique perdrait tout intérêt si l'on considérait que les données brutes obtenues par analyses automatisées ou par méthodes statistiques constituent des résultats de recherche en eux-mêmes, car la pratique de réflexion du linguiste serait vidée de sa substance :

There are really good reasons for building corpora, and as far as I'm concerned the bigger the better. But what I have been saying is probably not encouraging to people who want to do most of their analysis without expecting anyone to have to sit down and stare at the examples one at a time to try to work out just what is the intended cognitive experience of the interpreter, what are the interactional intentions of the writer, and so on. Should it ever come about that linguistics can be carried out without the intervention and suffering of a native speaker analyst, I will probably lose interest in the enterprise. (Fillmore 1992 : 59-60)

- 35 Aussi performantes soient-elles dans la description des données, les analyses statistiques automatiques ne pourront jamais supplanter le travail de théorisation qu'elles ont pour mission d'étayer.

BIBLIOGRAPHIE

Auroux S. (1994). *La Révolution technologique de la grammatisation*. Liège : Madraga.

Baude O. (éd.) (2006). *Corpus Oraux : Guide des Bonnes Pratiques*. Paris et Orléans : Éditions du CNRS et Presses universitaires d'Orléans.

Beupoil-Hourdel P. (2015). *Acquisition et Expression Multimodale de la Négation. Étude d'un corpus Vidéo et Longitudinal de Dyades Mère-Enfant Francophone et Anglophone*. Thèse de doctorat, Sorbonne Nouvelle.

Beupoil-Hourdel P. et Debras C. (2017). « Developing communicative postures : The emergence of shrugging in child communication », *Language, Interaction, and Acquisition*, 8(1) : 89-116.

Berger I. et Rae J. (2012). « Some uses of gestural responsive actions », *Journal of Pragmatics*, 44 : 1821-1835.

- Bertrand R., Blache P., Espesser R., Ferré G., Meunier C., Priego-Valverde B. et Rauzy S. (2008). « Le CID - Corpus of Interactional Data - Annotation et Exploitation Multimodale de Parole Conversationnelle », *Traitement Automatique des Langues, TAL* 49(3) : 105-133.
- Bigi B. (2015). « SPPAS – Multi-lingual Approaches to the Automatic Annotation of Speech », *The Phonetician* 111-112 : 54-69.
- Boersma P. et Weenink D. (2017). *Praat : doing phonetics by computer [Computer program]*. Version 6.0.28, retrieved 23 March 2017 from <http://www.praat.org/>.
- Boutet D., Sallandre M.-A. et Fusellier I. (2010). « Gestualité humaine et langues des signes : entre continuum et variations », *Langage et Société* 131 : 53-74.
- Brône G. et Sambre P. (2015). « Gestural specification and completion in multimodal construction grammar : A case study of instrumental causal actions involving Cut and Break », communication à l'*International Cognitive Linguistics Conference (ICLC 13), Newcastle-Upon-Tyne, 20-25 Juillet 2015*.
- Bressemer J. et Müller C. (2014a). « A repertoire of German recurrent gestures with pragmatic functions », in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill et S. Teßendorf (éd.) *Body - Language - Communication vol. 2*. Berlin et Boston : Walter de Gruyter, 1575-1591.
- Bressemer J. et Müller C. (2014). « The Family of away gestures: negation, refusal, and negative assessment », in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill et S. Teßendorf (éd.) *Body - Language - Communication vol. 2*. Berlin et Boston : Walter de Gruyter, 1592-1604.
- Bressemer J., Stein N. et Wegener C. (2015). « Structuring and highlighting speech – discursive functions of holding away gestures in Savosavo », in G. Ferré et M. Tutton (éd.) *Proceedings of GESPIN 4*, Septembre 2015, Nantes, France, 49-54.
- Calbris G. (2011). *Elements of meaning in gesture*. Amsterdam et Philadelphie : John Benjamins.
- CORLI. Consortium Corpus Langues et Interactions de TGIR (Très Grande Infrastructure de Recherche) Huma-Num, <http://www.huma-num.fr/>.
- Debras C. (2013). *L'expression multimodale du positionnement interactionnel (Multimodal Stance-taking)*. Manuscrit de thèse non publié, Sorbonne Nouvelle, Paris.
- Debras C. (2017). « The shrug: forms and functions of a compound enactment », *Gesture* 16(1) : 1-34.
- Debras C. et L'Hôte E. (2015). « Framing, metaphor and dialogue: a multimodal approach to party conference speeches », *Metaphor and the Social World* 5(2) : 177-204.
- Debras C. et Krawczak K. (2016). « Multimodal and cross-register profiling of epistemic stance in spoken English : A multifactorial quantitative account » Communication au colloque international ISGS7 (*International Society for Gesture Studies*), Paris, juillet 2016.
- Du Bois J. W., Chafe W. L., Meyer C., Thompson S. A., Englebretson R. et Martey N. (2000-2005). *Santa Barbara corpus of spoken American English, Parts 1-4*. Philadelphie : Linguistic Data Consortium.
- Ekman P. et Friesen W. V. (1969). « The Repertoire of Nonverbal Behavior : Categories, Origins, Usage, and Coding », *Semiotica* 1(1) : 49-98.
- Ferré G. (2011). « Analyse multimodale de la parole », *Rééducation Orthophonique* 246 : 73-85.
- Ferré G. (2014). « Multimodal Hyperbole », *Multimodal Communication* 3(1) : 25-50.
- Fillmore C. (1992). « "Corpus Linguistics" or "Computer-aided armchair linguistics" », in J. Sartvik (éd.) *Directions in Corpus Linguistics, Proceedings of Nobel Symposium 82*. Berlin et New York : Mouton de Gruyter, 35-60.

- Garfinkel H. (1984). *Studies in Ethnomethodology*. Cambridge : Polity Press.
- Glynn D. et Robinson J. (2014). *Corpus methods for semantics : Quantitative studies in polysemy and synonymy*. Amsterdam : John Benjamins.
- Goldman J.-P. (2011). « EasyAlign : an automatic phonetic alignment tool under Praat », *Interspeech'11, 12th Annual Conference of the International Speech Communication Association*. Firenze (Italy).
- Guez J., Boutet D., Hsieh C.-W., Tramus M.-H., Chen C.-Y., Vincent C., Chabalière C., Châteauvert J., Lubek J., Catteau F., Renna I. et Delacroix S. (2013). « Le projet CIGALE (Capture et Interaction avec des Gestes Artistiques, Langagiers et Expressifs) : une plateforme transdisciplinaire de création et d'exploration du sens » *Communication au Colloque international Le sujet digital : inscription, excription, téléscripture*, novembre 2013, Saint-Denis, France.
- Harrison S. (2015). « A modality-axis in gesture space ? The Vertical Palm and construal of negation as distance », in Ferré G. et Tutton M. (éd.) *Proceedings of GESPIN 4*, Septembre 2015, Nantes, France, 137-141.
- Hinnell J. (2014). « English aspect-marking constructions », *Communication à l'International Society for Gesture Studies Conference (ISGS 6)*, San Diego, July 2014.
- Janda L. (éd.) (2013). *Cognitive Linguistics : The Quantitative Turn*. Berlin and Boston : Mouton De Gruyter.
- Jefferson G. (2004). « Glossary of transcript symbols with an introduction », in G. H. Lerner (éd.) *Conversation Analysis: Studies from the First Generation*. Amsterdam : John Benjamins, 13-31.
- Jewitt C. (éd.) (2009). *The Routledge Handbook of Multimodal Analysis*. Londres : Routledge.
- Kendon A. (1980). « Gesticulation and speech: two aspects of the process of utterances », in M. Ritchie Key (éd.) *The Relationship of Verbal and Nonverbal Communication*. La Haye : Mouton and Co., 207-227.
- Kendon A. (2002). « Some uses of the head shake », *Gesture* 2(2) : 147-182.
- Kendon A. (2004). *Gesture : Visible Action as Utterance*. Cambridge : Cambridge University Press.
- Kipp M. (2014). « ANVIL : A Universal Video Research Tool », in J. Durand, U. Gut et G. Kristofferson (éd.) *Handbook of Corpus Phonology*. Oxford : Oxford University Press, 420-436.
- Krawczak K. (2015). « Near-synonymous epistemic stance predicates in English: A quantitative corpus-driven study of subjectivity », in D. Glynn et M. Sjölin (éd.) *Subjectivity and Epistemicity. Stance strategies in discourse and narration*. Lund : Lund University Press, 311-339.
- Ladewig S. H. (2011). « Putting the cycling gesture on a cognitive basis », *CogniTextes* 6. <http://cognitextes.revues.org/406>.
- Ladewig S. H. (2014a). « Recurrent Gestures » in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill et S. Teßendorf (éd.) *Body - Language - Communication vol. 2*. Berlin et Boston : Walter de Gruyter, 1558-1574.
- LHôte E. et Debras C. (2016). « The Party Conference Speech as a genre event: a multimodal approach », in N. Stukker, W. Spooren et G. Steen (éd.) *Genre in Language, Discourse and Cognition*. Berlin : Mouton De Gruyter, 361-394.
- Linell P. (2005). *The Written Language Bias in Linguistics : its nature, origins and transformations*. New York : Routledge.

- MacWhinney B. (2000). *The CHILDES Project : Tools for Analyzing Talk*. 3rd Edition. Mahwah, NJ : Lawrence Erlbaum Associates.
- Maxwell J. A. (2005). *Qualitative Research Design : an Interactive Approach*. Thousand Oaks et Londres : Sage.
- McClave E. (2000). « Linguistic functions of head movements in the context of speech », *Journal of Pragmatics* 32 : 855-878.
- McNeill D. (1992). *Hand and Mind. What Gestures Reveal about Thought*. Chicago : University of Chicago Press.
- Mittelberg I. (2007). « Methodology for multimodality: One way of working with speech and gesture data », in M. Gonzales-Marquez, I. Mittelberg, S. Coulson, M. Spivey (éd.) *Methods in Cognitive Linguistics*. Amsterdam et Philadelphie : John Benjamins, 225-248.
- Mondada L. (2007). « Multimodal resources for turn-taking : Pointing and the emergence of possible next speaker », *Discourse Studies* 9(2) : 195-226.
- Morgenstern A et Chang N. (2014). « The blossoming of three-argument verbal constructions in child language: a study of “give constructions” », in G. Girard-Gillet (éd.) *Autour du verbe anglais. Construction - lexique - évidentialité*. Paris : Presses de la Sorbonne Nouvelle, 107-120.
- Morgenstern A., Beaupoil-Hourdel P., Blondel M. et Boutet D. (2016). « A Multimodal Approach to the Development of Negation in Signed and Spoken Languages : Four Case Studies », in L. Ortega, A. E. Tyler, H. I. Park et M. Uno (éd.) *The usage-based study of language learning and multilingualism*. Georgetown : Georgetown University Press, 15-36.
- Mortensen K. (2012). « Conversation Analysis and Multimodality », in J. Wagner et K. Mortensen (éd.) *Conversation Analysis and Applied Linguistics. The Encyclopedia of Applied Linguistics (general editor C.A. Chapelle)*. Oxford : Wiley-Blackwell, 1061-1068.
- Müller C. (2004). « Forms and Uses of the Palm Up Open Hand: A case of a gesture family? », in C. Müller et R. Posner (éd.) *The semantics and pragmatics of everyday gestures, Proceedings of the Berlin conference, April 1998*. Berlin : Weidler Buchverlag, 233-256.
- Müller C. (2008). *Metaphors dead and alive, sleeping and waking*. Chicago : University of Chicago Press.
- Müller C., Ladewig S. H. et Bressemer J. (2013a). « Gesture and speech from a linguistic perspective: A new field and its history », in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill et S. Teßendorf (éd.) *Body - Language - Communication vol. 2*. Berlin et Boston : Walter de Gruyter, 55-81.
- Müller C., Bressemer J. et Ladewig S. H. (2013b). « Towards a grammar of gestures : A form-based view », in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill et S. Teßendorf (éd.) *Body - Language - Communication vol. 2*. Berlin et Boston : Walter de Gruyter, 707-733.
- Nosek B. et al. (2015) « Estimating the reproducibility of psychological science ». *Science*, 349 (6251).
- Norris S. (2004). *Analyzing Multimodal Interaction*. Londres : Routledge.
- O'Halloran K. L. et Smith B. A. (éd.). (2011). *Multimodal Studies: Exploring Issues and Domains*. New York et Londres : Routledge.
- ORTOLANG. Outils et Ressources pour un Traitement Optimisé de la LANGue, <https://www.ortolang.fr/>.

- Özyürek A. (2002). « Do speakers design their co-speech gestures for their addressees ? The effects of addressee location on representational gestures », *Journal of Memory and Language* 46(4) : 688-704.
- Priesters M. A. et Mittelberg I. (2013). « Individual differences in speakers' gesture spaces: Multi-angle views from a motion-capture study », in *Proceedings of the Tilburg Gesture Research Meeting (TiGeR), June 19-21, 2013*.
- Sacks H. (1992). *Lectures on Conversation, Volumes I and II*, Edited by G. Jefferson with an Introduction by E.A. Schegloff. Oxford : Blackwell.
- Sacks H. et Schegloff E. A. (2002). « Home position », *Gesture* 2(2) : 133-146.
- Sacks H., Schegloff E. A. et Jefferson G. (1974). « A Simplest Systematics for the Organization of Turn-Taking for Conversation », *Language* 50 (4 : 1) : 696-735.
- Schegloff E. A. (2007). *Sequence organization in interaction : A primer in conversation analysis*. Cambridge : Cambridge University Press.
- Schoonjans S., Brône G. et Feyaerts K. (2015). « Multimodalität in der Konstruktionsgrammatik: Eine kritische Betrachtung illustriert anhand einer Gestikanalyse der Partikel *einfach* », in J. Bücker, S. Günthner et W. Imo (éd.) *Konstruktionsgrammatik V. Konstruktionen im Spannungsfeld von sequenziellen Mustern, kommunikativen Gattungen und Textsorten*. Tübingen : Stauffenburg, 291-308.
- Sloetjes H. et Wittenburg P. (2008). « Annotation by category – ELAN and ISO DCR », in *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*.
- Stam G. et Ishino M. (éd.) (2011). *Integrating Gestures : The Interdisciplinary Nature of Gesture*. Amsterdam et Philadelphie : John Benjamins.
- Steen F. et Turner M. (2013). « Multimodal Construction Grammar », in M. Borkent, B. Dancygier et J. Hinnell (éd.) *Language and the Creative Mind*. Chicago : University of Chicago Press.
- Stivers T. (2008). « Stance, Alignment, and Affiliation During Storytelling : When Nodding Is a Token of Affiliation », *Research on Language and Social Interaction* 41(1) : 31-57.
- Stivers T. (2015) « Coding Social Interaction : A Heretical Approach in Conversation Analysis? », *Research on Language and Social Interaction* 48 (1) : 1-19.
- Stivers T. et Enfield N. J. (2010). « A coding scheme for question-response sequences in conversation », *Journal of Pragmatics* 42 (10) : 2620-2626.
- Svartvik J. (1992). « Corpus linguistics comes of age », in J. Svartvik (éd.) *Directions in Corpus Linguistics, Proceedings of Nobel Symposium 82*. Berlin et New York : Mouton de Gruyter, 7-16.
- Sweetser E. (2012). « Introduction : viewpoint and perspective in language and gesture, from the Ground down », in B. Dancygier et E. Sweetser (éd.) *Viewpoint in Language : A Multimodal Perspective*. Cambridge : Cambridge University Press, 1-24.
- Whewell W. (1840). *The Philosophy of the Inductive Sciences, Founded Upon Their History*. London : John W. Parker.
- Wu S. et Cienki A. (2014). « Types of Multimodality in Transitive and Intransitive Constructions », communication au colloque *Sixth International German Cognitive Linguistics Association Conference*. Erlangen, Allemagne, 30 Septembre – 2 Octobre 2014.
- Zima E. (2014). « English multimodal motion constructions. A construction grammar perspective », *Studies van de BKL - Travaux du CBL - Papers of the LSB*, vol. 8. <http://uahost.uantwerpen.be/linguist/SBKL/sbkl2013/Zim2013.pdf>.

NOTES

1. Page visitée le 3 mai 2017 : <http://corpus.byu.edu/corpora.asp>.
2. Müller (2008) remarque cependant qu'il existe des indices de la conscience que les locuteurs ont de leurs propres gestes, notamment lorsqu'ils produisent leurs gestes assez haut de manière à les faire entrer dans le champ de vision de l'interlocuteur, ou bien lorsqu'ils regardent leur propre geste.
3. « *What a rudimentary speech act theoretic analysis misses, and I suspect a sophisticated one will miss as well, is that parties to real conversations are always talking in some sequential context. I refer here not to social contexts like offices, classrooms, or families, but sequential contexts formulated in terms of more or less proximately preceding talk and the real jobs of projecting further talk which utterances can do, for which they can be inspected by their recipients, an inspection to which speakers must therefore be reflexively attentive. Such prior and prospective contexts are inescapably implicated in the real life projects, however humble or exalted, which are being prosecuted through the talk. These real life projects, and the sequential infrastructure of talk-in-interaction, are involved in the production and analysis of talk by the parties in such intimate detail that we are only beginning to understand it. But it is clear that temporality and sequentiality are inescapable; utterances are in turns, and turns are parts of sequences; sequences and the projects done through them enter constitutively into utterances like the warp in a woven fabric.* » (Schegloff 1988 : 61).
4. Nous reprenons ici la traduction proposée par Ferré, 2011.
5. MacWhinney (2000) : <http://talkbank.org/clan/>.
6. voir plus bas pour les références sur ces logiciels.
7. Boersma & Weeninck (2017) : <http://www.fon.hum.uva.nl/praat/>.
8. Bigi (2015) : <http://www.sppas.org/>.
9. Goldman (2011) : <http://latlcui.unige.ch/phonetique/easyalign.php>.
10. Sloetjes & Wittenburg (2008) : <https://tla.mpi.nl/tools/tla-tools/elan/>.
11. Kipp (2014) : <http://www.anvil-software.org/>.
12. ANR-08-COMM-0021, Communication Langagière chez le Jeune Enfant. <http://colaje.scicog.fr/>.
13. Projet *Towards a grammar of gesture*, page visitée le 4 mai 2017 : <http://www.togog.org/en/about/subprojects/p5a.html>.
14. <http://talkbank.org/>.
15. Child Language Data Exchange System : <http://childes.psy.cmu.edu/>.
16. *Common Language Resources and Technology Infrastructure* : www.clarin.eu.
17. Très Grande Infrastructure de Recherche pour les Humanités Numériques : <http://www.huma-num.fr/>.
18. Outils et Ressources pour un Traitement Optimisé de la LANGue <https://www.ortolang.fr/>.
19. Collections de corpus Oraux Numériques : <http://cocoon.huma-num.fr/>.
20. Multimodalité et nouvelles formes de communication : <https://corpuscomere.wordpress.com/2016/05/10/corli-groupe-multicom/>.
21. Corpus, Langues, Interactions : <https://corli.huma-num.fr/>.
22. <http://ircom.huma-num.fr/>
23. SBCSAE : <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>.
24. Corpus de langues parlées en interaction : <http://clapi.ish-lyon.cnrs.fr/>.
25. UMR 5191 Interactions, Corpus, Apprentissages, Représentations.
26. <http://www.redhenlab.org/>
27. Professeur, Département des sciences cognitives, Case Western Reserve, Cleveland, États-Unis.

28. Professeur assistant, Département d'études sur la communication, UCLA, États-Unis.

29. <http://newsscape.library.ucla.edu/>

30. Les questions de méthodologie en lien avec la taille des corpus et les complémentarités entre analyses qualitatives et quantitatives ont récemment été au centre de la réflexion épistémologique des chercheurs en linguistique anglaise en France, comme en témoigne l'organisation récente des colloques Engcorpora à l'Université Paris Est-Créteil en avril 2015 (*English linguistics and corpora*, organisé par Lionel Dufaye (UPEM, ALAES), Lucie Gournay (UPEC) et Philip Miller (Université Paris Diderot, ALAES), voir <https://engcorpora2015.sciencesconf.org/>) et NACL1 à l'Université d'Avignon en juin 2016 (Nouvelles approches du corpus en linguistique anglaise organisé par Aurélie Barrière, Annemarie Dinvaut et Graham Ranger à l'Université d'Avignon, voir <https://nacla1.sciencesconf.org/>).

RÉSUMÉS

Cet article interroge la question des corpus en analyse multimodale des corpus oraux d'interactions filmées dans le cadre d'une linguistique qui intègre l'étude des gestes à celle du discours pour rendre compte des pratiques langagières dans leur globalité. Ce type d'approche est traditionnellement associé à des petits corpus supports d'analyses qualitatives. À l'appui d'un tour d'horizon de recherches contemporaines en analyse multimodale des interactions orales, nous montrons que l'étude linguistique des gestes se prête à une grande variété d'approches innovantes au plan méthodologique. En pratique, il reste à ce jour difficile de construire des grands corpus pour l'analyse gestuelle, parce qu'il s'agit d'une entreprise longue et coûteuse (enregistrement, transcription, annotation) : les petits corpus sont nombreux parce qu'ils sont plus faciles à construire pour un chercheur seul ou une petite équipe de recherche aux moyens limités. Cette situation tend cependant à évoluer, et ce de plusieurs manières, grâce aux collaborations nationales et internationales et aux nouveaux outils numériques. Les « petits » corpus n'en restent pas moins sources de résultats précieux en analyse gestuelle. Pour diverses raisons, les standards internationaux de l'évaluation de la recherche tendent à valoriser l'analyse quantitative aux dépens de l'analyse qualitative, faisant oublier que ces deux types d'approches ont des objectifs certes différents mais complémentaires : ils permettent de répondre à des questions de recherche différentes (étroites ou larges) et de proposer des savoirs d'origines différentes (généralisabilité des résultats ou généralisabilité des processus observés), dont la légitimité est elle aussi de nature différente (réfutabilité des résultats ou consilience des experts). Nombre de travaux récents en analyse gestuelle encouragent à remettre en question la dichotomie traditionnelle entre analyses qualitatives et quantitatives, en intégrant les deux approches de diverses manières.

This article discusses the use of corpora in the multimodal analysis of videotaped data, as part of a linguistic approach that includes speech and gesture. This type of approach is traditionally based on the qualitative analysis of small corpora. I start with an overview of contemporary research in the multimodal analysis of spoken interaction, showing the variety of current innovative methodologies. In practice, creating large corpora for gesture analysis is still a challenge today, because recording, transcribing and annotating data are slow and costly processes; for scholars working alone, in small teams or on unfunded projects, small corpora are just easier to create. Things are starting to change, thanks to the development of national and

international shared databases and digital tools. “Small” corpora remain precious material for gesture analysis though. For multiple reasons, international standards of research evaluation tend to value quantitative over qualitative analysis, while the two approaches have different, but complementary aims, namely addressing different research questions (with a narrow or wide scope) and produce different forms of knowledge (generalizability of results vs generalizability of processes), based on different forms of scientific legitimacy (falsifiability of results or consilience of experts). Recent studies in gesture analysis question the traditional dichotomy between qualitative and quantitative analysis and integrate the two approaches in various ways.

INDEX

Mots-clés : multimodalité, gestualité, interactions orales, analyse qualitative et quantitative

Keywords : multimodality, gesture studies, spoken interactions, qualitative and quantitative analysis

AUTEUR

CAMILLE DEBRAS

Université Paris Nanterre