



HAL
open science

Le rôle des indices visuels dans la compréhension discursive en L2

Tsuyoshi Kida

► **To cite this version:**

Tsuyoshi Kida. Le rôle des indices visuels dans la compréhension discursive en L2. *Marges Linguistiques*, 2003, pp.260-285. hal-00132373

HAL Id: hal-00132373

<https://hal.science/hal-00132373>

Submitted on 21 Feb 2007

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.



Le rôle des indices visuels dans la compréhension discursive en langue seconde

Par Tsuyoshi Kida¹

Laboratoire Parole et Langage (CNRS-UMR 6057)

Université de Provence

(France)

Mai 2003

Hyakubun wa ikken ni shikazu

[Cent écoutes ne valent pas un coup d'œil]

Proverbe japonais

0. Introduction

On entend souvent un témoignage de ce genre chez les apprenants d'une langue seconde (L2) : « ce qui est le plus difficile, c'est de parler au téléphone ». Parmi les nombreuses causes pouvant être à l'origine d'une telle difficulté, on peut envisager le problème de la disponibilité de l'information visuelle. Ce problème se pose d'un côté en production et de l'autre en compréhension et, du côté production, la stratégie de compensation non verbale ou non vocale des apprenants, telle que le geste, la mimique ou le regard lors d'un trouble de communication en L2, a déjà fait l'objet de quelques travaux (voir Gullberg 1998 ; Kida & Faraco, sous presse). Néanmoins du côté compréhension, il n'y a que peu ou pas d'études sur la compensation visuelle. L'objectif du présent article est donc de s'attaquer à cette question, à savoir la compréhension du discours en L2 dans une perspective multisensorielle, notamment l'effet de l'indice visuel sur la compréhension discursive. Étant donné leur présence dans l'interaction interpersonnelle — et comme le laisse à penser le témoignage des apprenants ci-dessus —, on est en droit de supposer que les indices que fournissent les mains, les bras et la tête, c'est-à-dire le geste ou la mimique, jouent un rôle certain dans le processus de compréhension et, de ce fait, dans l'appropriation d'une L2. La présente étude se limitera à l'impact d'une telle information sur la compréhension du discours par des interlocuteurs non natifs en interaction avec un locuteur natif. Précisons que la visée de l'article sera moins théorique que descriptive sachant que le mécanisme d'appropriation ne pourrait être élucidé à travers la seule description de la compréhension du discours. L'objectif en est plutôt de présenter quelques exemples afin de fournir matière à réflexions pour les débats théoriques sur l'acquisition de L2.

L'article se déroulera de la façon suivante. La première section présentera une revue de quelques travaux relatifs au phénomène de compréhension en général chez les non-natifs. Cette section sera complétée par la section 2 qui se concentrera sur la dimension visuelle dans le processus de compréhension en discutant d'une question méthodologique relative à l'objet d'investigation du présent article. La section 3 présentera brièvement les informations concernant le corpus traité, qui fera l'objet d'une analyse approfondie dans la section 4. La méthode employée y est principalement l'étude comparative des conversations en face à face et en situation de non-visibilité mutuelle. Cette section traitera de quelques-uns des phénomènes remarquables observés dans le corpus, qui semblent suggérer une implication des indices visuels dans la compréhension discursive. Cette analyse inclut :

- la focalisation et l'ellipse que le locuteur emploie dans son discours en situation de privation des indices visuels ;
- la compréhension de l'expression métaphorique qui se caractérise souvent par la manipulation de l'aspect « visuel » du contenu propositionnel dans le discours ;
- l'implication visuelle dans le domaine de la syntaxe ;
- le rapport entre le geste et le domaine extradiscursif.

Enfin l'article finira par quelques remarques sur les études de l'appropriation de L2 en général et les implications didactiques de la perspective adoptée dans le présent essai.

¹ Je remercie les deux évaluateurs anonymes et A. Giacomi pour leurs commentaires.

1. Compréhension discursive en langue étrangère

La compréhension en L2 est moins souvent abordée que la production en L2 par les linguistes en général et ceux qui s'investissent dans l'étude de l'appropriation d'une L2. La problématique de la compréhension peut pourtant être importante pour cette appropriation. Par exemple, Labov (1996, pp. 246-247) a écrit à ce propos : « [...] one key to understanding the difference between second and first language speakers is to be found in the asymmetry of production and perception. The phenomena of interest are paradoxical cases, where L2 speakers use forms that are not used in either L1 or L2, and produce an output that is more remote from their mother tongue than the native L2 forms ». Les exemples démontrant l'effet de la perception sur la constitution linguistique ne nous manquent pas en effet. Il semble que le point central de la remarque de Labov concerne le décalage important entre le vocabulaire actif et le vocabulaire passif chez la plupart des locuteurs non natifs¹. Labov suppose qu'une perception imparfaite influe sur la façon dont les apprenants perçoivent les données extérieures en donnant lieu aux fossilisation et pidginisation. C'est-à-dire (Labov, 1996, pp. 250) : « [...] misperceptions may give rise to an asymmetry between the first and the second language learner's pool of probabilities upon which the learning process depends ». Labov explique ainsi le problème de l'appropriation des apprenants d'une L2 par une thèse « probabiliste » de la perception. Si on interprète la proposition de Labov au sens large, la compréhension est une piste de réflexion pour rendre compte du procès d'appropriation d'une L2.

Cependant, la compréhension discursive des locuteurs non natifs n'a pas constitué un objet d'étude systématique du domaine de l'acquisition de L2. Dans les années 70 (excepté Keller-Cohen, 1979), les chercheurs ont privilégié la description du discours des locuteurs natifs, considéré comme facilitateur de la compréhension pour les locuteurs non natifs (tant apprenants de langue qu'immigrés). Un des plus anciens thèmes en est l'ajustement linguistique (grammatical ou non) du discours adressé aux non-natifs, qu'on appelle communément *foreigner talk* (Ferguson, 1971). Les traits caractéristiques en sont, entre autres, la suppression (par exemple, article, copule, flexion, pronom nominatif). Pour ce qui est non grammatical, l'ajustement s'étend sur la longueur et la complexité syntaxique de l'énoncé, le choix lexical, le débit ralenti, l'usage inhabituel de formules, l'évitement de pronominalisation ou la répétition (cf. Larsen-Freeman & Long, 1991, pp. 117sq). Une telle description n'a, toutefois, pas pris en considération la compréhension des non-natifs.

D'autres ajustements non grammaticaux sont des modifications de l'input, appelées « input compréhensible » de Krashen (1982, entre autres), et de l'interaction, appelées « ajustements conversationnels » (Long, 1983). À la différence des travaux sur le *foreigner talk*, la notion d'input compréhensible accorde un rôle crucial à la compréhension dans le procès d'acquisition. Cette théorie s'inscrit plutôt dans un contexte éducatif. En relation avec l'acte de compréhension, Krashen propose deux conditions nécessaires (mais non suffisantes) pour que l'input se transforme en intake (ou saisie) : la compréhensibilité de l'input et la baisse du filtre affectif. Néanmoins, de nombreux points restent peu clairs. D'abord, le filtre affectif semble une notion « surajoutée » pour prévenir des cas où l'input compris ne deviendrait pas intake. La discussion de Krashen sur ce point n'a pas été approfondie. Ensuite, la notion d'input compréhensible n'est pas non plus expliquée par rapport au fait réel de compréhension, mais à la notion en tant que construction théorique. Pour Krashen, la compétence « i » de l'apprenant devient « i + 1 » grâce à l'input compréhensible du discours pédagogique « i + 1 ». Comment comprendre « i + 1 » avec la compétence de « i » ? Krashen répond que c'est grâce à la connaissance du monde et/ou à l'information extralinguistique (Krashen, 1982, pp. 21). La formule de la compréhension optimale de Krashen est alors : « i + 1 » (input compréhensible ou intake) = « i » (compétence de l'apprenant) + x (extralinguistique). Or, dans cette équation, « x » semble être considéré comme une donnée constante, laquelle n'a pas fait l'objet de réflexion, et, ainsi, le mécanisme de compréhension est resté inexpliqué. La théorie de Krashen a eu le mérite d'avoir déclenché un nombre important de travaux sur la compréhensibilité de l'input. Elle a néanmoins donné lieu à peu d'études sur la compréhension discursive en négligeant la place de l'apprenant au sein de l'interaction sociale (cf. Færch & Kasper, 1986, pp. 262).

En revanche, le modèle de l'interaction modifiée (ou l'input négocié ; Long, 1983) tient mieux compte de la place de l'interlocuteur non natif. D'après Larsen-Freeman et Long (1991,

¹ Cela n'est toutefois pas le cas de certains apprenants. Par exemple, l'aptitude productive de ceux qui sont hautement scolarisés en une langue étrangère est nettement supérieure à la compréhension.

pp. 120), les ajustements conversationnels sont plus nombreux que les ajustements linguistiques morpho-syntaxiques dans le foreigner talk. Ces ajustements ont été étudiés dans le contexte tant dyadique qu'éducatif. La notion de l'interaction modifiée présuppose la négociation avec un non-natif, considéré comme un interactant recevable, mais aussi la détection du trouble de communication, l'achèvement de l'intercompréhension, la notion d'étayage linguistique. Le discours de l'interactant natif, selon ce modèle, se décompose, à la suite du trouble, en des énoncés moins complexes au cours de la négociation, et cette modification facilite la compréhension de l'interactant non natif. En termes discursifs, ce sont la prise et les réparations comme la demande de confirmation, la reformulation, la reprise, la gestion de compréhension. Les ajustements conversationnels se divisent en deux types : « tactiques », pour réparer le discours après un trouble communicatif, et « stratégies », pour éviter le trouble communicatif (Long, 1983 ; « gestion » et « prévention » pour Giacomi & Hérédia, 1986). Du point de vue acquisitionniste, il a été avancé que tous ces traits discursifs présentent une vertu acquisitionnelle (De Pietro et al., 1989) puisque l'attention de l'apprenant est focalisée sur un problème linguistique précis (à la suite de quoi le discours peut s'ajuster) et que la compréhension est facilitée par ces traits eux-mêmes (Long, 1996). Néanmoins, ce modèle, privilégiant la description de la structure de l'interaction, se préoccupe moins du processus de compréhension. Les partisans de ce modèle distinguent mal, semble-t-il, la compréhension « en surface » (que la *structure d'échange* satisfait) de la compréhension substantielle (Aston, 1986 ; Færch, 1985 ; Faraco & Kida, 1999).

La problématique de (non-)compréhension fut, en effet, discutée depuis longtemps. Dans son *Essay concerning human understanding*, John Locke a déjà souligné l'« imperfection » du langage, car il ne transmet pas complètement des « *ideas* » qui sont « *unobservables* » (cf. Harris & Taylor, 1997). La compréhension est, en fait, un « cas particulier de malentendu » (Py, 2000, pp. 123, n. 6). Analyser la compréhension discursive est probablement d'autant plus difficile que l'auditeur est non natif. En effet, incompréhension partielle et malentendu ne se distinguent pas clairement dans les réactions des locuteurs non natifs (Vasseur et al., 1996) ; même divers signaux confirmatifs (par exemple le hochement de tête) ne garantissent pas que l'interlocuteur ait effectivement compris (Vion, 1985). De même, les non-natifs adoptent souvent une stratégie d'écoute (« *selective listening* », Larsen-Freeman (1983) ; « *wait and see* », Voionmaa, 1984) par laquelle le non-natif volontairement s'attribue le rôle d'auditeur lorsque la difficulté du discours à comprendre dépasse son niveau linguistique.

Pour rendre compte de la non-compréhension, il serait pratique de distinguer le malentendu, l'incompréhension totale et l'incompréhension partielle (Gass & Varonis, 1991 ; Vasseur et al., 1996)¹¹. Le malentendu présuppose une compréhension, mais implique une interprétation du message allant dans un autre sens que celui de l'intention du locuteur. L'incompréhension partielle surgit, en théorie, lorsque la source du trouble de compréhension est identifiée, tandis que l'incompréhension totale est un cas où l'interlocuteur n'identifie pas la localisation de cette source. Mais en pratique, on peut difficilement localiser et cerner toutes les non-compréhensions. Pour les analystes, l'incompréhension partielle serait plus facile à identifier puisque la source en est souvent signalée. Cependant, l'incompréhension partielle peut ne pas être traitée en tant que telle, mais signalée en tant qu'incompréhension totale même si l'interlocuteur fait face à un problème précis dans sa parole intérieure. De même, l'incompréhension totale peut être traitée, de façon stratégique, à partir de ce qui ne relève pas directement d'un problème de compréhension (voir l'exemple 10 ci-dessous). La déclaration de la non-compréhension dépend de facteurs psychologiques, tels que la confiance en soi, la personnalité, mais aussi du savoir stratégique. La non-compréhension réelle peut donc être différente du message de non-compréhension en tant que pratique interactionnelle. L'étude de la compréhension nécessite alors d'analyser les négociations dans le déroulement de l'interaction (Bremer, 1996 ; Vion, 1985). Sur la base de cette conception étendue de la compréhension, il y a quelques travaux sur le discours des non-natifs, mais peu nombreuses sont les analyses sur la compréhension chez les non-natifs (entre autres, Giacomi et al., 1984 ; Bremer, 1996). Bremer (1996) énumère comme causes du trouble de compréhension des non-natifs, la mauvaise discrimination phonologique (le malentendu au sens premier), le déficit du

¹¹ En sus de la branche « non-compréhension (*miscommunication*) » avec ces trois types, Gass et Varonis (1991) isolent une autre branche « non-engagement », composée de « non-communication » et « rupture de communication (*communication breakoff*) ». Ces deux branches sont désignées comme « communication problématique ».

savoir pragmatolinguistique, la complexité de la construction syntaxique de l'énoncé (sur laquelle s'appuie le modèle de l'interaction modifiée), les conduites spécifiques de parole (par exemple, ellipse, reprise, incise, glose). Mais ces résultats semblent ne pas avoir éclairci les problèmes auxquels se sont confrontés les travaux sur les input et interaction modifiés.

Par ailleurs, certains travaux inspirés de la psycholinguistique du discours montrent une corrélation intéressante entre les données et le modèle théorique. Le plus illustre renvoie à l'étude de Kasper (1984) sur la compréhension des actes de langage dans la conversation entre apprenant et natif. En s'appuyant au préalable sur la théorie cognitive du traitement et la théorie des actes de langage, elle propose comme variables, la force illocutoire, le degré de conventionnalité, le type de traitement et la procédure d'interprétation. L'analyse corrélatrice des malentendus et des indices prosodiques et pragmatolinguistiques du corpus montre que le malentendu des apprenants vient de la confusion entre échange informationnel et échange phatique. Kasper note que les apprenants échouent souvent dans l'activation d'un cadre d'interprétation pertinent dans le traitement de bas niveau (*bottom-up*) et dans l'utilisation de la procédure d'interprétation multiple du sens. Ce type de traitement est, suppose-t-elle, dû à la forme même de l'apprentissage de la classe, qui use et abuse du traitement de bas niveau, rendant inflexible la procédure de compréhension des apprenants. De façon intéressante, ces caractéristiques procédurales se retrouvent dans la compréhension du discours écrit (Carrell, 1984). Les traits typiques en sont le traitement littéral du texte, l'absence d'utilisation du contexte pour l'interprétation. En termes cognitifs, les lecteurs non natifs n'utilisent pas le contexte général comme traitement de haut niveau (*top-down*), non plus que la transparence des indices lexicaux comme traitement de bas niveau (*bottom-up*) pour interpréter le texte de façon adéquate (voir Carrell 1990 pour plus d'informations).

Comme Krashen l'a noté, les facteurs qui influencent la compréhension discursive ne seraient pas seulement d'ordre linguistique, mais aussi d'ordre non linguistique (tels que la connaissance du monde, le contextuel, la logique, etc.). Étant donné que l'information visuelle relève des indices non linguistiques, il serait nécessaire d'en tenir compte pour l'analyse de la compréhension des non-natifs. En fait, la thèse probabiliste du procès d'appropriation proposée par Labov (voir *supra*) achoppe sur ce point : il est vrai que, comme Labov le laisse entendre, la perception auditive est un acte actif ou analytique, mais repose aussi sur la procédure multimodale. Par exemple, le célèbre « effet de MacGurk » (MacGurk & MacDonald, 1976) a montré que le phonème /ba/ avec un mouvement articulatoire /ga/ donnait aux sujets l'illusion auditive de /da/, qui est proche de /ba/ sur le plan articulatoire. L'effet de MacGurk suggère alors que le mouvement articulatoire peut intervenir dans la perception auditive. Ainsi, la perception de certains phonèmes dépend du « geste » phonatoire et la perception est donc synthétique. Quelques propositions concernant l'implication de l'indice visuel dans la compréhension orale ont été formulées par des chercheurs dans le domaine de la didactique des langues étrangères (voir la revue de S. Kellerman, 1990, 1992). Les travaux en didactique ne sont, néanmoins, pas allés au-delà de la question sémasiologique (c'est-à-dire, de la forme au sens), en d'autres termes, la perception de phonèmes ou la compréhension linguistique non contextuelle du mot et du syntagme¹.

La question de savoir comment s'établit la compréhension mutuelle a davantage préoccupé ceux qui s'intéressent au rapport entre interaction et appropriation d'une L2 plutôt que la question de savoir ce qui est compris (cf. Roberts, 1996a, pp. 9). De la sorte, le processus d'intercompréhension dans l'interaction sociale a été un objet privilégié d'investigation. Néanmoins, les chercheurs, ayant porté peu d'attention à l'indice visuel et ne mesurant pas la portée de celui-ci, n'ont pas assez décrit le vrai processus de compréhension discursive, voire le procès d'appropriation. C'est ainsi que l'absence d'analyse systématique du non-verbal dans le projet pan-européen sur l'acquisition d'une L2 (ESF, European Science Foundation, voir Perdue, 1993) a été jugée regrettable (Roberts, 1996b, pp. 221), car ce type d'indice participe pleinement à la construction du contexte pour la compréhension. De même, E. Kellerman et Sharwood Smith (1981, cité par Sharwood Smith, 1986) ont fait remarquer que les apprenants sont entourés par le monde « iconique », qui fait partie du non-linguistique non négligeable pour leur interprétation du sens. Mais la portée de l'iconicité dans la compréhension du discours par les apprenants reste méconnue.

¹ De façon révélatrice, les travaux investis dans la compréhension soulignent l'avantage de la vidéo sur les matériaux audio et, recommandent son introduction dans la classe de langue, car l'indice visuel augmente le potentiel de la compréhension et, de ce fait, améliore la compréhension orale des apprenants.

2. Compréhension discursive et information visuelle

De même que la compréhension du discours a été moins étudiée que la production, les études non verbales ont aussi souffert d'un manque certain d'investigations. Bien que les travaux sur le geste soient à présent nombreux et que l'intérêt porté sur cette question soit, semble-t-il, grandissant¹, la plupart des travaux portent sur la production gestuelle. Par suite, il va sans dire que la compréhension du geste dans le discours, plus exactement, l'impact de l'indice visuel (geste ou mimique) sur la compréhension du discours n'a, lui, pratiquement pas été exploré (Kendon, 1994)². Les travaux empiriques sur ce thème et portant sur les sujets non natifs sont, à ma connaissance, inexistant³. Cependant, ce problème semble non négligeable pour la problématique acquisitionnelle, car la proposition de Labov concernant la supériorité de la perception sur la production chez les apprenants (voir *supra*) pourrait avoir un lien avec la contribution visuelle sur la compréhension. La réticence des investigateurs devant cette question proviendrait d'une difficulté méthodologique : on peut retenir quelques éléments du discours de son partenaire, mais on est, en principe, incapable de se souvenir, après coup, des gestes faits lors d'une conversation. Cela ne réfute, néanmoins, pas l'octroi d'une fonction aux gestes dans la procédure de compréhension puisque, comme Sapir l'a noté, le gestuel est un code inconsciemment partagé par l'ensemble de la communauté linguistique⁴.

Kendon (1985) a tenté une analyse avec la méthode qu'il appelle « observation-notation-sur-mémoire », c'est-à-dire le recueil de données au moyen de ses yeux, de sa mémoire et de sa propre compréhension, sorte d'observation participante de la compréhension. Une telle étude, note-t-il, permet au chercheur d'aborder les gestes sous des circonstances diverses et variées. D'après Kendon, le geste, par sa projection, joue le rôle de :

- moyen de relayer la parole dans des circonstances où le bruit rend difficile sa réception (contrainte physico-acoustique) ;
- moyen de compléter l'énoncé dans lequel quelque chose qui n'est pas complètement acceptable du point de vue social, est en train d'être dit (contrainte sociale) ;
- moyen de désambigüiser un mot potentiellement ambigu (contrainte sémantique) ;
- moyen de rendre clair ce dont on est en train de parler en pointant quelque chose dans l'environnement (contrainte pragmatique) ;
- moyen de donner une dimension supplémentaire du sens à ce qui est dit (forme, taille, localité relative de quelque chose pendant que le mot est utilisé pour le désigner) (contrainte langagière).

Le geste surmonte, selon Kendon, différentes contraintes auxquelles est assujéti le langage oral dans la procédure d'interprétation discursive, bien que l'inverse soit possible aussi (le modèle gestaltiste, voir Rimé & Schiaratuna, 1991). Séduisants, ces résultats méritent, comme le reconnaît Kendon, des examens détaillés. En outre, l'analyse de Kendon suscite une question : la compréhension a-t-elle été facilitée par la projection du contenu par le geste ou bien par l'inférence logique du déroulement thématique du discours ?

¹ En se limitant à la littérature francophone moderne, on pourra citer, entre autres, des auteurs comme J. Cosnier, M. de Fornel dans une visée conversationniste ; R. Muchembled, J.-C. Schmitt en histoire ; M. Jousse, A. Leroy-Gourhan, M. Mause en anthropologie ou sociologie de technique ; J. Montredon, L. Porcher, Ph. Rilley, Ch. Coupier dans la didactique des langues ou acquisition ; H. Marcos en acquisition du langage ; Ch. Cuxac, D. Bouvet, C. Maury-Rouan, J.-L. Nespoulous dans la pathologie du langage ou le langage des signes ; T. Slama-Cazacu, G. Calbris, J. Vaysse en sémiotique ; B. Rimé, P. Feyereisen, J.-D. de Lannoy, J. Proust en psychologie ; I. Guaitella et ses collaborateurs en phonétique. En linguistique générale, G. Mounin et E. Buysens se sont aussi intéressés au langage gestuel des moines cisterciens ou des Amérindiens. Les références ne sont pas précisées ici pour ne pas alourdir la bibliographie.

² « [...] only few studies have appeared since 1983 that deal directly with the question of what contribution gesture makes to the utterance from the recipient's point of view » (Kendon, 1994, pp. 176).

³ Un nombre important d'études sur la compréhension des gestes conventionnels par les non-natifs existent depuis longtemps mais elles ne s'inscrivent guère dans cette perspective (voir Faraco 2001).

⁴ « Prenons l'exemple des gestes. L'individu et le social s'y mêlent inextricablement ; néanmoins, nous y sommes extrêmement sensibles, et nous y réagissons comme d'après code, secret et compliqué, écrit nulle part, connu de personne, entendu par tous. [...] Comme toute conduite, le geste a des racines organiques, mais les lois du geste, le code tacite des messages et des réponses transmis par le geste sont l'œuvre d'une tradition sociale complexe. » (Sapir, 1967 [1927], pp. 46).

Plusieurs auteurs ont remarqué que l'énoncé était souvent « compris » avant d'être achevé. Beattie et Barnard (1979) soupçonnent que pour 34 % des énoncés de la conversation téléphonique, la fin de ceux-ci est anticipée par l'interlocuteur. Dans ce cas, il ne peut s'agir que de l'inférence discursive, puisqu'il y a non-prise en compte de l'indice visuel. Certains ont noté que la compréhension était facilitée par l'information visuelle. Pour l'achèvement interactif (Gülich 1989), considéré comme un exemple de l'activité d'inférence discursive, il y a au moins une étude (Bavelas & Coates, 1992) qui a montré l'existence d'une projection gestuelle de ce qui peut suivre dans l'énoncé inachevé du locuteur. Cette étude devra toutefois être complétée par une analyse sous condition de privation visuelle pour déterminer s'il s'agit de projectivité ou d'inférence. Il arrive aussi que l'interlocuteur hoche la tête avant le mot clé de l'énoncé et ce phénomène a été analysé par quelques conversationnistes (entre autres, C. Goodwin, 1986 ; M. H. Goodwin, 1980 ; Heath, 1992). Néanmoins, le hochement de tête ne constituerait pas, nous l'avons dit, un indice fiable pour mesurer la compréhension discursive et il semble que l'analyse de l'achèvement interactif soit un cadre de recherche plus fiable que le hochement pour s'interroger si cette compréhension est projective ou inférentielle.

Une autre notion qui mérite d'être étudiée est le malentendu ou la non-compréhension. Dans le cadre de l'interaction natif/non-natif, cette notion est souvent associée à la négociation du sens pour laquelle la contribution mimo-gestuelle a été décrite (Kida & Faraco, 2002 ; Kida, 2002). Mais comme l'étude portant sur la négociation se focalise sur un aspect de résolution de problème par des indices visuels (« reformulation intermodale » dans le but de la compréhension lexicale des apprenants), cette optique s'éloigne — bien que le problème lexical soit intimement liée au discours — de la problématique plus générale de la compréhension discursive. Il en va de même pour le geste projectif lors d'une recherche lexicale (par exemple, Streeck, 1995). Ce type de geste, produit au cours d'une pause discursive, a pour but de conduire l'interlocuteur à « traduire » le geste en expression linguistique à travers la focalisation de l'attention sur l'univers visuel, et non sur l'apport d'indices visuels *dans le discours*, et il s'éloigne donc de la présente discussion. Une alternative serait alors d'étudier le déroulement de négociation en se demandant comment et pourquoi le malentendu a surgi, ce qui a troublé la compréhension et, en particulier, si et dans quelle mesure le geste et d'autres indices visuels ont ou auraient pu contribuer à la compréhension (cf. Roulet, 2000). Une telle analyse contextuelle est indispensable pour l'étude de la compréhension des apprenants puisque le contexte est particulièrement sensible à l'interaction interculturelle, certains savoirs linguistique et socioculturel faisant défaut aux sujets non natifs (Roberts, 1996a, pp. 24).

Il y a deux types d'études qui ont exploité la notion de malentendu à travers un lien entre indices visuels et compréhension discursive. Le premier concerne le moment où la cohésion pronominale est ambiguë (dans la mesure où le pronom ne renvoie pas au référent le plus proche). McNeill et Levy (1993) découvrent qu'il y a souvent un geste qui signale une telle « violation », même si la forme de ce geste ne se réfère pas symboliquement à ce qui précède. Quant à Brown (1998, pp. 189-190), elle avance que la levée d'une telle ambiguïté nécessite l'adoption d'une même perspective par les participants. Les indices qui peuvent remplir cette fonction sont, en plus du geste, la prosodie, la posture, le jeu de contact visuel (Brown, 1998, pp. 191). Mais encore une fois, ce cadre de recherche ne répond pas à la question de savoir si la compréhension a été sous-tendue par la projection gestuelle ou par l'inférence discursive. Le second type d'études s'intéresse au cas où le geste ne correspondrait pas au contenu propositionnel du discours (*speech-mismatched gesture*, McNeill et al., 1994). Dans cette étude expérimentale, les sujets, qui ont visionné, sur écran, un discours narratif rempli de faux gestes, le racontent à leur tour à un autre sujet qui ne connaît pas l'histoire.

En d'autres termes, le sujet interprétant s'est confronté à la contradiction entre information visuelle et information verbale. Les résultats montrent que les sujets, ou choisissent l'une des informations, ou intègrent ces informations pour créer une nouvelle scène (p. 234), ou se souviennent mal du contenu (p. 231). Ainsi, la non-correspondance entre ce qui est vu et entendu génère des problèmes communicatifs et mnémoniques, ou la compréhension des sujets devient un état syncrétique des informations gestuo-verbales du discours. Cette étude suggère que l'indice visuel constitue une part du discours *pour l'interlocuteur* (McNeill, 1985), ce qui est une avancée considérable dans la compréhension de la notion de discours. Néanmoins, cette étude (et les autres travaux gestuels en général) ne touche pas à l'aspect procédural de la compréhension.

Dans une autre perspective, il y a des études qui ont analysé directement la portée visuelle sur l'interaction en supprimant la visibilité entre partenaires. Les premiers travaux de ce type (entre autres, Argyle et al., 1973 ; Cohen, 1977 ; Cohen & Harrison, 1977 ; Cook & Lalljee, 1972) ont porté principalement sur la différence du contenu dans la communication médiatisée (par exemple, la conversation au téléphone)¹. Deux études de ce genre ont trait à la portée du geste sur la compréhension discursive : celle de Graham et Argyle (1975) et celle de Walker et Nazmi (1979). Ces groupes se sont investies tous deux dans une comparaison entre des sujets de cultures italienne et britannique. La tâche des sujets dans leurs études consiste à dessiner des figures géométriques à partir d'une explication verbale donnée par un locuteur. Pour mesurer l'impact de l'information visuelle, sous une condition, le locuteur est autorisé à utiliser le geste, et sous une autre ne l'est pas. Graham et Argyle trouvent que, bien que la réalisation des Britanniques ne se différencie pas sous les deux conditions, les Italiens ont réalisé leur dessin nettement mieux lorsque l'interlocuteur voyait le locuteur-gesticuteur. Dans l'étude de Walker et Nazmi, les Italiens se rappellent mieux le dessin face au locuteur qui utilise le geste. Ces études indiquent que les Italiens tirent mieux profit du contenu visuel que les Britanniques. Autrement dit, il semble que les Italiens appuient plus leur compréhension sur le geste (du moins pour ce qui est du contenu pictural) que les Britanniques. De façon intéressante, ces résultats convergent avec ceux de l'étude ethnographique d'Efron (1972) sur les gestes des immigrés italiens et juifs de New York, étude qui postule une différence interculturelle du rôle que joue le geste dans la communication.

En somme, divers types de contribution visuelle à la compréhension discursive ont été pressentis, et l'existence de ce phénomène semble se confirmer. Cela suppose que l'étude de la compréhension discursive ne peut ignorer, comme ce fut le cas de la prosodie en linguistique générale, la place de l'indice visuel. Par ailleurs, la variation culturelle de cet apport permet une hypothèse que les interlocuteurs natifs et non natif sont différents quant à la portée qu'ils donnent à la contribution visuelle. Enfin, tout comme l'étude de la compréhension discursive, l'analyse de l'apport visuel dans l'interaction exige un examen soigné. Une exigence méthodologique sera l'analyse du discours dans le déroulement de l'interaction. Elle nécessite notamment une élaboration de cadres de recherche spécifiques pour avoir une vision cohérente et approfondie de la compréhension du visuel en tant que partie intégrante du discours. Tout cela constituera notre hypothèse de base pour l'analyse de la compréhension du discours par les sujets non natifs.

3. Cadre d'analyse

Les informateurs non natifs sont des Japonais de la jeune génération qui sont en France soit pour un séjour linguistique, soit pour des études universitaires (sauf un médecin de coopération internationale, ayant la quarantaine et apprenant le français pour ses missions). Leur niveau linguistique se situe entre débutant et intermédiaire pour les apprenants de langue et entre intermédiaire et avancé pour les étudiants inscrits dans le cursus universitaire français. Tous les informateurs n'ont qu'un projet de séjour limité (souvent moins d'un an), rendant impossible la réalisation d'une enquête longitudinale. Leur profil sociologique restreint donc les lieux et les moments de leur contact avec la L2 et les autochtones. Leur aire d'apprentissage se limite à la participation à la classe de langue ou au cours, à la lecture-écriture, à l'accès aux médias et à la conversation. En général, ils ont, à la différence des immigrés (par exemple, les informateurs du projet ESF) peu d'interactions institutionnalisées, si ce n'est dans des transactions commerciales limitées. Parfois il a été frappant de constater un écart considérable entre niveau linguistique élevé et connaissances socioculturelles pauvres chez certains de ces informateurs.

Ce profil sociologique des informateurs nous a conduit à adopter la conversation comme cadre majeur de recueil de données. Le partenaire de conversation était un locuteur natif (autre que l'expérimentateur). Toutes les conversations ont été enregistrées et filmées. Avant

¹ En ce qui concerne la production, contre toute attente, il y a peu (ou pas) de différences entre conditions de visibilité et de non-visibilité. De même, l'expérience de l'interaction non médiatisée avec un écran opaque (cf. Rimé, 1984) n'a pas non plus montré de différence notable quant au discours du locuteur (contenu, lexique, syntaxe, et structure grammaticale) ou quant aux aspects affectifs et interpersonnels. Le seul aspect ayant été affecté est la régulation de l'interaction (plus de carambolages, silences, pauses remplies).

l'enregistrement, l'auteur du présent article en tant qu'expérimentateur a donné aux participants quelques explications spécifiques à la situation, a mis en marche la caméra et a quitté la pièce d'enregistrement. Personne n'a ressenti de difficulté due à la présence de la caméra. Un autre dispositif supplémentaire a été de poser un miroir pour intercepter simultanément les partenaires de la dyade dans un seul champ de caméra. Le recueil de données a été effectué dans un local universitaire. Ce cadre est différent de celui de l'enquête de la sociolinguistique interactionnelle ou de la méthode de l'analyse de discours critique où l'interaction institutionnalisée est la situation privilégiée pour le recueil de données (voir Roberts, 1996b). Notre objectif est, nous l'avons dit, de sonder l'impact de l'information visuelle dans l'interaction. Les situations sociales ou quotidiennes sont en général difficiles à filmer pour des raisons pratiques et éthiques (voir pourtant Gumperz & Roberts, 1991). Il faudrait alors recourir au jeu de rôle, mais, d'après mes expériences personnelles, le jeu de rôle n'est pas fiable dans la mesure où nos informateurs ne sont pas tous familiers avec certaines situations sociales que propose le jeu de rôle, de même qu'il n'est pas aisé de contrôler leur motivation de participation. La seule exception en est la classe de langue, mais les données issues d'une telle situation ne supportent pas une analyse fine de l'interaction, laquelle est la visée du présent article. Ainsi notre cadre d'analyse est-il déterminé non seulement par des raisons sociologiques propres aux informateurs, mais aussi par des raisons pratiques et méthodologiques propres à l'enquête.

Bien que les informateurs aient participé à au moins une conversation libre, par souci de comparabilité, le corps principal du corpus est la conversation thématique. Le thème a été la cuisine du pays natal, sujet culturellement connoté. Les participants ont demandé à leur partenaire d'expliquer une recette puisqu'ils voulaient être à même de la reproduire chez eux. Ce thème a été choisi pour motiver les informateurs dans la participation à la tâche. Certains informateurs ont déclaré que cette situation leur était arrivée, d'autres ont manifesté leur difficulté à trouver des termes appropriés pour les activités culinaires, mais tous les informateurs étaient satisfaits de la participation¹. Pour améliorer nos conditions d'analyse, une consigne a été imposée aux participants : le demandeur de recette devait ré-expliquer ce qu'il avait appris ou compris après l'explication du locuteur. Cette consigne avait pour but de minimiser le risque de l'évitement ou de l'acquiescement feint et de maximiser la mise en évidence de la nature des problèmes de compréhension. Elle a obligé les participants à poser des questions en cours d'interaction lorsqu'ils n'étaient pas certains de quelque chose. Cette contrainte a présenté un double avantage pour évaluer la nature de la compréhension des participants : d'une part, il a été possible de contrôler la compréhension à travers des négociations locales et, d'autre part, la session de ré-explication a permis de porter un autre regard pour la vérification de l'état de la compréhension. La procédure d'identification de la non-compréhension dans le présent travail a été de partir de la pratique verbale et d'examiner ensuite le traitement du trouble et d'autres indices se trouvant dans l'ensemble des données.

Pour apprécier le rôle de l'indice visuel, un écran opaque a été installé dans un second temps, pour séparer visuellement les partenaires. Les partenaires de chaque dyade ont donc chacun exposé une recette sous deux conditions, en face à face et en privation visuelle. Presque tous les informateurs ont manifesté quelque difficulté dans la situation utilisant un écran par rapport à la situation en face à face. Par ailleurs, le corpus vidéo implique une activité kinésique du locuteur même dans la situation d'écran, ce qui permet d'inférer des rôles éventuels d'un geste sur la compréhension discursive. De plus, la constance du thème de conversation a permis la comparaison entre situations en face à face et privation visuelle². Pour le présent travail, les conversations sur une recette française en face à face (R1b) et avec écran (R2b) ont été analysées, puisque les sujets non natifs étaient confrontés à des phases de compréhension du discours français. Occasionnellement ont été abordées les données issues de la situation où les non-natifs parlent de leur recette japonaise (R1a et R2a). L'ensemble des données observées atteint environ 200 pages de corpus transcrit.

¹ Mais leur satisfaction ne nous a semblé que momentanée, puisqu'à la proposition d'une interview rétrospective, seule une minorité s'y est intéressée. Ces données n'ont donc pas constitué une donnée substantielle pour notre analyse.

² Nous sommes conscient qu'une simple comparaison ne mettrait pas en évidence l'impact réel de l'indice visuel sur la compréhension du discours. Le meilleur et le seul moyen pour cela est d'analyser l'interaction naturelle au cas par cas (Kendon, 1994). Notre hypothèse de travail était que l'écran préserverait mieux l'état naturel de l'interaction que d'autres cadres expérimentaux conduits en psycholinguistique.

La fonction métadiscursive d'équivalence est-elle alors exprimée par la mise en parallèle des gestes qui portent sur le discours et non pas par le connecteur ? Nous reviendrons sur ce point ultérieurement.

Le geste 3 forme un cercle qui renvoie sans doute au référent *morceau d'ail*, qui n'est pas verbalisé à ce moment, mais juste après. Le geste 4 est quelque peu différent des gestes 1, 2, 3, 5 et 7. Il indique le locus central du geste 3 (c'est-à-dire *morceau d'ail*). Le geste 4, qu'on peut traduire *ici*, ou *milieu*, est une sorte de locatif ou complément de lieu qui remplit une fonction de pointage déictique. Les gestes 3 et 4 forment une syntaxe MORCEAU + MILIEU, inverse à l'ordre syntaxique du français (Diderot a tort peut-être sur ce point), mais similaire à celui de certaines langues (par exemple, le japonais). Tout comme la deixis verbale, le geste 4 doit fixer le point de référence (« origo », Bühler, 1965 ou « centre d'instance énonciative », Kida, 2001), et le geste 3 aurait alors pour but de préparer l'expression verbale *au milieu de l'ail*. La syntaxe gestuelle peut être différente de la syntaxe verbale.

Le geste 5 représente une infime partie après les gestes MORCEAU et MILIEU, et semble correspondre à *germe* dans le discours. Ce geste est iconique et projectif, mais son mouvement est, comme la longueur du stroke l'indique, statique. Le geste 7 représente l'action d'enlever quelque chose et correspond donc à l'expression verbale en 047. On s'aperçoit que le verbe manque de son deuxième actant (complément). Par cette ellipse, l'intention du locuteur paraît focaliser l'attention de l'interlocuteur sur le segment important de l'énoncé. L'ellipse apparaît plus fréquemment dans le discours lorsque, par exemple, l'intercompréhension est en danger. Une telle focalisation discursive s'accompagne souvent d'un geste et/ou d'une emphase prosodique, probablement pour souligner l'effet de focalisation. Par ailleurs, on remarque une autre spécificité de l'activité gestuelle : le mouvement gestuel ENLEVER implique déjà l'objet, inhérent au mouvement. L'ellipse est peut-être rendue possible par cette spécificité du mouvement que présentent certains gestes.

À la différence des gestes 5 et 7, certains n'ont pas de corrélat verbal direct, mais viennent compléter l'énoncé. C'est le cas des gestes 6, 8 et 9. Ces gestes pointent tous du doigt (ici l'index) la bouche du locuteur.

et ça c'est très très fort /ça n'est pas très bien /pour l'haleine		
Geste B	Geste B	Geste B
(BOUCHE)	(BOUCHE)	(BOUCHE)

Si on attribue au geste B ce qui est pointé comme ci-dessus, le geste complète l'énoncé (*pour l'haleine dans/de la bouche*) ou bien le reformule (*ça n'est pas très bien pour la bouche*). Si telle était l'interprétation du geste, l'activité gestuelle anticiperait le référent qui aurait dû suivre mais qui n'est pas verbalisé. Il s'agit d'une complémentarité résultant d'une réserve de l'expression verbale, sans doute pour une raison culturelle. Comme le dernier geste est plus dynamique que les deux autres, on a l'impression que le locuteur se déculpabilise de l'expression par une focalisation sur le dernier segment de l'énoncé. Notons aussi l'absence du connecteur de causalité *donc* : *et ça c'est très très fort* (cause), *donc ça n'est pas très bien pour l'haleine* (effet). Il est possible que la manifestation anticipée et répétée du référent par le geste B ait pour fonction de visualiser la relation cause-effet avant même que l'énoncé ne s'achève et ce, en l'absence du marquage verbal (voir *infra*).

En résumé, diverses formes et mouvements traduisent un objet, une action ou un lieu. La relation entre le geste et son correspondant verbal peut être explicite ou implicite (complémentaire). D'autre part, certains gestes n'ont pas de corrélat verbal. La synchronisation entre action et contenu est assez variable, mais dans de nombreux cas, la projection gestuelle est nette. Par ailleurs, la combinaison de gestes, notamment la reprise à travers certains des énoncés, mais aussi la combinaison spécifique avec l'information verbale peuvent être liées à une fonction métadiscursive. Cette diversité gestuelle peut-elle alors aider à la compréhension discursive des non-natifs ?

4.2. Focalisation discursive et ellipse

L'activité gestuelle, nous l'avons dit, anticipe le discours. On peut *a priori* imaginer qu'une telle projection crée un indice qui fournit à l'interlocuteur un contexte immédiat à partir duquel le discours est prévisible. En réalité, il n'est pas aisé d'établir une relation entre projection visuelle et activité interprétative du discours.

Par exemple, il est difficile d'examiner si l'interlocuteur prête effectivement attention à l'indice visuel et de quelle façon l'interlocuteur le retient avant l'arrivée de l'information verbale et, après coup, l'utilise pour l'interprétation du discours¹. Ici supposons que l'interlocuteur tienne compte de l'indice visuel si son regard est dirigé vers le geste du locuteur.

Pour une raison technique (l'angle de la caméra), notre corpus vidéo ne permet pas de systématiquement mettre en évidence l'intérêt porté à l'information visuelle par l'interlocuteur, puisqu'il est difficile de déterminer si le focus du champ visuel est la face du locuteur ou bien l'action des mains. L'exemple suivant est un des rares exemples qui montre clairement que l'interlocuteur dirige son regard vers un geste du locuteur. LN explique, en face à face, comment faire cuire le couscous à AY, son partenaire non natif (niveau linguistique avancé). Le moment où AY regarde le geste de LN est représenté en gras.

(2) AY-R1b-couscous

123LN : ¹⁻¹{et:: quand on a fini:/ on met un peu de sel/ et on met un peu d'huile
d'olive\ ++ et: il faut:: euh}²⁻¹{:::++}\}

124AY : ((sourire))

125LN : ³⁻¹{euh:: ++ DÉcoller\ ++ un peu::}\ ⁴⁻¹ {le::[::]: }]

126AY : [(Hv)]

127LN : les GRAInS\ parce que ça se colle\ }]

128AY : [(Hv)] ((bas)) oui

129LN : tu vois ++

130AY : ((bas)) oui ++

Avec les gestes 1, 2 et 3, LN frotte ses mains et imite l'action de décoller les grains de couscous. Le geste 4 est une autre action. Le geste 1 implique 11 strokes, le geste 2, 3 strokes et le geste 3, 4 strokes. Dans le geste 2, LN opère un changement de la direction de ses mains (probablement dû à une hésitation en 123), mais le mouvement est identique. Les gestes 1, 2 et 3 peuvent être donc considérés comme une action solidaire qui renvoie à l'expression verbale *décoller les grains* (en 125 et 127). Quant au comportement de AY, son premier coup d'œil porte sur le geste 1 (en 123) et le sourire et le deuxième coup d'œil qui suivent témoignent de l'intérêt suscité par l'activité gestuelle de LN. Notons que l'anticipation de ce geste sur l'expression verbale est de 6 secondes.

Le geste 4 est tout différent. Il se réfère à l'action de deux choses qui sont en contact (donc *se coller*, en 127). Le focus du discours de LN est *décoller*, qui est marqué par une emphase prosodique (125). Le terme *grains* est aussi prosodiquement marqué, mais la valeur de focus s'estompe de par le fait que le geste 4 renvoie, non pas à ce terme mais à l'expression qui vient (le fait de se coller). Autrement dit, *grains* n'est pas marqué gestuellement. Une telle « défocalisation » annexe ce terme au premier focus *décoller* ou au second focus *se coller*. Si le contexte le permet — et si le locuteur le veut —, le terme annexé peut être omis comme en 123 (voir aussi l'exemple 1 pour un cas similaire).

Deux hochements de tête de AY (le premier au moment de l'hésitation sur l'expression cruciale et le second après la proposition *parce que*) et ses réponses *oui* témoignent suffisamment de sa saisie du discours. La focalisation seulement à travers la prosodie (voir Rossi, 2000) est partiellement non prédictible dans cet échange, puisque l'activité gestuelle peut participer à la hiérarchisation de focalisation.

Ainsi, le rapport entre ellipse, focalisation et indice visuel est assez complexe en situation de face à face. Par contre, lorsque les interactants ne disposent pas de biais visuel, ils doivent modifier leur mode de communication. Sous la condition de non-visibilité, les interlocuteurs sont privés d'une mise en focus subtile. Pour faciliter la compréhension de son partenaire, le locuteur abuse nettement de l'emphase prosodique et de l'ellipse discursive.

Plusieurs types d'ellipse sont attestés à l'intérieur de notre corpus (voir Appendices). Nous avons au préalable examiné la relation entre la non-visibilité et le discours elliptique du locuteur natif en interaction avec sept informateurs non natifs.

¹ Encore plus difficile est le mécanisme de la vision, à savoir comment l'interlocuteur utilise l'information non focalisée dans le champ visuel.

Tableau 1

	Avancé					
	AY-R1b/R2b		AK-R1b/R2b		YO-R1b/R2b	
tours	139	102	104	110	101	128
ellipse	4	8	3	15	11	9
moyen	34,8	12,8	34,7	7,3	9,2	14,1

	Intermédiaire				Débutant			
	HI-R1b/R2b		MR1-R1b/R2b		SA-R1b/R2b		MA-R1b/R2b	
tours	115	353	124	196	198	148	100	62
ellipse	4	24	10	19	11	7	9	4
moyen	28,8	14,7	12,4	10,3	18,7	20,9	11,1	15,5

(moyen = nb de tours/ellipse ; R1 = face à face ; R2 = écran)

Cependant une simple comparaison du nombre d'occurrences entre conditions de visibilité et de non-visibilité ne montre pas de régularité. Avec deux apprenants avancés (AY et AK), l'ellipse de LN devient plus fréquente en situation d'écran, probablement à cause de la difficulté de la situation, mais avec un autre apprenant avancé (YO), ceci n'est pas valable. Les apprenants intermédiaires et débutants se ressemblent relativement sauf HI : avec eux, l'ellipse est plus fréquente qu'avec les avancés et la comparaison visibilité/non-visibilité n'offre pas de grand contraste par rapport aux avancés. Il semble qu'avec les intermédiaires et débutants, la difficulté de la tâche qu'est la conversation en L2 pèse déjà sur les partenaires. HI se rapproche des avancés, mais ce qui ne veut pas dire que le type d'interaction soit similaire. Avec HI, peu motivé lors de la conversation sous visibilité (R1b), l'ellipse de LN est relativement peu fréquente ; par contre, lors de la conversation sous non-visibilité (R2b), l'augmentation de l'ellipse est liée, comme le nombre important de tours l'indique, à la présence massive de moments problématiques. La lecture du tableau ci-dessus suggère que la privation de l'information visuelle et le niveau linguistique de l'interlocuteur non natif sont probablement liés à l'utilisation de l'ellipse, mais cette proposition devra être considérée avec précaution.

Même si le niveau linguistique n'est pas un facteur déterminant de la fréquence de l'ellipse, il semble qu'il influence l'aspect qualitatif de l'échange. Une des différences de cette portée en cas de privation visuelle s'entrevoit dans le corpus. L'exemple qui suit vient de la conversation avec un apprenant avancé sous la condition d'écran. Le terme autour duquel l'échange s'articule est *laisser*. Le discours de LN devient elliptique lorsque la négociation se déclenche (les parties elliptiques sont en gras).

(3) AK-R2b-brandade

- 024LN : pas trop\ ++ du sel et du poivre/ + et:: puis::: on laisse la:: + la purée:: à l'atten [te]
- 025AK : [n:]
- 026LN : il faut: ++ il faut: quand même + PAS la laisser\ sinon elle du:rcit ++ et pendant ce temps::: [on (a)]
- 027AK : [donc::] on on on laisse::: ++ on LAISse\ ++
- 028LN : ouI:: on **laisse:: au CHAUD: quoi** ++
- 029AK : [n::]
- 030LN : [mais]:: + on on **fait pas cuire plus** + on:: + **on laisse tranquille**
- 031AK : n:\ ++
- 032LN : on **ne doit PAS faire le lendemain** je veux dire\ ++
- 033AK : n:\ ++

La négociation est déclenchée par l'intervention de AK sur le terme focalisé en 027. L'ellipse est utilisée par LN pour une confirmation du point soulevé par AK (*on laisse au chaud, on laisse tranquille*) ou pour une reformulation (*on fait pas cuire plus, on ne doit pas faire le lendemain*), accompagnées par une glose ou particule métaénonciative (...*quoi, ...je veux dire*). AK se montre capable de suivre chacune des ellipses par acquiescement chevauché ou accommodé. Sa gestion de moments problématiques repose davantage sur l'inférence contextuelle du discours et ce comportement se rapproche de celui d'un locuteur natif. Autrement dit, l'objet de focalisation est négocié par les partenaires sur la base du partage de la fonction des ellipses.

En revanche dans l'exemple suivant, un autre apprenant, débutant, affiche peu de participation, ce qui rend difficile d'appréhender la nature de son incompréhension. Notons d'abord le contexte général : LN a déjà expliqué à SA la préparation de la viande pour la daube (la viande découpée mise dans un bol rempli de vin rouge au frigo pendant une nuit) et en retour SA a déjà répété l'explication. Ce passage commence juste après ces épisodes et consiste en une confirmation nouvelle de la part de LN (en 221). Malgré l'apparent suivi de SA jusqu'à ce moment-là, l'incompréhension totale de SA en 222 est frappante. Ce fait oblige LN à reprendre l'explication à partir de 223.

(4) SA-R2b-daube

221LN : quand même je je crois que au début j'ai pas bien expliqué + la viande il faut faire RE:VE:NI::R un petit PEU::/ ++ et:: aPRÈS:/ ++ on on VERse + le vin ++ desSUS/ ++ et:: on **laisse ++ cui::re ++** TOUT doucement ++
 222SA : ah:::(h) par(h)don(h)\ je ne je ne comprends pas\ c'est un peu compliqué\ he [-he]
 223LN : [oui] ++ non:: non\ +++ la viande + on on PREND/ +
 224SA : [hum hum/]
 225LN : [la vian]de qui était dans le VIN::/ ++
 226SA : hum hum/ + ouI ++
 227LN : et:: on **MET: avec** les oiGNONS::: et les caROT:: [tes et]
 228SA : [hum hum/]
 229LN : des champiGNONS:: +
 230SA : hum hum/ ++
 231LN : et:: aPRÈS::/ ++
 232SA : hum hum/ ++
 233LN : on aJOU:te ++
 234SA : hum hum/ ++
 235LN : le vin\ ++
 236SA : hum hum/ ++
 237LN : tu comprends/ + le vin qui: était::: + >dans le frigo\< ++++
 238SA : n:::~::~:

Comme on l'a vu dans l'exemple 2, LN met en saillance quelques termes par la prosodie, auxquels SA réagit par acquiescement (*hum hum*). Mais cette réaction à l'emphase de LN paraît trop systématique :

LN	SA
la viande on on PREND/ +	hum hum
la viande qui était dans le VIN::/ ++	hum hum + oui
et on MET: avec les oiGNONS:::	∅
et les caROT:: [tes	hum hum
et des champiGNONS:: +	hum hum
→ et:: aPRÈS::/ ++	hum hum
on aJOU:te ++	hum hum
le vin\ ++	hum hum
tu comprends/ +	∅
le vin qui: était::: + >dans le frigo\<	n:::~::~:

Certes, le signal *hum hum* de SA constituerait la « stratégie d'écoute », permettant, par exemple, de laisser avancer le discours du partenaire pour sa propre compréhension. Notons toutefois que SA use excessivement d'un tel comportement durant l'ensemble de la conversation mais qui prend peu d'initiatives pour mettre en évidence la nature de l'information saisie. Le signe d'acquiescement utilisé par SA est alors une réaction mécanique à la rythmicité prosodique du discours (le bien-suivi du discours par SA se voit compromis notamment eu égard à sa réaction au connecteur, indiqué par une flèche ci-dessus, qui n'apporte pas d'information pertinente). Si bien que SA coopère plus au niveau phatique de l'échange qu'au niveau informationnel. Comme le partenaire de SA ne saurait point ce qu'est l'état actuel de compréhension et comment ajuster le discours pour une meilleure intercompréhension, ce comportement rend, en fin du compte, problématique sa propre compréhension.

Ainsi, on voit une différence qualitative chez les apprenants quant au comportement face à des moments problématiques sous la condition de non-visibilité mutuelle. Les apprenants avancés tendent à négocier, de façon stratégique, le problème de compréhension conjointement avec le partenaire. Ce comportement interactionnel, qui repose probablement sur l'inférence discursive, prolonge la séquence de focalisation discursive en faisant augmenter le nombre d'ellipses dans le discours du partenaire en privation visuelle par rapport à la situation en face à face. Par contre, les apprenants les moins avancés saisissent peu ce genre d'opportunité pour une entente mutuelle, mais adoptent plutôt une stratégie d'écoute. Ils tendent à laisser au partenaire natif le loisir de mener la focalisation. Avec un tel comportement, il y aurait peu de différences entre situation d'écran et face à face. Autrement dit, l'ellipse est moins bien exploitée par les débutants que par les avancés, différence comportementale pour les ellipses qui semble apparaître dans le tableau ci-dessus.

À propos, on peut se demander quelle est la source du blocage de SA dans l'exemple précédent. De façon spéculative, nous dirons qu'il y a d'autres ellipses non focalisées. Concrètement, les énoncés 221 et 223-237 et suivants impliquent les ellipses suivantes (inférées du contexte) :

221

la viande, il faut faire REVENIR* un petit PEU [lieu, ø1] /et aPRÈS, on VERse le vin [ø2, complément] dessus [ø3, anaphore] /et on laisse cuire [ø4, lieu] tout doucement

- ø1 = dans la casserole
- ø2 = (le vin) du bol
- ø3 = sur la viande dans la casserole
- ø4 = dans la casserole
- * sans doute l'expression est méconnue

223sq

la viande, on PREND /la viande qui était dans le VIN [ø5, métonymie] /et on MET [ø6, complément] avec les oiGNONS et les caROTtes et les champIGNONS [ø7, lieu] /et aPRÈS, on aJOUte le vin [ø8, complément] [ø9, lieu] /(tu comprends) le vin qui était dans le frigo [ø10, métonymie]

- ø5 = dans le bol rempli de vin
- ø6 = la viande (complément focalisé)
- ø7 = dans la casserole
- ø8 = du bol
- ø9 = dans la casserole
- ø10 = dans le bol avec la viande

Cette analyse montre que la localisation (bol et casserole) manque dans le discours du LN. Il est à noter que l'information concernant la localisation est exprimée, derrière l'écran, par un certain nombre de gestes. Peut-être pour SA, l'action complexe de transvaser la viande ou le vin d'un récipient à l'autre, ou la désignation indirecte de ceux-ci ont constitué, faute d'indice visuel et de capacité d'inférence contextuelle, un obstacle à la compréhension discursive. Dans ce cas, la mise en saillance prosodique, seul moyen de focalisation faute de moyen visuel, ne fonctionnerait pas comme une compensation de compréhension discursive.

4.3. Métaphore et focalisation

Dans la section précédente, nous avons vu que la mise en emphase prosodique était souvent utilisée comme focalisation discursive en situation de non-visibilité mutuelle. La focalisation s'accompagne non seulement de l'ellipse, mais aussi de l'expression métaphorique et de l'omission de certains termes non focalisés. Lakoff (1987) a souligné le rôle des métaphores et métonymies dans la cognition des expériences humaines. La conceptualisation est pour lui une perception topologique qui s'appuie sur la logique de complémentarité (cf. Taylor, 1993, pp. 212). La grammaire est, pour Lakoff, constituée « spatialement ». Les représentations schématiques du langage, sur lesquelles le modèle de la grammaire cognitive repose (entre autres, Langaker, 1991), sont une conventionnalisation de cette conceptualisation. Il est possible que la compréhension de la métaphore par les locuteurs non natifs repose souvent sur l'information visuelle, qui est proche de la cognition naturelle, puisque l'expression métaphorique qu'utilisent les locuteurs natifs est une conventionnalisation spécifique à la L2. Sans cette information, l'emploi ou non de la métaphore peut différencier leur compréhension.

Deux exemples sous la condition d'écran permettent de comparer la portée de la métaphore sur la compréhension. L'un implique une expression métaphorique qui semble être la cause du trouble de communication, l'autre n'implique pas d'expression métaphorique et ne présente pas de trouble. Dans l'exemple suivant, touchant au premier cas, LN explique la soupe au pistou à HI, un apprenant intermédiaire. C'est au moment où on met la sauce de pistou dans la casserole. La métaphore y est représentée en gras.

(5) HI-R2b-soupe au pistou
 241LN : [...] maintenant les hariCOTS:::: donc::::: je:::: METS:::/ +
 242HI : oui/ ++
 243LN : ce méLAN:ge/ + /de toMA:te et d'huile d'oLI:ve et de basiLI:c + **DANS les haricots** +++
 244HI : ((suspect)) e:::::/
 245LN : qui sont TOUjours + en TRAIN de CUIre\ ++ tu comprends/ ++
 246HI : ((bas)) non\ +++ e:::::/ +
 247LN : qu'est-ce que tu comprends pas:\ +
 248HI : e::::\ je ne peux:::: PAS:: imaginer\ ++ e::::: +++

Effectivement HI ne peut pas imaginer ce à quoi renvoie l'expression métaphorique, *dans les haricots*, qui peut être traduite, dans ce contexte, comme *dans la casserole où les haricots sont en train de cuire*. A priori, cette métaphore (plus exactement, la métonymie), employée inconsciemment par LN pour focalisation, pose problème à HI. Notons que LN a tenté une nouvelle fois l'explication par reformulation (cette fois, *dans les haricots, dans l'eau + des haricots*), mais HI n'est toujours pas parvenu à la suivre (voir *infra*). En revanche dans l'autre exemple avec un autre apprenant intermédiaire, MR, qui a expérimenté une même situation, mais sans métaphore, il en va autrement :

(6) MR1-R2b-soupe au pistou
 167LN : je PRENDS::: et j'écRA:::se j'écRA:::se j'écRA:::se ++
 168MR : hum hum ++
 169LN : et::::: ++ QUAND j'ai: fi:NI:::/ ++
 170MR : hum hum
 171LN : je:::: VER:::se ++ tu compRENDS ++
 172MR : je ++ [(ver-)]
 173LN : [je: METS:::/] ++ le méLAN:::ge ++
 174MR : oui ++
 175LN : **DANS:: la SOU::pe + de haricots** ++
 176MR : [ah oui]
 177LN : [je mets] dedans\ ++
 178MR : ouI: ouI: ouI: ++

Dans cet échange, le discours de LN n'incluant pas de métaphores ne pose pas de problème. Ces deux exemples laissent supposer que la métaphore, qui est une ellipse sémantique, peut provoquer un trouble de communication chez certains partenaires non natifs.

Le propos tenu jusqu'ici est cependant général. Le plus remarquable dans la comparaison des deux exemples ci-dessus, c'est que LN utilise la même activité gestuelle (mais derrière l'écran). Le geste ajouté ci-dessous est une action bimanuelle avec laquelle LN est en train de verser quelque chose vers le bas (c'est-à-dire, dans la casserole).

avec HI
 maintenant les hariCOTS:::: donc::::: {je:::: METS:::/} + [...] ce méLAN:ge/ + de toMA:te et d'huile d'oLI:ve et de basiLI:c + {DANS les haricots\} +++ [...] qui sont TOUjours + en TRAIN de CUIre\

avec HI (reprise)
 et QUAND je finis/ de préparer le méLANge/ ++ [...] {je METS:::/} ++ {DANS les hariCOTS/} ++ {DANS L'EAU/ + [...] des hariCOTS/} [...] qui sont\ ++ TOUJOURS\ TOUJOURS sur-\ ++ en train de cuire↓

avec MR
 et::::: ++ QUAND j'ai: fi:NI:::/ ++ [...] je:::: {VER:::se ++ tu compRENDS} ++ [...] {je: METS:::/} ++ le méLAN:::ge ++ [...] **DANS:: la SOU::pe + de haricots** ++ [...] je mets dedans\

Dans ces trois cas, le geste et le discours se synchronisent. De plus, cette gesticulation implique non seulement l'action concernée, mais aussi l'objet à verser et sa destination. Autrement dit, elle préfigure deux agents du verbe (complément d'objet et complément de lieu). Il est possible qu'une telle projection prépare un contexte visuel dont tient compte l'interlocuteur pour l'interprétation du discours. Notons que, en face à face, l'explication de la même recette avec la même action a été suivie avec succès par un autre apprenant débutant.

La métaphore est une figure de rhétorique qui s'appuie sur la comparaison entre deux entités. Bien que la métaphore ait été considérée comme source de trouble dans les exemples ci-dessus, il n'est pas nécessaire de mettre en cause la comparaison elle-même en tant que technique communicative. La circonlocution est souvent utilisée à des fins explicatives au cours de la négociation du sens, comme dans la conversation avec un apprenant avancé (YO) sous la condition d'écran. Le terme reformulé est *peau*.

(7) YO-R2b-ratatouille

067LN : [...] les auberGI:nes\ + ça c'est plus long à: à CUI::re/ +
 068YO : ah ha::/
 069LN : on peut laisser la peau +++
 070YO : ((bas)) oui ++
 071LN : tu SAIS/ ++
 072YO : laisser:: la/ ++
 073LN : la peau ++
 074YO : [po] ++
 075LN : tu compRENDS/ +
 076YO : [PO]\ c'est:: casse [RO:le\]
 077LN : [la PEAU/] c'est::\ ce qui est autour de:: ++ du
 légume\ ++ COMMENT je peux expliquer:\ + tu SAIS les pommes de
 TER:re/ ++
 078YO : oui/ ++
 079LN : les pommes de TER:re/ pour les manger/ + il faut: ENlever la peau\ ++
 080YO : ah::: ++ d'ac:co::rd\ ++

La circonlocution que LN utilise (*ce qui autour de...*) implique une certaine imagerie et LN gesticule, derrière l'écran, comme s'il montrait l'illustration de l'expression *peau*. L'indisponibilité visuelle, telle que LN l'a vécue ici, a donné lieu à une augmentation de la fréquence de la circonlocution dans toutes les interactions. Si l'interactant pouvait voir l'autre, il aurait pu utiliser sans doute d'autres moyens de négociation, comme le geste par exemple.

Dans cet exemple-ci, LN ménage le temps de latence comme s'il laissait au partenaire non natif (HI) le temps de conceptualiser le référent problématique *pulpe* pour ce dernier :

(8) HirR2b-soupe au pistou

209LN : j'ajoute::: + la:: ++ PULPE + de toma:te/ ++
 210HI : pu- ++
 211LN : n la tomate\ mais pas la peau\ ++
 212HI : [po/] ++
 213LN : tu comprends/ ++ dans la toMA:te
 214HI : [oui]
 215LN : [il y] A:: la la:: ++ la pulpe\ ++ c'est:: comment je peux dire\ + la
 toMAte ++
 216HI : [n/]
 217LN : [SANS] +++
 218HI : ou[i]
 219LN : [la] peau\ ++
 220HI : AH:: ah + okay\ d'acc [ord\]
 221LN : [N:/] ++ et j'écrase avec ÇA/ ++
 222HI : hum

La circonlocution exploite l'aspect visuel des connaissances du monde des locuteurs autrement que d'autres types de métaphore. La circonlocution est une description littérale et directe du visuel dans le discours alors que la métaphore est une expression suggestive et indirecte du visuel. Ainsi, le discours implique l'univers visuel. Il n'est pas étonnant de constater que l'expression métaphorique, qui entretient un lien dissimulé avec l'indice visuel dans le discours, pose problème aux non-natifs pour la compréhension lorsqu'ils ne disposent pas d'indice visuel.

4.4. Syntaxe et indice visuel

De façon relative, de nombreux cas problématiques sont liés à la méconnaissance lexicale du partenaire non natif. Elle cause normalement, si le trouble a été détecté, la négociation du sens comme dans l'exemple précédent. À comparer différents exemples, les causes du trouble ne semblent pas identiques. Il y a ignorance totale de la forme linguistique du référent ; dans ce cas, le moyen visuel est souvent une base sur laquelle la négociation s'appuie, mais l'incertitude peut persister chez les partenaires. L'inférence du sens d'une expression inconnue par le seul indice visuel au cours du discours est rare, du moins difficile à mettre en évidence pour les analystes. Ce cas de figure se confond aussi avec le problème de vocabulaire passif. Par exemple, les apprenants avancés se montrent aptes à saisir certaines expressions comme *faire revenir du riz, lier la sauce en face à face*, mais lors de leur production, ils utilisent *cuire du riz, mélanger la sauce*. Cela veut dire que même si l'expression verbale n'était pas dans leur vocabulaire actif, ils la traitent. L'information visuelle a été peut-être efficace dans ces cas-là, mais une telle saisie peut provenir d'une inférence discursive aussi.

Quelques exemples fournissent des indices de l'implication visuelle dans la compréhension au niveau syntaxique. Cela peut apparaître avec les spécificatifs ou des expressions similaires. Les spécificatifs sont une partie d'un syntagme nominal. Certains apportent seulement une information unitaire (par exemple, *gousses d'ail*), d'autres une information formelle ou quantitative (par exemple, *morceau de morue, tranches de porc, un verre d'huile d'olive, un grand pot de pistou*) d'autres encore une dénomination spécifique d'une partie de l'entité (par exemple, *pulpe de tomate, germe d'ail, filets d'anchois, feuilles de basilic*). En général, les non-natifs sont peu accoutumés aux spécificatifs, même si la langue française n'en possède pas beaucoup. Leur compréhension peut s'appuyer largement sur l'indice visuel correspondant si l'activité gestuelle du partenaire se montre congruente avec l'expression verbale concernée. Le fait de savoir qu'une chaîne sonore recouvre seulement un objet facilite probablement d'inférer de quoi il s'agit. Sans ce savoir, l'interprétation d'un supécificatif inconnu reviendrait à la problématique de la compréhension du mot inconnu. Une telle suggestion métalinguistique à travers l'indice visuel peut promouvoir l'inférence discursive lors de la compréhension.

L'exemple suivant (en face à face) illustre ce type d'implication gestuelle dans la compréhension. LN est en train d'expliquer à un apprenant débutant comment utiliser l'ail pour la soupe au pistou. Le spécificatif en question est *gousse*.

```
(9) SA-R1b-soupe au pistou
105LN : on aJOU:te::: une::1-1{:: + TÊte + d'ail\} +++
106SA : [ (xxxx) ]
107LN : [ 2-1{de l'a}il\ ++}
108SA : ah:: l'ail/ ++ [ oui/ ]
109LN : [ 3-1{mais } GROS ++}
110SA : hum hum/ ++
111LN : 4-1{entier\} ++
112SA : hum hum/ ++
113LN : 5-1{donc:: un:: deux:: trois::}
114SA : [ oui ha-ha-ha ]
115LN : [ 6-1{peut être:: six} ] ++ six GOUSses/ +
116SA : hum hum/ ++
117LN : 7-1{on éCRA:se l'ail+ [ +++ ]
118SA : [ hum hum/ ]
119LN : comme ça/[ ++} ]
120SA : [ oui ]
```

Avec les gestes iconiques 1 et 2, LN présente la configuration générale de l'ail. Les gestes 3 et 4 renvoient respectivement à l'adjectif qualificatif concerné. Le geste 5 ressemble à l'action de séparer une à une les gousses. Tous ces gestes sont donc iconiques.

Toutefois, comme l'exemple 10 (en face à face) le montre, le terme *gousse* n'a pas été repris par SA dans sa ré-explication. Dans cet extrait, le moment délicat est le regard de SA dirigé vers LN en 272. La signification de son regard est en fait ambigu. On pourrait le considérer comme demande de confirmation à propos du nombre de gousses d'ail, mais à regarder et écouter le corpus à plusieurs reprises, il semble que SA exprime indirectement le doute sur le

terme *gousse*. Et à la suite de la réaction de LN en 273, différente de son attente, SA abandonne son but initial en s'ajustant au déroulement de l'échange (évitement). En conséquence, la saisie du terme par SA dans l'exemple 9 a été probablement suspendue par SA qui se contente d'induire le sens à l'aide de l'indice visuel. L'important pour la présente discussion, c'est que SA ait compris du moins grâce à l'information visuelle que le terme en question ferait partie du syntagme nominal¹.

(10) SA-R1b-soupe au pistou
 270SA : on [KRA::z/] ah bon\ ++ on [KRA(h)::z/] +++ uh::: de: [LEJ/] +
 271LN : de l'ail/ +
 272SA : de [laj] +++ ah:: peut être + quaTRE/ (regard vers LN) ++
 273LN : (hochement de tête) ++ CINQ\ +
 274SA : qua[tre cinq]
 275LN : [cinq gousses\] [oui]
 276SA : [oui] b(h)on ho-ho-ho (H-) /uh::: ++ /n::: ++
 après/ ++

Un problème similaire a été observé dans l'exemple 1, mais différent puisqu'il inclut une reformulation verbale (*la tête d'ail, les gousses d'ail, les morceaux d'ail*). L'exemple suivant (en face à face) montre comment YO réagit pour sa part lors de la ré-explication.

(11) YO-R1b-aioli
 140YO : et::: + un:: morceau de:: ++ [LAjy]\ ++
 141LN : oui ++
 142YO : [et:::]
 143LN : [presque]:: une °tête\ ++
 144YO : une tête
 145LN : n::\ ++

On constate une divergence entre le schéma que LN suit dans son discours de l'exemple 1 et le schéma que YO utilise à son tour :

LN (ex. 1) : une tête d'ail = les gousses = les morceaux
 YO (ex. 11) : une tête d'ail = un morceau (= une gousse ?)

Il semble que dans cet échange, il y a quelques ambiguïtés qui s'articulent autour du terme *gousse* : 1) la réponse de LN en 141 par *oui* est-elle d'ordre informationnel ou phatique ? ; 2) la réparation de LN en 143 est-elle une reformulation ou une correction ? ; 3) lors de la reprise de YO en 144, comment a-t-il compris *gousse d'ail* par rapport à *tête* et *morceau* ? En fait, le problème réside dans la compréhension du déterminant (ici l'article), puisque la fausse interprétation de YO s'expliquerait si on supposait que YO n'a pas entendu la pluralité de l'article *les* en se focalisant trop sur l'entité *ail* dans l'activité gestuelle de LN de l'exemple 1. Il y a, certes, peu de différence entre *gousse* ou *morceau d'ail* et *tête d'ail* au niveau de l'expression gestuelle, et la pluralité est difficile à exprimer gestuellement en même temps que l'objet concerné. Probablement, le statut morphosyntaxique de la pluralité exprimée par le geste (ici, le non-marquage comme dans certaines langues) est différent de celui par le langage oral². Ce problème évoque l'inversion syntaxique des gestes 3 et 4 de l'exemple 1. La compensation visuelle a donc une limite dans la mesure où l'expressivité du système gestuel (ici pour le syntagme nominal) diffère de celle du langage verbal.

Notons toutefois que le corpus implique de nombreux exemples de la contribution visuelle pour indication du statut de données verbales. Par exemple, avec l'écran, un apprenant intermédiaire a pris un verbe (*tu fais bouillir les haricots*) pour un substantif (*[dubui], c'est quel légume ?*) ; également un autre apprenant débutant a pris le même terme pour un autre substantif *bouillon*, qui existe dans sa langue maternelle comme mot d'emprunt (*buiyon*).

¹ « Some information coming from non-linguistic input may be supplementary rather than complementary. [...] The gesture will provide about the nature of the NP and create the possibility of learning of a new word (for example). [...] We also have to assume that the learner registered the fact that there was an NP that needed to be account for » (Shawood Smith, 1986, pp. 254, note 14).

² Dans le langage des signes, on répète le même geste pour exprimer la pluralité.

Tous ces problèmes, ayant pour cause une confusion de fonctions métalinguistiques des mots, peuvent être compensés éventuellement par l'imagerie que les gestes laissent transparaître.

Outre une simple confusion métalinguistique entre objet (nom) et action (prédicat), il y a des cas encore plus subtils. L'exemple suivant (écran) montre la confusion de l'apprenant intermédiaire MR sur la fonction du modificateur (représenté en gras).

(12) MR1-R2b-soupe au pistou
 067LN : CENT cinquante grammes de haricots ROU::ges/ ++
 068MR : [oui\]
 069LN : [enSUIte] + CENT cinquante grammes de + ha:ri:cots ve::rts\ +++
 070MR : ouI:: ++
 071LN : **TRÈS TRÈS fins** + tu SAIS:/ +++
 072MR : ah + oh + ah + (attends\
 073LN : les haricots verts\ c'est **TRÈS fin** + **FIne** + **TRÈS TRÈS fin** ++
 074MR : ah:: ce n'est PAS:: normal [aliko ve::\] + ((rires)) PHU-hu-hu-hu ++
 075LN : euh::: + SI: SI: + quand quand tu MANGes:: des + haricots ve:rts ++
 076MR : ouI: ++
 077LN : par exemple avec de la viande\ ++
 078MR : oui\ ++
 079LN : c'est les haricots **TRÈS:: MIN::ces** ++
 080MR : ah c'est **le COUpé** ++
 081LN : oui ++ je vais les couper\ ++
 082MR : ah oui D'ACcord D'ACco::rd\ ++
 083LN : tu comPRENDS ++
 084MR : oui j'ai compris\ +++

Pour le segment *haricots verts*, LN, anticipant son aspect problématique pour MR, initie la négociation du sens en 073. L'explication donnée par LN consiste à décrire la forme générale du légume avec l'adjectif *fin* en 071 et 073, dont la fonction est apposition, et il en va de même pour la reformulation *très minces* en 079. Le corpus vidéo montre en effet que les gestes de LN indiquent cette relation. Or, connaissant le mot et n'ayant pas accès à l'activité gestuelle de LN, MR suppose d'après ces deux adjectifs que LN est en train de lui fournir une information supplémentaire, et les interprète autrement. Pour MR, la fonction de l'adjectif est déterminative. La divergence des interlocuteurs vient alors de la fonction du qualificatif :

LN :	haricots, qui sont fins (ou dont la forme est fine)	→ apposition
MR :	haricots qui sont coupés fin	→ fonction déterminative

Dans cet exemple, supposons-le, l'indice visuel aurait désambiguïté une telle dualité fonctionnelle de l'adjectif. À l'écrit, ce type de divergence est différencié, comme ci-dessus, par la ponctuation de la proposition relative. Bien évidemment, l'équivalent gestuel de la ponctuation n'existe pas. Mais, en combinaison avec le discours, l'indice visuel peut quelquefois remplir une fonction qu'assume le système de ponctuation du discours écrit. De même, du côté de LN, l'évaluation de l'état de compréhension de son partenaire non natif aurait pu être plus aisée s'il pouvait observer le comportement de MR.

4.5. Structure discursive et domaine extradiscursif

La référence de l'information visuelle se limite apparemment à des choses concrètes comme des entités, des actions ou des lieux, mais l'examen approfondi de l'activité gestuelle suggère qu'au-delà du rapport signifiant/signifié, les signes entretiennent des relations invisibles *dans le discours*. Cette logique s'applique non seulement aux signes non linguistiques, mais aussi aux signes linguistiques. Le propre visuel du geste peut donner à l'interlocuteur, mais aussi aux analystes, un avantage considérable pour appréhender certaines des fonctions discursives non marquées. L'exemple 1 a laissé entendre que cette implication visuelle pouvait s'étendre sur le rapport entre les énoncés, c'est-à-dire sur la connexion discursive (telle que l'équivalence, la causalité, la cohérence). Étant donné que les interlocuteurs non natifs montrent souvent un problème lexical, le lien d'équivalence serait le cas le plus fréquent, puisqu'elle concerne la reformulation. Toutefois, il n'y a pratiquement pas d'exemples qui indiquent que la non-visibilité cause la compréhension de ce genre chez l'interlocuteur.

Deux arguments peuvent être avancés pour cela. D'une part, le locuteur ajuste le discours à la situation écologique pour que la non-visibilité ne pose pas problème à la compréhension de l'interlocuteur. Certes, n'ont pas été trouvés (du moins pour l'instant) une expression gestuelle de causalité, prétendue dans l'exemple 1, ou des formes apparentées dans le discours de LN en situation d'écran. Cela n'est pas contradictoire avec le fait que le locuteur a utilisé plus de stratégies préventives, de répétitions et de demandes de confirmation en privation visuelle qu'en face à face. C'est dire que l'intention communicative du locuteur dépend, dans une certaine mesure, du contexte d'interaction qui influence les modalités d'expression.

D'autre part, on peut se dire que l'indice visuel n'est pas indispensable pour la compréhension de la relation discursive, qui peut être implicite et parfois non marquée même en face à face. Il se peut que le traitement cognitif des marquages discursifs chez l'interlocuteur soit différent du traitement des autres types d'information en face à face. Les connecteurs sont multifonctionnels en nécessitant une interprétation complexe en relation avec d'autres informations concrètes dans le contexte discursif. Comme les relations discursives qui apparaissent dans le corpus sont des connecteurs fréquents (*après, d'abord, et, parce que, etc.*), qui ne posent pas de problème immédiat, l'information verbale a peut-être suffi pour que l'interlocuteur opère une interprétation de ce genre. Le troisième argument possible est que les interlocuteurs ont adopté la stratégie d'attente-écoute (de type *wait and see*) pour ne pas réagir immédiatement à la non-compréhension de la relation discursive, mais ce cas de figure s'avérerait inanalysable par la présente étude.

Cependant, la compréhension de l'anaphore semble affectée dans une certaine mesure par la privation visuelle. L'analyse de l'exemple 4 (rappelons qu'il s'agit de l'utilisation d'un bol rempli de vin et de la préparation de la viande pour la daube, avec écran) a montré que le non-marquage de localisation avait gêné la compréhension du débutant. Le corpus implique quelques exemples similaires, notamment lorsque l'interlocuteur est débutant. L'analyse de ces exemples révèle une similitude de la source du trouble, venant de la manutention de deux ou plusieurs récipients avec divers ingrédients. En plus, la relation anaphorique est implicite à cause de l'absence du marquage anaphorique dû à la focalisation discursive. Cette relation est, en fait, exprimée soit par la spatialisation topologique, soit par l'« iconisation » des ingrédients par le biais visuel, dont l'interlocuteur n'a pas disposé sous condition d'écran. Ainsi, la visualisation spatiale et iconique peut rendre optimale l'interprétation de l'anaphore \emptyset ou du faux usage d'un pronom.

L'indice visuel semble contribuer à la compréhension de la relation spatiale et iconique plus que la relation logique. L'état actuel de nos analyses ne prouve que la réalité d'une contribution indirecte de l'information visuelle à la fonction métadiscursive. Peut-être parce que le corpus s'appuie sur l'échange informationnel, d'autres types de corpus (par exemple, l'interaction didactique ou la conversation spontanée, qui ne sont pas analysées pour la présente étude) montreraient autre chose, mais pour le moment, restons sur l'hypothèse différentielle de l'apport visuel. À ce propos, Clark (1997, pp. 591) parle d'une variation pragmatique de l'indice visuel dans le discours, alors que pour Kendon (1994, pp. 192-3), une telle variation est sujette au contexte où l'interaction est structurée à un moment donné.

De même, cette variabilité est aussi attestée pour la perception multimodale des phonèmes (l'effet de MacGurk), qui a été abordée plus haut. Par exemple, Mills (1987) rapporte que les enfants malvoyants de naissance font une acquisition problématique de certains phonèmes d'« articulation visible » (tels que /b/, /p/, /m/, /f/, /v/) par rapport aux enfants non handicapés. Autrement dit, la limitation de l'input visuel chez les enfants malvoyants semble causer un retard dans l'acquisition de certains phonèmes à mouvement articulatoire marqué. Dans le contexte de la présente étude, il est prudent de dire que l'apport de l'information visuelle à la compréhension est limité aux parties « spatialisables » du discours.

La structure du discours oral spontané est typiquement opaque et incomplète du point de vue linguistique. Pour une compréhension discursive optimale, tout interlocuteur doit résoudre au moins une partie de ces points implicites ou illogiques du discours. Ce processus requiert à la fois un traitement, au niveau contextuel, de bas niveau (*bottom-up*) et un traitement, au niveau plus textuel, de l'interprétation multiple du sens (Kasper, 1984). Pour les interlocuteurs non natifs dépendant du traitement de bas niveau (voir *supra*), l'information visuelle est probablement un élément important pour l'interprétation du discours.

La prépondérance du traitement de bas niveau chez les apprenants ne signifie pas qu'il n'existe pas de traitement de haut niveau (le savoir sur les structures et fonction discursives, l'information du contexte social de l'interaction et la connaissance du monde). Notamment quand l'interaction est interculturelle, c'est le cas où, par exemple, les termes donnent certaines présuppositions culturellement définies. Ici, nous nous contentons de citer quelques exemples. Nous avons parlé plus haut du problème que posent les termes renvoyant à des actions spécifiquement liées à la cuisine (faire revenir/bouillir/cuire, réserver, écraser, mélanger, lier la sauce, saisir la viande) lesquels ont posé problème dans le cas de privation d'information visuelle, ce qui était peu le cas en face à face. Comme autre exemple, un apprenant intermédiaire n'a pas saisi, lors de l'explication de la recette de la daube (écran), l'expression *enlever les herbes de Provence*, ce qui a donné lieu à une négociation inutilement longue. L'analyse montre que pour cet apprenant, les herbes de Provence sont une épice à ajouter, puisque dotées d'une valeur symbolique dans la cuisine provençale, et non pas quelque chose à enlever. Autrement dit, sa croyance personnelle en ce point aurait dû le rendre « sourd » à l'expression concernée. S'il pouvait voir le geste ENLEVER qu'accompagne le discours, ses idées reçues n'auraient pas été si persistantes. De même, la divergence culturelle sur la façon de préparer l'omelette dans les dyades franco-japonaise et hispano-japonaise a causé des problèmes d'intercompréhension avec l'écran, puisque la différence résidait dans l'ordre et le type d'ingrédients associés. En effet, la relation spatio-topologique et temporelle des agents d'une recette est sous-tendue par un schéma culturel de l'événement culinaire. Évidemment, la présupposition n'est pas seulement d'ordre culturel, mais elle peut être fondée — comme c'est souvent le cas du malentendu entre natifs — sur une croyance personnelle ou situationnelle. Mais lorsque l'interlocuteur est un sujet non natif, le socioculturel est un domaine très sensible. Dans tous les cas, l'indice visuel peut alors instantanément modifier l'a priori avec lequel l'interlocuteur envisage le discours de l'autre.

5. Remarques conclusives

Que faut-il retenir, en conclusion, de l'analyse de l'indice visuel par rapport à la compréhension du discours en L2 par les non-natifs ? En effet, il y a des exemples qui laissent penser que les non-natifs tirent profit d'un certain nombre d'informations visuelles qui accompagnent le discours du locuteur natif. Dépourvus de cette information, ils se trouvent en difficulté de compréhension. Cela indique qu'ils s'y réfèrent pour l'interprétation du discours et la procédure de construction du sens, même si l'implication visuelle est limitée à quelques aspects iconiques et spatio-topologiques du discours. Comme les apprenants sont des sujets qui dépendent souvent du traitement de bas niveau du contexte, il semble qu'ils utilisent les indices visuels comme stratégie d'interprétation. Notamment les apprenants moins avancés souffrent de l'absence de ces mêmes indices, qui sont utiles pour la compréhension de ce qui n'est pas verbalement exprimé comme l'ellipse, la métaphore, la fonction métalinguistique, la délimitation paradigmatique ou la relation syntagmatique des entités linguistiques. De même, il arrive que l'indice visuel permette à l'interlocuteur de modifier ses présupposés, culturels ou non, et d'ajuster la procédure d'interprétation à l'intention communicative du locuteur. Du fait que l'interlocuteur non natif n'est pas un sujet totalement familier à la convention de la langue, du discours et de la rhétorique et au savoir socioculturel de la communauté d'accueil, l'indice visuel peut constituer pour lui un recours efficace. Enfin, la vertu du visuel n'est pas seulement compensatoire, il peut être exploré dans le discours comme la circonlocution ou la métaphore.

La présente étude soulève un certain nombre de points à interroger sur les théories et analyses de la linguistique acquisitionnelle et les aspects didactiques de l'apprentissage d'une langue étrangère. Si l'information visuelle est une composante constitutive du discours oral dévouée à la compréhension, on peut se demander dans quelle mesure l'appauvrissement de l'input linguistique renvoie à l'innéisme de la grammaire linguistique. En fait, la dimension iconico-topologique qu'implique le « discours visuel » pourrait être une source importante de l'acquisition linguistique, comme les travaux sur l'acquisition du langage chez les enfants (par exemple, Marcos, 1998) semblent l'indiquer¹. Évidemment, reste ouverte la question de savoir si la compréhension discursive chez les sujets adultes non natifs fonctionne comme chez les

¹ Pourtant des enfants sourds de naissance nés des parents entendants, qui ont pour seule source un « pidgin gestuel » des parents, acquièrent un système « grammatical » du langage des signes malgré l'environnement pauvre d'input visuel « grammatical » (Goldin-Meadow & Mylander, 1990).

enfants tout à fait différemment, ou encore variablement selon la culture d'origine. Dans le processus d'appropriation d'une L2, il conviendrait alors de dégager l'importance de la gestualité. D'autre part, certains corpus oraux qu'utilisent les chercheurs pour la description du phénomène de l'appropriation, notamment pour l'étude des interactions en situation de contact, pourront être remis en question, puisque les non-natifs semblent être des sujets interprétants qui utilisent le visuel à leur compte sans que le partenaire de conversation ou l'analyste le sache. Si l'interaction sociale ou la production du discours est une source de l'appropriation discursive et pragmatique, l'activité gestuelle et d'autres indices visuels en sont consubstantiels.

Pour ce qui est de la pratique pédagogique des langues étrangères ou secondes, la présente étude laisse entrevoir que l'information visuelle contribue à la compréhension du discours par les apprenants, mais cette hypothèse demande à être envisagée avec précaution. Certes, les indices visuels dans l'interaction en face à face assureraient mieux l'intercompréhension entre enseignant et apprenant, mais il apparaît que le recours au geste souvent s'accompagne de l'ellipse sur le plan formel du discours. Un tel discours ne pourrait sans doute promouvoir l'aspect formel des acquisitions grammaticales et discursives. Si l'enseignant veut privilégier l'acquisition plutôt formelle, l'activité gestuelle ne devrait pas empêcher la production d'énoncés complets du point de vue grammatical. Néanmoins, on ne pourrait pas toujours réfuter la réduction formelle du discours pédagogique si on voulait promouvoir l'entraînement de l'aptitude de compréhension discursive par inférence chez les apprenants. En effet, le discours ordinaire des locuteurs natifs est elliptique, opaque, métaphorique, implicite, même agrammatical par rapport à la norme pédagogique. Le surétayage visuel par l'activité gestuelle de l'enseignant pourrait alors empêcher les apprenants de se familiariser à ce type de discours. L'interprétation du discours est une activité hautement complexe qui repose tantôt sur l'indice prosodique, tantôt sur l'indice mimogestuel, tantôt sur l'indice linguistique (Kida, 2001). Le bon déroulement de l'interaction dépend de l'aptitude de l'interlocuteur à utiliser différentes sources dans le flot d'informations communicatives. Dans ce sens, enseigner le geste aux apprenants ne serait pas une solution pour une compétence interprétative du discours¹. Jadis Corder (1975, pp. 102) notait : « Efficient interpretative behaviour is [...] behaviour which holds hypotheses *lightly* and is even prepared to entertain for as long as possible potentially conflicting hypotheses about meaning, that is, to accept ambiguity, contradiction and paradox, i.e. to keep our interpretative options open (sic) ». Il faudra alors que les apprenants sachent comment utiliser une multiplicité d'indices — qui sont parfois contradictoires — pour une interprétation appropriée du discours.

¹ Ce qui ne veut pas dire que l'enseignement du geste est inutile : comme le geste est constitutif du discours, il peut être traité en tant que culture discursive faisant l'objet d'un enseignement du savoir socio-culturel .

Références bibliographiques

- Argyle (M.), Ingham (R.), Alkema (F.) & McCallin (M.). 1973. « The different functions of gaze ». in : *Semiotica*. 7, pp. 19-32.
- Aston (G.). 1986. « Trouble-shooting in interaction with learners: The more the merrier ? ». in : *Applied linguistics*. 7, 2, pp. 128-143.
- Bavelas (J.-B.) & Coates (L.). 1992. « How do we account for the mindfulness of face-to-face dialogue ? ». in : *Communication monographs*. 59, 3, pp. 301-305.
- Beatie (G.-W.) & Baranrd (P.-J.). 1979. « The temporal structure of natural telephone conversations (directory enquiry calls) ». in : *Linguistics*. 17, pp. 213-229.
- Bremer (K.). 1996. « Causes of understanding problems ». in : Bremer (K.) & al. 1996, pp. 37-64.
- Bremer (K.), Roberts (C.), Vasseur (M.-T.), Simonot (M.) & Broeder (P.). 1996. *Achieving understanding: Discourse in intercultural encounters*. London : Longman.
- Brown (G.). 1998. « Context creation in discourse understanding ». in : Malmkjær (K.) & Williams (J.), (eds). *Context in language learning and language understanding*. Cambridge, UK : Cambridge University Press, pp. 171-192.
- Bühler (K.). 1965. *Sprachtheorie : Die Darstellungsfunktion der Sprache*. Stuttgart : Gustav Fischer Verlag [ed. originale en 1934].
- Carrell (P.-L.). 1984. « Evidence of a formal schema in second language comprehension ». in : *Language Learning*. 34, 1, pp. 87-112.
- Carrell (P.-L.). 1990. « Culture et contexte dans la lecture en langue étrangère : rôle des schémas de contenu et des schémas formels ». in : Gaonac'h (D.), (ed). *Acquisition et utilisation d'une langue étrangère. L'approche cognitive*. Paris : Hachette, pp. 16-29.
- Clark (H.-H.). 1997. « Dogmas of undersatnding ». in : *Discourse processes*. 23, pp. 567-598.
- Cohen (A.-A.) & Harrison (R.-P.). 1973. « Intentionality in the use of hand illustrators in face-to-face communication situations ». in : *Journal of personality and social psychology*. 28, 2, pp. 276-279.
- Cohen (A.-A.). 1977. « The communicative function of hand gesture. Journal of hand gestures ». in : *Journal of communications*. 27, 4, pp. 54-63.
- Cook (M.) & Lalljee (M.). 1972. « Verbal substitutes for visual signals in interaction ». in : *Semiotica*. 6, pp. 212-221.
- Corder (S.-P.). 1975. « Interpretative procedures: Seeing, reading and understanding discourse ». in : Corder (S.-P.) & Roulet (E.), (eds). *Some implications of linguistic theory for applied linguistics*. Brussels : AIMAV ; Paris : Didier, pp. 95-103.
- De Pietro (J.-F.), Matthey (M.) & Py (B.). 1989. « Acquisition et contrat didactique : les séquences potentiellement acquisitionnelles dans la communication exolingue ». in : Weil (D.) & Fugier (H.), (eds). *Actes du 3e colloque régional de linguistique*. Strasbourg : Université des Sciences Humaines et Université de Louis Pasteur, pp. 99-124.
- Efron (D.). 1972. *Gesture, race and culture*. The Hague : Mouton [ed. originale en 1941].
- Færch (C.) & Kasper (G.). 1986. « The role of comprehension in second language learning ». in : *Applied linguistics*, 7 (3), pp. 257-271.
- Færch (C.). 1985. « Meta talk in LF classroom discourse ». in : *Studies in second language acquisition*. 7, 2, pp. 184-199.
- Faraco (M.) & Kida (T.). 1999. « Acquisition de L2 et séquences d'apprentissage : degré de stabilité interactive ». in : *International review of applied linguistics for language teaching (IRAL)*. 38, 3, pp. 215-230.
- Faraco (M.). 2001. « Que peut apporter l'étude du non-verbal à la didactique de L2 ? ». in : Cavé (C.), Guaitella (I.) & Santi (S.), (eds). *Oralité et gestualité. Interaction et comportements multimodaux dans la communication*. Paris : L'Harmattan, pp. 531-536.
- Ferguson (C.-A.). 1971. « Absence of copula and the notion of simplicity : a study of normal speech, baby talk, foreigner talk ». in : Hymes (D.), (ed). *Pidginization and creolization of languages*. New York : Cambridge University Press, pp. 141-150.
- Gass (S.-M.) & Varonis (E.-M.). 1991. « Miscommunication in nonnative speaker discourse ». in : Coupland (N.), Giles (H.) & Wieman (J.-M.), (eds). « *Miscommunication* » and *problematic talk*. New Burypark (CA), London, New Delhi : Sage, pp. 121-145.
- Giacomi (A.) & De Hérédia (C.). 1986. « Réussites et échecs dans la communication linguistique entre locuteurs français et locuteurs immigrés ». in : *Langages*. 84, pp. 9-24.

- Giacomi (A.), Houdaifa (E.-T.) & Vion (R.). 1984. « Malentendu et/ou incompréhension dans la communication interculturelle : « à bon entendeur, salut ! » ». in : Noyau (C.) & Porquier (R.), (eds). *Communiquer dans la langue de l'autre*. Paris : Presses universitaires de Vincennes.
- Goldin-Meadow (S.) & Mylander (C.). 1990. « Beyond the input: the child's role in the acquisition of language ». in : *Language*. 66, 2, pp. 323-355.
- Goodwin (C.). 1986. « Gestures as a resource for the organization of mutual orientation ». in : *Semiotica*. 62, 1-2, pp. 29-49.
- Goodwin (M.-H.). 1980. « Processes of mutual monitoring implicated in the production of description sequences ». in : *Sociological inquiry*. 50, 3-4, pp. 303-317.
- Graham (J.-A.) & Argyle (M.). 1975. « A cross-cultural study of the communication of extra-verbal meaning by gestures ». *International review of psychology*. 10, 1, pp. 57-67.
- Gülich (E.). 1986. « L'organisation conversationnelle des énoncés inachevés et de leur achèvement interactif en situation de contact ». in : *Documentation et recherche en linguistique allemande et française de Vincennes (DRLAV)*. 34-35, pp. 161-182.
- Gullberg (M.). 1998. *Gestures as a communication strategy in second language discourse: A study of learners of French and Swedish*. Lund, Sweden : Lund University Press.
- Gumperz (J.) & Roberts (C.). 1991. « Understanding in intercultural encounters ». in : Blommaert (J.) & Verschuren (J.), (eds). *The pragmatics of international and intercultural communication*. Amsterdam, Philadelphia : John Benjamins, pp. 51-90.
- Harris (R.) & Taylor (T.-J.). 1997. *Landmarks in linguistic thought I: The western tradition from Socrates to Saussure*. London : Routledge [2nd ed.].
- Heath (C.). 1992. « Gesture's discreet tasks : Multiple relevancies in visual conduct and in the contextualisation of language ». in : Auer (P.) & Di Luzio (A.), (eds). *The contextualization of language*. Amsterdam, Philadelphia : John Benjamins, pp. 101-128.
- Kasper (G.). 1984. « Pragmatic comprehension in learner-native speaker discourse ». in : *Language learning*. 34, 4, pp. 1-20.
- Keller-Cohen (D.). 1979. « Repetition in the nonnative acquisition of discourse ». in : Freedle (R.-O.), (ed). *New directions in discourse processing*. Norwood, NJ : Ablex, pp. 245-272.
- Kellerman (E.) & Shawood Smith (M.). 1981. « Iconicity and the case of the missing lexical item ». Workshop discussion paper given at the *Sixth International Congress on Applied Linguistics*, Lund, August.
- Kellerman (S.). 1990. « Lip service: the contribution of the visual modality to speech perception and its relevance to the teaching and testing of foreign language listening comprehension ». in : *Applied linguistics*. 11, 3, pp. 272-280.
- Kellerman (S.). 1992. « « I see what you mean »: the role of kinesic behaviour in listening and implications for foreign and second language learning ». in : *Applied linguistics*. 13, 3, pp. 239-258.
- Kendon (A.). 1980. « Gesticulation and speech: Two aspects of the process of utterance ». in : Key (M.-R.), (ed). *The relationship of verbal and nonverbal communication*. The Hague, Paris, New York : Mouton, pp. 207-227.
- Kendon (A.). 1985. « Some uses of gestures ». in : Tannen (D.) & Saviile-Troike (M.) (eds.), *Perspectives on silence*. Norwood (NJ) : Ablex, pp. 215-234.
- Kendon (A.) 1994. « Do gestures communicate ? A review ». in : *Research on language and social interaction*. 27, 3, pp. 175-200.
- Kida (T.) & Faraco (M.). 2002. « Gesto y iniciación de la interacción didáctica en L2 ». in : *de-Signis*. 3, pp. 113-131.
- Kida (T.) & Faraco (M.). sous presse. « Gesture in second language discourse: metacommunication and perlocution ». in : Rector (M.), Poggi (I.) & Trigo (N.), (eds). *Gesture : meaning and use*. Porto, Portugal : Universidad Fenando Pessoa Press.
- Kida (T.). 2001. « À propos du dynamisme dans le système de prise de parole ». in : *Travaux Interdisciplinaires de Parole et langage d'Aix-en-Provence (TIPA)*. 19, pp. 73-89 [Laboratoire Parole et Langage, Université de Provence].
- Kida (T.). 2002. « Rôle cognitif et social des indices visibles dans la négociation du sens en situation de contact ». in : *Travaux interdisciplinaires de Parole et langage d'Aix-en-Provence (TIPA)*. 21, pp. 87-100.
- Krashen (S.-D.). 1982. *Principles and practice in second language acquisition*. Oxford, UK : Pergamon.

- Labov (W.). 1996. « Some notes on the role of misperception in language learning ». in : Bayley (R.) & Preston (D.-R.), (eds). *Second language acquisition and linguistic variation*. Amsterdam, Philadelphia : John Benjamins, pp. 245-252.
- Lakoff (G.). 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, BO : University of Chicago Press.
- Langaker (R.-W.) 1991. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin, New York : Mouton de Gruyter.
- Larsen-Freeman (D.) & Long (M.-H.). 1991. *An introduction to second language acquisition research*. London : Longman.
- Larsen-Freeman (D.). 1983. « Assessing global second language strategies ». in : Seliger (H.-B.) & Long (M.-H.), (eds). *Classroom-oriented research in second language acquisition*. Rowley, MA : Newbury House.
- Long (M.-H.). 1983. « Native speaker/non-native speaker conversation and the negotiation of comprehensible input ». in : *Applied linguistics*. 4, 2, pp. 126-112.
- Long (M.-H.) 1996. « The role of the linguistic environment in second language acquisition ». in : Ritchie (W.-C.) & Bhatia (T.-K.), (eds). *Handbook of second language acquisition*. San Diego : Academic Press, pp. 413-468.
- MacGurk (H.) & MacDonal, (J.). 1976. « Hearing lips and seeing voices ». in : *Science*. 218, pp. 746-8.
- Marcos (H.). 1998. *De la communication prélinguistique au langage : formes et fonctions*. Paris : L'Harmattan.
- McNeill (D.) & Levy (E.-T.). 1993. « Cohesion and gesture ». in : *Discourse processes*. 16, 4, pp. 363-386.
- McNeill (D.). 1985. « So you think gestures are nonverbal ? ». in : *Psychological review*. 92, 3, pp. 350-371.
- McNeill (D.), Cassel (J.) & McCullough (K.-E.). 1994. « Communicative effect of speech-mismatched gestures ». in : *Research on language and social interaction*. 27, 3, pp. 223-237.
- Mills (A.). 1987. « The development of phonology in the blind child ». in : Dodd (B.) & Campbell (R.), (eds). *Hearing by eyes : Psychology of lip-reading*. London : Laurence Erlbaum Associates.
- Perdue (C.), (ed). 1993. *Adult language acquisition : Cross-linguistic perspectives. Vol. I: Field methods*. Cambridge, UK : Cambridge University Press.
- Py (B.). 2000. « Le discours comme médiation. Exemples de l'apprentissage et des représentations sociales ». in : Bertoud (A.-C.) & Mondada (L.), (eds). *Modèles du discours en confrontation*. Berne : Peter Lang, pp. 117-130.
- Rimé (B.) & Schiaratura (L.). 1991. « Gestures and speech ». in : Feldman (R.-S.) & Rimé (B.), (eds). *Fundamentals of nonverbal behavior*. New York, Cambridge, UK : Cambridge University Press, pp. 239-281.
- Roberts (C.). 1996a. « A social perspective on understanding: some issues of theory and method ». in : Bremer (K.) & al. 1996, pp. 9-36.
- Roberts (C.). 1996b. « Taking stock: contexts and reflections ». in : Bremer (K.) & al. 1996, pp. 207-238.
- Rossi (M.). 2000. *L'intonation, système du français. Description et modélisation*. Gap : Ophrys.
- Roulet (E.). 2000. « Énoncé, tour de parole et projection discursive ». in : Berthoud (A.-C.) & Mandada (L.), (eds). *Modèles du discours en confrontation*. Berne : Peter Lang, pp. 5-22.
- Sapir (E.). 1967. « L'influence des modèles inconscients sur le comportement social ». in *Anthropologie*. Paris : Minuit, pp. 35-54 (tr. fr. par C. Baudelot et P. Clinquart ; original : « The unconscious patterning of behavior in society » in : Dummer (E.-S.), (ed). *The unconscious : A symposium*. New York : Knopf, pp. 114-142].
- Sharwood Smith (M.). 1986. « Comprehension versus acquisition: Two ways of processing input ». in : *Applied linguistics*. 7, 3, pp. 239-256.
- Streeck (J.). 1995. « On projection ». in : Goody (E.-N.), (ed). *Social intelligence and interaction. Expressions and implications of the social bias in human intelligence*. Cambridge, UK : Cambridge University Press, pp. 87-110.
- Taylor (J.-R.). 1993. « Some pedagogical implications of cognitive linguistics ». in : Geitman (R.-A.) & Rudzka-Ostyn (B.), (eds). *Conceptualisation and mental processing in language*. Berlin, New York : Mouton de Gruyter, pp. 201-222.

- Vasseur (M.-T.), Broeder (P.) & Roberts (C.). 1996. « Managing understanding from a minority perspective ». in : Bremer (K.) & al. 1996, pp. 65-108.
- Vion (R.). 1985. « Compréhension et comportement communicatif ». in : *G.R.A.L. Papiers de travail*, 2, pp. 67-95. Aix-en-Provence : Publications de l'Université de Provence [Acquisition du français par des travailleurs marocains].
- Voionmaa (K.). 1984. « Lexikal överföring och rationaletet ». Paper presented to the *4th Scandinavian Conference on Bilingualism*, Uppsala.
- Walker (M.) & Nazmi (M.-K.). 1979. « Communicating shapes by word and gestures ». in : *Australian journal of psychology*, 31, pp. 137-143.

Appendices

A- Conventions de transcription

++	: pause
-	: rupture au cours de la prononciation d'un mot
::	: allongement vocalique
/	: intonation montante
\	: intonation descendante (et/ou rupture neutre d'une courbe mélodique)
MOt	: emphase
(mot)	: mot inaudible inféré
[aaa]	: transcription phonématique simple (souvent indicative)
aaa	: alternance codique (langue maternelle ou autre)
((bas))	: commentaire du transcripteur
>aaa<	: débit rapide
a :a :a :	: débit ralenti
(Ç)	: claquement de langue
HA-HA-HA	: rires audibles
[aaa]/[aaa]	: chevauchement
{aaa}	: période de gesticulation (soulignement = stroke)
(Hv)/(Hh)	: hochement vertical/mouvement horizontal de tête

B- Types de focalisation par ellipse sous condition d'invisibilité mutuelle

<i>Pronom avec référence au geste :</i>	on fait comme ça ; on fait ça
<i>Article défini (anaphore) :</i>	(aubergine) tu peux laisser/on enlève la peau (de ø)
<i>Ellipse quotidienne (manger) :</i>	on mange (ø) avec des pains grillés
<i>Verbes typiques de la cuisine : (ellipse de genre discursif)</i>	on sort (ø) du feu ; on fait (ø) à la vapeur ; pour mettre au four (ø) ; on laisse (ø) /au four/au chaud/tranquille ; on met (ø) dans un bol en bois ; on fait revenir (ø) dans une poêle ; on fait bouillir (ø) dans une grande casserole ; tu laisses cuire (ø) /comme un gratin/longtemps/un petit peu/ ; tu mélanges (ø) ; on ajoute (ø) dans la purée
<i>Emphase-répétition :</i>	j'écrase-j'écrase-j'écrase (ø) complètement ; on écrase (ø) bien-bien-bien
<i>Métaphore :</i>	j'ouvre les (ø de) haricots et je prends des petits (ø de) haricots ; je mets cette préparation dans les haricots (ø, métonymie)
<i>Expansion :</i>	(les feuilles de basilic) et j'ajoute (ø) dans le bol avec l'ail, l'huile d'olive, la tomate et le basilic
<i>Contraction :</i>	tu coupes les petits morceaux de poisson (pour dire « le poisson en morceaux »)

NB. () est une information contextuelle, qui n'est pas forcément l'antécédent immédiat.