

L'intégration entre gestes et prosodie : une vision incarnée

Maíra Avelar Miranda
PUC Minas¹ / CNPq
<mairavelar@gmail.com>

Résumé

Cet article a comme objectif de discuter l'intégration de la prosodie et des gestes, en démontrant comment ces deux types de phénomènes peuvent être vus comme un phénomène incarné dans les interactions en direct [online]. Pour démontrer cela, nous présenterons d'abord les concepts de base d'« esprit incarné » et de « blending expérientiel ». Ensuite, nous présenterons deux relations basées sur le corps: la relation entre gestes et métaphore, et celle entre prosodie et gestes. Pour terminer, nous prendrons une série d'échantillons d'un débat du second tour de l'élection présidentielle brésilienne, afin d'appliquer les concepts théoriques discutés auparavant.

1. Base conceptuelle

1.1. *Esprit incarné*

L'approche d'*esprit incarné*² est une des principales approches de ce travail, puisqu'il s'agit d'une des notions clés dans l'étude de la linguistique cognitive contemporaine. Un des principaux obstacles à la compréhension de ce concept est dû à la grande tradition scientifique, philosophique et même de sens commun qui dichotomise la question de *corps x esprit*. Cette tradition est nommée la *théorie représentationnelle* de l'esprit par Johnson (2007). Selon la théorie représentationnelle, il y aurait un « intérieur », représenté par l'esprit et le raisonnement abstrait, et un « extérieur », représenté par le corps et sa relation avec le monde : « The representational theory claims that the « mind » operates on « internal representations » (ideas, concepts, images) that can re-present (and thereby 'be about') external objects and events » (Johnson, 2007: 112). De cette façon il y aurait une séparation de nature ontologique, entre ce qui compose le corps et ce qui compose l'esprit.

¹ Cet article est une partie d'une recherche de doctorat qu'on développe à PUC Minas (Brésil), en partenariat avec l'Université de Genève. Je remercie David Vermeersch pour son aide précieuse pour la traduction et la mise au point du texte final.

² « Embodied (Mind, Cognition) » est traduit ici par « incarné » ; une alternative serait « incorporé ». Ni l'un ni l'autre ne sont pleinement satisfaisants.

Dans le cadre de la Linguistique Cognitive, à l'opposé de ce que clament les approches cognitivistes et connexionnistes, l'approche de la dynamique incarnée ne conçoit pas les pensées – ou l'esprit en général – comme des objets mentaux fixes, essayant désespérément d'entrer en contact avec le monde extérieur, les pensées étant des formes d'action et d'interaction. Elles seraient, de cette façon, dans et du monde, au lieu d'être au sujet du monde, elles sont en fait des processus d'expérience. Dès lors :

« The key to this reconceiving of mind is to stop treating percepts, concepts, propositions and thoughts as quasi-objects (mental entities or abstract structures) and see them instead as patterns of experiential interaction. They are aspects or dimensions or structures of the pattern of organism-environment coupling » (Johnson, 2007:177).

En bref, la cognition, et ensuite la construction de sens, émerge de notre expérience corporelle. Le cerveau est la clé du standard de l'interaction organisme-environnement :

« Thus, according to our neuronal model, we would predict that the abstract inferences are « computed » using sensorimotor neural structure and activated as target-domain inferences because there are neural connections from sensori-motor areas of the brain to other areas that are responsible for so-called high cognitive functions » (Johnson, 2007:177).

Partant des considérations de Johnson, il est possible de comprendre qu'au lieu d'exécuter un processus inférentiel au niveau sensori-moteur et ensuite exécuter un processus inférentiel totalement différent pour des concepts abstraits, nous utilisons les patrons inférentiels localisés dans les régions sensori-motrices du cerveau pour construire notre raisonnement abstrait. Ceci nous donne alors une vision continualiste, pour laquelle il n'est pas nécessaire d'introduire un nouveau type de raisonnement abstrait – comme une base ontologique distincte de notre standard sensori-moteur – pour expliquer le raisonnement logique ou les concepts abstraits.

Dès lors, l'esprit n'est pas situé dans le cerveau, car il est un processus continu qui possède divers niveaux d'interaction, y compris: un cerveau qui opère dans un corps, qui, à son tour, est continuellement situé dans des environnements complexes (pas seulement physiques, mais aussi sociaux et culturels). Ensuite, les structures projetées [*mapped*] du cerveau, résultant de l'activation de groupes [*clusters*] et de patrons neuronaux, découlent de l'expérience sensori-motrice et notre orientation corporelle :

« The central idea is that image schemas, which arise recurrently in our perception and body movement, have their own logic, which can be applied to abstract conceptual domains, via primary and higher level conceptual metaphors. Image-schematic logic then serves as the basis for inferences about abstract entities and operations ». (Johnson, 2007: 181).

En conséquence, la construction du sens dépend de trois dimensions : le cerveau, le corps et l'interaction organisme-environnement. Si nous excluons l'une de ces dimensions il n'y a pas d'émergence de sens : « *Thus mind emerges* » (Johnson, 2007: 151). L'esprit émerge en tant que résultat de l'interaction entre organisme et environnement – physique, social et culturel. Il s'ensuit que nous ne naissions pas avec un esprit prêt au raisonnement : l'esprit est un processus que nous acquérons et formons au cours de notre vie. Jeunes, nous possédons des capacités cognitives primitives qui, à travers l'interaction avec le monde et les personnes, deviennent un esprit fonctionnel. De ce point de vue, l'esprit peut être vu comme une question de degrés, qui s'étend de l'esprit primitif d'un nouveau-né à l'esprit développé d'un adulte. Donc, « l'esprit est un aboutissement et non une faculté inée »³ (Johnson, 2007:152).

1.2. *Blending expérientiel*

Un autre concept pertinent de ce travail est celui du *blending expérientiel*. Le *blending expérientiel* est l'intégration de la représentation conceptuelle abstraite et de l'expérientation interne sensori-motrice des caractéristiques matérielles de l'énonciation (Auchlin 2003). Selon Auchlin (à paraître), le *blending expérientiel* est considéré comme un cas spécifique d'intégration, au sens de Fauconnier & Turner (2003): deux espaces d'entrée substantiellement différents sont intégrés, appartenant à deux niveaux d'intégration différents. Un espace d'entrée est l'expérience du 'ici-et-maintenant' de l'énonciation et ses propriétés matérielles – construites d'un point de vue perceptif. L'autre espace d'entrée est l'intégration linguistique et conceptuelle. D'une façon générale, le *blending* consiste alors en un moulage humain de la parole. Comme nous pouvons le voir, le *blending expérientiel* comporte un ancrage matérialiste. Ce type d'ancrage a été proposé par Hutchins (2005). Selon l'auteur, « *Projecting material structure into a blended space can stabilize the conceptual blend. I call an input space from which material structure is projected into a blend a 'material anchor' for the blend* » (Hutchins, 2005: 1555). Dans le cas du *blending expérientiel*, l'ancrage matériel n'est cependant pas occasionnel mais systématique.

Le phénomène de *blending* peut aussi être appliqué à la dimension de la perception, comme le notent Fauconnier & Turner (2003). Selon Auchlin (2010: 10), à la suite de Bache (2005) :

³ « *Mind is an achievement, not a pre-given faculty* » (Johnson, 2007:152).

« Perception is a 1st level blending phenomena. Conscious, deliberate, dis-integration of distinct sources (input spaces) of information is not possible in (ordinary) perception; we cannot reverse the binding process in order to isolate input spaces' (bound) information. We cannot for example mentally dis-integrate saturation, luminosity, environment reflection and contrast, out of a « green spot ». It *is* a green spot. In Fauconnier & Turner's terms, the (blended) effect cannot be reversed back to its (input) causes ».

L'élaboration perceptive doit donc être comprise comme une action incarnée : une élaboration, interne à l'organisme, résultant de l'interaction de l'organisme avec le monde. Selon Auchlin (2008), cette élaboration perceptive lie les unes aux autres les différentes propriétés de l'environnement, d'une telle façon qu'elles ne peuvent perceptivement être dissociées les unes des autres. De cette façon, l'intégration perceptive (*binding* – Bache 2005), peut être considéré comme le niveau fondamental du processus de *blending* conceptuel.

En bref, nous pouvons suggérer qu'il y a deux phases dans le *blending* expérientiel. La première étant structurelle, un *binding*, pour lequel il y a une intégration automatique des espaces d'entrée et duquel il est impossible de désintégrer l'information de ces espaces. Cette phase est liée à l'élaboration quadridimensionnelle de l'expérience interne, de l'espace et du temps. Nous pouvons donc considérer l'information comme phénoménologiquement transparente. La seconde phase est liée à la fusion des deux espaces d'entrée et leurs propriétés ; entre ce qui est dit et comment c'est dit. De cette façon, la seconde phase consiste à l'intégration du contenu linguistique et de l'énonciation (cf. fig. 7 *infra* 3.1.).

2. Gestes

2.1. *Les gestes comme simulation incarnée*

D'un point de vue incarné, il y a une relation de continuité entre perception et action : la capacité de percevoir a évolué d'un besoin d'interagir avec le monde. « Perception enables us to know the affordances of objects in our environment and thus enables us to guide our actions in a goal-directed way » (Hostetter & Alibali, 2008: 495). Nous pouvons donc considérer qu'il y a différentes sortes de gestes qui remplissent plusieurs fonctions communicatives. De plus, les neuro-nes miroirs ont un important lien à la perception et à l'action : ils sont activés aussi bien lorsque nous effectuons une action que lorsque nous assistons à l'exécution d'une action. En conséquence, « The fact that the very same cells are involved in both action and perception suggests that activating potential actions may be an automatic consequence of perception. » (Hofstetter & Alibali, 2008: 496). Dès lors, la perception active automatiquement les régions du cerveau impliquées dans la production des actions correspondantes.

Cependant, ce n'est pas seulement la perception qui détermine les actions possibles; les actions déterminent aussi ce que, et comment, nous percevons. Perception et action sont indissolublement liées, chacune influençant et déterminant l'autre dans un système de coordination plus large. Aucun événement ne peut être pensé comme étant purement perceptuel ou purement moteur. Ce cycle constant de perception déterminant l'action et l'action déterminant la perception rend possible l'interaction avec le monde.

La vision selon laquelle les images motrices sont analogues aux actions physiques est parallèle à la vision, discutée plus haut, que les images visuelles sont analogues aux perceptions visuelles. Comme dans le cas des images visuelles, les images motrices sont des représentations « hors-ligne » qui utilisent des systèmes perceptifs et moteurs. En bref, l'imagerie mentale (tant visuelle que motrice) est alors perçue comme un processus incarné qui repose sur la simulation de la perception et de l'action. Tout comme nous l'avons vu dans la discussion de la compréhension du langage, l'imagerie mentale est possible parce que nous sommes capables de recréer nos expériences avec le monde. Hofstetter et Alibali (2008) soutiennent que l'usage de systèmes sensori-moteurs pendant aussi bien la production de langage que l'imagerie mentale donne naissance aux gestes.

Nous comprenons des textes, créons des images mentales et planifions des mouvements en utilisant les mêmes processus de perception et d'action que nous utilisons pour interagir avec le monde. Le traitement du langage et de l'imagerie mentale sont donc accomplis via des simulations de perception et d'action – des simulations qui activent ou rétablissent des états de perception et d'action. Ces deux simulations impliquent l'activation de régions neurales qui sont respectivement impliquées dans la prévision d'actions physiques et dans la perception d'objets physiques. En général, la simulation implique l'activation d'états pré-moteurs ; cette activation a le potentiel de se diffuser à des régions motrices et de se concrétiser en actions. Quand cette diffusion d'activation a lieu, un geste est né.

Selon le cadre des Gestes comme Action Simulée [*Gestures as Simulated Actions*] GSA, postulé par Hofstetter et Alibali (2008), il y a trois facteurs qui contribuent au fait que l'activation impliquée dans la stimulation se concrétisera en mouvement effectif, comme un geste (Hofstetter & Alibali, 2008: 503) : (1) la force de l'activation de l'action simulée : cette activation doit être suffisamment forte pour être diffusée des régions pré-motrices aux régions motrices et dépasser le seuil de geste – le seuil d'activation à partir duquel le locuteur ne peut inhiber la manifestation d'actions et gestes simulés. (2) La hauteur du

seuil de gestuelle du locuteur ; et (3) l'engagement simultané du système moteur pour la parole.

La force de l'action simulée peut également dépendre de la mesure où les pensées sous-jacentes à la parole concernent des objets et des événements concrets. Quand nous parlons d'idées abstraites, les locuteurs peuvent s'appuyer sur des représentations verbales et propositionnelles plus que sur des représentations imagées.

Cependant, comme décrit plus haut, il y a des preuves que notre compréhension de nombreuses idées abstraites a ses racines dans notre compréhension du monde physique. Les schémas d'images basiques (ex. CONTENEUR, OBJET, ATTRACTION, ÉQUILIBRE) concernant les caractéristiques d'entités physiques deviennent des métaphores pour le raisonnement concernant des objets abstraits. Ils peuvent être perçus (Je *vois* ce que vous voulez dire), déplacés dans l'espace (Mettions cette idée *de côté* quelques instants) et évalués (le rapport a *équilibré* les deux parties de la discussion).

2.2. *Métaphores et gestes*

D'après les postulats théoriques de Cienki et Müller (2008), nous soutiendrons l'hypothèse que la métaphore est un principe cognitif général et que les *mappings* métaphoriques peuvent être traités en ligne. Dès lors, nous analyserons la métaphoricité comme une forme dynamique inhérente à la cognition incarnée. Cependant, il est important de souligner que les gestes métaphoriques ne sont pas seulement des représentations d'un domaine abstrait. En fait, ce que Cienki et Müller (2008) appellent gestes métaphoriques sont ceux qui ont le potentiel d'engager un *mapping* actif trans-sectoriel, qui est le processus cognitif de comprendre quelque chose en tant que quelque chose d'autre.

Selon Cienki et Müller (2008), les gestes métaphoriques ont généralement été conçus comme des mouvements de main qui représentent ou indiquent le domaine d'origine de la métaphore. De plus, les gestes sont largement reconnus, non seulement comme des « illustrateurs », comme établi par Sonesson (2001), mais comme des parties intrinsèques de l'énoncé. Ensuite, en accord avec les observations de Cienki, « gestures appear as an articulatory independent mode of expression which is used flexibly, and not only to illustrate the semantic content expressed verbally, nor only to treat abstract discourse objects metaphorically » (Cienki & Müller, 2008: 493). Cela signifie que certaines métaphores peuvent être exprimées en gestes sans être présentes dans la parole. Alors les métaphores peuvent être déclenchées ou non par le contenu verbal de l'énoncé.

De plus, les gestes peuvent décrire des éléments spatiaux du domaine source de la métaphore, ce qui n'est pas possible pour des

expressions métaphoriques dans les langues parlées. Beaucoup de gestes métaphoriques ont leur origine dans des actions incarnées quotidiennes, les recréant de façon iconique. Si nous observons la façon dont les métaphores gestuelles et verbales sont intégrées dans des énoncés en cours, nous constatons que la métaphoricité a une gradation et possède des propriétés dynamiques, et non statiques. « The argument is an iconic and an interactive one: the more cues that direct the attention of the interlocutors to the metaphoric quality of a verbal metaphoric expression, the higher the degree of cognitive activation of metaphoricity in the speaker (and also potentially the addressee) » (Cienki & Müller, 2008: 497).

Quand nous pensons à la relation entre le langage parlé et les gestes, nous constatons que les métaphores peuvent être instanciées de façons différentes, sans se baser nécessairement sur la même métaphore conceptuelle. « Imagistic, embodied, and propositional modes of thinking interact during speaking (witness McNeill's concept of the growth point) and gestures may trigger new verbal metaphoric expressions » (Cienki & Müller, 2008: 498). De plus, l'analyse syntaxique, sémantique, prosodique et le contexte incarné d'expressions métaphoriques verbales et gestuelles révèle que la métaphoricité est dynamique et non statique, ce qui est la propriété d'objets linguistiques, qui peuvent plus ou moins être dégagés et recevoir alors davantage d'attention. Dans la section suivante, nous analyserons comment des propriétés iconiques prosodiques se manifestent dans la parole et quels sont leurs rôles dans les interactions en direct.

3. Prosodie

3.1. *La prosodie comme trait basé sur le corps*

Selon les principes d'analyse de la prosodie comme une propriété basée sur le corps (*body-based feature*), Auchlin (à paraître) propose que l'iconicité prosodique n'inclut pas seulement des schémas d'images et des icônes motrices, mais possède également un rôle dans notre façon directe d'expérimenter le monde : « the import of prosodic iconic formation is in modelling speech experience itself, not only in the elaboration of an abstract conceptual interpretation (though not excluded) : its use and function is to instantiate substantial experience of verbal content, hence creating some kind of stereophonic effect in meaning experience » (Auchlin, à paraître: 1). Selon l'auteur, les propriétés prosodiques, comme le rythme et le débit de parole, ne sont pas médiées symboliquement ou conceptualisées. Elles sont, plutôt, expérimentées de façon directe, immédiate.

Par l'analyse de quelques exemples, Auchlin démontre comment la

prosodie modèle directement notre expérience, comme dans l'exemple ci-dessous :

Les TGV Aquitaine arrivent jusqu'en gare de Biarritz

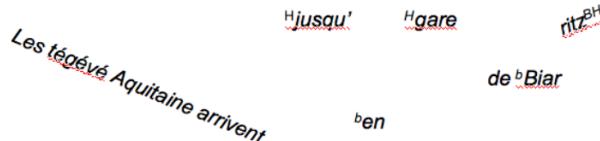


Figure 1: *Représentation graphique de la prosodie d'une phrase (Auchlin, à paraître : 7)*

Dans la première partie de l'exemple, comme souligné par Auchlin (à paraître), nous pouvons constater une descente mélodique qui se connecte iconiquement avec des parties de l'énonciation représentant le trajet en train jusqu'à Biarritz. La courbe mélodique descendante fournit une preuve perceptible que le trajet en train est : doux, sans effort et confortable. Dans le déroulement temporel de l'énonciation, l'arrivée à l'information pertinente lors de l'énonciation coïncide avec l'arrivée à destination, la gare de Biarritz. Dans la seconde partie de l'exemple, nous avons une répétition mélodique (trois fois) qui peut être iconiquement interprétée comme un frein pour ralentir le véhicule en mouvement. L'interprétation iconique des deux parties de l'énonciation appartient à un second niveau d'intégration, le *blending* expérientiel. Ceci implique un transfert analogique, iconique, du domaine source (énonciation) au domaine cible (trajet). Dans les deux parties de l'énonciation, nous pouvons construire des icônes dynamiques complexes qui sont composées d'une formation iconique liée au confort de voyage et une autre qui repose sur un cadre général d'un véhicule s'immobilisant.

Dès lors, nous pouvons inférer que l'émergence du *blending* expérientiel dépend crucialement de : 1. l'association locale avec le signifiant de l'énonciation, et 2. du signifiant de l'énonciation dans son ensemble.

Nous pouvons alors inférer que les icônes prosodiques peuvent être combinées avec le contenu verbal des énonciations : « If prosodic icons are in some way material anchors for blending linguistic and conceptual input, the emergent and creative part of experiential blending output was claimed to be experiential » (Auchlin, à paraître: 8). Selon Auchlin, (à paraître), il est important de souligner qu'aussi bien l'expression linguistique que l'iconicité prosodique doivent recevoir une attention commune, en raison de leur contigüité matérielle. D'un point de vue métaphorique, la différence entre les

deux agit de la même façon que la différence (minimale) entre les canaux audio gauche et droit dans l'audition stéréophonique : cette différence est responsable d'un changement qualitatif, ajoutant une nouvelle dimension (audition spatialisée) qu'aucun input ne peut créer séparément.

De plus, concernant l'information visio-gestuelle, la vision et les gestes jouent des rôles complémentaires dans la perception active de la parole, pas seulement des rôles redondants. Selon Auchlin, l'observation d'Arbib (2006) montre que l'implémentation neurale est tellement pertinente, que la modalité visuelle joue un rôle bien plus important dans le déclenchement de systèmes pré-moteurs que la modalité auditive, dans des expériences de perception de parole observées (comme, par exemple l'effet Mc Gurck-Mac Donald). Pour conclure, Auchlin propose de définir la communication comme un processus de co-expérience ; de cette façon, la communication est l'émergence et le guidage de l'expérience du partage progressif du sens dans le temps. En ce qui concerne les questions relatives à l'iconicité, elles sont analysées lors d'interactions en ligne ; dans l'usage de la langue. Dès lors, la question importante ici n'est pas ce qu'est l'iconicité mais comment elle émerge comme un produit basé sur le corps (*body-based*) dans les interactions en direct. Dans ce travail nous adopterons ce point de vue interactionnel, essayant d'analyser le rôle de l'iconicité prosodique dans les interactions face-à-face en ligne.

4. Analyse des données

4.1. Sélection du corpus

À partir d'un corpus de six débats, nous avons choisi un débat politique, transmis par la chaîne brésilienne « Record », lors du second tour de l'élection présidentielle brésilienne de 2010, opposant Dilma Rousseff (*Dilma*) à José Serra (*Serra*). On a alors choisi quatre phrases d'une séquence canonique, constitué de « question-réponse-réplique-tréplique » : deux appartiennent à la parole de Dilma et deux à la parole de Serra.

4.2. Méthodologie

Pour analyser l'intégration entre prosodie et gestes dans notre corpus, nous réaliserons la description et catégorisation gestuelle et, ensuite, la description et la catégorisation prosodique.

4.3. Description des gestes

Pour réaliser la description gestuelle, nous récupérons la catégorisation des gestes proposée par Hofstetter et Alibali (2008), présenté à la sous-section 3.3, où les auteurs divisent les gestes en cinq catégories :

- Gestes déictiques : gestes représentationnels qui englobent des mouvements qui représentent le contenu de la parole en pointant vers un référent dans environnement physique ;
- Gestes iconiques : mouvements qui dessinent un référent avec le mouvement ou la forme des mains ;
- Gestes métaphoriques : mouvements qui dessinent un référent concret ou indiquent une localisation spatiale d'un champ abstrait.
- Gestes rythmique [*beat*] : mouvements qui soulignent la prosodie ou la structure de la parole ;
- Gestes interactifs : mouvements utilisés pour contrôler le changement de tour de parole et autres aspects de l'interaction.

4.4. *Description de la prosodie*

Pour réaliser la description prosodique nous devons d'abord faire une transcription orthographique du corpus. Ensuite, nous devons faire un alignement manuel du son avec la transcription. Nous avons décidé de créer un TextGrid dans Praat (Boersma & Weenink, 2011) avec des silences pour mesurer les pauses. Après l'alignement, le Prosogramme (Mertens, 2011) doit être fait. Le Prosogramme fournit une stylisation de la fréquence fondamentale (F0) appliquée aux voyelles. La stylisation crée une segmentation du continuum sonore qui peut être réalisée à partir d'une base strictement acoustique, au travers de l'identification des pics des noyaux syllabiques.

Pour conclure, nous analysons le profil prosodique généré par le Prosogramme, qui fournit les mesures statistiques de chaque échantillon analysé. Selon Mertens (2010) :

« The stylization uses a segmentation of the speech signal into a sequence of nuclei, ideally corresponding to syllabic nuclei. The availability of this segmentation allows for the computation of statistical data about the prosodic properties of these nuclei, as well as of the sequences of nuclei ».

En conséquence, il est possible d'établir une comparaison entre les profils prosodiques de Dilma et Serra, ainsi que de faire une description de chacun de ces profils. Les paramètres prosodiques considérés sont les suivants⁴:

- Temps de phonation estimé [*estimated phonation time*], qui est la somme du temps intra- et inter-noyaux vocaliques moins la somme du temps de pause de l'échantillon analysé;

⁴ On a décidé d'adapter la table originale fournie par le Prosogramme. Les termes sont en anglais, traduits ici.

- Débit de parole [*speech rate*], qui est le nombre de syllabes par seconde;
- Taux d'articulation [*articulation rate*], qui est le nombre de syllabes sur le temps de phonation estimé [*estimated phonation time*];
- Agitation mélodique moyenne [*average melodic agitation*], qui est la trajectoire totale sur le temps internucleus (*internucleus time*);
- Dynamicité de la parole [*dynamicity of speech*], qui cumule toutes les valeurs liées à la trajectoire de la parole (trajectoire des intervalles inter-syllabiques et intra-syllabiques du *pitch*, trajectoire totale de la parole (*total trajectory of speech*), trajectoire de la parole sur le temps de parole (*trajectory of speech per speech time*) et trajectoire de phonation sur le temps de phonation (*trajectory of phonation per phonation time*).

4.5 Résultats et discussion

4.5.1. Échantillon 1

O senhor não tem coragem de assumir a sua posição, que fica falando que num se influencia pelos outros

[Vous n'avez pas le courage d'assumer votre position, vous ne faites que dire que vous ne vous laissez pas influencer par les autres.]



O senhor não tem coragem (...)
[Vous n'avez pas le courage (...)]

Main droite fermée avec l'index pointant à droite, qui va de haut en bas (geste rythmique (beat))



(...) de assumir a sua posição
[d'assumer votre position]

Main droite fermée qui frappe la main gauche ouverte (geste metaphorique).



e que fica falando que num se (...)
 [vous restez en parlant que vous ne
 vous laissez pas (...)]

Mains qui tombent à la droite de
 Dilma (geste métaphorique et
 iconique).



influencia pelos outros
 [influencer par les autres]

Mains qui tombent à gauche de Dilma
 (geste métaphorique et iconique)

4.5.2 *Échantillon 2*

Que hoje apoia a Dilma, fanaticamente, eh, assim, com muita força, o Collor é cabo eleitoral dela.

[Qui aujourd’hui soutien Dilma, fanatiquement, eh bien, avec beaucoup de force, Collor est son démarcheur politique]



Que hoje apoia a Dilma
 [Qui aujourd’hui soutien Dilma]

Geste déictique qui pointe vers Dilma



Fanaticamente
[Fanatiquement]

Geste rythmique [beat] ascendant-descendant qui empharise la prosodie.



Eh, assim, com muita força
[Eh, bien, avec beaucoup de force]

Geste iconique où Serra ferme sa main, qui prend un format de poing. Ce geste est renforcé par l'expression faciale de Serra qui fronce les sourcils.



O Collor é cabo eleitoral dela.
[Collor est son démarcheur politique]

Combinaison du geste de poing avec le geste iconique (qui pointe Dilma) et qui renforce l'emphase de la parole.

4.5.3. Résultats comparatifs entre les deux locuteurs

Nous avons l'intention de croiser les données analysées et de réaliser une analyse comparative entre les deux locuteurs. Nous pouvons trouver une similarité fondamentale entre l'échantillon 1, qui appartient à la parole de Dilma, et l'échantillon 2, qui appartient à la parole de Serra. Dans les deux échantillons, la parole des candidats présente des effets perlocutoires liés aux fautes de l'opposant, que chacun des candidats essaie de dénoncer auprès des téléspectateurs. Ces essais de disqualification de l'interlocuteur sont incarnés par la prosodie, les gestes et le contenu de la parole. Pour tester nos hypothèses avec une fiabilité quantitative, nous comparerons les résultats statistiques des

deux échantillons, pour voir si l'analyse qualitative peut être soutenue par les données quantitatives.

Comme nous pouvons observer, presque tous les résultats sont similaires, excepté ceux liés aux trajectoires : toutes les trajectoires de F0 de la parole de Serra présentent une valeur supérieure aux trajectoires de la parole de Dilma (contrairement aux attentes normatives concernant la parole féminine *vs.* masculine). Une autre comparaison qui a attiré notre attention est liée au débit de parole : les deux candidats présentent un débit de parole plus bas dans les échantillons dont on peut trouver les effets perlocutoires qui disqualifient l'interlocuteur, que dans les échantillons où ils donnent des explications (non présentés ici pour cause de taille). Cette contradiction apparente peut être attribuée à deux hypothèses : la première est l'intention de capter l'attention de l'interlocuteur, en soulignant l'information d'un point de vue incarné : en parlant plus lentement. La seconde hypothèse est liée à un essai de récupérer leur « accord intérieur » (Auchlin 2003), ou « auto-stabilité » du locuteur à travers les événements interactifs.

Dès lors, en même temps que les candidats mettent de l'emphase dans les dénonciations qu'ils font contre leur opposant, ils essaient de démontrer leur équilibre, leur stabilité. Donc, la diminution du débit de parole peut être considérée comme une tentative de compenser le *pathos* du discours de dénonciation, à travers la construction d'un *ethos* de la rationalité et du judicieux – en d'autres mots, de l'*ethos* de la *phronesis* (Amossy, 2005). Malgré la gravité des dénonciations, incarnées à travers les gestes et la prosodie, les locuteurs ne perdent pas le contrôle à cause de la haine ou de l'indignation. Cependant, cette tentative de construire l'*ethos* de la *phronesis* peut donner lieu à un discours perçu comme monotone.

5. Conclusion générale

Nous reprenons quelques-unes des informations théoriques les plus importantes que nous avons faites dans les sections initiales, en essayant de démontrer comment elles peuvent être appliquées aux interactions en direct analysées dans ce travail.

Les gestes métaphoriques sont typiquement conçus comme des mouvements de mains qui représentent ou indiquent le domaine source de la métaphore.

Nous pouvons soutenir cette affirmation lorsque nous reprenons les métaphores basées sur le corps que nous avons trouvées dans notre *corpus*, comme : « LA LÂCHETÉ C'EST TOMBER D'UN CÔTE À L'AUTRE ». Dans cet exemple, nous avons le domaine source basé sur une activité motrice (« tomber d'un côté à l'autre »), lié à un domaine

cible abstrait (« lâcheté »).

Plutôt que de transférer des propriétés spatiales ou motrices en codes verbaux ou propositionnels, les gestes expriment ces propriétés directement (Hofstetter & Alibali, 2008) et les propriétés prosodiques sont expériencées d'une façon directe (Auchlin, à paraître).

La prosodie et les gestes ne sont pas médiés symboliquement : ils sont expérimentés directement et sont dès lors des expériences vécues qui ont une activation sensori-motrice et, par conséquent, appartiennent à la dimension de la perception. Selon Auchlin et al (2006 : 228) « La prosodie ne se contente pas d'indiquer des changements de contexte, mais elle contribue à les créer, à les faire émerger (c'est là le sens d'enacter), et contribue ainsi à l'interprétation conjointe du discours et des actions ».

Les icônes prosodiques sont combinées avec le contenu verbal des énoncés et fonctionnent comme des ancrages matérielles pour les blends (Auchlin, à paraître)⁵.

Quand nous prenons en considération le cadre théorique expérientiel et incarné, cela implique une redéfinition de la communication (verbale) comme la construction coordonnée des expériences de parole entre les participants, qui peut aussi être appelée « co-expérience ». Donc, la construction de sens n'a pas seulement une dimension linguistique, mais aussi une dimension matérielle et perceptuelle.

Références

- ARBIB, M. (2006), « The Mirror System Hypothesis on the linkage of action and languages », in: *Action to language via mirror neuron system*. Cambridge: Cambridge University Press, 3-47.
- AMOSSY, R. (2005), « Da noção retórica de ethos à Análise do Discurso ». In: *Imagens de si no Discurso*. São Paulo: Contexto.
- AUCHLIN, A. (2003), « Compétence discursive et co-occurrence d'affects: 'blends expérientiels' ou (con)fusion d'émotions? », in: Colletta J.-M. & A. Tcherkassof (éds) (2001), LPS, Université Pierre Mendès France, Grenoble II, & LIDILEM, Université Stendhal, Grenoble III, 11-25.
- AUCHLIN ET AL (2006), « (Én)action, expérienciation du discours et prosodie », *Cahiers de Linguistique Française* 26, 217-249.
- AUCHLIN, A. (2008), « Du phonostyle à l'éthos, les prosodies comme interfaces entre sens et corps ». Actes du III Symposium sur L'Analyse du Discours, Belo Horizonte, Brésil, mimeo.

⁵ « Prosodic icons are combined with verbal content of utterances functioning as material anchors for blendings » (traduction libre).

- AUCHLIN A. (à paraître), « Prosodic Iconicity and Experiential Blending », in HANCIL S. (ed.), for series *Iconicity in Language and Literature*, Amsterdam, John Benjamins.
- BACHE (2005), « Constraining conceptual integration theory: Levels of blending and disintegration ». *Journal of Pragmatics* 37, 1615–1635.
- BOERSMA, P. & WEENINK, D. (2009), *Praat: doing phonetics by computer (Version 5.1.05)* [Computer program]. Retrieved May 19, 2011, from <http://www.praat.org/>.
- CIENKI, A. & MÜLLER, C. (2008), « Metaphor, Gesture, and Thought », in: Gibbs, R. (ed.) *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press.
- HOFTETTER, A. & ALIBALI, M. (2008), « Visible embodiment: Gestures as simulated action ». *Psychonomic Bulletin & Review*, 15 (3), 495-514.
- FAUCONNIER & TURNER (2003), « Conceptual blending, form and meaning ». *Recherches en communication* 19, 57-86.
- HUTCHINS, E. (2005), « Material anchors for conceptual blends ». *Journal of Pragmatics* 37, 1555–1577.
- JOHNSON M. (2007), *The meaning of the body: aesthetics of human understanding*. Chicago and London: Chicago University Press.
- MERTENS, P. (2011), « Tutoriel Prosogram », 1-12. URL: http://bach.arts.kuleuven.be/pmertens/prosogram/tutoriel_prosogram.pdf Date of access: 23/05/12.
- MERTENS, P. (2010), « User guide » URL: <http://bach.arts.kuleuven.be/pmertens/prosogram/userguide.html> Date of access: 23/05/12.
- SONESSON (2001), « De l'iconicité des images à l'iconicité des gestes », in: Cavé, C., Guaitella, I., & Santi, S. (eds), *Oralité est gestualité: Interactions et comportements multimodaux dans la communication*. Actes du colloque ORAGE 2001, Aix-en-Provence, 18-22 juin 2001. Paris: L'Harmattan , 47-55.