



**HAL**  
open science

## Transcrire le geste

Tsuyoshi Kida

► **To cite this version:**

Tsuyoshi Kida. Transcrire le geste. Travaux interdisciplinaires du Laboratoire Parole et Langage, 2006, 25, pp.127-150. hal-00314518

**HAL Id: hal-00314518**

**<https://hal.science/hal-00314518>**

Submitted on 27 Aug 2008

**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.

# TRANSCRIRE LE GESTE

Tsuyoshi Kida

*Ingres soll die Ruhe geordnet haben ;  
ich möchte über das Pathos hinaus  
die Bewegung ordnen.*  
Paul Klee, *Tagebuch* (1914)<sup>1</sup>

Cette contribution est le fruit d'une réflexion menée sur la méthodologie de la transcription de corpus audio et vidéo, réflexion axée plus précisément sur la capture de l'aspect mimo-gestuel, dans le cadre d'un travail sur l'appropriation du geste par les apprenants japonophones de français (Kida, 2005). Cette transcription du geste, perfectible certes, est abordée ici de façon plus large pour pouvoir être utilisée pour transcrire les mouvements de tout locuteur/gesticulateur francophone ou non francophone.

Dans une première partie, seront précisées les conventions générales de transcription. La seconde partie permettra de dégager les caractéristiques gestuelles essentiellement aux plans typologique et kinésique. La troisième s'attachera à la présentation matérielle du corpus. Enfin, dans une quatrième partie, quelques illustrations permettront de montrer les possibilités qu'offre une telle méthodologie de transcription.

## 1. Conventions générales de transcription

Ces conventions de transcription générale concernent l'interaction (1 et 2), la localisation de la gesticulation (3), la prosodie (conventions 4 à 13), la transcription de l'énoncé (14 à 18), les phénomènes annexes (19 à 22), les phénomènes mimo-gestuels, hors mouvements des mains et des bras (23 à 26) :

---

KIDA, Tsuyoshi (2006), Transcrire le geste, *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, vol. 25, p. 127-149.

---

1. « Ingres, dit-on, aurait organisé le repos. J'aimerais, au-delà du pathos, organiser le mouvement », Paul Klee, *Théorie de l'art moderne*, Paris, Denoël, Folio-Essai (traduction P.-H. Gonthier), 1964, p. 46.

1. [aa][bb] : début/fin de chevauchement
2. = : contiguïté du discours d'un tour à l'autre (NB 1 pour une précision)
3. {aa} : début/fin de gesticulation
4. + : pause ("+" = 0,1 sec) (voir NB 2)
5. - : allongement vocalique (durée approximative :: = 0,1 sec)
6. :: : intonation montante discontinue (voir NB 3)
7. // : intonation descendante discontinue
8. \\ : intonation montante continue (voir NB 3)
9. / : intonation descendante continue ou discontinuité neutre (NB 4)
10. \ : accent attendu actualisé devant double flèche ou simple flèche (NB 5)
11. ° : accent non attendu jamais réalisé devant flèche (voir NB 6)
12. MOt : débit rapide
13. >aaa< : débit ralenti
14. a:a:a : rupture au cours de la prononciation d'un mot
15. (mot) : mot inaudible inféré ("x", syllabe non inférée)
16. [aaa] : transcription phonétique large (souvent indicative)
17. \*aa ((=bb))\* : alternance codique (langue maternelle ou autre), avec traduction
18. "le" maison : utilisé librement pour signaler un problème linguistique
19. (Ç) : claquement de langue
20. HA-HA-HA : rires
21. p(h)eut (h)être : rires en cours d'élocution
22. ((bas)) : commentaire du transcripteur
23. Tv/Th/Tx : mouvement vertical/horizontal/particulier de la tête (voir NB 7)
24. F ou Fx : expression particulière de la face (NB 7)
25. Fs/F#/Fsl : sourire/froncement de sourcils/haussement de sourcils (NB 7)
26. C/Cepl : mouvement remarquable du corps/haussement des épaules

NB 1. Le signe "=" est utilisé pour indiquer une continuité discursive ou textuelle de l'énoncé ; ne pas tenir automatiquement "=" pour achèvement interactif, lequel est repérable quand "=" est placé entre deux tours consécutifs dont chacun est produit par un locuteur différent.

NB 2. En lecture, le défilement temporel du logiciel correspond à un intervalle de 0.03. Comme notre définition temporelle est de 10 images/sec, 30 + = 1 seconde.

NB 3. Le terme discontinu renvoie au contour intonatif se terminant par un intonème conclusif ou terminatif (Rossi, 1999). Réalisé à la fin de l'unité tonale, cet intonème peut être soit montant (/ /), soit descendant (\\) (montée ou chute de la F0). Le terme continu renvoie au contour intonatif se terminant par un intonème continuatif (Rossi, 1999). Réalisé à la fin de l'unité tonale, cet intonème peut être soit montant (/), soit descendant (\) (montée ou chute de la F0, mais moins marquée que pour la discontinuité). Dans ce cas, le locuteur a probablement l'intention de poursuivre son énoncé.

NB 4. Le transcripteur ne juge pas cet intonème comme étant conclusif ou continuatif, ce contour intonatif neutre se réalise souvent avec une pause remplie (par exemple, euh\ ) ou avec une rupture en cours de prononciation d'un mot (par exemple, écha-\ échapper/).

NB 5. Les accents attendus seront notés ainsi : c'était la première °fois/ pour °lui/ au ja°pon//

- NB 6. En français, l'accent attendu à la fin du groupe rythmique se réalise sur la dernière voyelle du groupe intonatif, autre que [ə] caduc. Quand il y a emphase accentuelle réalisée sur le schwa à la fin du groupe rythmique, l'accent est non attendu (°sourCES\). Il s'agit d'un cas où l'accent non attendu se réalise devant une frontière tonale (matérialisée par une flèche simple ou double). On le notera donc en majuscules pour rendre compte de cette particularité assez partagée par les japonophones parlant français.
- NB 7. Les signes ('Tv', 'Th', 'F') sont signalés soit entre (()), soit entre ( ), la première notation indiquant que le mouvement est concomitant à la production verbale, la seconde que ce mouvement s'insère exactement à l'endroit de transcription.

Dans ce qui suit, il nous faut expliquer la genèse des méthodes d'analyse qui seront adoptées ici pour transcrire le geste. Deux angles principaux d'observation sont donc proposés : la typologie formelle, l'aspect kinésique<sup>2</sup>, en gardant à l'esprit la distinction entre gesticulation spontanée *vs* non spontanée<sup>3</sup>.

## 2. Caractéristiques gestuelles

Bien que notre but premier ne soit pas d'élaborer une classification gestuelle, nous proposons, par commodité, de classer les productions gestuelles, pour rendre visible le parcours de l'appropriation gestuelle chez les étrangers et la production gestuelle native. Les catégories permettront de caractériser effectivement le comportement des sujets, et, par le biais de la comparaison, d'élucider l'appropriation ou la production du geste. La catégorisation peut aboutir à une quantification du comportement (*cf.* chap. 9 dans Kida, 2005) et une telle quantification guide en général l'observation qualitative de données (voir *infra*, § 4)<sup>4</sup>.

Pour les différentes classifications, nous adoptons principalement un codage intrinsèque. Il faut avouer que cette méthode soulève un problème d'objectivité analytique et de comparabilité, puisqu'elle engendre des cas où une subjectivité de l'analyste est inhérente au jugement classificatoire d'un geste. Par suite, il arrive parfois que l'analyste soit incertain de tel ou tel étiquetage pour telle ou

---

2. Un autre critère est mis en avant, dans Kida (2005), il s'agit de la synchronie geste et discours qui est éliminée de la présente contribution.

3. Une autre distinction de l'activité gestuelle sera opérée, selon que la gesticulation s'accompagne d'un discours fluide « gesticulation spontanée » ou non « gesticulation non spontanée ». Bien entendu, ce critère de spontané *vs* non spontané est applicable à la production gestuelle d'un locuteur natif qui, lui aussi et tout naturellement, peut être aux prises avec des hésitations ou des recherches lexicales, par exemple.

4. Birdwhistell (1970, p. 82) fait une remarque similaire : « However, the specialist cannot determine how distinctively individual any particular performance is before he knows the structure of significant ranges of behavior for a particular behavioral area. When he makes such judgments in an *a priori* manner, he abdicates his professional role and interprets as an amateur, as do other members of society. Because such behaviors are communicational, as members of society, we use the pattern of another's behavior to anticipate and to react to him – but the scientific study of expressional behavior as a reliable test for determining underlying personality dynamics must await extensive experimentation before we can test productive value and reliability of clinical judgments. ».

telle production gestuelle. Pour notre étude, ce type d'incertitude analytique, relevant de la subjectivité de l'analyste, sera aussi considéré comme une donnée analytique. Par ailleurs, pour diminuer la subjectivité et rendre la subjectivité la plus « objective » possible, nous avons procédé à un étiquetage répété et une analyse du phénomène sous plusieurs angles.

## 2.1. Typologie formelle

Notre classification typologique a pour but premier d'analyser en détail la tendance du comportement gestuel sur le plan formel et elle a été élaborée à partir de celle d'Efron (1972), de celle d'Ekman et Friesen (1969) et de celle de Cosnier (1982), à quelques différences près. Mais à la différence d'Ekman et Friesen et de Cosnier, nous nous limitons à l'activité des mains et des bras, laquelle constitue le geste selon notre définition. Dans la recherche contemporaine, c'est la classification de McNeill (1992) qui semble la plus répandue, mais nous ne l'adoptons pas, puisque nous préférons nous doter d'une classification qui soit la plus large possible en fonction du critère de faisabilité analytique.

Nous pouvons ajouter à cela une autre raison. McNeill (1992) limite le cadre d'analyse aux gestes qui sont activés lors de l'énonciation en excluant les gestes qui sont produits pendant une pause discursive (c'est-à-dire les « mimes »). Cette perspective vise tout particulièrement à décrire la place du geste dans l'énoncé et de l'énonciation (c'est-à-dire la « gesticulation spontanée » selon notre terminologie). Bien que notre intérêt soit le même que celui de McNeill, nous ne pouvons pas exclure la « gesticulation non spontanée » de notre étude (pour poursuivre la réflexion dans le cadre de la L2<sup>5</sup>, voir Kida et Faraco (à paraître) et Kida (à paraître)).

D'après nos observations empiriques, certains gestes émis sans production verbale jouent un rôle certain dans l'échange, notamment dans l'interaction où au moins un des participants est un sujet non natif. En tout cas, les mimes sont identiques, sur le plan physique, aux gestes co-verbaux, et nous ne souhaitons pas les exclure du cadre d'analyse du geste discursif.

Voici notre classification formelle des gestes :

### A) indexical :

- geste déictique (concret et abstrait) ;
- geste physiographique :
  - geste iconographique : rapport direct entre référent-objet et geste ;
  - geste kinétographique : forme d'une action ou activité humaine et non humaine ;

---

5. Langue seconde.

- geste métaphorique : rapport indirect entre référent et geste ainsi que gestes spatio-topologiques ;

B) emblème

C) discursif :

- geste bâtonique : abstrait sur le plan formel, souvent utilisé pour rythmer ou mettre en emphase le discours ;
- geste idéographique : plus ou moins concret sur le plan formel, mais sémantiquement abstrait.

Dans ce qui suit, nous rendrons compte de quelques différences de notre classification par rapport à d'autres classifications.

Les déictiques sont des gestes de pointage, qui consistent à indiquer quelque chose par une direction du geste dans l'espace. Nous n'y inclurons pas le pointage discursif (l'index orienté vers le haut, par exemple) dont la direction n'est pas ciblée. La direction peut être indiquée par l'index ou le pouce. La main ouverte peut aussi indiquer une direction (comme, par exemple, le geste qui désigne poliment l'interlocuteur en disant vous), mais cela n'a pas été considéré ici comme geste déictique, car ce geste est difficile à distinguer d'un geste de la même forme à caractère discursif. La cible du geste déictique peut être présente *in situ* (geste dit « déictique concret » ou de « définition ostensive ») ou représentée de façon simulée (geste « déictique abstrait »). Les gestes déictiques s'appliquent aussi à la temporalité dans le discours (par exemple, *hier* ou *avant* est souvent un geste DERRIÈRE<sup>6</sup> ; voir Calbris, 1985). Le pointage pour énumération est aussi inclus dans le geste déictique.

L'emblème est utilisé au sens d'Ekman et Friesen (1969)<sup>7</sup> et se réfère à des gestes dont le sens est « arbitrairement » défini par la convention sociale d'une communauté donnée. La compréhension de ce type de gestes est en principe limitée aux membres d'une communauté. Certains emblèmes sont compréhensibles même par les étrangers (comme le geste OK), et d'autres moins.

Le « geste bâtonique » ou « bâton » est un mouvement répétitif d'une ou des mains lorsque le locuteur accentue son discours. Ce geste peut être combiné avec un autre geste dont la forme est figurative (gesticulation dite « superposée » ; voir *infra*). « Iconographique » est utilisé pour un geste qui rappelle directement la forme d'un objet dans le discours, et kinétographique pour un geste qui

---

6. Les majuscules renvoient au sens attribué à tel ou tel geste.

7. L'« emblème » au sens d'Ekman et Friesen (1969) est aussi appelé « geste quasi-linguistique » (Cosnier, 1982) ou « *quotable gesture* » (Kendon, 1993). Rappelons que la définition d'« emblème » par Efron (1972) est différente de celle d'Ekman et Friesen (1969), différence que certains chercheurs contemporains ignorent. Si, pour Ekman et Friesen (1969), la définition d'un emblème est « arbitraire », pour Efron, les « emblèmes » sont des gestes « conventionnellement » définis dans une communauté donnée. La définition d'emblème par Efron est donc plus large que la définition d'Ekman et Friesen, Cosnier ou Kendon.

décrit une activité ou une action de l'être humain ou non humain (par exemple, les gestes MANGER, ÉCRIRE, *etc.*).

Quant aux gestes métaphoriques, une précision s'impose. Dans les classifications d'Efron et d'Ekman et Friesen, ce type de geste n'apparaît pas. Le terme « métaphorique » a été employé par McNeill (1992), lequel se substitue au terme « idéographique » d'Efron, Friesen et Ekman. Pour comprendre la raison pour laquelle cette classe de geste métaphorique a été ajoutée dans les gestes physiographiques, considérons cet exemple :

(1)YO-Conv <sup>8</sup>		
240YO	:	c'est à [di:re\]
241LN	:	[((Tx) hm\] (TvTv) +
242YO	:	plutôt\ {les gens} {"normal"\ [+++++++]}
243LN	:	[((Tv) n\ + ((Tv) n\ ++]

Dans cet extrait, deux gestes sont présents : le premier est abstrait et de nature discursive et le second est plus concret et « physiographique », plus précisément métaphorique selon la classification présentée ci-dessus. Précisément, YO utilise, pour le premier geste, une forme plutôt abstraite, et, pour le second, un mouvement qui fait penser à « niveau bas ». Ce geste se distingue du geste iconographique dans la mesure où il ne dessine pas un objet, mais exprime une impression visuelle du sujet gesticulant. Il est, en effet, difficile d'exprimer visuellement l'expression GENS NORMAUX par une forme gestuelle.

McNeill (1989) utilise le terme « geste métaphorique », qui se définit comme geste qui n'a pas de rapport direct avec le discours mais qui fait apparaître une idée abstraite. Or, cette définition s'applique aux deux gestes concernés – idéographique et métaphorique –, qui sont nettement distincts aux yeux de l'analyste, lorsque ce dernier se réfère au contenu discursif. Si la terminologie de McNeill consiste à exprimer, sous le terme « métaphorique » – un geste idéographique dans notre classification –, il y aura une ambiguïté terminologique. Donc, nous les différencierons et utiliserons « gestes métaphoriques » pour les expressions manuelles de type GENS NORMAUX, qui illustrent une notion concrète ou une impression visuelle<sup>9</sup>.

8. AA (majuscules correspondant aux initiales du prénom du locuteur) – type de tâche (Conv : Conversation, Rec. : Explication d'une recette de cuisine en face-à-face). Dans l'extrait, les chiffres renvoient à la numérotation du tour de parole dans le corpus de Kida (2005), LN renvoie à locuteur natif.

9. La remarque de Rudolf Arnheim (1969/76) sur la représentation picturale illustre notre propos : « Il est facile de penser à des images qui représentent "maison", "lutte", voire à des rapports entre objets matériels comme "plus grand que" ou "compris dans". Mais que dire de "si, parce que, comme, bien que, ou bien... ou bien" ? » (p. 254). Pour nous, le premier type est iconographique, le second type métaphorique et le troisième type, idéographique ou discursif.

En principe, les gestes liés à bon nombre d'adjectifs ou d'adverbes exprimant une qualité (par exemple, *grand, petit, différent*), une quantité (par exemple, *beaucoup, un peu*), une durée (par exemple, *longtemps*), une exhaustivité (par exemple, *complètement, tout*) ou encore une approximation (par exemple, *à peu près*), relèvent des gestes métaphoriques dans notre classification typologique. De cette façon, le corrélat verbal de tous les types de gestes physiographiques maintiendra en général un type de représentation du monde visuel que sont objet, action et état du monde visuel ou visible.

Par ailleurs, l'équivalent gestuel de certaines locutions spatiales (*haut, bas, gauche, droit, dedans, dehors, à côté*) sera considéré comme métaphorique. En fait, ces rapports spatiaux sont appelés « relation topologique », mais comme ces gestes relèvent de l'espace mental des objets concrets (plus exactement de leur relation spatiale), ce sont tout d'abord des gestes physiographiques et non des gestes discursifs. Dans la mesure où ils expriment une direction, ces gestes pourraient être déictiques, mais l'iconicité dont ils s'accompagnent semble plus saillante dans le cas de ces gestes spatiaux. S'ils ne s'accompagnent pas de pointage, les gestes de l'indication de la spatialité seront donc considérés plutôt comme appartenant à la classe de gestes métaphoriques dans la présente étude. Nous incluons aussi dans la classe de gestes métaphoriques, les gestes exprimant un numéral (par exemple, *premier, second...*), un chiffre ou un nombre, qui peuvent difficilement être considérés comme discursifs.

Il existe une ambiguïté terminologique entre « métaphorique » et « iconographique ». Par exemple, pour désigner gestuellement une aubergine, un locuteur utilisera un geste qui trace une forme ovale, qui peut aussi s'appliquer à un concombre ou à un pain. C'est dire que tout geste iconographique est « métaphorique » ou « métonymique » au sens strict du terme, et que mis à part les termes géométriques, on peut difficilement représenter « iconiquement » un référent-objet par un geste. Notre distinction entre iconographique et métaphorique est ce à quoi renvoie la forme d'un geste : une forme ovale renvoie directement à la forme d'un légume ou autre, mais un geste métaphorique (par exemple le geste GENS NORMAUX) est motivée par une impression visuelle du sujet gesticulant et non par la forme concrète du référent. La différence entre « iconographique » et « métaphorique » correspond *grosso modo* à celle qui existe entre « objet visible » et « concept visible ».

Cet ajout d'une classe métaphorique ne provient pas seulement d'un souci d'exhaustivité. En fait, nous supposons que l'objet, l'action et le qualificatif fonctionneraient différemment dans la cognition du sujet. Notre classification formelle repose donc sur un « modèle linguistique ». Par exemple, un sujet appuie sa gesticulation sur une action, un autre sur un objet, et un autre encore sur un qualificatif ; comme l'ordre syntaxique en français et en japonais est différent, il peut y avoir



une différence comportementale à ce niveau<sup>10</sup>. De plus, lorsque le sujet gesticule, il sera d'un intérêt certain de savoir précisément quel type de rapport il met en œuvre avec les composantes du discours. Une gesticulation peut être redondante ou complémentaire par rapport au discours au niveau fonctionnel, et ces principes qu'utilisent les étrangers peuvent évoluer au cours de l'appropriation. Pour sonder ce type de rapport, la distinction entre qualificatif, action et objet contribuera *a priori* à une précision analytique. C'est la raison pour laquelle la classe métaphorique, correspondant au qualificatif dans le discours, a été introduite dans notre classification gestuelle.

En ce qui concerne les gestes idéographiques, on peut dire d'abord que ce sont des gestes qui ne sont pas liés à l'aspect informationnel du message. À regarder le mouvement, ces gestes sont plus ou moins saillants sur le plan figuratif et, dans une certaine mesure, proches des gestes conventionnels et physiographiques sur le plan formel, puisque la forme des mains est quelque peu « figurative » par rapport aux gestes bâtoniques ou aux gestes déictiques. Mais on sent que ces gestes ne sont pas liés à des référents objets ou actions, mais plutôt à la structure discursive et à l'aspect pragmatique du discours, sans pour autant pouvoir énoncer leur sens et/ou leur fonction précise. Par exemple, un geste qui se produit au début d'un énoncé correspond probablement à un connecteur d'introduction discursive.

Généralement, nombre de gestes idéographiques semblent remplir une fonction présentative comme *voici* ou *voilà* ou *ceci*, comme si le geste soulignait un segment du discours. Cette fonction est exprimée par plusieurs mouvements. D'abord il y a des gestes qui ressemblent au geste DONNER, c'est-à-dire un mouvement bi- ou mono-manuel dirigé vers le partenaire (*cf.* P<sup>11+</sup>). Certains autres sont dirigés vers soi-même (*cf.* P-) ou s'infléchissent vers le bas (*cf.* P±). Quelquefois, pour ces gestes, le poing est fermé ou mi-fermé mais par le mouvement identique, ils assurent une fonction présentative.

Les gestes idéographiques remplissent d'autres fonctions. Il y a des gestes qui semblent remplir une fonction convocative, comme si le sujet gesticulant était en train d'attirer l'attention de son partenaire en montrant le poing ouvert d'une ou des mains (*cf.* convocatif). On observe souvent ce geste pour dire *arrête* ou *attends*, mais aussi pour la négation ou le refus. Le pointage discursif a aussi pour fonction d'attirer l'attention du partenaire afin de prédisposer l'interlocuteur à un contenu important du discours subséquent (*cf.* pointage discursif). L'implication du partenaire est primordiale dans tous ces gestes à fonction convocative.

---

10. Par exemple, Luís Fróis (1532-1597), un des premiers missionnaires jésuites arrivant au Japon au XV<sup>e</sup> siècle au Japon, vit un certain « verbo-centrisme » dans le comportement langagier des Japonais. Voir son traité écrit en 1585, réédité récemment.

11. P = Paume.

Il y a des gestes qui semblent être moins directement liés au partenaire qu'au discours. Ces gestes sont plus figuratifs et plus proches des gestes physiographiques que d'autres gestes idéographiques. Par exemple, des gestes dont le mouvement vertical est fait du tranchant de la main, soulignent l'assertion ou la précision dans le discours (*cf. chop*). Ce type de gestes ressemble au geste kinétographique COUPER. Par ailleurs, il y a aussi des gestes dont la forme n'est pas différente de celle des iconographiques, mais ne renvoyant pas au référent-objet (*cf. IP, icono-présentatif*). À les regarder, ces gestes semblent impliquer une nature discursive en assumant une fonction présentative ou d'autres fonctions liées au discours. D'autres formes spécifiques observées dans notre corpus sont une forme ronde produite avec les doigts comme le geste de OK (*cf. rond*), ou une montagne (*cf. pyramide*), qui sont typiques dans la gesticulation méditerranéenne (*cf. de Jorio, 1832*)<sup>12</sup>.

Quant aux gestes bâtoniques, ce sont des gestes dont la forme et le mouvement sont assez simples, par exemple les gestes emphatiques ou phonogéniques. Nous incluons dans cette catégorie un certain nombre de gestes discursifs qui ne trouvent pas d'item correspondant dans notre classification. La distinction entre idéographique et bâtonique pourra être importante dans la mesure où même si les participants non natifs peuvent facilement employer les gestes bâtoniques, il se peut que les gestes idéographiques, eux, fassent l'objet d'une appropriation (voir la réflexion menée sur les raisons de mettre en relation l'acquisition d'une langue seconde et les gestes, par M. Gullberg, 2006).

Il faudra aussi préciser la façon dont une catégorie gestuelle de notre classification sera attribuée à telle ou telle production gestuelle dans le corpus réel. D'abord, la classification que nous avons adoptée n'est pas purement fondée sur un système sémiotique de l'activité gestuelle (codage extrinsèque), mais considère l'interaction entre activité gestuelle et langage verbal lors de l'interprétation (codage intrinsèque). Les gestes DONNER ou PAYER dont la forme est identique à celle des gestes idéographiques (voir ci-dessus), seront considérés comme kinétographiques si le sujet gesticulant veut signifier l'action de donner ou payer. On pourrait, en effet, établir un système purement sémiotique, à partir de différents types de forme et mouvement sans tenir compte du langage verbal. Or, il est établi qu'un geste peut donner lieu à de multiples interprétations<sup>13</sup>. On observe un geste dont la forme est iconographique ou kinétographique (par exemple, DONNER

---

12. D'après de Jorio, la montagne et le pyramide signifient *chiedere* (demander) à Naples (p. 474, édition 2000).

13. Par exemple, Reinhard Krauss (1930) a déjà démontré l'impossibilité de transmettre une idée ou une émotion à autrui à travers le seul dessin (Gombrich, 1982, p. 159).

ou PAYER), mais fonctionnant clairement comme un geste discursif<sup>14</sup>. Si le point de vue formel était strictement respecté, ce geste ne pourrait pas être étiqueté comme un geste discursif. Il nous a semblé qu'une telle méthode ne décrirait pas une réelle activité gestuelle du « sujet gesticulant ». De même, ce ne serait pas compréhensible pour le partenaire de conversation, qui regarde le comportement gestuel consciemment ou subconsciemment avec le discours. Il semble que ce point sur le principe classificatoire n'ait pas été suffisamment précis dans les études gestuelles précédentes (Efron, 1972, par exemple). La base analytique de notre étude peut être différente : ce qui est important pour nous, c'est le point de vue de celui qui produit un geste et le fait de porter un regard analytique sur la production gestuo-verbale, plutôt que l'optique puriste de l'analyste d'un geste.

Efron (1972, p. 96-98) explicite la différence entre physiographique et discursif (ou idéographique selon sa terminologie) de la façon suivante : le geste discursif se caractérise comme étant celui qui « traces or sketches out in the air the « paths » and « directions » of the « thought-pattern », qu'il appelle « course of the ideational process » (p. 96) ou « logical process », ou encore « mental process ». Efron utilise aussi les termes « logico-topographique » et « logico-pictorial » comme synonymes d'idéographiques. Il mentionne aussi que sa notion d'idéographique ou de logico-pictorial correspond à ce que Cicéron (*Oratoire*, Livre III) appelle « *significatio* », et à ce qu'Engel (1785-86/1795) appelle « analogique ». Par contre, le geste physiographique renvoie à « mental object » (Efron, 1972, p. 97). La différence cognitive entre gestes physiographiques et idéographiques est ce à quoi renvoie le geste comme objet ou comme processus.

Par ailleurs, le mouvement du geste (ou de la tête) qui exprime un « rejet » ou une « négation », par exemple, est dans une certaine mesure idéographique, mais pour Efron, ce type de geste est, non pas un processus, mais plutôt une « attitude logique (*logical attitude*) » (p. 98)<sup>15</sup>. Pour la présente étude, ces gestes seront considérés comme des emblèmes. Les gestes de l'hésitation (par exemple, l'index sur la bouche avec la pause remplie *euh:::* ou les mains croisées au moment de la réflexion, les deux mains ouvertes sur les genoux) sont exclus de notre taxonomie du geste. Ces gestes remplissent sans doute une fonction interactive (par exemple, empêcher l'intervention du partenaire ou maintenir son tour de parole)<sup>16</sup>. Dans une certaine mesure, ils seraient d'ordre conventionnel, mais nous limitons le terme « emblème » à des gestes dont le mouvement ou la forme sont explicitement culturels et arbitraires.

---

14. Geste et discours sont dans le processus de « contextualisation réciproque » (Bouvet, 2001, p. 140), c'est-à-dire, le sens du geste est complété par le discours et le discours est complété par le geste.

15. De même, Sapir dit du froncement des sourcils qu'il est un « symbole d'attitude » (1927, p. 60).

16. Cosnier (1982) qualifie ce type de gestes d'« extracommunicatifs ».

En ce qui concerne le geste kinétophique, une remarque s'impose : ce type de geste, à la différence des gestes iconographiques et métaphoriques, est souvent à l'œuvre avec la répétition de plusieurs mouvements identiques. La délimitation d'un geste pose alors problème. Par exemple :

(2) YS-Rec.		
035LN	:	[{++} <u>et:: on\ on enlève la::</u>
036YS	:	[((Tv)) n]
037LN	:	[ <u>la</u> ] <u>peau::/ +</u>
038YS	:	[((Tv)) n]
039LN	:	[{++}] + et:::\ {on:: <u>écrea::se</u> } {°très} {°très} {bien:::/ ++}
040YS	:	[((Tv)) n]
041LN	:	[{++}] ((...))

Dans cet exemple, sur deux énoncés s'activent deux groupes de gestes : celui qui est produit avec l'énoncé *on enlève la peau (de l'ail)* (035-039) et celui qui accompagne l'énoncé *on écrase très très bien* (039). Pour le premier, le locuteur répète le mouvement manuel de l'action d'éplucher et chaque mouvement est si rapide et si court qu'on ne pourrait découper ce mouvement en plusieurs gestes. Pour le second, le locuteur répète le même geste d'écraser – cette répétition gestuelle correspondrait au modificateur adverbial –, et tous les gestes seraient encore délimitables.

## 2.2. Aspect kinésique

Le mouvement d'un geste a son début et sa fin, et cette période pourrait définir le geste. Dans sa forme canonique, un geste est constitué par trois périodes : préparation, stroke et repos<sup>17</sup>. Or, le mouvement des gestes n'est pas identique et varie de l'un à l'autre. Kendon (1980) note plusieurs variantes de l'activité gestuelle. Le retour peut être partiel (*partial recovery*, p. 213), la gesticulation peut être figée pendant un moment (*hold*, p. 213), le stroke peut être multiple (*complexe stroke*, p. 215). De même, l'analyse de l'unité gestuelle doit tenir compte de la « manualité (*bandedness*) », à savoir l'activation de la main droite, de la main gauche ou des deux mains. Mais cette dernière précision kinésique risque de rendre peu lisible la transcription, notamment quand un mouvement commence

---

17. « In forelimb gesticulation the limb is typically lifted away from the body as it performs one or more complex movement patterns, and then it is returned to what may be called its *rest position*. [...] Within such an excursion the limb may perform one or more *phrases* of gesticulation. A phrase of gesticulation, or G-Phrase is distinguished for every phase in the excursionary movement in which the limb, or part of it, shows a distinct peaking of *effort* [...]. Such an effort peak, or less technically, such a moment of accentuated movement, is termed the *stroke* of the G-Phrase. It is usually preceded by a *preparation* – that is, by a phase in which the limb moves away from its rest position to a position at which the stroke begins. The stroke is then succeeded by a *recovery* or *return* phase in which the limb is either moved back to its rest position or in which it is readied for another stroke. » (Kendon, 1980, p. 212).

pendant que le geste de l'autre main est déjà activé. Dans notre corpus, la manualité n'est pas prise en compte, et le début de ce deuxième geste sera considéré comme la fin du premier.

Ainsi, une catégorisation simplifiée de l'activité gestuelle a été retenue. Nous avons élaboré l'aspect kinésique sur la base de la catégorie de Harmant-Dammien (1897) et de Kendon (1980), comme ci-dessous :

- 1) gesticulation simple ;
- 2) gesticulation entière ;
- 3) gesticulation enchaînée ;
- 4) gesticulation complexe ;
- 5) gesticulation superposée ;
- 6) échec gestuel ;
- 7) geste statique « l'arrêt sur image ».

### 2.2.1. Gesticulation simple

{préparation/stroke/retour}

C'est une gesticulation canonique dont la période consiste en trois phases. Elle implique un stroke statique, donnant une image fixe sur une durée courte, laquelle est associée à un corrélat verbal. Le moment de ce stroke est marqué par un soulignement plein qui représente une phase statique du geste.

### 2.2.2. Gesticulation entière

{préparation/stroke/retour} ou {stroke/retour} ou {stroke}

Cette gesticulation a lieu quand la totalité du mouvement est significative. Le soulignement en pointillés est introduit pour représenter cet aspect dynamique de la gesticulation entière (forme en mouvement), distinguée de la gesticulation simple (forme statique). Cette gesticulation peut prendre sens sans préparation, c'est-à-dire dès le début de son mouvement, et/ou finir sans retour, c'est-à-dire que la fin du stroke coïncide avec la fin de la gesticulation.

- (3) YO-Conv  
 030YO : [parce que] c'est: {;NA°TU::re\} c'est +++=  
 031LN : [((bas)) oui\]=  
 032YO : [= {TRAN]°QUI::[LLE} ]

### 2.2.3. Gesticulation complexe

{préparation/stroke/.../stroke/retour}

Ce type de gesticulation concerne une multiplication de strokes partiels, qui sont difficiles à considérer chacun comme un geste autonome. C'est souvent le cas pour un certain nombre de gestes kinétophiques, mais aussi pour des gestes iconographiques. Par exemple, pour

l'expression *carré*, la gesticulation nécessite le marquage d'au moins un trait vertical et un trait horizontal, qui donnent lieu respectivement à un mouvement. Dans ce cas, il est difficile de considérer ces deux mouvements comme deux gestes différents. Ce cas de figure correspond à la gesticulation associée à l'expression *on enlève la peau* (voir exemple 2, ci-dessus) :

(4) YS-Rec. : soupe au pistou  
 035LN : [ {++} et:: on\ on.enlève.la::  
 036YS : [ ((Tv) n ]  
 037LN : [ la.l.peau::/.±  
 038YS : [ ((Tv) n ]  
 039LN : [ ++} ] + ((...))  
 • geste=ÉPLUCHER

#### 2.2.4. Gesticulation enchaînée

{préparation/stroke} {stroke}... {stroke/retour}

C'est un type de gesticulation qui se situe entre gesticulation entière et gesticulation complexe, dans la mesure où les strokes multiples sont considérés ici comme des gestes autonomes. La distinction entre gesticulations enchaînée et complexe est parfois difficile à établir, notamment lorsque la même forme est répétée. En voici un exemple.

(5) YS-Rec. : soupe au pistou  
 039LN : [ ++} ] + et::: \ {on:: écras::se} {°très} {°très} {bien:::/ ++  
 040YS : [ ((Tv) n ]  
 041LN : [ ++} ] ((...))  
 • geste = ÉCRASER ; répétition gestuelle = modificateur adverbial

La distinction entre gesticulation complexe et gesticulation enchaînée, même si elle paraît superflue, peut avoir son importance pour tenir compte de la différence entre répétition gestuelle à des fins emphatiques et répétition inhérente à un geste (notamment physiographique). Cette distinction se fait en général à l'œil nu, en fonction de la vitesse, de la rythmicité, de la forme, de la présence ou de l'absence d'une pause potentielle du mouvement manuel.

#### 2.2.5. Gesticulation superposée

{préparation/ {stroke} }... {stroke} /retour}

Traditionnellement, ce type de gesticulation concerne le geste dont la forme se combine avec un ou plusieurs mouvements bâtoniques<sup>18</sup>. Dans certains cas, un geste se superpose à un geste autre que le geste bâtonique. Une gesticulation est considérée également comme superposée lorsque les deux mains présentent différents gestes.

### 2.2.6. Échec gestuel

{préparation}

C'est le cas où le stroke n'est pas identifiable. Un autre geste peut s'enchaîner éventuellement. Parfois, le type de gestes qui devrait suivre peut être inféré (notamment certains gestes kinétographiques).

### 2.2.7. Gesticulation statique « arrêt sur image »

{préparation/stroke\_\_\_\_/retour} ou {préparation/stroke\_\_\_\_/retour}

Il s'agit d'une cristallisation prolongée de la forme gestuelle d'un stroke dynamique ou statique avant le mouvement de retour. Cela peut concerner un geste tant physiographique qu'idéographique. Quand il s'agit d'un geste kinétographique ou idéographique, l'arrêt sur image est une sorte d'« iconisation » du mouvement (par exemple, MANGER).

En principe, c'est une caractéristique supplémentaire de toute gesticulation, et l'arrêt sur image ne peut être considéré comme une classe kinésique autonome. Cette gesticulation peut apparaître, par exemple, au cours d'une gesticulation simple. Voici un exemple :

- (6) YO-Rec. : aïoli  
 169LN : [lc{presque}::\ une °tête\ +++[+++++}]  
 170YO : [une tête]  
 • le geste (statique) correspond à *tête* dure jusqu'à la réaction du partenaire

Comme le montre l'exemple, l'arrêt sur image se trouve à l'intérieur de la gesticulation simple (d'un geste iconographique), et l'allongement du stroke dans le corpus peut évoquer à peine la présence de l'arrêt sur image. Puisque ce geste est déjà classé comme gesticulation simple, le trait de l'arrêt sur image est noté en plus dans le fichier d'analyse.

En général, ce type de gesticulation présente une certaine fonction interactionnelle comme le contrôle de la compréhension du partenaire ou la régulation de l'intercompréhension. Comme dans l'exemple précédent, on observe souvent que la forme des mains reste figée jusqu'à la

---

18. Selon la terminologie de McNeill (1992), il s'agit de « super-imposed beat », mais ce terme ne concerne que le geste bâtonique. Nous pensons qu'il y a d'autres combinaisons de gestes.

production discursive subséquente du locuteur ou jusqu'à la réaction du partenaire, comme si celui qui gesticule attendait une réaction de l'autre. Pour ne pas confondre un tel geste statique avec un simple arrêt gestuel (échec gestuel), nous précisons que l'arrêt sur image qui nous intéresse ici concerne un geste empreint d'une iconicité certaine. Ce geste montre certes la volonté communicative qui se focalise sur le segment sur lequel porte le geste, mais aussi le souci de l'intercompréhension ou la conscience de l'effet perlocutoire (Kida et Faraco, 2003).

Toutefois, une telle fonction peut tout aussi bien être remplie par l'arrêt d'une gesticulation qui n'implique pas d'iconicité, mais il semble que la gesticulation non iconisée s'appuie moins sur une volonté communicative. Un tel comportement peut avoir pour fonction une élicitation d'un message implicite pour le maintien du tour de parole, la demande « couverte » d'un étayage linguistique. Par exemple :

- (7) YS-Conv  
 102YS : ((Tv)) ((haut)) oui/ ++ mais\ [s::]\ mais:/ y a <sup>M</sup>{beaucoup de:::}\  
<sup>B</sup>{[enta:]puliz:::}\} ++ <sup>Id<sup>P</sup></sup>{[ede:::]} ((=aider)) un peu <sup>Id+D<sup>P</sup></sup>{:::}\  
<sup>B</sup>{petit}\} + <sup>B</sup>{petit}::: \ ++ uh:: ++ <sup>B</sup>{[ba::dje]-t::}\  
 ++[+++++++]  
 103LN : [((Tv)) oui::\ \]  
 104YS : <sup>B</sup>{ci-}  
 105LN : [petit budget\ \ ((Tv)) oui::\ \]  
 106YS : [<sup>B</sup>{ciné}ma <sup>Id</sup>{(petit} [by]dje}]}  
 107LN : [((TvTv)) oui oui ++]

L'arrêt d'une gesticulation non iconique s'observe autour de l'expression *petit budget* en 102 (signalé en gras) et continue jusqu'en 106, où l'activité gestuelle s'arrête à la suite de l'obtention du mot recherché. Bien que ce geste manque de précision formelle, il suscite une réaction de la part du partenaire (105). On peut supposer qu'il s'agit d'une demande d'assistance et non d'une recherche lexicale, puisque le regard est dirigé vers LN<sup>19</sup>. Ainsi, la gesticulation statique non iconique fonctionne aussi comme une stratégie d'interaction.

Cette stratégie peut être déployée plus explicitement par l'arrêt iconique d'une gesticulation. L'arrêt sur image dans l'exemple suivant (qui débute vers la fin de 188 et qui finit en 192, signalé en gras) a déclenché un étayage linguistique du partenaire.

- (8) SA-Conv

---

19. Lorsque le partenaire est non natif, il est moins probable qu'il s'agisse d'une demande lexicale, mais la fonction du geste serait plutôt un acte implicite pour demande de confirmation du message.



188SA	:	<sup>Id</sup> {et puis::\ uh::\ B{[filto]}:r}\ cinq// ++} <sup>Éc</sup> {c'est très::} M{:\ comment dit/ c'est\ uh::\ n::\ beaucoup} <sup>Éc</sup> {de +++} M{beaucoup différent de ((Tx)) contraste//
189LN	:	((Tv)) oui/+++ [++]
190SA	:	[((Tx)) n::/] [+++++]
191LN	:	[+ très::\ oui\ \ +++ <sup>Id</sup> {contraste.\ très très fo::rt\}] [oui\ (TvTvTv)]
192SA	:	[((Tv)) ah/ ((Tv)) oui\ oui\]
193LN	:	[(TvTvTvTvTvTvTv)]
194SA	:	[ <sup>Id</sup> {contraste est::: très} M{très fort\}] [(TvTv) ++]

Cet arrêt sur image est, comme on le voit, de très longue durée ; en 189, LN semble avoir compris le message de SA en 188, mais ce geste persiste ; en 190, SA produit une interjection et un mouvement de tête (Tx) qui signalent à LN un doute par le biais d'un geste statique ; en 191, LN devine enfin que l'intention cachée derrière le geste statique de SA est en fait une demande de réparation, et propose une réparation ; en 192, SA n'arrête pas sa production du même geste jusqu'à ce qu'elle passe enfin à un autre geste au milieu du tour de parole en 194. En effet, l'extraordinaire persistance de l'arrêt sur image par SA, mais aussi la densité de son regard, font comprendre qu'elle n'est pas satisfaite de la réaction de LN (189), en indiquant qu'elle attend une autre intervention de la part de LN (191). On voit de nouveau le même geste à la fin de 194 (signalé en gras) pour la dernière confirmation (à noter deux hochements de tête).

L'arrêt d'une gesticulation non iconographique fonctionne donc comme une stratégie interactionnelle. Le cas du geste iconographique « arrêt sur image » présente de surcroît une volonté communicative de par la visualisation de l'élément problématique dans le discours. Notons enfin qu'il y a un cas spécial de l'arrêt sur image : le geste fige sa forme représentative avec une main et un autre geste commence à être effectué par l'autre main. Notre système de transcription ne permet pas de montrer un tel rapport complexe entre deux gestes successivement produits, et il faudra en signaler les cas dans l'analyse.

### 3. Présentation du corpus

Nous indiquons trois possibilités de présentation du corpus. Premièrement, un corpus brut dans lequel le geste est transcrit à partir du corpus vidéo :

(9) YO-Conv	
136YO	: [Id <sup>2</sup> {++++ <u>donc::\</u> }] Id <sup>0</sup> { <u>très::\</u> } M{si <u>on</u> marche dans la <u>D</u> {[ly//] ((=rue)) [+++++]}]
137LN	: [hum ++]
138YO	: M <sup>2</sup> {tout le <u>monde::\</u> } D <sup>0</sup> {::\ <u>++ nous</u> } Id <sup>2</sup> {voit// <u>++</u> [+++++]}
139LN	: [((Tv)) oui\\] +++++ [VOUS remarque\]=

Nous rappelons que l'activité gestuelle est délimitée par des accolades. Le type de geste est spécifié selon la typologie formelle que nous avons établie (D = déictique, Ic = iconographique, K = kinétographique, M = métaphorique, Id = idéographique, B = bâtonique, Éc = échec gestuel, Em = emblème, ? = ambiguïté gestuelle), et le stroke du geste est souligné. Rappelons aussi que le soulignement en pointillés traduit un stroke dynamique (notamment pour des gestes kinétographiques, idéographiques ou bâtoniques), et que le soulignement plein renvoie à un stroke statique (pour des gestes iconographiques, par exemple). L'observation de traits du stroke permet d'appréhender partiellement l'aspect kinésique de la gesticulation. Par exemple, pour une gesticulation simple, il y a au moins un stroke statique dont le soulignement est relativement court, pour les gesticulations enchaînée et entière, le stroke dynamique est étendu, le stroke statique prolongé représente l'arrêt sur image. La gesticulation superposée est représentée par des accolades à l'intérieur d'un geste. Tous ces traits présentés s'appuient sur notre interprétation personnelle, c'est-à-dire un codage « intrinsèque ».

Deuxièmement, un corpus pouvant être présenté sous forme d'un discours suivi du sujet, comme suit :

(10) YO-Conv : 136-138
[Id <sup>2</sup> {++++ <u>donc::\</u> }] Id <sup>0</sup> { <u>très::\</u> } M{si <u>on</u> marche dans la <u>D</u> {[lu//] ((=rue)) [+++++]}] M <sup>2</sup> {tout le <u>monde::\</u> } D <sup>0</sup> {::\ <u>++ nous</u> } M <sup>2</sup> {voit//... <u>++</u> [+++++]}

Pour ce qui est de ces deux types de formes, nous présentons soit l'une, soit l'autre, soit les deux et ce, variablement selon que le contexte d'analyse porte sur l'interaction ou sur le contexte discursif.

Enfin, l'interprétation des gestes sera présentée, avec quelques précisions formelles. Par exemple :

[disc.]	donc	très	si on	marche dans la rue	tout le monde	nous	voit
[geste]	P+?	2M-Px	rotation	LÀ-BAS	VERS MOI	MOI	conclusif?

Cet extrait implique sept gestes. Comme ci-dessus, lorsque l'interprétation verbale d'un geste est possible, nous la notons en majuscules comme nous l'avons fait jusqu'à présent ; si ce n'est pas le





Dans cet exemple, MS (niveau D)<sup>20</sup> semble avoir un souci lexical en 034 et 036. Ce que MS veut dire est, supposons-le, « j'ai trouvé une station de radio qui répète toute la journée les actualités », dans lequel « la station » et « répéter » sont problématiques. Pour exprimer « la station », il utilise deux gestes kinétographiques puis il répète un même geste iconographique trois fois en 034. Ces gestes kinétographiques, dont le mouvement est identique,

peuvent être considérés comme GSV ; pour les trois derniers gestes, le premier est une GSV, et les deux derniers, iconographiques, sont envisagés avec un mot d'origine étrangère, à savoir *channel*. Ces deux groupes de gesticulation sont différents sur le plan formel : le second groupe est une sorte de « reformulation » du premier ou « reformulation intra-codique » (Kida et Faraco, 2003). Avec ces gestes en 034, mais aussi pour le cas du geste RÉPÉTER en 036, LN est parvenu à trouver les mots que cherchait son partenaire MS. Ces gestes sont tous comptés comme gesticulation non spontanée.

Les traits typiques de la gesticulation sans verbal sont : 1) le détournement du regard ; 2) la *disfluence* de la conduite verbale, caractérisée par une pause silencieuse, une pause remplie, un allongement vocalique et/ou une prononciation incomplète de l'expression ; 3) la gesticulation. Dans certains cas, le sujet – s'il est non natif surtout – peut procéder à une alternance codique, comme *channel* dans l'exemple ci-dessus. Il arrive aussi que, comme pour les deux gestes ci-dessus, le sujet suspende le mouvement gestuel, en figeant sa gesticulation (arrêt sur image), comme si les sujets attendaient une réaction de leur partenaire : étayage linguistique, une confirmation, une ratification, une proposition de correction.



Souvent, le partenaire de conversation se propose de dire quelque chose à la place de celui qui produit la GSV pour faire avancer le déroulement de l'échange. Mais cette chance – mais aussi la chance que coïncide l'intention communicative du gesticulateur et la proposition de son partenaire – dépendra de plusieurs éléments. Il faut, avant tout, que le partenaire dirige son attention sur la production

20. Niveau A > niveau B > niveau C > niveau D.

gestuelle, afin de détecter la présence du trouble de la communication et de l'analyser. Ensuite, il faudra que la forme gestuelle soit exempte d'ambiguïté pour le partenaire. Mais il se peut que ce dernier point soit aléatoire et dépende d'autres facteurs. Le contexte discursif et l'aspect sémantico-textuel, le partage du savoir socioculturel, ainsi que le rapport interpersonnel peuvent être des facteurs qui alimentent la réflexion du partenaire. Enfin, il est tout à fait possible qu'après une GSV, le sujet résolve de lui-même le problème en prononçant l'expression qu'il cherchait, avant même que son partenaire ne réagisse à celle-ci.

Il est à noter que la GSV s'observe chez tous les sujets, y compris les sujets avancés en français et LN. Par ailleurs, certaines des GSV des sujets avancés semblent être utilisées à des fins « rhétoriques », et sont souvent mises en œuvre avec un regard dirigé vers le partenaire. Par exemple, considérons un geste de RE, sujet non japonais de niveau B. La GSV apparaît dans le dernier tour de parole.

(12) AS/RE-Rec.  
 230RE : ((bas)) mais\ (Th) ++++++ ((bas))  
 ((Tx)) ah:: mais\ je ((ThTh)) je n'aime pas beaucoup ce  
 plat [par ce que::\  
 231AS : [((Fx)) POURQUOI::\  
 ((...))  
 236RE : Id{((Tx)) ce::\ca c'est::\Id{+(voilà::)/} Id{c'est très. froi-  
 \. c'est=  
 237AS : [((Tv)) oui::\  
 238RE : [=très. froid/}  
 239AS : [((Tv)) oui::\  
 240RE : [Éc{alo::rs/ "et"\] +++++ uh::\  
 Éc{c'est les::\  
 c'est\ c'est\ {l'heu::re de déjeuner:://  
 241AS : [(Fs) ((TvTv)) hah\  
 242RE : [+++++} et]::\  
 K{manger::/ les [nui::]/.++}  
 M{((Tx)) comme ça:::.....\ ++[+++++]}  
 [disc.] et manger les nouilles comme ça  
 [geste] MANGER Tx 2M-approximatif

Tout le long de la séquence, RE, qui a déjà goûté le plat qu'est en train de lui décrire AS, exprime sa réticence à manger ce plat, puisque ce sont des nouilles froides, et qu'il trouve cela bizarre. Voici comment RE se comporte : s'il exprimait directement son sentiment négatif, RE risquerait de blesser AS, pour qui il s'agit d'un de ses plats favoris ; il préfère alors ne pas dire directement ce qu'il pense, en utilisant un geste, moyen indirect de communication. Le geste métaphorique en 242, mais aussi le mouvement spécifique de la tête entrent dans cette logique d'atténuation. Le geste est celui de l'approximation (mouvement bi-manuel en alternance), qui exprime ici son

malaise à l'égard de ce plat. Ainsi, RE évite une expression verbale directe à l'aide de la GSV. Quant au regard dans l'exemple ci-dessus, il est en effet dirigé vers le partenaire, comme si RE invitait son partenaire à interpréter son geste.



La fonction de cette gesticulation est d'éviter une communication qui risque de porter atteinte au cadre social de l'interaction<sup>21</sup>. Un tel geste rhétorique s'observe notamment chez les sujets avancés, mais peu chez les sujets moins avancés. La gesticulation rhétorique est-elle l'objet d'une appropriation ? Il serait difficile d'affirmer qu'il s'agit d'un comportement spécifiquement culturel, et on peut tout à fait supposer que tous les sujets savent utiliser une telle gesticulation.

## 5. Conclusion

La transcription à codage intrinsèque du geste que nous avons élaborée obéit à deux critères : lisibilité et cohérence. À cette fin, le regard de l'analyste porte sur la production gestuo-verbale en tentant d'adopter, autant que faire se peut, le point de vue de celui qui produit le geste. L'application d'une telle transcription à des productions de locuteurs non natifs a permis de donner à voir, par effet de loupe, quel type de rapport le locuteur établit entre ses gestes et les composants de son discours, d'une part, ou entre ses gestes et son interlocuteur, d'autre part. Pour ce qui est de ces deux types de formes, deux présentations du corpus sont utilisées variablement selon le contexte d'analyse. La première présentation permet d'apprécier le contexte de l'interaction, et la seconde a pour but d'attirer l'attention sur le rapport entre discours et gestes. Une telle quête pour rendre la dimension visuelle « en mouvement » dans la transcription manuelle de corpus pourrait sans doute contribuer à rendre compte globalement des aspects contextuel, discursif, typologique, chronologique, kinésique, symbolique et/ou fonctionnel du comportement gestuel des sujets dans leur langue maternelle ou leur langue seconde.

## 6. Bibliographie

ARNHEIM, R. (1969). *Visual thinking*. Berkeley : University of California Press (trad. fr. : *La Pensée visuelle*, Paris, Flammarion, Nouvelle bibliothèque scientifique, 1976, Flammarion-Poche, 1997).

---

21. Bien évidemment, le geste rhétorique présente d'autres fonctions que l'évitement. Le discours politique est, par exemple, un domaine discursif qui est propice à l'observation de plusieurs autres fonctions.

- BIRDWHISTELL, R. L. (1970). *Kinesics and context: Essays on body motion communication*, Philadelphia : University of Pennsylvania Press.
- BOUVET, D. (2001). *La dimension corporelle de la parole. Les marques posturo-mimo-gestuelles de la parole, leurs aspects métonymiques et métaphoriques, et leur rôle au cours d'un récit*, Collection « Linguistique » Société de Linguistique de Paris, Paris : Peeters.
- CALBRIS, G. (1985). Espace-temps : Expression gestuelle du temps, *Semiotica*, 55, 1/2, p. 43-73.
- CICERON (1956). *De l'orateur* (édition de H. Bornecque, E. Courbaud), Paris : Les Belles Lettres.
- COSNIER, J. (1982). Communications et langages gestuels, in J. Cosnier, A. Berrendonner, J. Coulon, C. Orecchioni, *Les voies du langage. Communications verbales, gestuelles et animales*, Paris : Dunod, p. 255-304.
- DE JORIO, A. (1832). *La mimica degli antichi investigata nel gestire Napolitano*, Naples : Fibreno (seconde édition, préface de G. Cocchiara : Naples, Associazione Napoletana per i Monumenti e il Paesaggio, 1964 ; réimpression anastatique de l'édition originale : Bologna, 1979 ; trad. avec une introduction et notes par A. Kendon : *Gesture in Naples and gesture in classical antiquity*, Bloomington and Indianapolis : Indiana University Press, 2000).
- EFRON, D. (1972). *Gesture, race and culture*. The Hague : Mouton (1ère édition sous le titre de : *Gestures and environment: A tentative study of some of the spatio-temporal and linguistic aspects of the gestural behavior of Eastern Jews and Southern Italians in New York City, Living under similar as well as different environmental conditions*, New York: King's Crown Press, 1941).
- EKMAN P. ; FRIESEN, W.V. (1969). The repertoire of nonverbal behavioral categories: Origins, usage, and coding. *Semiotica*, 1(1): 49-98 (repris in A. Kendon (ed.), *Nonverbal communication, interaction, and gesture*, The Hague: Mouton, p. 58-105, 1981).
- ENGEL, J.J. (1785-86). *Ideen zu einer Mimik*. Berlin, Auf Kosten des Verfassers und in Commission by August Mylius (trad. fr. : *Idées sur le geste et l'action théâtrale*. Paris : H.J. Jansen et Comp, 1795 ; reproduction originale de l'édition française, Genève : Slatkine Reprints, 1979).
- FROIS, L. (1993). *Traité de Luis Fróis, S.J. (1585) sur les contradictions de mœurs entre Européens et Japonais*. (trad. fr. par J.-M. Garcia et X. de Castro), Paris : Chandeigne.
- GOMBRICH, E. H. (1982). *The image of the eye. Further studies in the psychology of pictorial representation*, Oxford: Phaidon.
- GULLBERG, M. (2006). Some reasons for studying gesture and second language acquisition (Hommage à A. Kendon). *IRAL, special issue : Gestures and second language acquisition*, edited by M. Gullberg, vol. 44, 2, p. 103-124.
- HARMANT-DAMMIEN (1897). *Du geste artistique dans l'action oratoire*. Abbeville : C. Paillart.
- KENDON, A. (1980). Gesticulation and speech: Two aspects of the process of utterance, in M. Ritchie Key (ed.), *The relationship of verbal and nonverbal communication*. The Hague: Mouton, p. 207-227.

- KENDON, A. (1993). Human gesture, in Kathleen R. Gibson, Tim Ingold (eds), *Tools, language and cognition in human evolution*. Cambridge: Cambridge University Press, p. 43-62.
- KIDA, T. ; FARACO, M. (2003). Gestures in second language discourse: metacommunicative function and perlocution, in M. Rector, I. Poggi, N. Trigo (eds), *Gestures: Meaning and use*. Porto : Universidad Fernando Pessoa, p. 305-318.
- KIDA, T., FARACO, M. (à paraître). A few remarks on gesture in the L2 classroom, in G. Stam, S. Mac Cafferty (eds), *Second Language Acquisition and Classroom Research*, Mahwah (NJ): Lawrence Erlbaum.
- KIDA, T. (2005). *Appropriation du geste par les étrangers : le cas d'étudiants japonais apprenant le français*. Thèse de Doctorat en Sciences du Langage : Université Aix-Marseille I, sous la direction de A. Giacomi.
- KIDA, T. (à paraître). Does gesture aid discourse comprehension in L2 ?, in G. Stam and S. Mac Cafferty (eds), *Second Language Acquisition and Classroom Research*, Mahwah (NJ): Lawrence Erlbaum.
- KRAUSS, R. (1930). Über den graphischen Ausdruck, in *Beihfte zur Zeitschrift für angewandte Psychologie*, 48.
- MCNEILL, D. (1989). A straight Path-to where? Reply to Butterworth and Hadar. *Psychological review*, 96 (1), p. 175-179.
- MCNEILL, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: Chicago University Press.
- ROSSI, M. (1999). *L'intonation. Système du français : description et modélisation*. Gap-Paris : Ophrys.
- SAPIR, E. (1927). The unconscious patterning of behavior in society, in E.S. Dummer (ed.), *The unconscious: A symposium*. New York: Knopf, p. 114-142 (repris in D.G. Mandelbaum (ed.), *Selected writings of Edward Sapir in language, culture and personality*. Berkeley: University of California Press, p. 544-559, 1954 ; trad. fr. par Ch. Baudelot, P. Clinquart : L'influence des modèles inconscients sur le comportement social, in *Anthropologie*, 1967, Paris : Minuit, p. 35-54).





**CORPUS DU PARLER RUSSE SPONTANÉ :  
ANNOTATIONS ET OBSERVATIONS SUR LA  
DISTRIBUTION DES FRONTIÈRES PROSODIQUES**

Irina Nesterenko

## 1. Introduction

La présente contribution a pour objectif de démontrer l'intérêt qu'il y a à mener un travail de description prosodique à partir d'un corpus doté d'une annotation élaborée.

Nous insisterons tout particulièrement sur le fait que la méthode de corpus permet de conjuguer les démarches analytiques descendante (*top-down*) et ascendante (*bottom-up*) : une fois les entités prosodiques définies sur la base des critères formels, la tâche est de mettre en évidence des contraintes de haut niveau (contraintes linguistiques), influant sur leur formation et la variabilité observée. En d'autres termes, les unités prosodiques formelles doivent être mises en relation avec des unités linguistiques et discursives.

Le travail sur une formalisation des analyses prosodiques s'impose d'autant plus que nous travaillons sur un corpus en langue russe. Les paradigmes de la phonologie intonative et de la phonologie des domaines, standards dans les études prosodiques actuelles, n'ont pas été appliqués jusqu'à présent dans les descriptions de la prosodie du russe.

Le présent article sera organisé en deux parties. Nous allons d'abord proposer une description de notre corpus d'études ; dans cet objectif, nous discuterons des modalités de recueil de données avant de détailler les annotations que nous avons effectuées pour en préparer des traitements postérieurs. Dans un deuxième temps, nous nous servirons des annotations établies dans le but d'aborder l'épineuse question du phrasé prosodique.

## 2. Travailler sur la parole spontanée

Nous avons fait le choix de travailler sur la parole spontanée. Il importe de remarquer qu'une grande partie des recherches antérieures en prosodie (quelle que soit leur finalité) se sont fondées principalement sur de la parole lue, et ceci également pour la langue russe. En effet, en lecture, les

---

NESTERENKO, Irina (2006), Corpus du parler russe spontané : annotations et observations sur la distribution des frontières prosodiques, *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, vol. 25, p. 151-172.

variations prosodiques sont contrôlées et restreintes, ce qui permet d'appréhender au mieux les rapports entre les dimensions et les structures prosodique et syntaxique de l'énoncé. Or, même dans ce cadre simplifié, il apparaît que la gestion de certains phénomènes (tels l'emplacement des frontières prosodiques et le degré de saillance des éléments structurels de l'énoncé) n'est pas aussi aisée qu'il y paraît. Dans la suite des recherches prosodiques appliquées, il a été attesté que le transfert direct des connaissances acquises pour le style lu à de la parole non-contrainte se trouve être inadéquat, surtout dans le contexte d'une prise de conscience sur les mécanismes régissant la production de la parole spontanée.

Notre deuxième choix méthodologique a été de nous positionner dans la perspective analytique guidée par les données. Dans le contexte d'une confrontation (*cf.* Lickley *et al.*, 2005) entre les méthodes de la phonologie de laboratoire (Pierrehumbert *et al.*, 2001 ; D'Imperio, 2005) et de la linguistique de corpus, que, pour notre part, nous considérons comme complémentaires, nous jugeons souhaitable, à l'étape initiale, de fonder notre étude sur l'observation des usages que font des locuteurs des moyens mis à leur disposition par le système prosodique de la langue.

Pour argumenter notre position, il nous semble nécessaire de fournir quelques précisions sur la dichotomie établie entre la parole de laboratoire et la parole spontanée.

Par parole de laboratoire, on entend généralement des enregistrements effectués dans une chambre anéchoïque, ce qui assure une bonne qualité du signal : le matériau linguistique représente des listes de mots ou de phrases, construites au préalable par l'expérimentateur, avec un contrôle systématique de diverses sources de variabilité. En plus d'une bonne qualité de l'enregistrement, cruciale dans les études acoustiques, un atout majeur de ce type de corpus réside dans la possibilité d'élaborer des stimuli en fonction des hypothèses testées. En revanche, en travaillant à partir d'un corpus authentique, on pourrait avoir besoin de nombreuses heures d'enregistrement pour aboutir à une quantité représentative d'exemples du phénomène étudié ; dans un corpus construit ce phénomène est présent de manière condensée.

Pourtant, le désavantage des corpus construits réside dans la restriction stylistique qui leur est propre : très fréquemment, il s'agit de textes/stimuli/listes lues, et non pas de parole produite naturellement. Précisons également que ces corpus représentent de la parole produite dans un but non communicatif. Qui plus est, dans le travail à partir des corpus construits, chacun est amené à adopter un raisonnement descendant (théorie ~ empirie) sur les conséquences des principes sous-jacents d'une théorie linguistique.

À l'opposé, les corpus de parole attestée relèvent d'une parole authentique. Ici, la parole n'est pas préparée à l'avance par l'expérimentateur, la parole est produite dans l'intention de communiquer un message et, selon la situation, les processus de planification et de production se réalisent

conjointement (et on en trouve des marques dans le signal sous la forme de différents phénomènes de disfluences).

Dans ce débat sur le statut et la validité des données dans les études phonétiques expérimentales, il est légitime de poser la question des objectifs auxquels se rapportent ces données. Ainsi, de manière générale, tout investigateur se penche sur les données observables soit pour construire une hypothèse de recherche, soit pour valider une hypothèse issue de l'analyse d'autres types de données et/ou de réflexions théoriques. On distingue donc les objets d'observation et les objets de validation : tous deux peuvent provenir de corpus fabriqués comme de corpus de parole attestée. Le seul contexte dans lequel la parole naturelle est préférable est le cadre d'une étude discursive et conversationnelle avec l'accent mis sur les aspects interactionnels. Ainsi Beckman (1997) conclut-il son plaidoyer en faveur de la parole spontanée :

*« I have spent so much time motivating why we want to look at spontaneous speech for a better modelling of how focus domain relates to accentuation in English because this is a good example of how lab speech has served us well and of how it fails to serve our needs completely. »*

Dans la même lignée, Lickley *et al.* (2005) défendent les approches de la phonologie de laboratoire :

*« Doubts are sometimes raised about the “ecological validity” of conclusions based on the analysis of intonation in controlled sentences read aloud under laboratory conditions, and more generally about the intonation of “decontextualised” sentences. We argue that these concerns, though by no means entirely misplaced, are not sufficiently serious to discredit laboratory phonology methods. »*

Les auteurs remarquent, entre autres, que les différences entre la parole de laboratoire et la parole naturelle, spontanée, sont levées dès que l'on atteint les niveaux d'implémentation phonologique et phonétique de la parole (*cf.* le modèle de Levelt, 1989). Cette conclusion intrinsèquement liée à l'approche modulaire entre en conflit avec les observations formalisées sous la forme de la règle de variabilité (*variability rule*), formulée par Cedegren et Sankoff (1974). Selon cette règle, la probabilité d'une forme phonétique, donc d'une réalisation concrète, est une fonction (dans le sens mathématique du terme) d'un ensemble de contextes.

Pour notre part, nous considérons que l'attention portée aux données contextualisées, à la parole spontanée, ne remet pas en cause les principes fondateurs de l'approche de la phonologie de laboratoire. Dans l'analyse du corpus, on valide les mêmes hypothèses phonologiques (ou autres) à travers l'observation d'une masse considérable de productions langagières, nous pouvons même nous faire une impression sur la fréquence d'un phénomène donné, et, donc, sur sa productivité et sa valeur cognitive (c'est-à-dire le poids des connaissances probabilistes dont disposent les locuteurs), points omis dans les travaux à partir des données construites. Le point méthodologique

crucial consiste donc à disposer de corpus suffisamment grands, dotés d'une annotation multi-niveaux et d'outils automatisés d'annotation et d'analyse : l'exploitation de cette piste permettra de lever la controverse entre parole de laboratoire et parole naturelle. Nous aborderons donc la question de l'annotation adoptée après avoir fourni des détails sur le corpus de travail.

### 3. Présentation du corpus

L'intégralité du corpus, dont une partie a été utilisée dans le cadre du présent travail, a été recueillie au laboratoire de phonétique de l'université d'État de Saint-Petersbourg lors d'un travail sur un projet de recherche international « *Spontaneous Speech of Typologically Inrelated Languages (Russian, Finnish and Dutch) : Comparison of Phonetic Properties* » (projet subventionné par le programme INTAS # 00-0915). Ce projet a été conçu dans une perspective comparative des propriétés phonétiques de la parole spontanée et de la lecture : cette orientation du projet a déterminé les modalités de recueil des données orales.

Le recueil des données comportait deux phases : d'abord, des dialogues spontanés ont été déclenchés chez dix locuteurs (5 hommes et 5 femmes) et enregistrés dans une chambre anéchoïque. Les locuteurs avaient pour interlocuteur une personne connue pour obtenir de la parole plus familière ; un sujet neutre leur a été proposé préalablement (« Les dernières vacances »), ensuite, ils étaient libres du choix des thèmes de conversation. Les enregistrements ont été faits avec un équipement professionnel ; les locuteurs portaient un micro-casque individuel AKG HSC 200 et le mixeur Behringer 802A a été utilisé. À la base des enregistrements de parole spontanée, ont été sélectionnés des fragments monologiques sans disfluences (pauses d'hésitations, reprises, allongements non-structuraux) qui ont été convertis en un texte homogène. Ce texte a été proposé aux locuteurs en lecture lors de la deuxième phase d'enregistrement. Les deux phases d'enregistrement ont été espacées d'une semaine. Les enregistrements ont eu lieu entre janvier et juin 2002.

Une telle méthodologie a été adoptée dans l'objectif de disposer d'enregistrements de qualité permettant une analyse acoustique fine. Conjointement, le contenu segmental est le même dans les enregistrements des deux types de discours ; les caractéristiques acoustiques des différents phénomènes prosodiques et des segments peuvent donc être étudiées sans recours à une procédure de normalisation. Dans le développement du projet, seuls les passages choisis des enregistrements spontanés et leurs homologues en lecture ont été soumis à une segmentation phonémique et ont reçu une annotation et une analyse. Pour notre part, nous avons choisi de

travailler sur la totalité des propos, recueillis lors de la phase spontanée, d'une locutrice, N, âgée de 20 ans, étudiante à la faculté de philologie de l'université d'État de Saint-Petersbourg.

La totalité du matériau spontané recueilli couvre 29,13 minutes de parole. Pourtant, nous avons restreint le matériau analysé dans le cadre de la présente étude sur la base des considérations suivantes : le corpus dont nous disposons est un corpus de parole spontanée contrainte. Certes, il s'agit d'une interaction entre locutrices et le discours est produit dans un objectif communicatif. Pourtant, les conditions d'enregistrement sont relativement artificielles, il n'est donc pas évident de parler sans à-coups ou même de trouver un sujet de conversation. Nous avons par conséquent exclu de l'analyse des passages que nous avons jugés artificiels. Également, de longs passages de silences, correspondant aux interventions prolongées de l'interlocutrice n'ont pas été pris en compte. Au final, nous disposons d'un corpus de 17 minutes (temps de parole, pauses comprises) sur lequel nous avons fondé la présente étude.

Nous avons choisi de travailler sur la totalité des enregistrements en espérant aboutir à une base de données enrichie comportant des annotations multiniveaux. Généralement, le terme d'« annotation linguistique » couvre le procédé d'une notation descriptive ou analytique des données linguistiques brutes : ainsi il s'agit d'un travail de dotation d'une description symbolique à un objet linguistique préexistant. Dans un contexte plus général, il peut s'agir de différents types de transcription, du marquage des parties du discours, de la structure syntaxique, annotation des cas de co-référence, *etc.* Dans la suite de notre travail analytique, le corpus annoté constitue la source de référence pour validation d'hypothèses.

#### 4. Annotation du corpus

Nous considérons que l'annotation du corpus représente une valeur ajoutée importante, toute l'originalité de la démarche résidant dans les principes qui la sous-tendent. Pourtant, l'annotation, telle que nous la concevons, doit communiquer aux données brutes une structuration particulière, dont les fondements doivent être les plus neutres possibles vis-à-vis des théories linguistiques et/ou phonologiques. Simultanément, cette annotation initiale doit mettre en œuvre un cadre dynamique, permettant l'ajout d'informations nouvelles ou bien la formalisation des théories alternatives à celles qui sous-tendent les annotations de base. Pour exploiter alors toute la structure des données, nous avons besoin de disposer de la possibilité de formuler des requêtes ou de récupérer des données brutes observables pour des traitements statistiques postérieurs (par exemple, sous la forme des scripts de PRAAT).

Nous nous sommes limitée ici à l'annotation prosodique du corpus de travail. Nous avons procédé notamment à l'annotation de l'organisation tonale et du phrasé prosodique, l'alignement temporel

à chaque niveau retenu de la constituance prosodique rendant exploitable la dimension de l'organisation rythmique et temporelle.

#### 4.1. Segmentation phonémique

Nous avons commencé le traitement de notre corpus en accomplissant une tâche d'alignement phonémique, considérant que le niveau phonémique était le pivot et la base pour l'alignement d'autres niveaux de l'analyse linguistique que nous proposons pour le corpus (*cf.* l'interprétation du niveau de squelette en phonologie autosegmentale).

Faute de disposer d'un aligneur automatique, nous avons procédé à une segmentation manuelle du corpus en phonèmes bien que cette tâche prenne beaucoup de temps. Ce choix méthodologique est fondé sur les considérations que les méthodes d'alignement automatique n'ont pas été exhaustivement testées sur de la parole spontanée.

Dans un discours spontané informel, le contenu prime sur la forme, et ceci entraîne un renforcement des processus de coarticulation : on constate alors des réalisations allophoniques non-conformes au modèle de la langue fondée sur la prononciation de mots isolés. Remarquons que les caractéristiques acoustiques sont modifiées pour les allophones, non seulement dans des positions faibles mais également dans des positions fortes.

Nous avons suivi les principes de segmentation décrits par Skrelin (1999), la démarche étant fondée sur une observation systématique de spectrogramme et d'oscillogramme, accompagnée d'une vérification auditive.

Les cas les plus litigieux sont :

- les occlusives sourdes à l'initiale : nous avons appliqué une contrainte temporelle stipulant la durée de l'occlusion supérieure à 30 ms, mais ne dépassant jamais 100 ms ;
- la séquence d'une voyelle suivie d'une consonne nasale ou liquide : nous nous sommes appuyée, dans un premier temps, sur des impressions auditives fondées toutefois sur l'examen du spectrogramme à bande large où nous avons cherché des micro-variations dans le comportement des formants et de l'intensité ;
- les désinences atones des adjectifs peuvent comporter une ou deux syllabes et subissent une réduction systématique. Très fréquemment l'identité du phonème vocalique n'est pas claire et nous avons choisi de transcrire toute la rime comme [ə] ([@] en SAMPA) ; de même, le /j/ final de la désinence du genre masculin n'était transcrit et segmenté que quand nous pouvions en observer des traces physiques sur les supports (sonagramme à bandes étroites, courbes d'intensité ou de  $f_0$ , oscillogramme) ;
- les consonnes /v/ et /l/ dans les positions intervocaliques. Nous avons pourtant appliqué le critère temporel, considérant qu'un phonème ne pouvait avoir une durée inférieure à 25 ms.

- Nous avons pu également constater des cas d'insertion de voyelles atones dans des groupes consonantiques, ce phénomène impliquant une resyllabation.

En vue d'un traitement automatique des données ainsi que d'une transférabilité et d'échanges de données, nous avons utilisé dans les annotations l'alphabet SAMPA (*Speech Assessment Methods Phonetic Alphabet*), l'adoption de l'API posant un problème technique, étant donné l'exigence de l'installation d'une police spéciale. De son côté, l'alphabet SAMPA est constitué uniquement de symboles ASCII et donc, compatible avec toutes les plate-formes sans installation préalable d'une police spécifique.

Dans les études menées à partir des corpus oraux, on retrouve plusieurs types de transcription de données : le choix d'un type particulier de transcription dépend directement du type d'étude que l'on envisage. Notamment, une distinction est faite entre une transcription phonémique large et une transcription phonémique étroite, cette dernière fixant avec exactitude les réalisations du locuteur. La transcription phonémique étroite étant un travail intéressant en soi mais très coûteux en temps, nous avons décidé de procéder plutôt à une transcription phonétique large, quasi allophonique : nous parlons d'une transcription quasi-allophonique car nous nous sommes fiée plus à notre perception du son, qu'aux principes phonologiques généraux. Nous adoptons donc 7 classes de voyelles (les variations contextuelles des consonnes étant moins problématiques à annoter, cf. tabl. 1).

**Tableau 1**

*Liste des sept classes de symboles vocaliques utilisées lors de l'annotation du corpus au niveau phonémique*

GROUPE	RÉALISATIONS
/a/	[ɑ, æ, a, ʌ, ɤ, ɐ], combinaisons /a/+j, /a/+V, /a/+ sonnante (en fonction de la réalisation)
/i/	[ɪ, i, j], combinaisons /i/+j, /i/+V, /i/+ sonnante (en fonction de la réalisation)
/o/	[o, ø, œ, ɔ], combinaisons /o/+j, /o/+V, /o/+ sonnante (en fonction de la réalisation)
/u/	[ʊ, u, ʉ, ɤ, y], combinaisons /u/ + V
/e/	[ɛ, e], combinaisons /e/ + V
/ɨ/	[i, ʉ], certaines réalisations de /a/, /u/
/ə/	[ɜ, ə, ɐ], certaines réalisations de /a/, /u/, /o/, /i/, /e/



Une fois l'alignement réalisé au niveau phonémique, nous avons procédé à un regroupement des phonèmes en unités supérieures, à savoir en syllabes, en mots phonétiques, en unités intonatives. Nous avons également appliqué l'algorithme MOMEL et le codage des points-cibles en sortie avec les étiquettes de l'alphabet INTSINT.

#### 4.2. Regroupement en syllabes

Le niveau immédiatement supérieur au niveau phonémique est le niveau syllabique. Nous évoquons donc les principes qui ont guidé notre démarche dans le regroupement de phonèmes en syllabes.

Pour la langue russe, il existe deux approches et deux ensembles de principes de syllabation qui ont été proposés respectivement par Avanessov (1954) et Scherba (1963) : précisons que ces principes de syllabation font référence explicitement aux frontières de mots. En même temps, certaines études expérimentales attestent d'une forte tendance du russe à former des syllabes ouvertes (nous pouvons y voir l'actualisation du principe de l'attaque maximale) et une tendance à la restructuration des mots phonologiques à l'oral.

Nous avons décidé de suivre au maximum le principe de la formation des syllabes ouvertes lorsque nous avons procédé au regroupement des segments en syllabes. Selon les hypothèses phonologiques, les groupes de consonnes résultants doivent obéir au principe de sonorité croissante : ce principe n'a pas toujours été respecté dans l'annotation du corpus. Nous avons plutôt respecté le principe phonotactique selon lequel un groupe de consonnes est autorisé au début de syllabe s'il est autorisé au début de mot. Un deuxième critère a été de nature perceptive : la syllabe résultante devait sonner juste, sans rupture interne perceptible.

Nous avons appliqué un traitement spécifique aux cas de chute des voyelles atones, notamment dans des positions post-accentuelles et dans le contexte des consonnes sibilantes (*cf.* fig. 1). Pourtant, ce processus ne donne pas lieu à une resyllabation au niveau du mot, du moins en perception : la consonne devient syllabique et les paramètres temporels (durée de la syllabe) sont conformes aux modèles standard.

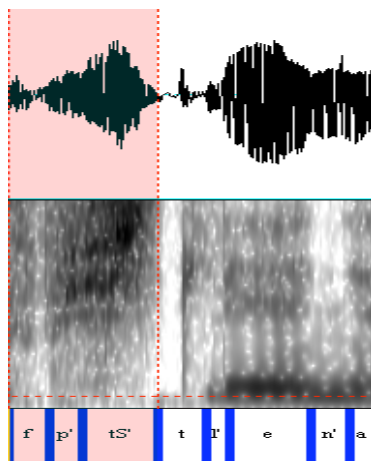


Figure 1

*Omission des voyelles atones dans le mot /fp'itʃətʲenʲɪ/*

#### 4.3. Les mots

Avant de procéder à l'annotation du corpus au niveau du mot, nous avons dû décider quels critères guideraient notre démarche. Rappelons que la phonologie prosodique classique (Nespor et Vogel, 1986) établit un parallèle entre le constituant du mot dans la hiérarchie prosodique et le mot morphosyntaxique. À cette étape, aucune distinction n'est faite entre les mots grammaticaux, privés de l'accent lexical dans les positions syntagmatiques faibles, et les mots lexicaux : cette différence est traitée lors de la formation des groupes clitiques, niveau immédiatement supérieur de la hiérarchie. À cette structure bi-niveaux la pensée phonétique russe oppose la distinction systématique entre un mot phonologique et un mot phonétique (ce terme désignant la structure comportant un mot lexical et tous les clitiques qui lui sont associés au niveau phonétique).

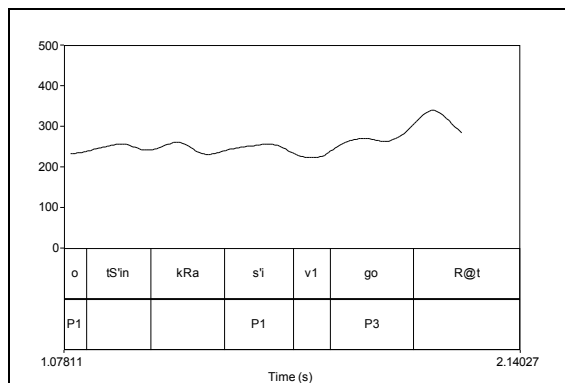
Nous avons également souligné la forte soumission (révélée dans les travaux de Kassevitch, 2002) de l'organisation rythmique du mot à un contexte rythmique et intonatif plus large. Un troisième paramètre à ne pas négliger fait référence aux critères qui fondent la formation des syllabes, à savoir le principe des syllabes ouvertes. Cependant, persiste le risque que les frontières de mots ne coïncident pas avec les frontières de constituants syllabiques. Une solution possible serait l'utilisation du concept du mot phonologique restructuré, proposé par Peperkamp (1999).

En conséquence, au niveau du mot nous disposons d'une annotation en termes d'unités proches des groupes clitiques de la phonologie prosodique classique (Nespor et Vogel, 1986 ; Hayes,

1989) : de ce choix méthodologique, des cas de non-coïncidence des frontières des mots et de celles des syllabes sont attestés dans la représentation multi-niveaux.

#### 4.4. Annotation du degré de proéminence

L'annotation du degré de proéminence nous permet de rendre compte de différents niveaux de l'organisation métrique de notre corpus. Nous avons adopté une annotation qui permet de reconstruire une grille métrique à trois niveaux (dans la lignée de Beckman et Edwards, 1990). Trois degrés de proéminence sont distingués et se résument en deux oppositions : premièrement les rapports d'opposition se créent entre les syllabes inaccentuées et les syllabes porteuses de l'accent lexical ; dans un deuxième temps, parmi les syllabes métriquement fortes au niveau du mot, nous différencions celles qui sont porteuses de la proéminence mélodique. Dans cette annotation, nous défendons le principe selon lequel tout mot lexical n'est pas forcément associé à une proéminence mélodique. Le concept de la proéminence mélodique, tel que le conçoit la tradition métrique-autosegmentale, n'a pas été entièrement assimilé dans les études sur l'intonation du russe. Lors de l'annotation, nous nous sommes principalement fiée à notre perception. Ainsi la figure 2 correspond à l'annotation du degré de proéminence pour un syntagme nominatif « **очень красивый город** » ([une] *très belle ville*) : nous considérons que c'est seulement le dernier mot (« **город** », *ville*) qui porte la proéminence mélodique ; aucune autre proéminence mélodique (accent mélodique pré-nucléaire) n'est perçue.



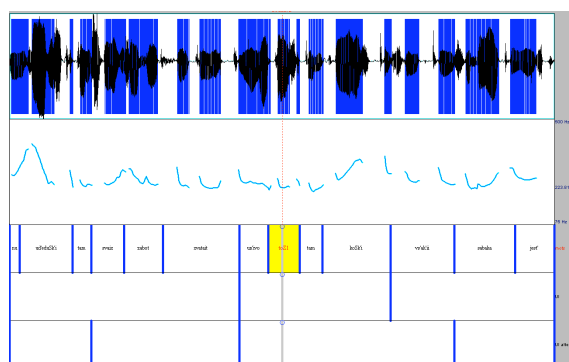
**Figure 2**  
*Exemple d'annotation du degré de proéminence*

#### 4.5. Unités intonatives

L'unité intonative paraît être une unité universellement reconnue dans l'analyse de l'organisation sonore du langage, le concept étant étroitement lié à la fonction structurelle de la prosodie. Il a été proposé une définition formelle par référence aux tons de frontières et à la présence d'une proéminence mélodique nucléaire (centre intonatif). On insiste également sur le rôle de la structure prosodique en tant qu'interface entre d'autres modules langagiers (syntaxique, informationnel, discursif). Dans le même esprit, on constate un lien étroit entre la structuration prosodique du discours et le domaine de la phonostylistique (cf. Zlatoustova et *al.*, 1986 : 79). Ainsi, selon les conditions de production, le même texte peut recevoir une structuration différente. Ces considérations nous font adopter une attitude de réticence vis-à-vis des approches qui présupposent l'identification de la structure prosodique sur la base du seul texte.

Une partie de notre corpus a reçu une annotation en syntagmes prosodiques dans le cadre du projet INTAS, mentionné ci-dessus. Nous disposons également d'une annotation du corpus par un expert francophone qui n'a aucune notion de russe : il s'agit donc d'une annotation fondée sur des propriétés formelles (mélodiques et temporelles) des unités résultantes. Nous avons pu constater des cas de divergence dans les représentations résultantes. La figure 3 reprend l'annotation d'une unité interpausale « *Ну у дедушки там своих забот хватает у него тоже там кошки всякие собака есть* » (« *Eh bien mon grand-père ne manque pas de soucis lui non plus, il a lâ-bas de nombreux chats, et il a un chien* »). L'expert français signale trois unités intonatives, il y en a quatre dans l'annotation des spécialistes russes. Une frontière est consensuelle, laquelle paraît séparer deux énoncés ; la dernière frontière paraît se décaler d'un mot dans les deux annotations : du point de vue de la grammaire l'adjectif « *КОШКИ* » (*chats*) forme un syntagme nominal avec le nom « *всякие* » (*différents*) qui porte la proéminence mélodique ; du point de vue de l'organisation prosodique, aucun indice acoustique n'est associé avec la frontière syntaxique en question. L'annotation des experts russes est alors guidée par les connaissances linguistiques générales. Une annotation alternative aurait pu être proposée, établissant une distinction entre la Macro-unité Intonative, composée de l'unité intonative et le segment d'unité intonative (Di Cristo et Hirst, 1996). Enfin, nous observons une seule frontière discordante. La justification de l'annotation de cette frontière peut être linguistique : un syntagme prépositionnel au début de phrase induit la présence d'une frontière. Il est évident qu'une proéminence mélodique est associée avec le premier mot de cette unité de performance ; or, pour respecter la contrainte structurelle de la formation du syntagme intonatif, il est nécessaire qu'un des mots du syntagme résultant « *СВОИХ ЗАБОТ ХВАТАЕТ* » (*il ne manque pas de soucis*) soit associé à un accent mélodique. Nous avons évoqué les controverses

sur l'annotation du degré de proéminence dans le paragraphe consacré à ce sujet : à notre avis, il est donc impossible, à l'étape actuelle, de prendre une décision sur la présence d'une frontière prosodique dans la position désignée.



**Figure 3**  
*Exemple d'annotation d'une unité interpausale en unités intonatives  
 (palier 2 : expert français ; palier 3 : experts russes)*

#### 4.6. Unités interpausales

Au niveau le plus concret (et le plus bas dans la hiérarchie adoptée) nous disposons de la segmentation de notre corpus en unités de performance : unités interpausales (*inter pausal units*, IPU, Koiso et al., 1997). Le terme IPU désigne un bloc de parole délimité du reste du discours par des pauses silencieuses d'une durée supérieure à 200 ms. Notons que cette unité est introduite dans une étude portant sur l'analyse de l'organisation discursive du message, et que le choix est argumenté par son caractère objectif. Ainsi, dans la transposition de la méthodologie du corpus, c'est l'unité interpausale qui peut servir de pivot du transfert des annotations pour tous les autres niveaux.

#### 4.7. Annotation mélodique : MOMEL et INTSINT

Afin d'identifier l'organisation tonale de notre corpus, nous avons appliqué les traitements suivants : la courbe mélodique a reçu une modélisation par application de la procédure MOMEL et a été annotée par la suite à l'aide de l'alphabet INTSINT.

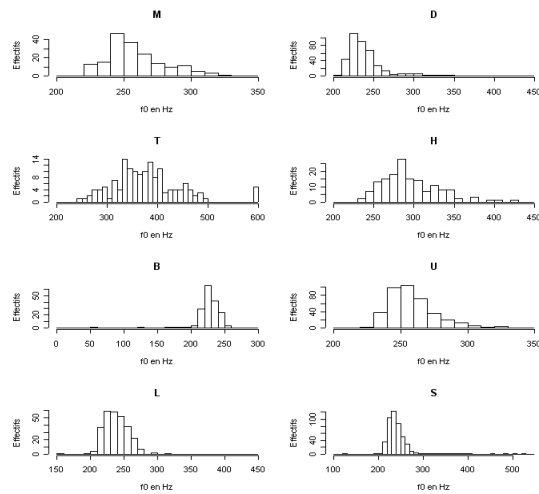
Cette annotation est semi-automatique. Nous nous sommes servie d'une implémentation de l'algorithme MOMEL sous la forme d'une fonction externe en C qui est appelée par un script de

l'intérieur du logiciel PRAAT (Boersma et Weenink, 2005). Lors de son application, ce script prévoit une possibilité de correction manuelle des points-cibles détectés. Comme nous l'avons précisé, lors de la correction manuelle, les points-cibles peuvent être supprimés, ajoutés ou déplacés.

Or, se pose la question de l'unité d'analyse pour le codage. Dans la version implémentée de l'algorithme, les points-cibles reçoivent leur codage sur la base de deux paramètres : clé (*key*) et étendue (*range*). Ensemble, les paramètres de clé (niveau) et d'étendue définissent le registre mélodique, concept qui n'est pas négligeable dans les analyses prosodiques (*cf.* Laver, 1994, sur cinq lectures possibles de ce terme). Le registre tonal permet d'évaluer les caractéristiques globales de la  $f_0$  d'un locuteur ou d'un contour mélodique. La dimension linguistique du phénomène du registre tonal se retrouve associée à l'organisation informationnelle du message (*cf.* Ladd, 1992, travaux sur l'organisation mélodique au niveau du paragraphe vocal, paraton, domaine du développement d'un topique Yule, 1980 ; Brown et Yule, 1983). Pour des raisons de facilité de traitement, notre corpus est segmenté en fichiers d'une minute environ. Or, compte tenu des remarques faites ci-dessus, pour le traitement MOMEL-INTSINT nous avons choisi de travailler avec des unités plus petites correspondant aux interventions de la locutrice. De plus, Hirst (2004) propose un traitement unifié pour le bloc de cinq énoncés, formant un tout du point de vue discursif dans le corpus Eurom-1. Compte tenu du choix de l'unité de traitement pour l'annotation de la composante tonale, nous nous sommes intéressée aux valeurs fréquentielles pour toute catégorie d'étiquettes. En effet, surtout pour les catégories absolues de points cibles qui caractérisent l'étendue tonale, nous nous sommes posé la question de savoir si ces catégories étaient bien délimitées. Nous présentons également sur la figure 4 les histogrammes des distributions de la  $f_0$  selon la catégorie d'INTSINT. L'examen de ces histogrammes nous permet de voir que la variabilité la plus élevée est propre aux étiquettes attribuées aux sommets des pics mélodiques ; les points-cibles associés à des vallées ont un comportement plus stable dans le domaine fréquentiel (*cf.* les observations similaires chez Patterson, 2000 ; Portes et Di Cristo, 2003). Néanmoins, même pour des catégories absolues (T, B, M)<sup>1</sup> nous voyons que les bandes de fréquence qui leur sont propres se chevauchent.

---

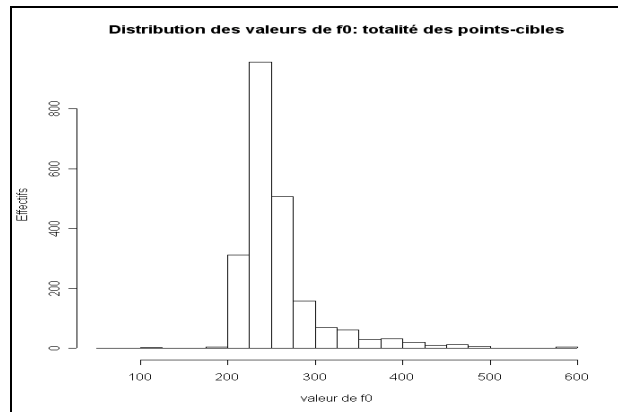
1. L'alphabet INTSINT comporte les symboles suivants : T(*op*), B(*ottom*), M(*iddle*), H(*igh*), L(*ow*), U(*pstepped*), D(*ownstepped*), S(*ame*).



**Figure 4**  
*Histogrammes des fréquences de  $f_0$  selon la catégorie INTSINT*

Nous avons également recouru à des mesures de dissymétrie et d'aplatissement pour caractériser les distributions résultantes. Le coefficient de dissymétrie permet de porter un jugement sur l'égalité de répartition autour de la moyenne. Nous prêtons attention surtout aux valeurs observées pour les symboles absolus (M, T, B) : pour les symboles relatifs, les valeurs éloignées de celles de la distribution normale paraissent cohérentes, vu leur caractère relatif et la dépendance des valeurs (*settings*) des paramètres initiaux. Nous constatons tout d'abord un coefficient de dissymétrie très élevé pour la catégorie B. En vérité, il est dû à quelques valeurs sporadiques (nombre d'observations égal à 6) dans la bande de fréquence de 100-150 Hz. Il s'agit des valeurs infra-basses pour notre locutrice, associées à des syntagmes à valeur pragmatique conclusive. Une fois ces valeurs exclues, le coefficient de dissymétrie affiche la valeur de 0,408, le coefficient d'aplatissement étant de -0,411. Prenons en compte également les coefficients de dissymétrie positifs pour les étiquettes M et T : une dissymétrie positive est révélatrice d'une longue queue à droite dans la distribution des valeurs de  $f_0$  (fig. 4). Ces tendances témoignent du fait que la locutrice « préfère » utiliser le registre élevé ; nous pouvons y voir l'effet des stratégies discursives adoptées (souci de ne pas être interrompue), mais également l'effet de l'organisation discursive en relation avec le développement de l'argumentation. Cependant, ces observations n'ont que le statut d'hypothèses et nécessitent une vérification dans une étude expérimentale spécifique.

Nous constatons également une dissymétrie positive importante pour la totalité des points-cibles (fig. 5) : la distribution a la forme d'une distribution de Weibul, plutôt que celle de la loi normale. Cette tendance peut être caractéristique de la locutrice de notre corpus. Or, elle fait penser également aux observations faites à propos de la dimension tonale de la parole spontanée avec une dominance des mouvements mélodiques ascendants. Il peut également s'agir d'une stratégie de la locutrice qui consiste à relever la ligne basse (ligne des vallées mélodiques) en raison de l'influence des contraintes discursives et des variations du débit de parole.



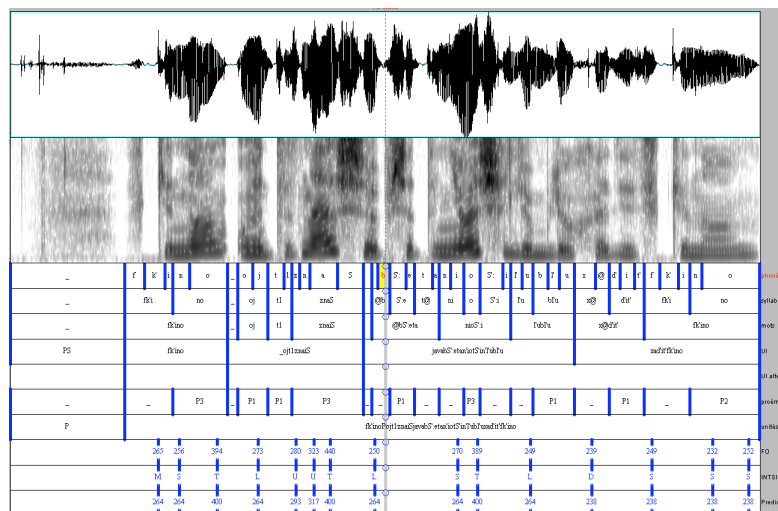
**Figure 5**  
*Histogramme des valeurs de f0 de tous les points-cibles*

## 5. Représentation multiniveaux : bilan

Le format des fichiers d'annotation que nous avons développé correspond au modèle TextGrid pour le logiciel PRAAT. Ce format d'annotation permet d'isoler les différentes sources d'informations sur des paliers différents (« tiers »), en se rapprochant ainsi des représentations typiques de la phonologie multilinéaire ainsi que des représentations hiérarchiques de la phonologie prosodique. Simultanément, le logiciel PRAAT permet une formulation simple des requêtes, ce qui facilite et automatise la récupération des données phonétiques (nous évoquerons ces possibilités lors de la discussion dans les chapitres suivants des analyses effectuées à partir du corpus).



Nous disposons donc d'un corpus doté d'une annotation multi-niveaux, comportant les niveaux des phonèmes (1), des syllabes (2), des mots (3), des unités intonatives (4 et 5)<sup>2</sup>, du degré de proéminence (6), des unités-interpausales (7), du codage INTSINT (9) avec les valeurs en Hz récupérées initialement du corpus (8) et prédites (10). Un extrait est présenté sur la figure 6.



**Figure 6**  
*Exemple d'un TextGrid d'annotation du corpus*

## 6. Observations préliminaires sur le phrasé prosodique

L'annotation multiniveaux s'avère être une ressource inestimable pour explorer les rapports entre les différents niveaux linguistiques, chacun représentant une partie des informations disponibles aux locuteurs en communication. Nous avons par conséquent décidé de proposer dans cette dernière partie de notre contribution quelques réflexions et observations portant sur la perception du phrasé prosodique en russe.

Rappelons que, dans l'annotation du phrasé prosodique, les experts avaient pour tâche de marquer les frontières des unités intonatives. Ce concept s'inscrit dans le paradigme de la phonologie prosodique, cadre de référence dans les travaux sur la structuration prosodique du flux de parole. Notamment, le phrasé prosodique reçoit dans le cadre de la phonologie prosodique une

<sup>2</sup> 2. Annotation de l'expert français pour la totalité du corpus et annotation des experts russes pour une partie du corpus.

modélisation formelle sous la forme d'une hiérarchie de constituants. Des hiérarchies multiples fondées sur des critères variés ont été proposées (Nespor et Vogel, 1986 ; Hayes, 1989 ; Selkirk, 1984 ; Beckman et Pierrehumbert, 1986). L'introduction de la hiérarchie des constituants prosodiques permet de rompre avec le parallélisme entre les structures syntaxique et phonologique. De même, la matérialité de la constituance prosodique a été démontrée dans de nombreuses études phonétiques, ainsi que dans les études neuro- et psycholinguistiques. Les chercheurs s'accordent sur l'interprétation de la constituance prosodique comme étant un cadre gérant l'implémentation des propriétés acoustiques du signal de parole.

Dans la hiérarchie classique des constituants, le syntagme intonatif est un des rares constituants définis par des propriétés acoustiques (donc, physiques) : entre autres, c'est le domaine d'un contour mélodique perceptivement cohérent et unifié. D'autre part, dans la phonologie prosodique de Nespor et Vogel (1986), l'algorithme de formation des unités intonatives s'appuie sur deux types d'informations : les informations morpho-syntaxiques et les informations sur la structure prosodique des niveaux inférieurs, cette structure étant supposée être connue, au moins partiellement, au moment de la dérivation de la structure intonative. Du point de vue de la morpho-syntaxe, il est nécessaire de disposer d'une partition de l'énoncé en termes de clauses et d'éléments appositionnels. Il est également considéré que seulement les frontières des clauses radicales (*root clauses*)<sup>3</sup> coïncident avec les frontières intonatives. Pour les clauses imbriquées, le marquage de la frontière gauche (et droite, selon la position dans la clause radicale) n'est pas obligatoire. Cependant, dans le travail de Halliday (1967), la structure clausale généralement détermine la distribution des frontières des unités intonatives.

Nous avons donc décidé, dans cette étude préliminaire, de procéder à une observation qualitative de l'articulation entre la structure clausale et la structure intonative perçue. Nos observations sont fondées sur l'annotation proposée par l'expert francophone ; cette segmentation a été réalisée sur la base de critères prosodiques formels, sans être influencée par les informations morpho-syntaxiques. L'unité initiale de notre analyse est l'unité inter-pausale (*cf. supra*). Dans un premier temps, pour chaque unité inter-pausale, nous avons déterminé si elle comprend une ou plusieurs clauses. Dans le cas des unités comportant plusieurs clauses, les frontières prosodiques internes signalées ont été classées en fonction de leur correspondance ou non-correspondance avec la frontière inter-clauses. Nous avons également recueilli des informations sur l'éventuelle présence d'autres frontières prosodiques internes, les unités comportant des pro-phrases et des marques de communication ayant été exclues de l'analyse.

---

3. On définit la clause radicale comme une clause qui n'est pas contenue dans une autre clause.

Nous résumons les informations recueillies dans les tableaux 2 et 3 (voir *infra*). Nous constatons que seulement un tiers des unités inter-pausales comportent plusieurs clauses. Ce résultat semble être la conséquence de la nature de notre corpus. En effet, en parole dialogique spontanée, les périodes d'échanges rapides de répliques alternent avec des périodes de développement étendu d'un sujet. En même temps, l'unité inter-pausale à la base de notre analyse étant une unité définie par des critères formels, elle ne correspond pas non plus à une unité de la structure discursive. Notamment, une réplique dans le dialogue peut comporter plusieurs IPU, les pauses qui séparent les IPU ayant une nature diverse : pauses structurelles *vs* pauses d'hésitation. Nous avons également pu constater que 38% des unités mono-clausales contiennent d'autres frontières prosodiques internes.

**Tableau 2**  
*Distribution des IPU en fonction du nombre de clauses qu'elles contiennent*

IPU comportant 1 clause	66%
IPU > 1 clause	34%

Si nous nous limitons aux IPU comportant plusieurs clauses (tableau 3), nous constatons que dans 38% des cas les frontières morpho-syntaxiques entre deux clauses n'ont pas reçu de marquage intonatif particulier. Compte tenu du fait que d'autres frontières prosodiques internes ont été perçues par l'expert, nous avons pu constater que connaître la structure clausale du matériau verbal permet de prédire correctement environ 45% de frontières intonatives, et ceci dans des conditions où les informations sémantico-syntaxiques ne sont pas disponibles pour l'auditeur.

**Tableau 3**  
*Marquage prosodique de la structure clausale à l'intérieur des IPU  $\geq 2$  clauses*

Marquage prosodique d'une frontière clausale	62%
Absence de marquage prosodique	38%

D'une part, le résultat obtenu corrobore l'hypothèse de la non-correspondance entre la structure prosodique et la structure syntaxique, pivot de la phonologie des domaines. Les études portant sur l'articulation prosodie  $\sim$  syntaxe constatent, que, entre autres, en présence d'indices grammaticaux, le marquage prosodique de la structure est optionnel. Des réflexions similaires sont menées dans le

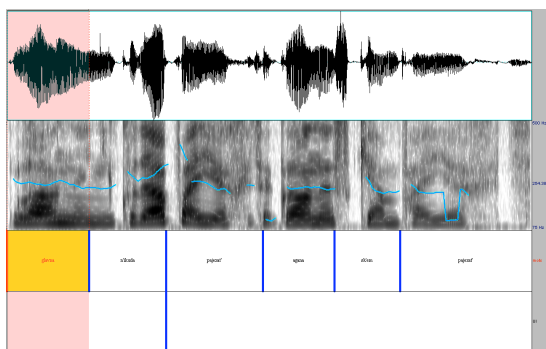
cadre de l'approche de compétition de la génération du sens avec les informations apportées par différentes composantes linguistiques (Blache et Di Cristo, 2002), où les auteurs se penchent sur la question de savoir pourquoi la prosodie présente plus de variabilité dans certains contextes par rapport à d'autres. On rejoint également la problématique de la désambiguïsation des structures morpho-syntaxiques par des moyens prosodiques. Cependant, nous constatons que plus de 60% des unités non-ambiguës reçoivent un marquage prosodique. Nous observons donc un certain parallélisme entre les structures prosodique et syntaxique, assuré par le jeu d'indices acoustiques. Certes, le matériau sur lequel nous avons mené nos analyses n'est pas suffisamment large, par suite cette hypothèse mérite une exploration approfondie dans le futur.

Des recherches futures, fondées sur d'autres corpus et menées dans le cadre de l'approche de la phonologie de laboratoire, s'avèrent d'autant plus importantes que nous avons fondé tout d'abord ces observations préliminaires sur les données de l'expert francophone, chez qui nous avons constaté une sensibilité plus marquée à la distribution de focus. Ainsi la figure 7 reprend l'annotation proposée pour l'IPU « Главное не куда поехать а главное с кем поехать » (*Le plus important ce n'est pas où aller, mais avec qui y aller*). La frontière entre deux clauses est absente dans cette annotation, cette frontière est même déplacée pour être alignée avec le mot porteur de l'accent nucléaire (en l'occurrence, l'adverbe куда, « où »). D'autre part, il semble difficile de recourir aux jugements des experts russophones, car dans les travaux expérimentaux, on dispose de peu de précisions sur la structuration du message en syntagmes : on s'appuie le plus souvent sur la structure syntaxique et on exige que les syntagmes soient des unités de sens, mais cette exigence n'est formulée qu'en termes très abstraits. La recherche d'un paradigme objectif d'annotation des frontières des unités intonatives s'impose. D. Hirst<sup>4</sup> propose une méthodologie objective, compatible avec l'organisation prosodique en lecture : comme en lecture, toute pause a une motivation structurelle et accompagne systématiquement une frontière prosodique, les positions ainsi marquées sont des sites inconditionnels de frontières prosodiques ; par conséquent, leurs réalisations acoustiques peuvent être analysées afin d'aboutir à un modèle acoustique. Cette approche est innovante dans la mesure où la méthodologie habituelle consiste à confronter l'annotation du phrasé par des linguistes selon les principes théoriques, souvent à la base des transcriptions de texte, et les productions des locuteurs. À l'étape suivante, l'algorithme d'annotation automatique pourrait être appliqué à l'annotation du corpus spontané et les

---

4. Projet de recherche 145 : Typologie prosodique : développement d'un protocole objectif d'analyse de la prosodie des langues.

annotations résultantes pourraient être confrontées aux structures morpho-syntaxique et discursive.



**Figure 7**

*Phrasé prosodique perçu pour l'IPU « Главное не куда поехать а главное с кем поехать »  
(Le plus important ce n'est pas où aller, mais avec qui y aller)*

## 7. Résumé et conclusion

Dans notre contribution, nous avons présenté un corpus de parole spontanée en langue russe. Notre corpus étant enrichi d'une annotation multiniveaux et aligné temporellement pour chaque niveau d'annotation, nous avons spécifié des principes qui ont guidé l'annotation à chaque niveau retenu. Enfin, nous avons procédé à une discussion critique de la performance de l'algorithme MOMEL-INTSINT dans l'annotation de ce corpus de parole spontanée en langue russe.

Par la suite, nous nous sommes penchée sur la question du phrasé prosodique d'autant plus actuelle pour le russe, les formalismes de la phonologie des domaines n'ayant pas encore été appliqués dans les analyses prosodiques pour cette langue. Nous avons notamment analysé l'articulation entre l'organisation clausale du corpus et le phrasé prosodique perçu par un expert francophone. Nous avons pu constater que la structure clausale (définie par référence au concept de prédictivité) ne détermine que 45% de frontières d'unités intonatives annotées. Dans nos analyses, nous avons cherché à démontrer tout l'intérêt qu'il y a à mener un travail de description formelle de l'organisation prosodique à partir d'un corpus enrichi d'annotations appropriées. Simultanément, les résultats obtenus nous poussent à nous interroger sur le statut de l'unité de *синтагма* (« syntagme »), concept de référence dans les analyses prosodiques du russe.

## 8. Références bibliographiques

- AVANESSOV, R.I. (1954). Аванесов, Р.И., О словоразделе и строении слога в русском языке. *Вопросы языкознания*, 6.
- BECKMAN, M.E. (1997). A typology of spontaneous speech, in Y. Sagisaka ; W. N. Campbell & N. Higuchi (eds) *Computing Prosody*, New York: Springer-Verlag, p. 7-26.
- BECKMAN, M.E. & PIERREHUMBERT, J.B. (1986). Intonational Structure in Japanese and English, *Phonology Yearbook*, 3, p. 255-309.
- BECKMAN, M.E. & EDWARDS, J. (1990). Lengthenings and Shortenings and the Nature of Prosodic Constituency, in Kingston, J. ; Beckman, M.E. (eds), *Papers In Laboratory Phonology I. Between the Grammar and Physics of Speech*, p. 152-178.
- BLACHE, Ph. & DI CRISTO, A. (2002). Variabilité et dépendances des composants linguistiques, *Actes de la Conférence Traitement Automatique du Langage Naturel (TALN-2002)*.
- BOERSMA, P. & WEENINK, D. (2005). *Praat: a system for doing phonetics by computer*, <<http://www.fon.hum.uva.nl/praat/>>.
- BROWN G. & YULE G. (1983). *Discourse analysis*, Cambridge: Cambridge University Press.
- CEDERGREN, H. & SANKOFF, D. (1974). Variable rules: Performance as a statistical reflection of competence, *Language*, 50, p. 333-355.
- D'IMPERIO, M. (2005). La Phonologie de Laboratoire : finalités et quelques applications, in Nguyen, N. ; Wauquier-Gravelines, S. ; Durand, J. (éds), *Phonologie et phonétique: Forme et substance*. Paris : Hermès, p. 241-264.
- DI CRISTO, A. & HIRST, D. (1996). Vers une typologie des unités intonatives du français, *Actes des XXIIes Journées d'Étude sur la Parole*, Avignon, p. 219-222.
- HALLIDAY, (1967). *Intonation and grammar in British English*, The Hague: Mouton.
- HAYES, B. (1989). The prosodic hierarchy in meter, in P. Kiparsky & G. Youmans (eds), *Rhythm and meter*, Orlando: Academic Press, p. 201-260.
- HIRST, D. (2004). Lexical and Non-lexical Tone and Prosodic Typology, *Proceedings of International Symposium on Tonal Aspects of Language*, 28-30 mars 2004, Beijing, China, p. 81-88.
- KASSEVITCH, V. (2002). Касевич, В.Б., О фонетическом слове. *Вопросы фонетики*, 4.
- KOISO, H. ; HORUICHI, Y. ; TUTIYA, S. ; ICHIKAWA, A. & DEN, Y. (1998). An analysis of turntaking and backchannels based on prosodic and syntactic features in Japanese map task dialogs, *Language and Speech*, 41, p. 295-321.
- LADD, D.R. (1992). An Introduction to Intonational Phonology (Prosody), in Docherty, G.J. ; Ladd, D. R. (eds), *Papers in Laboratory Phonology II. Gesture, Segment, Prosody*, p. 321-334.
- LAVER, J. (1994). *Principles of Phonetics*, Coll. Cambridge Textbooks in Linguistics, Cambridge: Cambridge University Press.

- LEVELT, W. (1989). *Speaking. From Intention to Articulation*, Cambridge: Cambridge University Press.
- LICKLEY, R. ; SCHEPMAN, A. & D.R. LADD (2005). Alignment of “phrase accent” lows in Dutch falling-rising questions: Theoretical and methodological implications, *Language and Speech*, 48, p. 157-183.
- NESPOR, M. & VOGEL, I. (1986). *Prosodic Phonology*, Dordrecht: Foris Publication.
- PATTERSON, D. (2000). *A linguistic approach to pitch range modelling*, Thèse de doctorat, Edinburgh University.
- PEPERKAMP, S. (1999). Prosodic words, *GLoT International*, 4(4), p. 15-16.
- PIERREHUMBERT, J. ; BECKMAN, M. E & LADD, D. R. (2001). Conceptual Foundations of Phonology as a Laboratory Science, in Burton-Roberts, N. ; Carr, P. and Docherty, G. (eds) *Phonological Knowledge*, Oxford: Oxford University Press, p. 273-304.
- PORTES, C. & DI CRISTO, A. (2003). Pitch range in spontaneous speech: semi-automatic approach versus subjective judgement, *Proceedings of the 15th ICPbS*, 3-9 août, Barcelone, Espagne, p. 583-586.
- SCHERBA, L.V. (1963). Щерба, Л.В., *Фонетика французского языка*. 7-ое изд., Москва.
- SELKIRK, E. (1984). *Phonology and Syntax : The Relation Between Sound and Structure*, Coll. Current Studies in Linguistics Series, 10, Cambridge: The MIT Press.
- SKRELIN, P.A. (1999). Скрелин, П.А., *Сегментация и транскрипция*. Санкт-Петербург.
- YULE, G. (1980). Speaker’s topic and major paratones, *Lingua*, 52, p. 33-47.
- ZLATOUSTOVA, L. ; ПОТАРОВА, R. ; ТРОУНИНЕ, V. (1986). *Phonétique générale et appliquée*, Moscou : Éditions de l’université de Moscou (en russe).

# UNE BASE DE DONNÉES SUR LES TRONCATIONS INVOLONTAIRES DE MOTS EN FRANÇAIS PARLÉ

Berthille Pallaud

## 1. Introduction

C'est à partir d'études sur la production et la perception « d'erreurs de langage » (Pallaud, 2001, 2002) dans des corpus de français parlé que s'est imposée la nécessité d'une clarification de ces phénomènes si fréquents à l'oral et de la constitution d'une base de données propre aux troncations de mots. Cette base de données a permis une analyse approfondie des amorces de mots identifiées dans des énoncés oraux transcrits et a également permis de préciser les apports de la notion de disfluence<sup>1</sup> et les effets du développement, dans ce secteur, des nouvelles technologies, en particulier lors des dix dernières années.

## 2. Du lapsus aux disfluences

Le lapsus, que le sens commun (du moins si on en juge par ce qu'en disent les médias, par exemple) relie sans hésiter au domaine de la psychanalyse, est un terme présent dans tous les dictionnaires de psychanalyse mais absent des lexiques ou dictionnaires de linguistique actuels. Pourtant, comme on va le voir, certains linguistes se sont intéressés aux lapsus, en particulier depuis les années 80 aux États-Unis, et quelque dix ans plus tard en France. Ce renouveau d'intérêt tient aux évolutions théoriques dans le domaine de la linguistique (surtout en phonétique) et celui de la psycholinguistique, majoritairement sous l'impulsion du cognitivisme. En revanche, la conception des études et spécifiquement la constitution de la base de données permettant les analyses, sont restées conformes à la tradition : elles sont fondées sur un recueil patient et souvent sporadique (dit « à la volée ») des phénomènes recherchés.

---

PALLAUD, Berthille (2006), Une base de données sur les troncations involontaires de mots en français parlé, *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, vol. 25, p. 173-184.

---

1. Le terme de disfluence, employé pour décrire les trébuchements de la parole dans les énoncés de français parlé *standard* est à l'origine un anglicisme.



C'est ainsi que le philologue Meringer et le psychiatre Mayer avaient procédé en leur temps (1895, 1908, 1923) à la collecte des matériaux de leur étude. Ils sont considérés comme ayant été les premiers à avoir véritablement étudié les lapsus et constitué à la fin du XIX<sup>e</sup> siècle un corpus de 8800 lapsus, dont la moitié sont des *lapsus linguae* (les autres sont des *lapsus calami* et *lectionis*)<sup>2</sup>. Il n'est pas une étude traitant des lapsus ou ratés de langage qui ne fasse référence aux travaux de ces auteurs allemands.

Les chercheurs qui ont étudié les lapsus en langue française, dans les années 90, ont constitué des recueils limités au cadre strict de la production langagière orale. Ils ont utilisé une méthode qui s'apparente au recueil à la volée<sup>3</sup> puisqu'ils s'appuient sur la manipulation de longs corpus (Rossi et Peter-Defare, 1998) et non sur l'extraction systématique de phénomènes à partir de transcriptions de corpus.

Au Laboratoire Parole et Langage, une série de recueils et d'études a abouti à plusieurs publications dont les principales sont celles de Rossi et Peter-Defare (1995, 1998) qui portent sur environ 4000 exemples d'erreurs de langage. Le travail d'E. Peter-Defare (1993) est la première étude systématique et linguistique dans ce domaine en langue française. Cette base de données est accessible sur le site du LPL<sup>4</sup>.

Arnaud (1997) a publié également une première analyse sur son propre recueil de lapsus qu'il qualifie de « naturels » (pour les distinguer des lapsus déclenchés en laboratoire, ou suscités artificiellement). La notation de ces 2400 exemples a été faite sur le mode « systématique » (*sic*) (« tout lapsus entendu a été noté », p. 308) et cela dans les secondes qui suivaient leur production (« afin d'éviter la rapide dégénérescence de la trace mémorielle », p. 308).

Fénoglio (1997), dans une approche quasi clinique pour aborder l'énonciation en discours oral, s'est appuyée sur un corpus de 25 lapsus relevés dans des entretiens enregistrés. Le lapsus apparaît comme une donnée d'articulation entre discours et parole.

Les travaux de Pillon (1998), à l'aide de lapsus déclenchés artificiellement, montrent l'implication de la morphologie dérivationnelle des mots dans les processus de production de la parole.

---

2. *Lapsus linguae* : lapsus oral ; *lapsus calami* : lapsus écrit ; *lapsus lectionis* : lapsus en lecture.

3. « Il s'agit de lapsus relevés soit dans les conversations de la vie courante, soit dans des dialogues et tables rondes diffusés sur les médias entre 1992 et 1996. Nous n'avons retenu que les lapsus dont nous étions sûrs et pour lesquels nous pouvions obtenir un contexte suffisant où pouvait être identifiée l'origine éventuelle. » (Rossi & Peter-Defare, 1998, p.17).

4. Liens : Outils/Bases de Données/Lapsus en français : <<http://www.lpl.univ-aix.fr/index.php?id=240>>.

Une quarantaine de lapsus sont accessibles. Les collègues qui souhaiteraient avoir accès à la totalité de la base pour des besoins de recherche sont invités à contacter Mario Rossi.

Ainsi, mis à part les travaux sur les lapsus déclenchés artificiellement, les études sur les lapsus ont en commun d'avoir été construites sur un ensemble d'exemples recueillis plus ou moins à la volée avec, la méthode l'implique, une définition *a priori* de ce que signifie un lapsus : à savoir une erreur involontaire de langage, se manifestant (sous l'effet d'un processus inconscient) par la production d'un mot pour un autre, incongru ou déformé.

- Un mot pour un autre :  
*oh mon écharpe + je l'ai prêtée à Gisèle parce qu'elle avait froid et je l'ai gardée eh elle l'a gardée*  
(corpus Pallaud 1997, (56))
- Un mot incongru :  
L1- *tu as entendu le Sud du Portugal est relié par le plus grand port d'Europe*  
L2- *tu veux parler d'un pont je suppose*  
L1- *oui bien sûr* (corpus Pallaud 1997, 125)
- Un mot déformé :  
*ma mère a dit surtout cache ce livre que Poupette ne le lise pas eh bien ce livre je l'ai [li] caché dans le lit*  
(corpus Pallaud 1997, 124)

Cette méthode prévoit de constituer des recueils *chemin faisant* au gré des conversations ou émissions de radio écoutées par les collecteurs<sup>5</sup>. Ne reposant pas sur des corpus oraux transcrits, elle ne prétend pas être exhaustive. Par ailleurs, l'identification même de ces phénomènes relève tantôt d'une analyse sémantique (l'incongruité d'un terme) tantôt de la détection des transgressions de règles morphologiques (un mot déformé).

Dans une étude entreprise sur les lapsus (Pallaud, 1999) une comparaison a pu être conduite sur deux types de recueils de disfluences : l'un à la volée, l'autre rassemblant les achoppements extraits (de façon systématique et exhaustive) d'un corpus enregistré (corpus Bertuzzi, 1997<sup>6</sup> et transcrit selon les conventions du Groupe Aixois de Recherche en Syntaxe (GARS)<sup>7</sup>. Dans les deux cas, les énoncés sont des propos tenus dans le cadre d'entretiens ou de conversations. L'étude du corpus Bertuzzi (1997) s'est révélée essentielle car la possibilité de lire et de relire la transcription des paroles prononcées a permis d'éviter de construire une analyse sur des faits déjà définis (par

---

5. Le relevé de lapsus est plus facile à faire lors d'écoutes d'émissions radio. Il n'est, en effet, pas facile d'adopter une attitude d'observateur tout en participant à une conversation avec d'autres personnes (Gadet, 2000, 2003).

6. Il s'agit du corpus Bertuzzi 97 qui est la transcription intégrale de trois enregistrements d'une personne évoquant les conditions d'exercice de son métier de fleuriste à Paris. Cette étude a été faite dans le cadre d'une soutenance de maîtrise au département de linguistique française à l'université d'Aix-Marseille I.

7. Ce qu'Arnaud (1997) nomme recueils extensif (à la volée) et intensif (sur corpus enregistré).

exemple, du type « un lapsus est un mot pour un autre »<sup>8</sup>). La lecture de l'énoncé permet de constater des phénomènes qui sont bien sûr perceptibles pour une écoute attentive et avvertie mais souvent impossibles à noter à la volée. L'étude sur corpus présente l'avantage de permettre surtout un relevé systématique de faits qui autrement paraissent rares (parce que discrets et rapides). Il est évident que le recueil à la volée n'est pas systématique (puisque'il est impossible de noter tout achoppement entendu), mais, de plus, il est apparu qu'il a évolué sous l'influence des constats faits lors de la comparaison des recueils. En d'autres termes, si au départ cette activité de recueil semblait simple, l'analyse du corpus enregistré a révélé des phénomènes insoupçonnés. Cela ne manqua pas de modifier le recueil à la volée ; ces phénomènes, passés inaperçus dans la vie quotidienne, devenaient faciles à relever. S'il y a une certitude méthodologique, c'est évidemment qu'un tel recueil de faits linguistiques ne peut être ni constant ni systématique (c'est bien ce qu'a souligné Gadet dans ses réflexions « *Derrière les problèmes méthodologiques du recueil des données* », 2003).

En particulier, bien que la fréquence des amorces de mots puisse beaucoup varier d'un corpus à l'autre, cette étude (Pallaud, 1999) a montré le biais qui peut, vraisemblablement, être introduit par un recueil à la volée et la nécessité qu'il peut y avoir à bénéficier de corpus enregistrés si on veut étudier la place de ces phénomènes dans l'élaboration des énoncés oraux. En effet, 59% des erreurs de langage dans le corpus enregistré sont des amorces de mots alors que dans ces types de recueils, le pourcentage tombe à 10,8%. En conséquence, les erreurs du type « un mot pour un autre » sont, sans doute, très surestimées dans les recueils à la volée, alors que les hésitations se traduisant par des amorces (amorces inachevées, corrigées et complétées) se trouvent sous-estimées.

### 3. Du recueil à la volée à l'extraction des phénomènes linguistiques

Les nouvelles technologies, en facilitant les enregistrements et le stockage des données sonores, ont permis le développement d'une culture des ressources langagières<sup>9</sup>. Il n'a plus été question de recueil à la volée mais d'extraction automatique ou non et de descriptions systématiques dans des corpus souvent volumineux (évalués en millions de mots). Le terme de *lapsus* disparaît alors au profit de ce qui a été nommé disfluences c'est-à-dire toutes les marques de trébuchement dans les énoncés. En renonçant à une définition *a priori* du phénomène lapsus, on peut réinsérer ce type de

---

8. La langue fourche de bien des façons et les achoppements ne se limitent pas au glissement d'un mot vers un autre. De plus, à l'oral, la notion de mot (sans parler de la notion de la phrase) soulève également des problèmes d'identification.

9. La Délégation Générale à la Langue Française, par exemple, a financé, dans les années 90, la constitution d'un Corpus de Référence de Français Parlé sur toutes les régions de France (équipes aixoises dirigées par Claire Blanche-Benveniste puis Jean Véronis).

phénomènes dans le champ des spécificités de l'oral. Ce faisant, on en revient à l'origine du terme latin *lapsus* qui, tout comme celui qu'utilisent les anglo-saxons, *slip (of the tongue)* évoque le trébuchement, l'achoppement.

C'est qu'en effet, les études sur les corpus oraux montrent combien les vers de Boileau :

*Ce qui se conçoit bien s'énonce clairement.  
Et les mots pour le dire vous viennent aisément.*

ne décrivent pas la situation d'énonciation. Le locuteur construit son énoncé au travers de différents procédés qui montrent à l'évidence que l'écoulement de la parole n'est pas fluide.

Si la fluidité d'un énoncé oral se mesure à la régularité rythmique dans sa production, il est clair que les énoncés oraux ne sont pas fluides mais disfluents (Shriberg, 1999 ; Pasdeloup, 1992). Tout locuteur produit de la parole avec une certaine variabilité dans le débit des mots, des pauses (silencieuses ou non, Duez, 2001), des allongements d'éléments linguistiques, *etc.* De façon générale, ces phénomènes, peu perçus par les locuteurs en présence (y compris par celui qui parle), déclenchent très rarement des commentaires. Certaines de ces disfluences dans la parole, au demeurant très fréquentes, se caractérisent par une interruption (notée IP) dans l'énoncé, que ce soit au niveau morphologique (l'amorce de mot) ou à la frontière de mots (ce qui est le cas du phénomène de bribe suivi ou non de répétition de mots ; Schriberg, 1999 ; Henry, 2002) :

- (1) Amorce de mot : *c'est vrai que c'est pas b-(IP) beau d'associer les deux* (Arborign, 5, 14)
- (2) Bribe suivie de répétition : *ils auront leur propre (IP) leur propre langage* (Laurent, 1, 2)

Le résultat de la transcription d'un énoncé oral diffère en bien des points d'un énoncé écrit, du moins de la forme achevée d'un énoncé écrit<sup>10</sup>. Les locuteurs qui se sont prêtés au jeu de l'enregistrement et qui ont eu accès par la suite au texte transcrit ont été au mieux surpris mais le plus souvent choqués par « la façon dont ils ont parlé ».

Les avatars spécifiques à l'énoncé oral (donc, absents, sauf exception, des textes écrits) et repérables dans une transcription se limitant au texte du discours sont de divers ordres :

- les reprises ou faux départs  
*J'avais beaucoup de mal à à gérer ce ce genre de difficulté(s)* (corpus C5aBelfo, 3,1)
- les pauses remplies ou silencieuses  
*C'est comme toute euh tout ce qu'on a envie de faire découvrir* (corpus C5bBelfo, 16,1)
- les incises autonymiques ou interjections

---

10. Les études sur les manuscrits ou les brouillons de textes publiés témoignent du patient travail d'écriture et de réécriture fourni par l'écrivain (*cf.* Lebrave).

**Ouais disons** *que ça a quand-même beaucoup beaucoup* **enfin** *pas ma- em-* **comment dire** *ça a énormément évolué* (corpus C41Cstra, 41,3)

- les interruptions de mots ou de syntagmes  
*C'est quand même tout à fait intéressant à ra-* **rappeler en passant** (corpus C41Astg, 3,1)  
*Comment ils appellent ça le liquide donc c'est mh le comptage du + le ni-* **comment dire ça le léchage euh** (corpus C41Cstg, 20,5)
- les ruptures syntaxiques  
**Là c'était c'était c'était-** *il se trouvait que euh la fondation / était, est/ en Allemagne* (corpus Pariscen, 9,1)
- les remplacements de mots ou les néologismes souvent nommés *lapses*  
*Si j'ai tant attendu pour* **attendre** *ce livre pour* **écrire** *ce livre* (Pallaud, 1997, 142)  
*j'ai l'impression que je vais* **apprendre** *que je vais* **apprendre** *à vivre avec toi* (Pallaud, 1997, 28)

Comme l'ont souligné Clark et Wasow (1998), ces phénomènes, qui témoignent d'une difficulté dans la fluidité verbale et que les linguistes ont appelé disfluences (par opposition aux phénomènes dits pathologiques de disfluence)<sup>11</sup> ont en commun la même structure de fonctionnement : une interruption se produit dans la fluidité de la parole que le locuteur va dépasser de diverses manières. Cette interruption est marquée en son point de rupture par des phénomènes phonétiques (travaux de Schriberg, 1999, par exemple).

Les transcriptions qui adjoignent au texte la description de différents éléments prosodiques et phonétiques (comme l'intonation, l'accent, la durée des syllabes, par exemple) montrent que les paroles du locuteur diffèrent en certains points de ce qui est décrit pour un locuteur idéal dit « standard ». Ces marques acoustiques accompagnent en particulier, on l'a vu, les points d'interruptions des énoncés. Il reste qu'en dehors de ces passages, il est possible de trouver des disfluences intonatives ou phonétiques qui introduisent des changements de sens pour l'énoncé comme dans l'exemple suivant que l'on transcrirait et ponctuierait alors ainsi selon l'intonation adoptée :

*Ne reviens pas trop tard*  
*Ne reviens pas ; trop tard.*

#### 4. La base de données sur les tronctions de mots

##### 4.1. Les recueils de corpus

La base de données ne concerne que les phénomènes de troncation de mots. Ils ont été identifiés dans une partie du « Corpus de référence de français parlé ». La constitution de ce recueil de production langagière répondait à une requête de la Délégation à la Langue Française (ministère de

---

11. Par définition les *disfluences* sont la marque d'un dysfonctionnement dans la parole alors que les *disfluences* ne sont pas liées à une pathologie.

la Culture), qui l'a totalement financée. La réalisation de ce projet avait été confiée en 1998 à l'URA 6060 du CNRS (sous la responsabilité de Claire Blanche-Benveniste), et à partir de 2000, le projet a été pris en charge par la Jeune équipe DELIC (DEscription Linguistique Informatisée sur Corpus), dirigée par Jean Véronis. L'objectif de ce corpus était de mettre à la disposition de la communauté de linguistes des témoignages de la langue française parlée aujourd'hui, dans les principales villes de l'hexagone. Des enregistrements (132) de locuteurs choisis selon des critères précis (âge, niveau d'éducation, *etc.*), dans une quarantaine de villes de France, représentant des types de parole variés (parole publique, professionnelle ou privée) sont transcrits selon les conventions du GARS. La totalité est estimée à environ 50 heures de parole, soit environ 400000 mots<sup>12</sup>. La transcription et la segmentation ont été réalisées majoritairement à l'aide du logiciel TRANSCRIBER 1.4.2.<sup>13</sup>, programme développé par Claude Barras spécialement pour la transcription et l'alignement de corpus oraux.

Depuis les années 70, l'équipe du GARS d'Aix a développé pour ses études une méthode de transcription des corpus de français parlé contemporain (Blanche-Benveniste et Jeanjean, 1987). Les deux conventions principales de cette méthode, à savoir une transcription orthographique et l'absence de ponctuation, ont été adoptées par la plupart des concepteurs de méthodes actuelles pour la transcription et l'édition de grand corpus (Blanche-Benveniste, 1997). Le souci d'établir un document linguistique a prévalu et avec lui, le refus d'apurer le texte de l'oral de ses soi-disant scories, même si la lecture de ces documents peut sembler au premier abord inconfortable pour un lecteur non averti :

*C'est l'accumulation des procédés caractéristiques des « avant-textes » de l'oral qui rend la lecture des transcriptions si inconfortable : retouches de toutes sortes, « hésitations », « énoncés inachevés », « ruptures de construction », pauses fréquentes, etc. (Blanche-Benveniste et Jeanjean, 1987, p. 162).*

Indications prosodiques et transcriptions phonologiques ne sont employées que dans certains passages (par exemple pour les énoncés ne pouvant recevoir une interprétation). Encore ne sont-elles pas portées dans le texte mais renvoyées en note (à l'exception des transcriptions d'enregistrement d'aphasiques à propos desquelles cette difficulté d'interprétation est si fréquente parfois que la lecture en serait trop inconfortable).

---

12. Pour le français parlé, il offre une base de comparaison avec les corpus de français parlé hors hexagone, le corpus Valibel en Belgique (dix fois plus important) et le corpus d'Ottawa-Hull au Canada ; il représente l'amorce de la réalisation d'un corpus comparable à celui du *British National Corpus*, pour l'anglais parlé.

13. Logiciel téléchargeable gratuitement sur le site : <<http://www.etca.fr/CTA/gip/Projets/Transcriber/>>

Selon les conventions du GARS, qui prévoient une transcription orthographique des énoncés oraux, les amorces de mots sont notées par un trait d'union collé au fragment du mot et donc identifiable automatiquement de façon univoque.

Tous ces enregistrements, sauf deux, ont été conduits en privé, selon la méthode de l'interview la moins directive possible, et ne rassemblaient que deux locuteurs adultes (l'interviewer et l'interviewé). Les deux corpus faits en public étaient des exposés improvisés devant un groupe de 40 personnes.

#### 4.2. Les troncations involontaires de mots

A l'heure actuelle, les troncations de mots ont été extraites d'un sous-ensemble de 20 corpus recueillis, à l'exception de l'un d'entre eux, dans le cadre de l'enquête précitée. Il comporte 105000 mots. Ces corpus ne sont pas étiquetés. Afin de pouvoir analyser les énoncés du seul locuteur interviewé (et non ceux du locuteur interviewer), ils ont été extraits des 20 corpus-sources pour constituer un deuxième sous-ensemble de 20 corpus interviewés. Chaque corpus correspond alors à un seul locuteur. C'est dans ce deuxième sous-ensemble qu'ont été recherchées les amorces de mots, à l'aide du logiciel CONTEXTES (établi par J. Véronis). Ce logiciel a permis de recenser les amorces de mots présentes, entourées d'un contexte antérieur et postérieur de 10 à 30 mots chacune. Les énoncés contenant ces fragments de mots sont au nombre de 441.

Si l'on se fonde sur un débit moyen de 200 mots/mn, le nombre total de mots étant de 105000 mots, la durée totale de l'ensemble de ces corpus est de 7h51 mn. La longueur moyenne des corpus est de 3080 mots, soit une durée moyenne de 16 mn. ; les valeurs extrêmes de ces durées sont de 1307 et 4931 mots. On retrouve, dans ce sous-ensemble de corpus, la fréquence moyenne d'apparition des amorces de mots dans un énoncé (1 amorce/57s. ; variation individuelle de 1/23 s. à 1/8 mn.).

Les 441 énoncés extraits de ces corpus interviewés sont rassemblés dans un tableur EXCEL. L'étiquetage (non automatique) est fait, dans un deuxième temps : les 441 énoncés comportant les troncations de mots sont disposés dans une même colonne (événements). Chaque ligne ne contient qu'un exemple de troncation resitué dans son contexte. Les colonnes à droite de cette première colonne sont utilisées pour permettre l'analyse des phénomènes. Chacune correspond à la description de l'amorce de mot en fonction d'un paramètre (*cf.* tabl. 1). Par exemple, le premier paramètre qui a été étudié dès le début des travaux sur ce type de phénomène (Pallaud, 2002) concerne le type d'interruption lexicale. En effet, il a été possible de distinguer trois sortes d'interruptions lexicales selon que le fragment de mot se trouve complété, modifié ou laissé inachevé :

- Fragment complété :  
Tropr102, 3,3 on va **com-** on va **commencer** à les + à le houspiller un peu
- Fragment modifié :  
Tropr101, 1, 3 euh et c'est vrai en plus on (n') a **jam-** on (n') a **pas** compris
- Fragment inachevé :  
Pariscen, 1, 5 alors il y a les O.N.G. + et puis les les **gouver-** les qui qui qui donnent des livres gratuitement

**Tableau 1**

*Exemple d'extrait du tableau obtenu sur EXCEL*

n°	Corpus	Page, ligne	Fragment de mot et son contexte	Type de troncation	Catégorie grammaticale	Type de mot	Rupture syntaxique
298	Pariscen	2,2	donc j'ai j'ai proposé un plan de formation + <i>euh</i> <b>se-</b> <i>euh</i> <b>segmenté</b> + par euh avec tous les différents savoir-faire de l'éditeur	complétée	verbe	mot plein	continuité
276	Tropr102	3,7	L'Italie a d'ailleurs a fait une loi - pour les repentis -- qui est bon qui était très contestée etc. mais qui a porté <b>beauc-</b> qui a <b>bien</b> porté ses fruits	modifiée	adverbe	mot-outil	continuité
274	Tropr102	3,5	il y avait un reportage à la télévision là que j'avais vu ça à la <b>tél-</b> c'était vachement intéressant	inachevée	nom	mot plein	rupture syntaxique

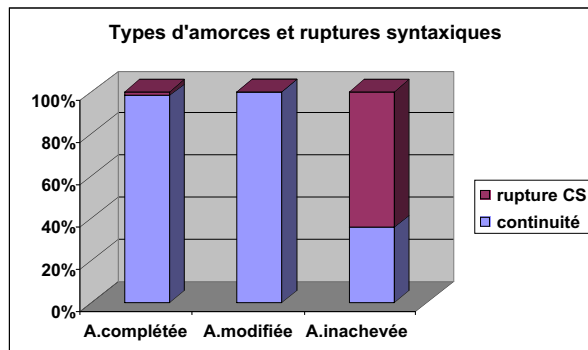
Pour chaque énoncé, plusieurs paramètres peuvent être distingués et leurs interactions étudiées. Par exemple, pour l'étude des effets dans le contexte droit par la troncation de mot (Pallaud, 2006), les paramètres observés ont été les suivants :

- la position syntaxique du fragment et sa localisation dans le syntagme
- la localisation de la reprise, quand il y en a une, après l'interruption lexicale
- la poursuite ou l'interruption de l'énoncé après la reprise

Le tableur EXCEL permet des quantifications et des représentations graphiques rapides. Un des objectifs de cette étude était de préciser l'influence du type de troncation (amorce complétée, modifiée et inachevée) sur la poursuite ou l'interruption de l'énoncé après la reprise. La fonction-



filtre de ce tableur permet de croiser ces deux variables et de déterminer par un  $\text{Khi}^2$  <sup>14</sup>, par exemple, quels sont les effets de l'une sur l'autre. A quelques exceptions près, les trois-quarts des amorces de mots complétées ou modifiées ne sont pas suivies d'une rupture syntaxique. Il n'en est pas de même pour le quart restant, laissé inachevé : 65% de ces troncations sont suivies d'une rupture de construction verbale (fig. 1).



**Figure 1**

*L'influence du type de troncation sur le devenir de la proposition*

Par ailleurs, l'étude de la position syntaxique de la troncation de mot a consisté à déterminer sa localisation avant (ou sur le verbe) ou après le verbe recteur. Cette variable a également une influence sur l'impossibilité pour le locuteur de maintenir la continuité de son énoncé. Les cas de rupture de construction verbale sont deux fois plus nombreux lorsque la troncation de mot se place avant et sur le verbe qu'après le verbe. Autrement dit, plus la troncation a lieu tôt (avant et sur le verbe), plus l'énoncé risque de s'interrompre<sup>15</sup>. En revanche, les cas de continuité syntaxique sont aussi nombreux, quelle que soit la localisation du fragment de mot.

14. Le test du  $\text{Khi}^2$  permet de décider avec une certaine probabilité si des échantillons (ici fragments de mot complétés, modifiés ou inachevés) appartiennent à une même population au regard du type de mot, par exemple. Les résultats de cette analyse sont négatifs : un mot plein a autant de chance d'être complété qu'un mot-outil. En revanche, le mot plein sera plus souvent modifié et moins souvent laissé inachevé que le mot-outil (Pallaud, 2004). Ces deux types de mot ne sont donc pas, statistiquement, de la même population.

15.  $\text{Khi}^2=4,40$  ;  $p<.05$  ; d.d.l.=1 ; le résultat du test montre que la différence est significative avec un risque d'erreur de 5%.

## 5. Conclusion et projet

Le commentaire qui accompagne la production d'un *lapsus* souligne la surprise du locuteur : « ça m'a échappé ». Or, ce que montrent les études actuelles sur les corpus oraux est que la parole semble bien échapper le plus souvent au locuteur et que cela n'a rien de pathologique. L'énoncé ne peut être produit que de cette façon : le locuteur égrène en le découvrant son énoncé telle l'araignée qui sécrète au fur et à mesure le fil auquel elle est suspendue. Parler témoigne du maillage des mots qui surgissent à l'appel de ce qui est déjà dit et va être dit. Le locuteur poursuit son énoncé tout autant en le complétant qu'en le rattrapant mot après mot et même en l'anticipant. L'équivocité du terme *poursuivre*, souvent employé à propos de l'énonciation (« poursuis ce que tu as commencé à dire ») convient bien à cette situation dont la caractéristique première est d'être équivoque.

Le projet à court terme est d'augmenter la base de données des troncations involontaires de mots grâce aux vingt autres corpus disponibles et surtout de lier cette base au contexte de son corpus et à la source sonore. Ce montage permettrait de pouvoir, à tout moment, réécouter l'énoncé choisi (l'extrait contenant le fragment de mot) sans avoir à revenir au corpus initial. Cette base de données contiendrait donc les extraits transcrits comportant une amorce de mot et la version sonore de ces extraits.

## 6. Références bibliographiques

- ARNAUD, P.J.L., (1997), Les ratés de la dénomination. Typologies dans les lapsus, in Boysson & Thoinon (éds), *La dénomination*, PUL, Lyon, 307 p.
- BLANCHE-BENVENISTE, C. ; JEANJEAN, C. (1987). *Le français parlé. Transcription et édition*, Paris : Didier Erudition.
- CLARK, H. ; WASOW, T. (1998). Repeating words in spontaneous speech, *Cognitive Psychology*, 37, p. 201-242.
- DUEZ, D. (2001). Signification des hésitations dans la production et la perception de la parole spontanée, *Revue Parole*, 17-18-19, p. 113-138.
- FENOGLIO, I., (1997). La notion d'évènement d'énonciation : le « lapsus » comme une donnée d'articulation entre discours et parole, *Langage et société*, 80, p. 39-71.
- GADET, F. (2000). Derrière les problèmes méthodologiques du recueil des données, *Les Cahiers de l'Université de Perpignan*, 31, « Linguistique sur corpus », p. 30-43.
- GADET, F. (2003). Derrière les problèmes méthodologiques du recueil des données, *Texte !* juin-septembre, <[http://www.revue-texto.net/Inedits/Gadet\\_Principes.html](http://www.revue-texto.net/Inedits/Gadet_Principes.html)>, consultée le 18/07/2006).

- HENRY, S. (2002). Étude des répétitions en français parlé spontané pour les technologies de la parole, *RECITAL'02*, Nancy, France, p. 467-476.
- MERINGER, R. & MAYER, K. (1895). *Versprechen und Verlesen: eine Psychologisch-linguistische Studie*. Stuttgart : Göschen.
- MERINGER, R. (1908). *Aus dem Leben der Sprache : Versprechen, Kindersprache, Nachahmungstrieb*, Berlin : Behr's Verlag.
- MERINGER, R. (1923). Die täglichen Fehler im Sprechen, Lesen und Handeln, *Wörter und Sachen*, 8, p. 122-140.
- PALLAUD, B. (1999). Lapsus et phénomènes voisins dans la langue parlée : problèmes d'identification, *Recherches sur le Français Parlé*, 15, p. 1-33.
- PALLAUD, B. (2001). Les lapsus : des pierres dans le champ linguistique, in M. Arrivé et C. Normand (éds), *Linguistique et Psychanalyse*, Colloque de Cerisy-la-Salle, 1-8 septembre 1998, IN Presse, p. 47-66.
- PALLAUD, B. (2002). Erreurs d'écoute dans la transcription de données orales. Actes du colloque *Transcription de la parole normale et pathologique*, Tours, 8-9 décembre 2000, *Revue Parole*, 22-23-24, p. 267-294.
- PALLAUD, B. & HENRY, S. (2004). Amorces de mots et répétitions : des hésitations plus que des erreurs en français parlé, in *Le poids des mots, Actes des 7es Journées Internationales d'Analyse statistique des Données Textuelles*, Louvain-la-Neuve, 10-12 mars 2004, Louvain : PUL, vol. 2, p. 848-858.
- PALLAUD, B. (2006). Troncations de mots, reprises et interruption syntaxique en français parlé spontané, in *JADT 2006, 8es Journées internationales d'Analyse statistique des Données Textuelles*, 20-22 avril, Besançon, p. 707-715.
- PASDELOUP, V. (1992). A Prosodic Model for French Text-to-Speech Synthesis: A Psycholinguistic Approach, in Bailly, G. ; Benoit, Chr. ; Sawallis, Th.R. (eds), *Talking Machines. Theories, Models and Designs*, p. 335-348.
- PILLON, A. (1998). Morpheme units in speech production evidence from laboratory-induced verbal slips, *Language and cognitive processes*, 13, 4, p. 465-498.
- PETER-DEFARE, E. (1993). *Aspects phonologiques, syntaxiques et phonologiques de l'empan des erreurs de langage*, DEA de Phonétique, Université de Provence, Aix-en-Provence, 166 p.
- ROSSI, M. & PETER-DEFARE, E. (1995). *Lapsus linguae: word errors or phonological errors?*, *International Journal of Psycholinguistics*, 11, 1[30], p. 5-38.
- ROSSI, M. & PETER-DEFARE, E. (1998). *Les lapsus ou comment notre fourche a langué*, Paris : PUF.
- SHRIBERG, E. (1999). Phonetic Consequences of Speech Disfluency, Symposium on the Phonetics of Spontaneous Speech (S. Greenberg, P. Keating, organizers), *Proc. International Congress of Phonetic Sciences*, San Francisco, 1, p. 619-622.