



## Les processus inférentiels dans l'interaction multimodale en L2

*Inferential Processes in Multimodal L2 Interactions*

Saghie Sharifzadeh and Cédric Sarré

---



### Electronic version

URL: <http://journals.openedition.org/tipa/3746>

DOI: 10.4000/tipa.3746

ISSN: 2264-7082

### Publisher

Laboratoire Parole et Langage

### Electronic reference

Saghie Sharifzadeh and Cédric Sarré, "Les processus inférentiels dans l'interaction multimodale en L2", *TIPA. Travaux interdisciplinaires sur la parole et le langage* [Online], 36 | 2020, Online since 01 June 2020, connection on 26 January 2021. URL: <http://journals.openedition.org/tipa/3746> ; DOI: <https://doi.org/10.4000/tipa.3746>

---

This text was automatically generated on 26 January 2021.



La revue *TIPA. Travaux interdisciplinaires sur la parole et le langage* est mise à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

---

# Les processus inférentiels dans l'interaction multimodale en L2

*Inferential Processes in Multimodal L2 Interactions*

Saghie Sharifzadeh and Cédric Sarré

---

Looking at both learners' speech and gesture holds promise for understanding the second language acquisition process, learners' interlanguage systems, and their thinking for speaking patterns<sup>1</sup> (Stam, 2007 : 122).

- 1 Si la littérature portant sur le rôle du corps dans les discours multimodaux à visée didactique commence à s'accroître, comme en témoignent les nombreuses recherches récentes sur la question (Aden, 2017 ; Azaoui, 2015 ; Sime, 2001, 2008 ; Tellier, 2014, 2016), les travaux consacrés à la multimodalité des interactions apprenants-apprenants en langue seconde (désormais L2) restent extrêmement moins nombreux que ceux portant sur les interactions entre enseignant et apprenant(s) ou encore entre locuteur(s) natif(s) et apprenant(s) (Kurhila, 2006). Peu d'études ont été publiées à ce jour sur la mobilisation de ressources autres que verbales dans les échanges apprenants-apprenants (Gullberg, 2010). Si Gullberg (1998) étudie les interactions apprenants-apprenants en tandem, l'utilisation de gestes par l'énonciateur-apprenant ou encore la manière dont ces gestes sont traités et interprétés par ses co-énonciateurs apprenants d'une même L2 sont autant de questions qui restent en suspens.
- 2 Notre contribution vise à combler certains de ces manques par une étude empirique exploratoire portant sur les processus inférentiels mis en œuvre par des interactants-apprenants de l'anglais L2 pour interpréter les ressources autres que verbales mobilisées par leurs co-énonciateurs-apprenants. Il ne s'agit donc pas de discours didactiques *stricto sensu*, dans la mesure où les interlocuteurs-apprenants n'ont pas pour objectif premier de faire apprendre la L2 à leurs partenaires, mais plutôt d'échanges relativement symétriques (entre pairs et non entre expert et non-expert, même si l'on peut considérer que les compétences langagières variées des apprenants introduisent une certaine forme d'asymétrie au sein de chaque groupe). Ces échanges

visent à répondre à la consigne d'une tâche donnée en contexte institutionnel et relèvent donc d'une intention didactique. L'énonciation résultant d'un processus cognitif complexe (Goodwin & Goodwin, 1986 ; Kerbrat-Orecchioni, 1992, 1994 ; Mondada, 2008b) intégrant ressources paraverbales (volume de la voix, débit, intonation) et ressources « posturo-mimo-gestuelles » (Cosnier, 1985), nous tenterons de mesurer l'impact de ces dernières dans les échanges apprenant(s)-apprenant(s) en présentiel. Dans un premier temps, nous présenterons le cadre théorique de notre recherche avant d'aborder la méthodologie adoptée pour constituer le corpus à l'étude. Nous proposerons enfin une analyse qualitative d'extraits du corpus qui permettra de mettre en lumière les multiples fonctions de certaines ressources PMG mobilisées par les énonciateurs-apprenants d'une L2.

## 1. Cadre théorique

### 1.1. L'inférence

- 3 L'inférence consiste, dans son acception traditionnelle, à interpréter les éléments non littéraux, non-dits ou encore implicites d'un énoncé donné (Chernyshova & Traverso, 2017 : para. 5). Elle constitue en cela un procédé central à l'accès au sens des énoncés :

A brief reconstruction of the steps [is] necessary to derive primary illocution from literal illocution [...]. Unless a hearer has some inferential strategy for finding out when primary illocutionary points differ from literal illocutionary points, he has no way of understanding indirect illocutionary acts<sup>2</sup> (Searle, 1979 : 34-35).

- 4 Notre étude se fonde ainsi sur le fait que le sens d'un énoncé, qui est le produit de l'énonciation d'une phrase en contexte, dépasse largement celui d'une phrase, simple combinaison de mots dans une structure grammaticale (cf. Austin, Searle ou Grice). Les phrases, dotées d'une signification qui leur est donnée par le code linguistique, sont augmentées d'un sens produit dans un contexte particulier (Sperber & Wilson, 1986). Mise en contexte, une phrase devient donc un énoncé qui donne lieu à des inférences de la part du co-énonciateur :

There are **two sorts of contextual information**, one much more restricted in scope than the other. Information that plays the limited role of combining with linguistic information to determine content (in the sense of fixing it) is restricted to a short list of variables, **such as the identity of the speaker and the hearer and the time and place of an utterance**. Contextual information in the broad sense is **anything that the hearer is to take into account to determine (in the sense of ascertain) the speaker's communicative intention**<sup>3</sup> (Bach, 1997 : 39).

- 5 Grâce à un processus pragmatique d'interprétation de ces indices de contextualisation (cf. Gumperz 1989), le co-énonciateur pourra ainsi accéder à l'intention de l'énonciateur. Pour opérer ces inférences, le co-énonciateur peut recourir à des éléments du contexte aussi bien intra-textuels, c'est-à-dire présents dans l'énoncé de son interlocuteur, qu'extra-textuels :

Most verbal utterances are a complex of linguistic, paralinguistic, facial and vocal gestures, which appear to function as a signal receiving a unified interpretation<sup>4</sup> (Carston, 2000 : 824).

- 6 Ainsi, le paraverbal et les ressources posturo-mimo-gestuelles (PMG) participent de l'intention communicative de l'énonciateur, plutôt que du « locuteur » puisque nous considérons qu'un énoncé n'est pas constitué que de mots. C'est précisément parce que

la communication est multimodale par nature qu'on n'a pas besoin de s'appuyer nécessairement sur le verbal pour opérer des inférences. En effet, comme le rappelle David (2017 : para. 4) :

les processus inférentiels [...] peuvent porter sur les éléments de la langue, les attitudes, les comportements, les pensées, l'argumentation, les émotions, les attentes, les intentions ou encore la culture de l'interlocuteur.

## 1.2. La multimodalité des échanges avec des apprenants d'une L2

- 7 Pour l'énonciateur, la multimodalité de l'énonciation est fondée sur l'association de sons articulés à des mouvements corporels ; pour le co-énonciateur, la multimodalité implique donc non seulement des mots, mais aussi la prosodie, les gestes, les mimiques faciales et les regards, autant de ressources PMG qu'il va devoir traiter et interpréter.

La parole<sup>5</sup> est multimodale. Plus exactement, la parole est un comportement qui associe des sons articulés à des mouvements corporels, et lorsque vous percevez une conduite langagière, vous percevez celle-ci dans les différentes modalités sensorielles dont vous êtes équipés. Autrement dit, non seulement vous reconnaissez les mots prononcés, mais vous les entendez avec leur musique (ce qu'on appelle la prosodie), et vous percevez en même temps les gestes, les mimiques, les regards... l'ensemble des mouvements corporels produits par le locuteur (Colletta, 2004 : 15).

- 8 Ainsi les gestes, et les modalités PMG plus généralement, déclenchent-ils des processus inférentiels permettant de reconstruire le sens des énoncés, *a fortiori* lorsqu'ils sont difficiles à mettre en mots. Dans le cas d'énoncés ambigus, erronés ou incomplets, les gestes coverbaux, c'est-à-dire produits concomitamment à l'acte de parole (McNeill, 1992, 2005 ; Goldin-Meadow, 2003), peuvent générer un travail d'inférence chez le co-énonciateur et lui faciliter l'accès au sens (Tellier, 2014). On peut supposer que le caractère multimodal de la communication interpersonnelle sera d'autant plus perceptible dans les échanges entre apprenants du fait de leur *déficit de compétences (compréhension parcellaire du discours de l'autre, manque de lexique, différence culturelle...)* (David, 2017 : para. 5).
- 9 Nous souhaitons vérifier cela dans la présente étude car, si les gestes pédagogiques, produits par l'enseignant, ont fait l'objet de nombreux travaux (Tellier, 2008, 2010 ; Azaoui, 2015 ; Tellier & Cadet, 2014), les connaissances portant sur la mobilisation de ressources PMG par les énonciateurs-apprenants dans les échanges en L2 (et sur les fonctions de ces ressources) sont insuffisantes, qu'il s'agisse du point de vue de l'énonciateur ou de celui du co-énonciateur. Ce que l'on sait aujourd'hui a généralement été formalisé dans le cadre d'études portant sur des échanges entre apprenants et natifs ou bien entre apprenants et enseignant, études qui révèlent que les interactants-apprenants ont davantage tendance à recourir aux indices contextuels posturo-mimo-gestuels qu'au cours d'échanges en L1 : l'importance du regard dans la gestion des tours de parole a par exemple pu être notée dans le cadre d'interactions en L2 entre apprenants et enseignant (Foerster, 1990) et nombre de travaux (Gullberg, 1998 ; Hadar *et al.*, 2001 ; Nobe, 2001 ; Yoshioka, 2005) montrent que le recours aux gestes est également plus marqué lors d'échanges en L2 qu'en L1, en particulier le recours aux gestes déictiques (Sherman & Nicoladis, 2004). Cependant, contrairement aux locuteurs natifs, les apprenants d'une L2 ont rarement recours aux gestes « compensatoires » se substituant à la parole. Gullberg (1998 : 64) présente deux types de gestes compensatoires :

In most studies, the term signifies those gestures occurring as substitution for words when the verbal-vocal channel collapses. Gesture is then either relied on entirely for the transmission of the message, or used to replace single items in an utterance, a process referred to by Slama-Cazacu (1976) as 'mixed syntax'.

Compensation is also used in a more general sense. Some compensatory gestures are said to complement the verbal-vocal message such that they augment, supplement or enhance it. Such complementing gestures occur simultaneously with speech<sup>6</sup>

- 10 Les apprenants d'une L2 recourent majoritairement au deuxième type de gestes compensatoires ainsi définis, ce qui peut s'expliquer par la difficulté d'interprétation du premier type entre locuteurs d'une langue seconde (Feyereisen *et al.*, 1988).
- 11 Les gestes sont utilisés comme des stratégies de communication au service du co-énonciateur (*addressee-directed*) et ont avant tout une fonction communicative (Gullberg, 1998) pouvant servir de béquille au verbal. Le but du recours aux gestes est pluriel : expliciter des éléments lexicaux, clarifier des problèmes de co-référence, signaler une recherche lexicale, réguler les tours de parole, etc. (Gullberg, 2010). Certaines études récentes insistent néanmoins sur le fait que les gestes peuvent aussi être au service de l'énonciateur (*speaker-directed*) et qu'ils ont donc également une fonction cognitive avec de multiples objectifs : aider à la planification conceptuelle (Hostetter & Alibali, 2004) ou linguistique (Alibali *et al.*, 2000), réduire la charge cognitive (Goldin-Meadow *et al.*, 2001), ou encore mobiliser du lexique (Frick-Horbury, 2002 ; Ravizza, 2003).
- 12 Dans la mesure où tout énonciateur adapte son discours à la situation d'énonciation et aux co-énonciateurs en présence, il est légitime de penser qu'il mobilise également des ressources PMG en fonction de ces mêmes paramètres. Si l'adaptation de l'action multimodale de l'enseignant selon le contexte et ses co-énonciateurs a fait l'objet de nombreuses études (Azaoui, 2014, 2015 ; Tellier & Stam, 2012), qu'en est-il des échanges en L2 entre apprenants ? Nous pouvons partir du postulat selon lequel l'apprenant d'une L2 mobiliserait des ressources PMG spécifiques à un échange en L2 avec d'autres apprenants de cette L2 et ces ressources pourraient être à l'origine de processus inférentiels, peut-être même en plus grand nombre si l'on en croit David (2017 : para. 15) : *Si l'inférence est au cœur de nos interactions orales en L1, elle est d'autant plus présente dans les interactions exolingues.*
- 13 Notre étude vise à examiner le rôle des ressources PMG dans l'interaction entre apprenants de l'anglais L2 pour celui qui les mobilise, mais aussi et surtout à analyser la manière dont ces ressources sont perçues par le co-énonciateur qui engage des processus inférentiels afin d'accéder au sens visé. Nous nous intéresserons aux interactions entre apprenants de la même L2 (l'anglais) et non aux interactions entre apprenants de L2 différentes dans un dispositif de type tandem (Gullberg 1998), ni aux interactions entre locuteur(s) natif(s) et apprenant(s) ou encore entre enseignant et apprenant(s), déjà maintes fois étudiées par ailleurs. Nous tenterons d'apporter une réponse aux questions suivantes : comment et dans quel(s) but(s) les ressources PMG sont-elles utilisées dans des échanges entre apprenants d'une L2 ?

## 2. Cadre méthodologique

### 2.1. Le corpus

- 14 Les interactions que nous avons étudiées se sont déroulées dans le cadre d'un cours obligatoire d'anglais L2 à destination d'étudiants de première année de Master (M1) spécialisés dans les sciences du vivant. Il s'agit donc d'un cours du secteur LANSAD (Langues pour Spécialistes d'Autres Disciplines que les langues), caractérisé par le niveau très hétérogène des étudiants, mais de niveau B1 (CECRL 2001) pour la plupart (selon les résultats obtenus au test d'autopositionnement Dialang (2003)). L'objectif visé du cours était le niveau B2. Les échanges constituant le corpus se sont déroulés en présentiel, au sein de 12 groupes de 3 ou 4 étudiants, dans le cadre de la réalisation de deux tâches collaboratives :
- 15 1. La première tâche, dans le cadre du scénario pédagogique *Becoming a biological scientist*, invitait les apprenants à réaliser une brochure à destination de futurs étudiants, dans le but de faire la promotion des cursus dans leur domaine de spécialité (les sciences biologiques);
- 16 2. Dans le cadre d'un scénario pédagogique sur la nourriture génétiquement modifiée (*GM food*), la seconde tâche consistait à endosser le rôle d'étudiants en thèse devant assister à un colloque sur cette thématique et à élaborer, à l'issue du colloque, une brochure de vulgarisation scientifique destinée au grand public et visant à expliquer ce qu'est la nourriture génétiquement modifiée et à en présenter les risques.
- 17 Notons que les deux tâches intervenaient à la fin de leur séquence pédagogique respective. La constitution des groupes avait été faite suite aux résultats obtenus au test d'autopositionnement de manière à parvenir à des groupes volontairement hétérogènes : si les étudiants constituant chaque groupe se connaissaient, ils n'avaient donc pas forcément l'habitude de travailler ensemble. L'enregistrement des interactions s'est déroulé selon la même configuration pour l'ensemble des groupes : les interactants étaient assis autour d'une table, la caméra leur faisait face, l'enregistrement était lancé par un technicien qui quittait ensuite la salle ; les apprenants se retrouvaient donc seuls lors de leurs échanges. La participation à l'étude s'est faite sur la base du volontariat (la totalité du groupe d'étudiants ayant accepté d'y participer). Les consentements éclairés des participants ont été recueillis en leur fournissant le minimum d'informations sur l'étude (il s'agit d'une étude sur l'apprentissage de l'anglais de spécialité à partir de tâches concrètes) de manière à ne pas influencer leurs comportements et productions et ainsi à ne pas biaiser les résultats.
- 18 Les interactions entre apprenants générées lors de la réalisation de ces deux tâches ont été filmées, puis transcrites avec la suite logicielle EXMARaLDA (Schmidt *et al.*, 2013) et, en particulier, l'éditeur pour la transcription de la langue parlée au format partition (*Partitur Editor 1.4.5*). Le corpus comporte 24 enregistrements vidéos d'une dizaine de minutes chacun, pour une durée totale cumulée de 251 minutes. Chaque enregistrement correspond au travail préparatoire (de type prise de décision et/ou résolution de problème) à la réalisation de la tâche conçue par l'enseignant (qui se trouve être également l'un des chercheurs). Les apprenants étaient habitués à travailler sur ce type de tâches, l'ensemble de leur formation étant constituée d'une série de séquences conduisant chacune à la réalisation d'une tâche ou plusieurs tâches de ce

type. Le corpus a ensuite fait l'objet d'un étiquetage en lien avec les phénomènes que nous avons souhaité étudier. Dans chaque extrait, les énonciateurs sont numérotés de gauche à droite pour faciliter la description de l'allocation des tours de parole : le premier énonciateur à gauche est donc E1, le deuxième E2, et le dernier à droite de la capture d'écran E3 ou E4 selon le nombre de participants. Le numéro de l'énonciateur n'est spécifié sous les gros plans qu'en cas d'ambiguïté. Une capture d'écran présentant la configuration générale du groupe est proposée en début de chaque exemple<sup>7</sup>.

## 2.2. Les phénomènes étudiés

- 19 Dans cette étude exploratoire, nous avons adopté une approche éémique afin d'examiner les ressources PMG utilisées pour communiquer un message (énonciateur) et décoder un message (co-énonciateur), ainsi que les traces de processus inférentiels liés à ces ressources. Nous avons donc procédé à un relevé d'indices composé d'une paire « ressource mobilisée par l'énonciateur / réaction du co-énonciateur », comme nous y invitent Chernyshova & Traverso :

L'enjeu est de relever les indices renvoyant à ce qu'un locuteur a dit et à ce que l'interlocuteur a pu inférer, aux différents moments de l'échange. C'est précisément grâce à ces indices que l'analyste peut retracer les inférences effectuées par les interactants (2017 : para. 19).

- 20 Les ressources PMG ont fait l'objet d'un étiquetage sur une ligne de transcription spécifique dans le logiciel EXMARaLDA à partir d'une première lecture "flottante" du corpus. Le codage a été effectué par un premier chercheur, puis par le deuxième chercheur. Ce contre-codage a permis d'atteindre un taux d'accord proche de 95 %. Les conventions de transcription adoptées (transcription orthographique et ressources PMG illustrées par captures d'écran) sont présentées en annexe. Notre point de départ a été la typologie de McNeill (1992 et 2005) et sa classification en gestes déictiques (gestes de pointage), iconiques (gestes représentant un concept concret), métaphoriques (gestes représentant un concept abstrait), battements (gestes ponctuant le verbal sans y ajouter de contenu sémantique) et Butterworth (gestes symbolisant une recherche lexicale), auxquels nous avons ajouté les expressions faciales et les regards, ainsi que les mouvements corporels autres que les gestes au sens mcneillien (changements d'orientation corporelle, par exemple). Aucune occurrence d'emblème ne figure dans notre corpus.
- 21 Dans une approche interactionniste et selon les principes de l'analyse conversationnelle, nous avons abordé les processus inférentiels comme des *phénomènes descriptibles* (Chernyshova & Traverso, 2017 : para. 19) permettant aux interlocuteurs de se rendre mutuellement compréhensibles et d'échanger de manière efficace. Nous avons donc examiné l'ensemble des ressources interactionnelles PMG du corpus qui nous ont semblé être à l'origine d'un travail inférentiel de la part des participants, et nous sommes parvenus à trois catégories (ou modes de fonctionnement) liant ressources PMG et inférences, que nous allons examiner dans la section suivante.

## 3. Analyse des données

- 22 Notons en préambule à notre analyse que les ressources PMG sont mobilisées par un énonciateur donné pour des raisons variées et que l'interprétation qui en est faite, bien



que fondée sur des indices contextuels, ne peut que rarement être indiscutable (Eco, 1984). L'objectif d'une telle analyse est donc de déterminer quelles interprétations sont les plus plausibles pour une situation de communication donnée (Arndt & Janney, 1987 : 357), soulignant ainsi *la nature inférentielle des signes* (Eco, 1984 : 39).

### 3.1. Les inférences fondées sur des ressources multimodales au service de la tentative de mise en mots

- 23 Ce premier mode de fonctionnement correspond essentiellement à des cas de mobilisation de ressources multimodales lors d'une recherche lexicale.
- 24 Dans les exemples 1 à 3, les étudiants devaient imaginer une brochure permettant de promouvoir les études en sciences biologiques.
- 25 Dans l'exemple 1, l'énonciatrice E1 prend la parole et expose son point de vue.

#### Exemple 1



1a  
what we put on the brochure (0.287)  
mouvement de mains rectangulaire, geste iconique d'une brochure, tête/regard vers le bas



1b  
to erm  
geste Butterworth de doigts frottés, tête tournée vers les co-énonciateurs



1c  
(1.772)  
tapotement sur la table (Butterworth) + tête tournée vers l'autre côté



1d  
attract  
geste métaphorique pour exprimer *attract*, tête tournée vers les co-énonciateur

- 26 L'énonciatrice en place, E1, fait un geste iconique coverbal (1a) puis montre qu'elle cherche un mot par un geste de doigts frottés (1b) produit dans une pause remplie. Après ce premier geste Butterworth, elle tapote les doigts sur la table (deuxième geste Butterworth) tout en changeant l'orientation de sa tête (1c) pour éviter le regard de ses co-énonciateurs. Notons que ce détournement du regard (*gaze aversion / gaze withdrawal*) est typique de la recherche de mots (Goodwin & Goodwin, 1986 ; Fehr & Exline, 1987 ; Strömquist, 1987). En 1c, le tapotement de la table n'est ni coverbal au sens strict, ni compensatoire, ni même « extra-communicatif », c'est-à-dire *étranger à la fois à la communication et à sa stratégie* selon la définition de Cosnier (1982). Selon Kinsbourne (1972), l'orientation assez typique du regard de l'énonciatrice vers le côté est révélatrice de la recherche de solutions à un problème verbal. Si le geste a avant



tout ici une fonction cognitive, au service de l'énonciatrice qui tente de mobiliser le lexique qui lui fait défaut, il a aussi une fonction communicative, celle de faire savoir à ses co-énonciateurs qu'elle n'a pas terminé son tour de parole, vraisemblablement couplée à une fonction modératrice du stress de la recherche lexicale et de l'attente que celle-ci engendre (*procédé auto-calmant* chez Cosnier, 1994, notamment). Les ressources PMG caractéristiques d'une recherche lexicale ont en effet souvent pour objectif de montrer aux co-énonciateurs que l'énonciateur en place ne souhaite pas céder son tour de parole malgré son silence (Ekman, 1979). Une fois le terme idoine produit (1d) concomitamment au geste métaphorique exprimant *attract*, l'un de ses co-énonciateurs ratifie sa proposition et l'échange peut alors se poursuivre.

27 Dans l'exemple 2, c'est au tour de l'énonciateur E2 de s'exprimer sur le même thème.

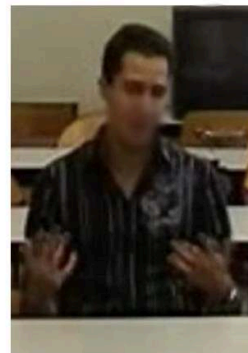
### Exemple 2



2a  
it's very important to  
show the different sec-  
tors = to euh  
geste Butterworth  
frotté



2b  
(3.69)  
mimique faciale  
(sourcils levés, regard  
vers le bas, bouche  
retroussée vers le bas) +  
début de battement  
(Butterworth)



2c  
(0.451)  
du battement au geste  
métaphorique

28 E2 cherche également ses mots : la recherche lexicale est accompagnée d'un geste de doigts frottés dans une pause remplie (2a), d'une mimique faciale (2b) et d'un mouvement des mains de type battement (2b + 2c) : la fonction de ces gestes est avant tout cognitive dans la mesure où ils semblent aider l'énonciateur à produire le terme qui ne vient pas (le geste frotté, en particulier) et apparaissent comme une extériorisation, sous forme de gestes, de sa recherche lexicale. Ils semblent cependant également remplir une fonction communicative : le geste en 2b est initié sous forme de battement – signalant que le processus d'accouchement du terme visé est en cours (cf. exemple 9) – mais son itération le rend métaphorique de l'accumulation (en l'occurrence, d'une multitude de propriétés) ; le geste permet alors d'exprimer un sens proche de *to show the diversity of the field*. Les co-énonciateurs comprennent que l'énonciateur en place est bloqué ; ils interviennent à tour de rôle pour dire avoir compris ce qu'il cherche à exprimer (*yeah I see*, puis *I know what you mean... I see where you want us to...*) mais sans jamais produire la fin de l'énoncé laissé en suspens. En contexte, ce « battement métaphorique » leur aurait ainsi permis de déduire l'intention de communication de l'énonciateur sans pour autant qu'ils soient capables de la mettre

en mots. Les mouvements de main de l'énonciateur en place s'accompagnent d'une pause verbale de plus de 4 secondes qui permet aux co-énonciateurs de comprendre qu'il est bloqué.

### Exemple 3



1.	E4	it's import = yes biology is important to = to understand euh
2.		the human being (0.567)
3.	E3	yes yes
4.	E4	and euh save people because erm # (0.462)# we could erm #
	im.	ooooooooooooooooooooooooooooioi#3aooooi#3booooooooooi#3c
5.		(3.822) # ermoo# (2.52) can you help me ? (1.134)
	im.	oooooo#3doooi#3e
6.	E1	what do you say ?
7.	E3	what do you want to say ? ((rires))
8.	E4	erm when erm people are sick (0.588) we could erm
9.	E3	to erm (2.0) cure
10.	E4	cure thanks



3a  
geste Butterworth  
circulaire de la  
main

3b  
tête et regard vers  
le haut

3c  
geste Butterworth  
frotté

3d  
tapotement Butte-  
rworth sur la table

3e  
fin du change-  
ment  
d'orientation de la  
tête (vers le bas et  
sur le côté) et  
direction du  
regard vers le bas

- 29 L'énonciatrice en place dans l'exemple 3, à savoir E4, tente d'expliquer que les connaissances en biologie peuvent sauver des vies mais elle ne trouve pas le mot pour compléter l'énoncé *biology is important to save people because erm we could erm....* Le recours à une série de ressources multimodales suit son énonciation, ressources mobilisées sur une période de plus de 6 secondes : geste circulaire de la main en 3a (qui pourrait signifier « c'est quoi le mot, déjà ? »), orientation de la tête et du regard vers le haut (3b), « regard à l'aide » vers ses camarades accompagné d'un geste Butterworth frotté (3c), puis tapotement Butterworth insistant sur la table (3d), et enfin orientation de la tête et du regard vers le bas (3e). Malgré ce besoin de renfort, les co-énonciateurs ne semblent pas en mesure de venir en aide à l'énonciatrice qui ne leur fournit pas assez d'informations. Elle finit donc par mettre en mots ce besoin : *can you help me ?* (ligne 5). Deux de ses camarades lui demandent d'éclaircir son propos (lignes 6 et 7), ce qui pousse l'énonciatrice à la glose *When people are sick, we could...* (ligne 8) accompagnée d'un autre geste circulaire de la main pour montrer la progression, geste quasi iconique, ce qui permet à tous les co-énonciateurs d'accéder au vouloir dire de l'énonciatrice en place et de le mettre en mot : *cure* (ligne 9), repris par l'énonciatrice E4 à la ligne suivante.
- 30 Dans la mesure où le détournement du regard et autres froncements de sourcils sont généralement des ressources PMG révélatrices d'un effort cognitif de la part de l'énonciateur (Arndt & Janney, 1987), on voit assez clairement ici la double fonction de cette série de ressources multimodales : si elles ont une fonction cognitive d'aide à la recherche lexicale du mot *cure* (geste circulaire de la main, détournement du regard pendant la recherche du terme qui fait défaut, geste frotté typiquement associé à ce type de recherche), elles ont également une forte dimension communicative en ce qu'elles permettent usuellement de signaler un processus de recherche en cours (tapotement sur la table pour dire « il manque quelque chose ») ou un besoin d'achèvement collaboratif (regard à l'aide, regard baissé pour exprimer un abandon, etc.).
- 31 En résumé, les inférences s'appuyant sur les ressources multimodales mobilisées lors de tentatives de mise en mots portent donc à la fois sur (1) le contenu sémantique de la forme manquant dans la structure syntaxique d'un énoncé (*attract* dans l'exemple 1, *diversity* dans l'exemple 2 ou encore *cure* dans l'exemple 3) et sur (2) l'allocation des

tours de parole (l'énonciateur en place souhaite-t-il céder son tour de parole pour qu'on lui vienne en aide ou bien cherche-t-il au contraire à garder la parole dans l'espoir que le terme idoine lui revienne ?). Aussi les ressources multimodales jouent-elles un rôle dans la régulation des tours de parole.

### 3.2. Les inférences fondées sur des ressources multimodales au service de la gestion des tours de parole

- 32 Les travaux sur la gestion des tours de parole s'intéressent à la manière dont les co-énonciateurs régulent collaborativement le flux de la conversation (Allwood *et al.*, 2007), processus qui implique la mobilisation de ressources verbales et PMG (Kendon, 1967 ; Duncan, 1972 ; Goodwin, 1981 ; Hadar *et al.*, 1984 ; Ford & Thompson, 1996 ; Navaretta & Paggio, 2013). Plus précisément, regards (Kendon, 1967), mouvements des mains (Duncan, 1972), mouvements de la tête (Hadar *et al.*, 1984) et expressions faciales (Navaretta & Paggio, 2013) font partie des ressources PMG généralement mobilisées dans la gestion des tours de parole. Certaines études ont pu montrer que près de 20 % des ressources PMG mobilisées lors d'une interaction ont une fonction régulatrice des tours de parole (Navaretta & Paggio, 2013 : 139). Pour jouer pleinement leur rôle, de telles ressources doivent cependant faire l'objet d'inférences de la part des co-énonciateurs, comme nous allons le montrer dans la suite du présent travail.
- 33 Les étudiantes de l'extrait suivant travaillent sur la promotion des études dans le domaine des sciences biologiques à travers l'élaboration d'une brochure. E3 a la parole.

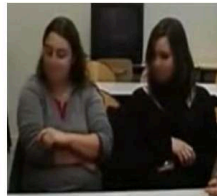
#### Exemple 4



4a

E3. what do you think about this?

orientation de la tête en direction de E2



4b

changement d'orientation de la tête et du regard de E2



4c

mimique faciale de E3 (sourcils levés + regard vers le bas + bouche retroussée vers le bas) et léger haussement d'épaules



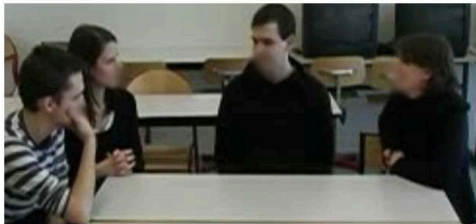
4d

changement d'orientation de la tête de E2 et E3 en direction de E4

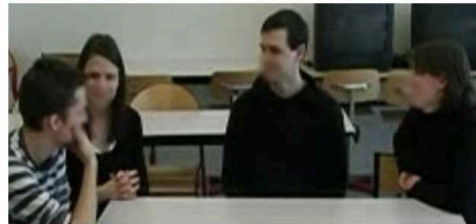
- 34 Dans cet exemple, E3 pose la question *What do you think about this?* en s'adressant au reste du groupe mais en se tournant plus particulièrement vers sa voisine de droite, E2 (4a), qui, pour exprimer son manque d'inspiration, détourne alors non seulement le regard mais aussi la tête tout en se frottant le bras (4b). L'énonciatrice en place s'étonne de cette réaction, comme en témoignent sa mimique faciale (sourcils levés, regard vers le bas et bouche retroussée vers le bas), le changement de direction de son regard et son léger haussement d'épaules (4c). La mimique faciale nous semble en effet tout à fait caractéristique d'une réaction de surprise si l'on se réfère au *Facial Action Coding System - FACS* (Ekman & Friesen, 1978), avec la combinaison des trois unités d'action (*action units*) que sont le lever de la partie intérieure des sourcils (*inner brow raiser*), le lever de la partie extérieure des sourcils (*outer brow raiser*) et un léger lever des paupières supérieures (*upper lid raiser*) rendant visible une plus grande partie des globes oculaires. L'unité d'action consistant à baisser la mâchoire (*jaw drop*) et conduisant à une ouverture relative de la bouche n'est pas systématique pour exprimer la surprise ; on observe simplement ici un abaissement du coin des lèvres (*lip corner depressor*), que nous avons nommé « bouche retroussée vers le bas ». Le malaise est de très courte durée puisque E4 comprend elle aussi que la co-énonciatrice E2 n'a pas du tout l'intention de s'exprimer et prend alors la parole (4d). La multimodalité de l'échange a alors permis de garantir une fluidité dans l'échange interpersonnel, à partir d'une double inférence : 1) inférence de la part de E2, fondée sur le regard appuyé de l'énonciatrice en place en sa direction, qui comprend qu'on lui donne la parole et exprime qu'elle ne la prendra pas par la position de son corps, par son regard et son frottement de bras ; 2) inférence de la part de E4, qui perçoit que E2 n'a pas l'intention de prendre la parole et décide alors de la prendre pour sortir de l'impasse communicationnelle décelable dans la réaction de E3 en 4c.

- 35 Dans l'exemple suivant, les étudiants doivent travailler à l'élaboration d'une brochure de vulgarisation scientifique destinée au grand public, expliquant ce qu'est la nourriture génétiquement modifiée et quels en sont les dangers. E4 demande à ses camarades s'ils estiment que davantage d'argent devrait être alloué à la recherche sur les organismes génétiquement modifiés (OGM). E3 répond qu'il faudrait plus de subventions non pas pour financer la recherche sur les dangers des OGM mais pour trouver de nouveaux OGM, moins nocifs. C'est alors au tour de E1 de s'exprimer.

## Exemple 5



5a  
E2. yes I [agree]  
E4. [me too]



5b  
(0.504) ((rires)) (2.499)

orientations corporelles et regards appuyés en direction du co-énonciateur E1 (au T-shirt rayé)

- 36 Suite à l'énoncé produit par l'énonciateur E3 et à sa ratification par les deux co-énonciatrices (5a), toutes les têtes et regards se tournent vers l'énonciateur E1 qui n'a encore rien dit (5b) : si le regard en direction du co-énonciateur associé à un silence est typique de la fin d'un tour de parole (Kendon, 1967), les trois co-énonciateurs invitent clairement E1 à réagir et à co-construire le discours. E1 regarde alors vers le bas et continue de toucher le bas de son visage dans un geste de soutien de la tête : on comprend alors à la fois l'anxiété typiquement exprimée par les gestes autocentrés (*self-touching*) (Mahl, 1956) et l'embarras traduit par le regard vers le bas (Exline *et al.*, 1970), tous deux vraisemblablement ressentis par l'énonciateur E1. En effet, pris au piège des regards convergents, celui-ci ne peut qu'inférer qu'il doit prendre la parole : difficile dans cette situation de détourner le regard et de faire mine de ne pas avoir compris. E1 essaie donc, tant bien que mal, de produire un énoncé : ce faisant, il recourt à son tour à une série de gestes pour montrer qu'il a du mal à s'exprimer, gestes exprimant un malaise parfaitement perçu par ses co-énonciateurs qui se mettent à rire pour désamorcer la situation. Dans cet exemple, les ressources multimodales utilisées par l'ensemble des énonciateurs en présence ayant pour fonction de donner la parole à E1 ont une fonction communicative si forte qu'il lui est impossible de se dérober.
- 37 À l'inverse, il arrive que l'énonciateur en place souhaite conserver la main et qu'il utilise alors un arsenal de ressources multimodales pour signifier à ses co-énonciateurs qu'il souhaite garder la parole. C'est le cas dans l'exemple 6, où les interlocuteurs évoquent le cas de papillons morts près de champs d'OGM ; ils s'interrogent sur le lien entre la sécrétion de protéines volontairement toxiques par ces plantes et la mort des papillons.



Exemple 6

1.	E3	for example if the same (0.3) erm example (0.44) a butterfly
2.		eats crops (0.42) and they die because the protein is erm toxic
3.	E1	yes
4.	E3	the erm a goal of this erm protein (0.28) I think # it's normal
.	im.	Oooo#6a
5.		but (0.26) it's erm a consequence (0.30) #
	im.	Ooooooooooooooooooooooooooooooooooooo #6b
6.	E1	erm erm measure is the first step so
7.	E3	# let's talk about the protein erm it's not erm specific
	im.	#6c



6a geste métaphorique de E3 pour normal



6b pause marquée par une position d'attente



6c geste métaphorique de mise en pause adressé à E1

38 Lorsqu'il a fini de produire son énoncé, l'énonciateur en place (E3), très expressif dans sa gestuelle lors des échanges (6a), marque une pause accompagnée d'une position du corps (mains croisées et bras au repos, reflet de la position d'attente adoptée par son co-énonciateur E1) qui semblent indiquer qu'il a terminé son tour (6b) et qu'il cède la parole à ses co-énonciateurs (et à E1 en particulier puisque son regard est tourné vers lui). En effet, le fait que l'énonciateur en place mette fin aux mouvements de mains produits en même temps que ses paroles est un fort indicateur de la fin de son tour de parole (Duncan, 1972 ; Duncan & Fiske, 1985). E1 débute alors un tour de parole *Erm erm measure is the first step, so...* mais E3 fait immédiatement un geste visant à mettre en attente son co-énonciateur (6c) pour poursuivre son propos (*let's talk about the protein, erm it's not erm specific and...*). La double inférence est parfaitement correcte ici : E1 a bien compris que E3 semblait avoir terminé son tour, comme tous les indices le laissaient penser, mais ce dernier souhaite reprendre la parole immédiatement, d'où le geste de mise en pause et l'ajout d'un argument ; le mouvement de doigt levé est correctement interprété par E1 qui lui rend la parole. Malgré sa tentative de formulation d'un début d'idée, E1 était resté dans une position d'attente (bras toujours au repos, mains croisées) sans véritablement s'installer dans son tour de parole. E3 s'est alors permis de ressaisir ce tour récemment cédé, se considérant peut-être encore légitime pour reprendre, voire conserver le tour de parole.



39 Un simple mouvement du regard (Goodwin & Goodwin, 1986) peut permettre de déduire que l'énonciateur en place appelle à un maintien ou à un changement de tour de parole, notamment suite à un silence intra-tour dont la longueur serait pesante, dépassant celle qu'on attendrait d'un silence dans un dialogue (dans un cadre ou une société donnée, avec ses habitudes conversationnelles). C'est le cas dans l'exemple 7, où E3 évoque le fait que seuls les pays riches peuvent recourir à la culture d'OGM alors même que les pays pauvres en auraient besoin. E2 rebondit sur ce propos en illustrant en quoi les OGM sont importants pour ces pays-là.

Exemple 7



1.	E2	gmo is really necessary for the future (0.5) because erm many
2.		people erm are hungry # (1.06) for example in a = africa (0.44) im.
	im	ooooooooooooooooooooi#7a
3.		erm #(4.66)
	im.	oooi#7b
4.	E1	they don't have any (0.38) = they don't have too much (0.86)
5.		erm (0.7) you want to say in africa we erm ((enfin)) = people
6.		don't have erm (0.82) enough erm money



7a

regard dans le vide (vers le bas) de E2



7b

changement d'orientation de la tête et du regard en direction de la co-énonciatrice de droite (E1)



7c

prise de parole de E1

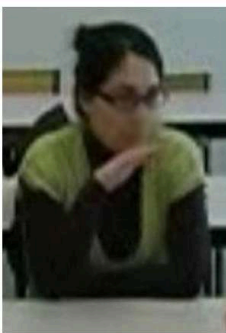
- 40 E2 débute son illustration par *For example, in Africa, erm...*, amorce qui est suivie de près de 5 secondes de "pause" intra-tour (Sacks et al., 1974 : 714-715) pendant lesquelles le regard de l'énonciateur est dans le vide (7a) puis celui-ci se tourne vers E1, scellant ainsi la demande de changement de tour de parole (Goodwin & Goodwin, 1986 : 71). La communication est ainsi faite de codes, d'éléments culturels faisant partie intégrante de celle-ci. Un silence de près de 5 secondes dans un échange interpersonnel peut rompre la fluidité conversationnelle. Lorsque la co-énonciatrice vient en aide à l'énonciateur en place, on comprend qu'elle était la mieux placée pour procéder à l'achèvement collaboratif un peu difficile à mettre en mots (7c) : *We don't have enough... we don't have much... He want to say in Africa **we (enfin) people** don't have enough money or...*
- 41 Ce que l'on peut d'ores et déjà déduire de ces exemples est le fait que, pour être correcte, il faut que l'inférence résulte de la combinaison de plusieurs éléments, dont une connaissance du monde, de conventions, d'habitudes de codage, etc. Si ces éléments divergent trop entre les différents énonciateurs en présence, cela peut constituer un obstacle au processus inférentiel et à l'échange conversationnel, ce qui est d'autant plus vrai dans un échange exolingue où le recours aux ressources PMG est plus marqué. On comprend ainsi qu'en cas de lacune lexicale, il sera beaucoup plus difficile d'inférer le sens du mot manquant sans contextualisation verbale suffisante, à plus forte raison si l'inférence est fondée sur des ressources multimodales compensatoires (c'est-à-dire se substituant au verbal).

### 3.3. Les inférences fondées sur des ressources multimodales se substituant au verbal

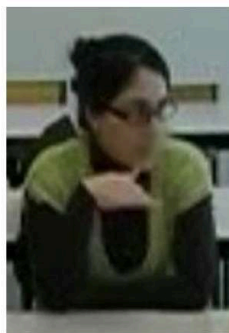
- 42 Rappelons tout d'abord les méthodes permettant des actes de communication formalisées par Clark et Gerrig (1990 : 765) à partir d'échanges en L1 : *indicating*, *describing* et *demonstrating*. Ces derniers illustrent chaque méthode à partir d'un contexte où Alice voudrait que Ben comprenne comment servait McEnroe :
- Alice can INDICATE, or point at, an actual serve by McEnroe (en le montrant)
  - Alice can DESCRIBE the serve (par le truchement du lexical)
  - Alice can DEMONSTRATE the serve (en l'imitant, geste iconique)

- 43 Et les auteurs d'ajouter : *Demonstrations are also special in the way they represent their referents : they depict rather than describe* (Clark & Gerrig, 1990 : 767).
- 44 En cela, les démonstrations sont d'autant plus utiles en L2 qu'elles ne servent pas uniquement à rendre le récit plus vivant ; elles peuvent être un véritable recours lorsque le verbal fait défaut, permettant d'instancier directement la place du mot non connu sans avoir à s'expliquer par de longues phrases du type « comment on dit déjà, quand etc. ». Ainsi, le fait de ne pas avoir d'accès au sens d'un mot isolé n'entrave pas le processus pragmatique d'inférence, si tant est que l'énonciateur fournisse suffisamment d'éléments contextuels pour que les gestes compensatoires puissent être interprétés (cf. Feyereisen et al., 1988). De la même façon que des stimuli paraverbaux (p. ex. l'intonation) viennent s'ajouter à la sémantique (c'est-à-dire à la composante purement linguistique, conventionnelle, de la signification d'une phrase), il apparaît dans nos échanges que les gestes ne se suffisent pas à eux-mêmes et viennent en appui au verbal. Les exemples 8 à 11 illustrent cela à travers des cas de démonstration et d'indication.
- 45 Dans le premier extrait, l'énonciatrice E3 revient sur la diminution drastique de papillons suite à la substitution des pesticides par des protéines toxiques pour les insectes.

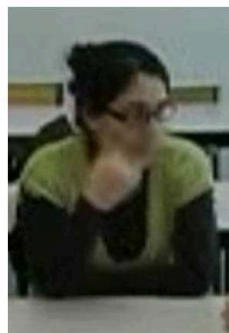
Exemple 8



8a  
E3. *erm the number of butterflies erm*  
geste métaphorique pour *decreased*



8b  
(0.80)



8c  
E4. *decreased*



8d  
E3. *decreased*  
fin du geste métaphorique

- 46 E3 débute son énoncé (*erm the number of butterflies erm*) et utilise un geste métaphorique coverbal pour exprimer la diminution (8a à 8c). L'un de ses co-énonciateurs complète

son énoncé avec le mot *decreased* que l'énonciatrice reprend tout en terminant son geste métaphorique (8d).

- 47 Un geste métaphorique peut parfois aider l'énonciateur lui-même à retrouver le terme approprié, ce qui semble être le cas dans l'exemple 9, où l'énonciatrice E3 évoque l'impact positif des OGM sur la quantité et sur la qualité des récoltes et l'importance de poursuivre des recherches dans ce domaine.

#### Exemple 9



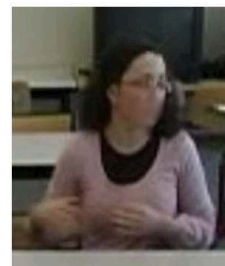
9a  
we should erm  
battement  
(Butterworth) ou geste  
métaphorique ?



9b  
(1.166)



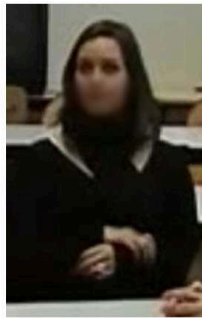
9c  
erm continue



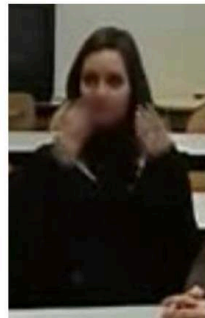
9d  
fin du geste

- 48 À la différence du geste métaphorique de l'exemple précédent, les mouvements de la main de l'énonciatrice semblent ici avoir pour vocation d'aider à l'accouchement du terme recherché. Si l'on peut croire de prime abord qu'elle mime *continue*, il se trouve que le même geste lui sert aussi de support à la recherche de mots tels que *knowledge* ou *organism* dans la suite de l'interaction. Ce battement pourrait être perçu comme étant métaphorique d'une forme de maïeutique propre à cette énonciatrice (les mouvements circulaires de ses mains mimant la progression dans la recherche du terme cible jusqu'à sa verbalisation). Un même geste peut ainsi très souvent avoir différentes dimensions et, par ricochet, différentes interprétations (McNeill, 1992) : un geste peut par exemple être à la fois métaphorique et iconique, et un geste métaphorique ou iconique peut lui-même être doublé de battements (McNeill, 1992 : 99 et 102).
- 49 L'*indication* peut également aider à produire le terme recherché, ou un autre suffisant à transmettre le sens visé. Dans l'exemple 10, les énonciatrices cherchent comment promouvoir les études en biologie.

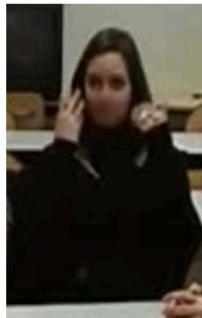
## Exemple 10



10a  
it may be very interesting for



10b  
(1.554)  
début du geste déictique



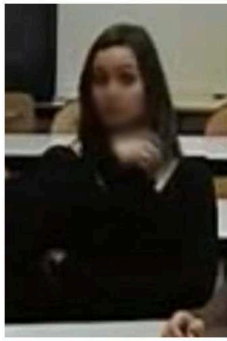
10c



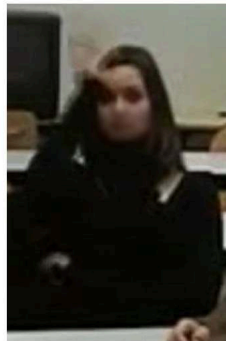
10d  
= for us  
fin du geste déictique

- 50 À la suite de l'énoncé laissé en suspens par E3, *you can work in different places, and we can meet other scientist who are in the world and it may be very interesting for...*, l'énonciatrice se désigne des deux mains (10b à 10d), réfléchissant à un groupe nominal élaboré. La suite de l'échange révèle qu'elle se signale elle-même comme chercheuse en biologie, et cet acte de monstration lui permet, après réflexion, de terminer son geste en l'accompagnant d'un simple ... *for us* (10d). Par l'intermédiaire de ce geste déictique, l'énonciatrice se rend compte qu'elle fait partie des chercheurs en biologie auxquels elle cherche à référer et qu'un simple *us* suffit; s'ensuit une mise en mot rendue possible par le geste qui agit comme une béquille pour le verbal. Si le geste est ici co-verbal, la suite de l'intervention de cette énonciatrice témoigne du fait que les ressources PMG peuvent parfois se suffire à elles-mêmes (exemple 11).

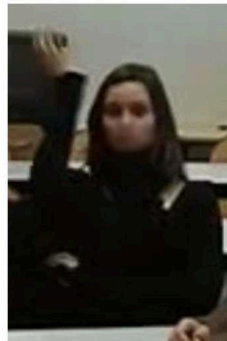
## Exemple 11



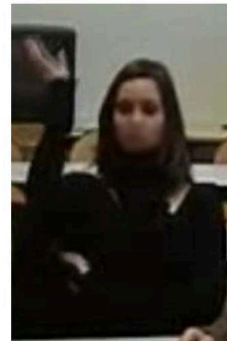
11a  
because they think it's very



11b  
erm



11c  
pouff



11d

début du geste métaphorique

fin du geste

- 51 E3 débute son énoncé par *Always people make a mistake speaking biology, because they think it's very...* puis marque une pause avant de recourir à un geste concomitant à la production du non-mot « pouff », même si le stimulus gestuel suffirait en contexte à instancier la place de l'adjectif. « Pouff » ne se substitue à aucun adjectif en particulier. La même interjection pourrait désigner quantité d'autres référents (une mauvaise odeur, une consternation, etc.). Le terme à lui seul n'aurait certainement pas suffi à transmettre le sens visé dans une phrase intonativement non marquée (*people think it's very pouff*). Le geste quant à lui pourrait signifier « qui part dans tous les sens » pour un texte, « rêveur/dispersé » pour quelqu'un, « haut en couleurs » pour une description par exemple, ou encore ici « complexe, impénétrable ». On comprend en effet grâce à la distribution que le geste occupe la place d'un adjectif (il suit directement la copule *be* et l'adverbe *very*). C'est donc la combinaison du geste en contexte, de l'intonation, et de la mimique faciale de l'énonciatrice qui permet d'en déduire le sens. Les co-énonciatrices rient de l'audace de ce recours à autre chose qu'un adjectif qualificatif et semblent parfaitement comprendre ce qui est signifié.
- 52 Le verbal ne constitue donc pas la seule issue en cas d'impasse communicationnelle (Sarré & Sharifzadeh, 2020). L'intonation et le contexte seuls peuvent même suffire à comprendre le sens visé en passant outre les erreurs lexicales. C'est le cas de l'énoncé suivant l'exemple 11, dans lequel l'énonciatrice marque la fin de l'échange par la phrase *It's finished* puis s'assure que tous les co-énonciateurs en sont d'accord en poursuivant par *Are you ok?*. Ce qu'elle pense exprimer par ce calque est « Etes-vous d'accord ? ». Grâce au contexte, au paraverbal et aux ressources PMG (intonation et orientation corporelle vers les autres), les co-énonciatrices infèrent correctement que leur camarade leur demande si elles sont d'accord. Les ressources paraverbales et posturo-mimo-gestuelles ont ainsi pu permettre la compréhension en contexte d'un énoncé parfaitement erroné.

## 4. Conclusion

- 53 Les études sur la gestuelle des apprenants dans des interactions apprenants-apprenants en présentiel sont rares. Le présent travail visait à étudier le rôle des ressources



posturo-mimo-gestuelles dans des échanges en L2 entre apprenants. Nous avons pu mettre en évidence leur apport essentiel dans les processus inférentiels et sommes parvenus à trois catégories (ou modes de fonctionnement) liant ressources PMG et inférences : les apprenants de l'anglais L2 tendent à faire des inférences fondées sur (1) des ressources posturo-mimo-gestuelles au service de la tentative de mise en mots, (2) des ressources PMG au service de la gestion des tours de parole et (3) des ressources PMG se substituant au verbal. Les inférences fondées sur les ressources PMG sont donc à la fois liées à l'intention de communication de l'énonciateur et à la gestion des tours de parole.

- 54 Cette étude qualitative exploratoire s'appuie sur un corpus d'interactions en L2 durant lesquelles les interactants sont assis autour d'une table, ce qui restreint vraisemblablement le type et l'amplitude des gestes qu'ils peuvent produire. Malgré ces limites, et comme nous en avons fait l'hypothèse, notre étude a pu révéler que l'énonciateur-apprenant d'une L2 mobilise des ressources PMG en fonction du contexte et des co-énonciateurs en présence. Les ressources multimodales ont souvent une double fonction dans les échanges apprenants-apprenants, à la fois cognitive et communicative, et peuvent être étudiées comme un système en constante évolution. Si, pour certains auteurs, les gestes appartiennent aux systèmes de l'interlangue de l'apprenant (Stam, 2007), ils constituent pour nous le pendant gestuel de l'interlangue (Gullberg, 2006), qu'on pourrait appeler l'*intergestuel* de l'apprenant d'une langue et qui correspondrait à un stade intermédiaire dans l'acquisition d'une gestuelle-cible correspondant aux codes culturels d'une langue et d'une aire géographique données, système dynamique qui évoluerait en même temps que l'apprenant gagnerait en compétence. Les auteurs du CECRL (2001 : 72) posent justement la question du *degré d'aptitude* auquel les apprenants devraient être *capables de joindre le geste à la parole*, ce qui laisse supposer que la compétence gestuelle augmenterait en même temps que les autres composantes de la compétence langagière (phénomène également reflété dans les échelles de compétence du CECRL). Si les gestes sont également mentionnés dans le volume complémentaire du CERCL (2018), ils ne semblent être intégrés qu'aux descripteurs des niveaux de compétence les plus bas (pré-A1, A1 et A2 essentiellement) car ils sont uniquement considérés comme des ressources compensatoires (*l'apprenant peut avoir besoin de pallier son répertoire limité par des gestes*, p. 114). Cela constitue incontestablement une approche réductrice de l'usage des ressources PMG, comme en témoignent les interactions de notre corpus, mettant en scène des apprenants de niveau B1, voire B2 pour certains. Le caractère évolutif de l'*intergestuel* de l'apprenant a ainsi plusieurs implications en termes d'enseignement-apprentissage d'une L2, la première étant la reconnaissance du fait que la compétence interactionnelle va au-delà de la simple compétence linguistique (Neu, 1990 ; Kellerman, 1992 ; Jungheim, 1995) et qu'elle ne saurait donc être réduite aux seules composantes linguistique, sociolinguistique et pragmatique. Nous rejoignons Berman & Slobin (1994) en considérant que la capacité d'un apprenant à faire la différence entre ce qui peut être inféré par ses co-énonciateurs (ce qu'expriment des non-dits, des silences, des gestes, etc.) et ce qui doit être explicitement exprimé verbalement s'affine en même temps que ses compétences langagières, ce qui tend à montrer que les différents types de ressources mobilisables en situation de communication sont partie intégrante de la compétence interactionnelle. Dans la mesure où l'apprenant d'une L2 doit être à la fois spectateur et producteur de gestes (Stam, 2013), sa capacité à inférer et à produire des indices « inférables » par ses co-énonciateurs ferait partie de sa compétence



interactionnelle, même si, comme le signale Gullberg (1998 : 35), les tentatives d'intégration des gestes aux cadres théoriques de la compétence communicative et interactionnelle restent embryonnaires à ce jour.

- 55 Une sensibilisation aux mécanismes des processus inférentiels pourrait être intégrée à l'enseignement des langues par des activités appropriées (voir par exemple les propositions de Piccoli, 2017), dans la lignée de certains auteurs qui ont prôné avant nous une sensibilisation à la kinésie, aux emblèmes ou encore à la proxémie en classe de langue (Stam, 2007). Une formation à la production et à l'interprétation de ressources PMG comme indices inférables dans une aire géographique et linguistique donnée nous semble en effet présenter une réelle valeur ajoutée. Ceci paraîtrait d'autant plus pertinent que nous avons montré l'importance des ressources multimodales dans les échanges interpersonnels : gestes, regards, silences et postures font partie intégrante de la communication et constituent des points d'appui indispensables au travail d'inférence. Nos exemples ont pu mettre en lumière le fait que le verbal est parfois secondaire dans l'interprétation de l'intention de communication des co-énonciateurs ; se pose alors la question des limites d'une hiérarchie entre les différentes ressources. L'existence de ressources verbales d'un côté et, de l'autre, de ressources co-verbales que seraient les ressources PMG, simples accessoires paralinguistiques au service du verbal, semble pour le moins artificielle. Les différents types de ressources résultent d'un même processus cognitif et font partie d'un système intégré unique. À la suite de Kress (2019) et de sa sémiotique sociale, nous considérons que c'est de leur combinaison que naît le sens. Les ressources PMG, les gestes en particulier, étant une source d'input supplémentaire favorisant la compréhension et l'apprentissage d'une langue, s'intéresser aux ressources multimodales mobilisées lors d'échanges en L2 reste une voie prometteuse pour la didactique et l'acquisition des L2.

---

## BIBLIOGRAPHY

- Aden, J. (2017) Langues et langage dans un paradigme enactif, *Recherches en didactique des langues et des cultures*, 14,1. Disponible sur : <http://rdlc.revues.org/1085>.
- Alibali, M. W., Kita, S. & A. Young (2000) Gesture and the process of speech production: We think, therefore we gesture, *Language and Cognitive Processes*, 15, p. 593-613.
- Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C., & P. Paggio (2007) The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing Phenomena, *Language Resources and Evaluation*, 41, 3-4, p. 273-287.
- Arndt, H., & R. W. Janney (1987) *InterGrammar: Toward an integrative model of verbal, prosodic and kinesic choices in speech*, Berlin/New York/Amsterdam: Mouton de Gruyter.
- Azaoui, B. (2014) Analyse multimodale de l'agir professoral et degré de granularité de traitement. Réflexions méthodologiques, *Lidil*, 49, p. 17-32.

- Azaoui, B. (2015) Fonctions pédagogiques et implications énonciatives de ressources professorales multimodales. Le cas de la bimanualité et de l'ubiquité coénonciative, *Recherches en didactique des langues et des cultures - Les cahiers de l'Acadle*, 12, 2, p. 225-253.
- Bach, K. (1997) The Semantics-Pragmatics Distinction: What It Is and Why It Matters, *Linguistische Berichte*, 8, p. 33-50.
- Berman, R. A. & D. I. Slobin (1994) *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cadet, L., & M. Tellier (2014) Avant Propos, in Tellier, M. & L. Cadet (eds) *Le corps et la voix de l'enseignant : théorie et pratique*, Paris: Editions Maison des Langues.
- Carston, R. (2000) Explicature and semantics, *UCL Working Papers in Linguistics*, 12, 1, p. 44-89.
- Chernyshova, E., & V. Traverso (2017) Inférences et processus d'intercompréhension dans les interactions quotidiennes : quelques questions méthodologiques, *Cahiers de praxématique*, 68. Disponible sur : <https://journals.openedition.org/praxematique/4536>.
- Clark, H. H., & D. I. Gerrig (1990) Quotations as demonstrations, *Language*, p. 764-805.
- Colletta, J. M. (2004) *Le développement de la parole chez l'enfant âgé de 6 à 11 ans : corps, langage et cognition*, Liège: Editions Mardaga.
- Cosnier, J. (1982) Communications et langages gestuels, in Cosnier, J., Coulon, J., Berrendonner, A. & C. Kerbrat-Orecchioni (eds.) *Les voies du langage, communications verbales, gestuelles et animales*, Paris: Dunod, p. 255-304.
- Cosnier, J. (1985) Méthodes d'étude de la posturo-mimo-gestualité, *Psychol. Méd.*, 17, 7, p. 955-960.
- Cosnier, J. (1994). *La psychologie des émotions et des sentiments*, Paris: Retz.
- David, C. (2017) Gestion de l'implicite dans l'interaction orale en L2, *Cahiers de praxématique*, 68. Disponible sur : <http://journals.openedition.org/praxematique/4534>.
- Duncan, S. (1972) Some Signals and Rules for Taking Speaking Turns in Conversations, *Journal of Personality and Social Psychology*, 23, 2, p. 283-292.
- Ekman, P. (1979) About brows: Emotional and conversational signals, in M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (eds.) *Human Ethology: Claims and Limits of a New Discipline*, Cambridge: Cambridge University Press, p. 169-202.
- Ekman, P. & W. V. Friesen (1978) *Facial Action Coding System*, Palo Alto, CA: Consulting Psychologists Press.
- Exline, R. V., Thibaut, J., Hickey, C. B. & P. Gumpert (1970) Visual interaction in relation to machiavellianism and an unethical act, in Christie R. & Geis F. L (eds) *Studies in Machiavellianism*, New York: Academic Press, p. 53-76.
- Fehr, B. J. & R. V. Exline (1987) Social visual interaction: A conceptual and literature review, in A. W. Siegman & S. Feldstein (eds) *Nonverbal behavior and communication*, Hillsdale, NJ: Erlbaum, p. 225-326.
- Feyereisen, P., Van de Wiele, M., & F. Dubois (1988) The meaning of gestures: What can be understood without speech, *European Bulletin of Cognitive Psychology*, 8, p. 3-25.
- Foerster, C. (1990) Et le non-verbal ?, in Dabène, L., Cicurel, F., Lauga-Hamid, M.-C. & C. Foerster (eds) *Variations et rituels en classe de langue*, Paris: Hatier, p. 72-93.

- Ford, C. & S. Thompson (1996) *Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Management of Turns*, in Ochs, E., Schegloff, E., & S. Thompson (eds) *Interaction and Grammar*, Cambridge: Cambridge University Press, p. 134-184.
- Frick-Horbury, D. (2002) The use of hand gestures as self-generated cues for recall of verbally associated targets, *American Journal of Psychology*, 115, p. 1-20.
- Goldin-Meadow, S. (2003) *Hearing gesture: How our hands help us think*, Cambridge, USA: Harvard University Press.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S.D. & S. Wagner (2001) Explaining math: Gesturing lightens the load, *Psychological Science*, 12, p. 516-522.
- Goodwin, C. (1981) *Conversational organization: Interaction between speakers and hearers*, New York: Academic Press.
- Goodwin, M. H., & C. Goodwin (1986) Gestures and coparticipation in the activity of searching for a word, *Semiotica*, 62, p. 51-75.
- Gullberg, M. (1998) *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*, Lund: Lund University Press. Disponible sur : <https://lup.lub.lu.se/search/ws/files/4825091/3912717.pdf>.
- Gullberg, M. (2006) Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon), *IRAL*, 44, p. 103-124.
- Gullberg, M. (2010) Methodological reflections on gesture analysis in second language acquisition and bilingualism research, *Second language research*, 26, 1, p. 75-102.
- Gumperz, J. J. (1989) *Sociolinguistique interactionnelle, une approche interprétative*, Paris: L'Harmattan.
- Hadar, U., Steiner, T. & F. C. Rose (1984) The timing of shifts of head postures during conversation, *Human Movement Science*, 3, 3, p. 237-245.
- Hadar, U., Rivka, D., & T. Amit (2001) Gesture during speech in first and second language: Implications for lexical retrieval, *Gesture*, 1, p. 151-165.
- Hostetter, A. B. & M. W. Alibali (2004) On the tip of the mind: Gesture as a key to conceptualization, in Forbus, K. D., Gentner, D. & T. Regier (eds) *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, Chicago: Cognitive Science Society, p. 589-594. Disponible sur : <https://cognitivesciencesociety.org/wp-content/uploads/2019/01/CogSci04.pdf>.
- Jungheim, N. O. (1995) Assessing the unsaid: The development of tests of nonverbal ability, in Brown, J. D. & S. O. Yamashita (eds) *Language testing in Japan*, Tokyo: The Japan Association for Language Teaching, p. 149-165.
- DIALANG (2003) Project for the development of diagnostic language test. Disponible sur : <https://dialangweb.lancaster.ac.uk>.
- Kellerman, S. (1992) 'I see what you mean': The role of kinesic behaviour in listening and implications for foreign and second language learning, *Applied Linguistics*, 13, p. 239-257.
- Kendon, A. (1967) Some functions of gaze-direction in social interaction, *Acta Psychologica*, 26, p. 22-63.
- Kendon, A. (2004) *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press.
- Kerbrat-Oreccioni, C. (1992) *Les interactions verbales*, tome II, Paris: Armand Colin.

- Kerbrat-Oreccioni, C. (1994) *Les interactions verbales*, tome III, Paris: Armand Colin.
- Kinsbourne, M. (1972) Eye and head turning indicates cerebral lateralization, *Science*, 176, 4034, p. 539-541.
- Kress, G. (2019) L'apprentissage en tant que travail sémiotique : vers une pédagogie de la reconnaissance, in V. Rivière & N. Blanc (eds) *Observer la multimodalité en situations éducatives. Circulations entre recherche et formation*, Lyon: ENS Editions, p. 23-48.
- Kurhila, S. (2006) *Second Language Interaction*, Amsterdam: John Benjamins.
- Mahl, G. F. (1956) Disturbances and silences in the patient's speech in psychotherapy, *Journal of Abnormal and Social Psychology*, 53, 1, p. 1-15.
- McNeill, D. (1992) *Hand and mind. What the hands reveal about thought*, Chicago: Chicago University Press.
- McNeill, D. (2005) *Gesture and thought*, Chicago: Chicago University Press.
- Mondada, L. (2008a) Projet CORINTE, conventions de transcription multimodale. Disponible sur : [http://icar.univ-lyon2.fr/projets/corinte/documents/convention\\_transcription\\_multimodale.pdf](http://icar.univ-lyon2.fr/projets/corinte/documents/convention_transcription_multimodale.pdf).
- Mondada, L. (2008b) Production du savoir et interactions multimodales, une étude de la modélisation spatiale comme activité pratique située et incarnée, *Revue d'anthropologie des connaissances* 2008/2, 2. Disponible sur <https://www.cairn.info/revue-anthropologie-des-connaissances-2008-2-page-219.htm>
- Navaretta, C., & P. Paggio (2013) Classifying multimodal turn management in Danish dyadic first encounters, in *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013)*, NEALT Proceedings Series 16 (No. 085), Linköping: Linköping University Electronic Press, p. 133-146.
- Neu, J. (1990) Assessing the role of nonverbal communication in the acquisition of communicative competence in L2, in Scarcella, R.C., Andersen, E. S. & S. D. Krashen (eds) *Developing Communicative Competence in a Second Language*, New York: Newbury House, p. 121-138.
- Nobe, S. (2001) On gestures of foreign language speakers, in Cavé, C., Guaitella, I. & S. Santi (eds) *Oralité et Gestualité (Orange'01)*, Paris: L'Harmattan, p. 572-575.
- Paggio, P., Allwood, J., Ahlsén, E., & K. Jokinen (2010) The NOMCO multimodal Nordic resource: goals and characteristics, in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valetta, Malta, May 19-21. European Language Resources Association (ELRA), p. 2968-2973.
- Piccoli, V. (2017) A la recherche des bons indices : inférences et recherches de mot entre locuteurs de langues romanes, *Cahiers de praxématique*, 68. Disponible sur : <https://journals.openedition.org/praxematique/4587>.
- Ravizza, S. (2003) Movement and lexical access: Do non-iconic gestures aid in retrieval ?, *Psychonomic Bulletin and Review*, 103, p. 610-15.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974) A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, p. 696-735.
- Sarré, C. & S. Sharifzadeh (2020) La correction dans l'interaction entre apprenants d'une langue seconde, in S. Raineri, M. Sekali & A. Leroux (eds) *La Correction en langue(s) - Linguistic Correction/Correctness*, Nanterre: Presses Universitaires de Paris-Nanterre, p 65-90.

- Schmidt, T., Wörner, K., Lehmborg, T. & H. Hedeland (2013) EXMARALDA (Partitur Editor version 1.5.2) [logiciel], Hamburg: University of Hamburg. Disponible sur : <http://www.exmaralda.org>.
- Searle, J. R. (1979) *Expression and Meaning*, Cambridge, Massachusetts: Cambridge University Press.
- Sherman, J. & E. Nicoladis (2004) Gestures by advanced Spanish-English second-language learners, *Gesture*, 4, 2, p. 143-156.
- Sime, D. (2001) The use and perception of illustrators in the foreign language classroom, in Cavé, C., Guaitella, I. et S. Santi (eds) *Oralité et gestualité. Interactions et comportements multimodaux dans la communication*, Paris: L'Harmattan, p. 582-585.
- Sime, D. (2008) 'Because of her gesture, it's easy to understand', Learners' perception of teachers' gestures in the foreign language class, in McCafferty, S. G. & G. Stam (eds) *Gesture: second language acquisition and classroom research*, New-York: Routledge, p. 259-279.
- Sperber, D., & D. Wilson (1986) *Relevance: Communication and cognition*, Oxford: Blackwell.
- Stam, G. (2007) Second Language Acquisition from a McNeillian Perspective, in S. D. Duncan, J. Cassell & E. T. Levy (ed.) *Gesture and the Dynamic Dimension of Language: Essays in Honor of David McNeill*, *Gesture Studies*, 1, Amsterdam: John Benjamins, p. 117-124.
- Stam, G. (2012) Gestes et recherche lexicale en langue seconde, in Vion, R., Giacomi, A., & C. Vargas (eds) *La corporalité du langage : Multimodalité, discours et écriture*, Aix-en-Provence: Presses Universitaires de Provence, p. 55-71.
- Stam, G. (2013) Second language acquisition and gesture, in C. A. Chapelle (ed.) *The Encyclopedia of Applied Linguistics*, Oxford: Blackwell, p. 225-251.
- Strömquist, S. (1987) Gaze aversion, code-switching, and search activities in route descriptions by six adult language learners, in Wande, E. Anward, J., Nordberg, Steensland, L. & M. Thelander (eds) *Aspects of multilingualism. Proceedings from the Fourth Nordic Symposium on Bilingualism*, Uppsala: University of Uppsala, p. 351-372.
- Tellier, M. (2008) Dire avec des gestes, *Le Français dans le monde, Recherches et applications*, 44, p. 40-50.
- Tellier, M. (2010) Faire un geste pour l'apprentissage : Le geste pédagogique dans l'enseignement précoce : impact sur le développement de la langue maternelle ?, in Corblin, C. & J. Sauvage (eds.) *L'enseignement des langues vivantes étrangères à l'école*, Paris: L'Harmattan, p. 31-54.
- Tellier, M. (2014) Quelques orientations méthodologiques pour étudier la gestuelle dans des corpus spontanés et semi-contrôlés, *Discours*, 15. Disponible sur : <https://journals.openedition.org/discours/8917>.
- Tellier, M. (2016) Prendre son cours à bras le corps, *Recherches en didactique des langues et des cultures* 13, 1. Disponible sur : <https://journals.openedition.org/rdlc/474>.
- Tellier, M. & L. Cadet (2014) *Le corps et la voix de l'enseignant : théorie et pratique*, Paris: Maison des langues.
- Yoshioka, K. (2005) *Linguistic and Gestural Introduction and Tracking of Referents in L1 and L2 Discourse*, thèse MS, Université de Groningen. Disponible sur : [https://www.rug.nl/research/portal/publications/linguistic-and-gestural-introduction-and-tracking-of-referents-in-l1-and-l2-discourse\(ed3d66a8-1714-48c5-a48f-1bffce7a38f5\).html](https://www.rug.nl/research/portal/publications/linguistic-and-gestural-introduction-and-tracking-of-referents-in-l1-and-l2-discourse(ed3d66a8-1714-48c5-a48f-1bffce7a38f5).html).

## APPENDIXES

### Conventions de transcription

#### Transcription orthographique

- erm : indique une hésitation
- ? : indique une intonation montante, dans le cadre d'une question, par exemple
- texte = texte : signale une opération de reprise (répétition, redémarrage, reformulation, autocorrection) ; indique le terme tel qu'il est produit suivi du signe = et de l'énoncé reformulé
- >texte< : indique un mot ou groupe de mot prononcé(s) avec un débit plus rapide
- <texte> : indique un mot ou groupe de mot prononcé(s) avec un débit plus lent
- TEXTE : indique un mot ou groupe de mots prononcé(s) avec une intensité supérieure du reste de l'énoncé
- [texte] : indique un chevauchement
- ((texte)) : indique un élément entendu par le transcripteur autre que les paroles (sonnerie, rires) ou des paroles dans une langue autre que l'anglais (le français, par exemple)
- (chiffre) : indique la durée d'une pause inter ou intra-tour

#### Annotation des ressources PMG

Nous avons annoté les gestes, expressions faciales et mouvements corporels autres que les gestes. Notons que la granularité choisie pour nos annotations ne comprend pas la trajectoire temporelle des gestes (segmentation en plusieurs phases : *preparation, stroke, hold, retraction*), ni leur intensité, mais simplement leur nature de façon essentiellement descriptive à l'échelle d'une unité gestuelle (*gesture unit*, Kendon, 2004).

- les gestes (d'après McNeill, 2005) : gestes déictiques, iconiques, métaphoriques, battements et Butterworth
- les expressions faciales (d'après MUMIN d'Allwood *et al.*, 2007)
  - visage (sourire, rire, moue)
  - sourcils (froncés, levés)
  - bouche (retroussée vers le haut, retroussée vers le bas, en avant, rentrée)
  - regard (vers le haut, vers le bas, sur le côté, vers le co-énonciateur)
- les mouvements corporels (d'après NUMIN d'Allwood *et al.*, 2007 et NOMCO de Paggio *et al.*, 2010)
  - mouvements de la tête (acquiescement, vers l'avant, vers l'arrière, penchée sur le côté, tournée vers le côté, secouée)
  - mouvements du buste (en avant, en arrière, sur le côté, vers le co-énonciateur, en retrait du co-énonciateur).

#### Captures d'écran (d'après CORINTE de Mondada, 2008a)

Lorsqu'un échange est transcrit (plusieurs interlocuteurs), des captures d'écran sont proposées et sont assorties de commentaires concernant les ressources PMG mobilisées.

Dans la transcription orthographique, des points de repère (#) sont insérés car ils indiquent le moment exact auquel renvoie une image (capture d'écran) de façon synchronisée avec la parole. Une ligne de transcription codée « im. » est dévolue à l'identification de l'image correspondant à chaque point de repère : l'étiquette de

l'image (#+chiffre+lettre) est alignée avec le repère # correspondant dans la transcription orthographique.

## NOTES

1. « Étudier les productions orales des apprenants en tenant compte de leurs gestes pourrait permettre de comprendre les processus qui sous-tendent l'acquisition d'une langue seconde, les systèmes de l'interlangue des apprenants et les processus cognitifs convoqués par l'activité de parole. » (Nous traduisons chaque citation en anglais).
  2. « Une brève reconstruction des étapes [est] nécessaire pour dériver l'acte illocutoire primaire de l'acte illocutoire littéral [...]. À moins qu'un auditeur adopte une stratégie inférentielle pour savoir si les contenus illocutoires primaires diffèrent des contenus illocutoires littéraux, il n'a aucun moyen de comprendre les actes illocutoires indirects. »
  3. « Il existe **deux types d'information contextuelle**, la portée du premier étant beaucoup plus restreinte que celle du second. L'information dont le rôle se limite à se combiner avec l'information linguistique pour en déterminer (dans le sens de repérer) le contenu se résume en une courte liste de variables, **telles que l'identité du locuteur et de l'auditeur ou l'heure et le lieu d'un énoncé**. L'information contextuelle au sens large est **tout ce que l'auditeur doit prendre en compte pour déterminer (dans le sens d'établir) l'intention de communication du locuteur**. ». C'est nous qui soulignons.
  4. « La plupart des énoncés verbaux combinent les dimensions linguistique, paralinguistique, faciale et vocale, qui semblent fonctionner ensemble comme un signal recevant une interprétation unifiée. »
  5. Au sens d'« énonciation » dans notre terminologie.
  6. « Dans la plupart des études, le terme renvoie aux gestes se substituant aux mots lorsque le canal verbo-vocal fait défaut. On peut alors s'appuyer sur le geste seul pour transmettre le message ou remplacer un item unique au sein d'un énoncé, processus appelé « syntaxe mixte » par Slama-Cazacu (1976).  
Le terme 'compensation' est également utilisé dans un sens plus général. On dit que certains gestes compensatoires complètent le message verbo-vocal de manière à l'augmenter, l'étoffer ou l'améliorer. Ces gestes complémentaires se produisent simultanément à la parole. »
  7. Dans les extraits requérant davantage de contextualisation, les tours de parole sont spécifiés entre la vue d'ensemble et les gros plans. Dans tous les autres cas, toute verbalisation apparaît directement sous chaque vignette.
- 

## ABSTRACTS

While published research on the role of the body in multimodal didactic discourse is growing at a fast pace, as shown by the numerous recent studies on the subject (Aden, 2017 ; Azaoui, 2015 ; Sime, 2001, 2008 ; Tellier, 2014, 2016), research on the multimodal nature of learner-learner interactions in a second language (L2) remains more confidential than that on interactions between teacher and learner(s) or between native speaker(s) and learner(s) (Kurhila, 2006). In addition, few studies have been published to date on the use of non-verbal resources in learner-



learner exchanges (Gullberg, 2010). As a result, our knowledge of the types of interactional resources involved in these exchanges and of the way they are used is still limited. This is particularly true of interactional resources other than the words uttered by interactants. Interacting results from a complex cognitive process (McNeill, 2005 ; Kendon, 2004) which combines verbal resources as well as vocal and visual ones (postures, facial expressions and gestures, cf. “ressources posturo-mimo-gestuelles” for Cosnier, 1985). We investigate the impact of the visual resources in face-to-face L2 learner exchanges so as to better understand their role in learner-learner interactions, as well as the way they are interpreted by the co-speakers through the inferential processes in play. It is only thanks to these pragmatic inferential processes, which are instrumental in the interpretation of contextual elements (cf. Gumperz 1989), be they intra-textual or extra-textual, that co-speakers can truly access the speaker’s intended meaning.

Gestures and more generally visual resources trigger inferential processes that make it possible to reconstruct the meaning of utterances, even more so when the latter are difficult to put into words. In the case of ambiguous, erroneous or incomplete utterances, coverbal gestures, i.e., gestures produced at the same time as speech acts (McNeill, 1992, 2005 ; Goldin-Meadow, 2003), can generate inferential work for the co-speakers and facilitate access to the speaker’s intended meaning (Tellier, 2014). The multimodal nature of interpersonal communication is claimed to be all the more perceptible in exchanges between L2 learners because of the deficits in their L2 system (David, 2017) : a number of studies (Gullberg, 1998 ; Hadar *et al.*, 2001 ; Nobe, 2001 ; Yoshioka, 2005) show that the use of gestures is more prevalent during exchanges in L2 than in L1, in particular the use of deictic gestures (Sherman & Nicoladis, 2004).

The exploratory empirical study presented in this paper aims to shed light on the role of visual resources for a given speaker in interactions between learners of L2 English, but also on the way these resources are interpreted by the co-speakers, who engage in inferential processes to gain access to the intended meaning. Our work is based on a corpus of face-to-face oral interactions between learners of L2 English in tertiary education. Our analysis highlights the essential role of inferential processes in the interpretation of visual resources in three different patterns : learners of L2 English tend to make inferences based on (1) visual resources used during word searches, (2) visual resources used in turn management, and (3) visual resources used as substitutes for verbal elements. Our results show that inferences based on visual resources are related to both the speaker’s communicative intent and the management of speech turns. Our study also sheds light on the multiple functions that a single visual resource can take on in a given context and leads us to advocate language classes with specific training in interpreting and resorting to visual resources as “inferrable” clues which are specific to a given geographical and linguistic area.

Les interactions entre apprenants d’une langue seconde (L2) restant encore insuffisamment étudiées, les connaissances sur le type de ressources interactionnelles étant en jeu dans ces échanges et sur la manière dont elles sont mobilisées sont dès lors limitées. La communication étant multimodale par nature, bon nombre de ressources posturo-mimo-gestuelles (PMG) sont mobilisées par les locuteurs apprenants d’une L2 et participent ainsi de l’intention communicative de l’énonciateur. Il semble donc essentiel de mieux connaître le rôle de ces ressources PMG pour les énonciateurs apprenants d’une L2 dans le cadre d’interactions apprenants-apprenants, mais également la manière dont ces ressources sont décodées par leurs co-énonciateurs, par la mise en oeuvre de processus inférentiels permettant d’accéder au sens visé. L’étude empirique exploratoire que nous présentons dans cet article se propose d’apporter un éclairage sur ces questions. À partir d’un corpus d’interactions orales entre apprenants de l’anglais L2, nous mettons en évidence l’apport essentiel des processus inférentiels dans l’interprétation de ressources PMG au service de la tentative de mise en mots et de la gestion des

tours de parole, ainsi que dans le décodage des ressources PMG se substituant au verbal. Il émerge également de notre analyse qu'une même ressource PMG peut avoir de multiples fonctions en contexte. Tous ces éléments nous poussent à plaider en faveur de l'intégration, dans l'apprentissage des langues, d'une formation à la production et à l'interprétation des ressources PMG comme autant d'indices « inférables » par les interlocuteurs dans une aire géographique et linguistique donnée.

## INDEX

**Mots-clés:** interaction multimodale apprenants-apprenants, inférence, ressources posturo-mimo-gestuelles, gestion des tours de parole

**Keywords:** multimodal learner interaction, inferential process, visual resources, postures, facial expressions, gestures, turn management

## AUTHORS

### SAGHIE SHARIFZADEH

Sorbonne Université - Centre de Linguistique en Sorbonne (CeLiSo)

saghie.sharifzadeh@sorbonne-universite.fr

### CÉDRIC SARRÉ

Sorbonne Université, INSPE de Paris - Centre de Linguistique en Sorbonne (CeLiSo)

cedric.sarre@sorbonne-universite.fr