



HAL
open science

Segmentation et annotation du geste : Méthodologie pour travailler en équipe

Marion Tellier, Brahim Azaoui, Jorane Saubesty

► To cite this version:

Marion Tellier, Brahim Azaoui, Jorane Saubesty. Segmentation et annotation du geste : Méthodologie pour travailler en équipe. JEP-TALN-RECITAL 2012 : Atelier DEGELS 2012 (Défi GEste Langue des Signes), Jun 2012, Grenoble, France. pp.41-55. hal-01510435

HAL Id: hal-01510435

<https://hal.science/hal-01510435>

Submitted on 8 Feb 2018

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.

Segmentation et annotation du geste : Méthodologie pour travailler en équipe

Marion Tellier¹, Brahim Azaoui², Jorane Saubesty¹

(1) LPL, UMR 7309, Université d'Aix-Marseille, France

(2) DIPRALANG, EA 739, Université Paul Valéry, Montpellier III, France

marion.tellier@lpl-aix.fr, brahim.azaoui@etu.univ-montp3.fr, jorane.saubesty@gmail.fr

RESUME

Dans cet article, nous proposons de décrire notre méthodologie de segmentation et d'annotation des gestes coverbaux. Nous avons travaillé en équipe de 3 gestualistes, ce qui nous a demandé à la fois de trouver une méthode pour coordonner notre travail, une méthode d'évaluation de l'accord inter-juge et une méthode d'annotation et de segmentation du geste qui fasse consensus. Le présent article a pour vocation d'expliquer notre démarche de travail afin de la partager avec la communauté de chercheurs travaillant sur le geste coverbal.

ABSTRACT

Gesture segmentation and coding : a methodology for teamworking

In this paper we intend to describe our methodology for segmenting and coding co-speech gestures. Working as a team of 3 gesture researchers required to find a method to coordinate our work, another to evaluate the interrater agreement and a last one to code and segment the gestures that would bring about consensus. This paper aims at making our approach clear so as to share it with the community of co-speech gesture researchers.

MOTS-CLES : segmentation, annotation, geste coverbal, méthode

KEYWORDS : segmentation, coding, co-speech gesture, methodology

1 Constitution de l'équipe et organisation du travail

1.1 Répartition du travail

Nous avons travaillé en équipe de 3. Lors de la précédente édition DEGELS, Tellier, Bigi et Guardiola (2011) avaient travaillé en équipe mais chacune annotait une modalité différente en fonction de ses compétences, il s'agissait donc d'une équipe à compétences complémentaires. Pour l'édition 2012, l'équipe est composée de 3 gestualistes. Nous allons présenter ici à la fois la méthodologie de l'annotation mais également une méthodologie de travail en équipe. Cette dernière a été construite *ad hoc* et présente probablement des imperfections. L'annotation gestuelle en équipe étant relativement rare car très coûteuse en temps, aucun d'entre nous ne l'avait

réellement expérimentée auparavant. Nous partagerons donc ici le fruit de cette démarche.

Les délais étant très courts, il a été proposé aux annotateurs un fichier Elan avec un template imposé et contenant déjà les phrases gestuelles (c'est-à-dire les différentes étiquettes où un geste est produit) sur la base de l'annotation de Tellier *et al.* (2011). Les annotateurs devaient tout d'abord vérifier que ces phrases étaient bien alignées dans le temps sinon les modifier. Chaque annotateur devait ensuite :

1. Annoter la manualité (Handedness)
2. Annoter les phases du geste
3. Annoter les dimensions du geste
4. Indiquer dans une piste « notes » les problèmes ou doutes rencontrés

Les trois premiers gestes avaient déjà été annotés par l'annotateur 1 (A1) pour montrer l'exemple de la marche à suivre. Les annotations ont été réalisées en aveugle, c'est-à-dire sans voir les annotations des autres membres de l'équipe. Les trois annotateurs (A1, A2 et A3) étant éloignés géographiquement ou peu disponibles en même temps, le travail post-annotation s'est fait à distance et de manière asynchrone. Les fichiers et les instructions étaient sauvegardés dans un dossier partagé sur service de partage en ligne, l'ensemble était coordonné par l'A1.

Nous avons travaillé en suivant les consignes de cet atelier mais également selon nos propres besoins en utilisant nos typologies habituelles et en annotant les aspects qui sont pertinents pour nous. Les termes utilisés pour l'annotation sont souvent en anglais, par habitude de travail et parce que la communauté des gestualistes utilise majoritairement l'anglais dans ses échanges et publications.

1.2 Schéma de codage

1.2.1 Template

Les deux pistes concernant la transcription verbale (une par locuteur) ont été fournies aux annotateurs. Elles s'appellent TOE_LocA et TOE-LocB puisqu'elles sont transcrites suivant les conventions de la TOE (voir Bertrand *et al.*, 2008 et Tellier *et al.*, 2011). La parole du locuteur B (masculin) a été tokenisée (en items lexicaux) automatiquement.

Le schéma de codage pour les aspects gestuels est présenté dans la Figure 1, les vocabulaires contrôlés sont listés dans les rectangles jaunes. Pour chaque phrase gestuelle, l'annotateur devait indiquer la manualité dans la piste [Handedness]. Cette piste, une piste enfant en *symbolic association*, dans la terminologie Elan, est dépendante de la piste [Gesture_Phrase], ce qui signifie que les bornes de l'annotation sont strictement les mêmes que celles de la phrase gestuelle (piste parent) et ne peuvent être déplacées. Chaque phrase gestuelle devait ensuite être découpée dans la piste [phase], elle aussi, piste enfant de [Gesture_Phrase] mais dont la dépendance est

included in. Cela signifie que les bornes extérieures des annotations sont fixes mais que l'intérieur de l'annotation peut être découpé en plusieurs morceaux.

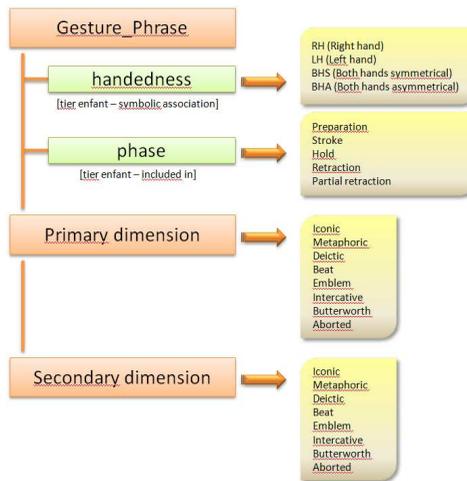
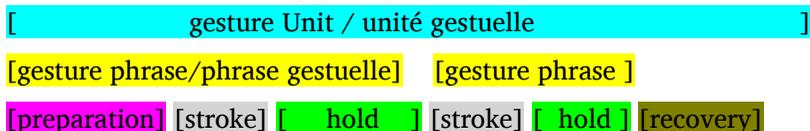


FIGURE 1 Schéma de codage

1.2.2 Unités, Phrases et phases

Il est peut-être bon ici de faire une petite précision terminologique. Kendon appelle *gesture unit*, un ensemble de mouvements (« excursion ») produits avec la parole. « This entire *excursion*, from the moment the articulators begin to depart from a position of relaxation until the moment when they finally return to one [position of relaxation], will be referred to as a *gesture unit*. » (2004 : 111). Il explique qu'une unité gestuelle peut contenir différentes phrases. Une phrase, quant à elle, est composée de différentes phases (preparation, stroke, hold...). Une phrase ne peut contenir qu'un seul stroke. La *retraction* (selon le terme de McNeill, 1992) que Kendon (2004) nomme *recovery* ne fait pas partie de la phrase mais de l'unité gestuelle. « The *recovery* movement, when the hand (or other body part) relaxes and is returned to some position of rest is not considered to be part of the *gesture phrase*, although it is, of course, part of the *gesture unit* which contains the *gesture phrase*. » (2004: 112). Le schéma ci-dessous tente de résumer la structuration de Kendon.



McNeill (2005: 31) souligne l'incohérence de cette terminologie et surtout la confusion qu'elle peut engendrer. En effet, on s'attendrait plutôt à ce que *phrase* soit l'ensemble le plus large, constitué de plusieurs groupes que l'on pourrait appeler

unités. L'utilisation de phase et de phrase n'est pas très heureuse non plus vue la proximité phonémique et orthographique des deux mots qui pourraient être confondus. Dans notre annotation, nous n'avons pas tenu compte des unités gestuelles. Nous avons découpé les gestes en phrases. Contrairement à Kendon, nous avons considéré que la phase *recovery* (que nous appelons *retraction* comme McNeill, 1992, 2005), faisait partie de la phrase gestuelle. Il nous semble, en effet, incohérent de la considérer à part car elle fait partie du même mouvement signifiant. McNeill explique d'ailleurs: «The retraction phase, especially its end, is not without significance, contrary to what I have written about it in the past (McNeill, 1992). » (2005: 33).

1.2.3 Dimensions

Les dimensions du geste ont été ensuite annotées sur deux pistes qui présentent les mêmes vocabulaires contrôlés. On a attribué *a minima* une dimension à chaque geste sur la piste [primary dimension]. La piste [secondary dimension] sert à indiquer lorsqu'un geste a deux dimensions (une dimension iconique et une dimension déictique, par exemple), on met dans la piste primaire la dimension qui nous semble primer. Elle est également utilisée pour faire apparaître les battements superposés.

Déictique	<i>Geste de pointage</i>
Iconique	<i>Geste illustratif d'un concept concret</i>
Métaphorique	<i>Geste illustratif d'un concept abstrait</i>
Battement	<i>Geste rythmant la parole, sans contenu sémantique</i>
Emblème	<i>Geste culturel, conventionnel</i>
Butterworth	<i>Geste de recherche lexicale</i>
Interactif	<i>Geste adressé à l'interlocuteur pour la gestion de l'interaction</i>
Avorté	<i>Geste esquissé mais avorté</i>

TABLE 1 Typologie des gestes utilisée pour l'annotation

Les dimensions sont des annotations qui portent sur le sens du geste. Elles ont été élaborées à partir de la typologie de McNeill (1992, 2005) c'est-à-dire déictique, iconique, métaphorique et battement. Elle a été enrichie par 4 autres types de gestes : les emblèmes, les Butterworth, les interactifs (Bavelas et al., 1995) et les gestes avortés (voir Table 1). Cette typologie a déjà été utilisée par Tellier et Stam (2010) et Tellier, et al. (2011).

2 Mise en commun et comparaison des annotations

Les fichiers de chacun ont ensuite été réunis en un seul en utilisant la fonction d'Elan *Merge transcriptions* pour pouvoir importer toutes les pistes dans un même fichier sans perdre les vocabulaires contrôlés et les autres attributs des pistes. Le résultat obtenu

de comparer deux juges tandis que le kappa de Fleiss est utilisé lorsqu'il y en a plus de deux (Santos, 2010). Pour interpréter les résultats, on utilise le barème de Landis et Koch (1977, cité dans Santos, 2010), présenté dans la Table 2. Il faut cependant relativiser ce barème car plus il y a de juges et plus il y a de stades (de « scores » à attribuer), plus le kappa aura tendance à être faible. « Ainsi, par exemple, un $\kappa = 0.40$ pourra être considéré comme très médiocre si deux juges avaient seulement à choisir entre deux scores A et B, mais pourra être perçu comme relativement honorable s'ils devaient choisir entre 10 stades différents. » (Santos, 2010).

Le logiciel Elan calcule automatiquement l'accord entre deux annotateurs en faisant un kappa de Cohen (fonction *Compare annotators*). Cette fonctionnalité d'Elan est particulièrement intéressante mais possède des limites : l'accord n'est calculé que sur la base des segmentations temporelles et non sur le contenu des annotations.

Sur la durée des phrases gestuelles, le taux d'accord inter-juges été très élevé : cela est dû au fait que le fichier Elan de départ contenait déjà les étiquettes des phrases et que peu de modifications ont été apportées par les annotateurs (Table 3).

Juges	Taux d'accord
A1 et A2	0,948
A1 et A3	0,863
A2 et A3	0,904

TABLE 3 Taux d'accord pour les unités gestuelles

Sur les dimensions du geste, les taux d'accord inter-juges ont été calculés avec un logiciel de statistiques. On constate que le taux d'accord est plus bas (Table 4).

Juges	Taux d'accord
A1 et A2	0,520
A1 et A3	0,347
A2 et A3	0,281
A1, A2 et A3	0,406

TABLE 4 Taux d'accord pour les dimensions primaires du geste

Cependant, si l'on considère le kappa de Fleiss entre les 3 juges (0,406), on est très proche d'un accord moyen, ce qui, considérant le fait que les trois annotateurs devaient choisir entre 8 catégories, est plutôt honorable. De plus, en ce qui concerne les dimensions, le désaccord est à relativiser car très souvent, les annotateurs ont attribué les deux mêmes dimensions à un geste mais n'étaient pas forcément d'accord sur la dimension qui devait être considérée comme primaire (voir 2.2).

2.2 Révision des annotations

Une fois les annotations réunies dans un fichier Elan, nous avons procédé par étapes en créant plusieurs fichiers de révisions. A chaque fois, le fichier Elan de travail était enregistré et partagé dans un service de partage en ligne afin de s’assurer que la même version était utilisée.

Dans un premier temps, nous avons commencé par les phrases gestuelles et la manualité en créant un premier fichier de révision [degels_modif1.eaf]. A1 a repéré tous les points de divergence entre annotateurs (12 en tout) et créé une piste spéciale pour discuter de ces divergences. Nous avons donc inventé un système de conversation asynchrone sur Elan (Figures 8, 11 et 13).

Dans un deuxième temps, nous avons comparé les points de divergence pour les dimensions primaires et secondaires des gestes avec la même méthode et en créant un deuxième fichier [degels_modif2.eaf]. Sur les 63 gestes annotés, il y avait 41 cas de divergence. Il s’agissait souvent de divergences sur les cas où un geste avait deux dimensions et où les annotateurs n’étaient pas d’accord sur l’attribution des caractères primaires ou secondaires (Figure 4). On peut d’ailleurs se demander s’il est pertinent de maintenir une hiérarchie dans l’attribution des dimensions.

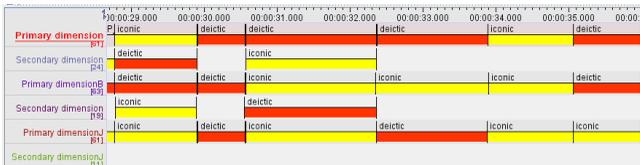


FIGURE 3 Exemple de divergence sur les dimensions primaire et secondaire

Après une première discussion et argumentation sur les dimensions, il ne restait que 11 gestes sur lesquels il y avait encore des différences de points de vue. Une deuxième session d’argumentation a été nécessaire afin d’obtenir une annotation unique qui fasse consensus.

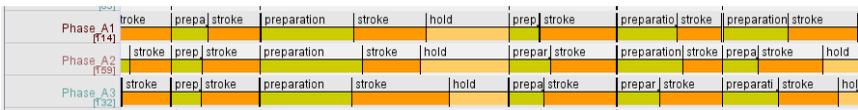


FIGURE 4 Comparaison des annotations des phases

Dans un troisième temps, nous avons travaillé sur les phases gestuelles. Pour cet aspect les divergences étaient considérables mais portaient davantage sur le découpage temporel que sur les phases en elles-mêmes (Figure 4). Il est à noter qu’aucun de nous n’est habitué à annoter les phases car nous n’avons jusqu’à présent pas eu besoin de cet aspect dans nos recherches respectives. Nous sommes donc assez inexpérimentés sur la question. La plus grande difficulté dans cette annotation réside dans le découpage temporel. En effet, le degré de précision est très élevé, on découpe

image par image (soit à 40 ms près), il est donc quasiment impossible pour différents annotateurs de produire le même découpage. Faute de temps et étant donné l'importance de l'écart entre nos annotations, nous avons été amenés à sélectionner certains gestes sur lesquels discuter (39 phases ont été corrigées).

3 Emergence de consignes méthodologiques

3.1 Manualité

Annoter la manualité n'est pas forcément une pratique courante en étude de la gestuelle, sauf si les hypothèses de recherche portent sur le sujet (comme les études sur la latéralisation par exemple ou pour définir le style gestuel d'un individu). Nous avons donc choisi de noter sur une piste si chaque phrase gestuelle était produite avec la main droite (RH), la main gauche (LH), les deux mains symétriques (BHS) ou les deux mains asymétriques (BHA). Nous avons considéré la symétrie/l'asymétrie sur 2 axes : vertical (hauteur)/horizontal (largeur). Un 3^{ème} axe aurait pu être rattaché à la profondeur du geste (mouvement avant/arrière), mais il n'y en avait pas d'exemple dans le corpus.

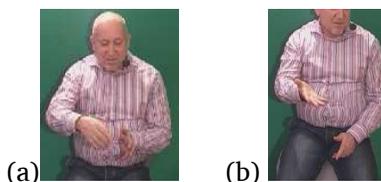


FIGURE 5 Deux exemples de gestes bimanuels

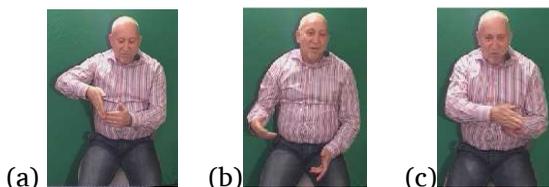


FIGURE 6 (a) BHA/(b) BHS : mains en miroir /(c) BHS : mouvement conjoint

Un geste fait avec les 2 mains (BH) est un geste où les mains sont en mouvement ensemble (Figure 5 a), ou lorsque les doigts d'une main ont un mouvement *significatif* en même temps que la main opposée (Figure 5 b). Un geste asymétrique (BHA) est un geste dans lequel les mains agissent de manière décalée principalement dans l'espace (Figure 6, a). Un geste symétrique (BHS) est un geste dans lequel les mains agissent en miroir (Figure 6,b) ou conjointement (Figure 6, c). Dans l'ensemble, la segmentation [handedness] ne posait pas de difficulté particulière (BH/RH/LH). Lorsqu'il en existait une, elle apparaissait à un des 2 niveaux suivants (1) Distinction BHA/BHS (2) Prise en compte des doigts de la main opposée comme éléments

significatifs dans la réalisation du geste global. L'exemple donné dans la Figure 7 illustre très bien le premier cas problématique. Les annotateurs n'étaient pas d'accord : A1 et A2 avaient codé ce geste BHA et A3 avait codé BHS. A première vue, les deux mains semblent bouger symétriquement (suivant un axe horizontal), cependant, la main gauche (celle du bas) s'arrête de bouger vers la fin du geste tandis que la main droite continue à dessiner la forme du référent. Le dialogue entre les annotateurs (Figure 8 révèle que cette décision n'a pas été facile à prendre et que c'est dans le travail d'équipe que les arguments se forment.



(01 : 02 : 600) [c'est un pouce géant et]

FIGURE 7 Cas de désaccord sur la symétrie

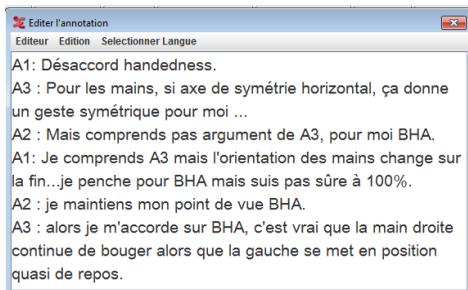


FIGURE 8 Dialogue entre les annotateurs (sur le geste Figure 7)

L'autre type de cas problématique apparaît lorsque certains gestes ont été codés main droite (RH) ou main gauche (LH) parce que le mouvement principal était effectué avec une main alors que l'autre main produisait un tout petit battement du pouce, parfois à peine perceptible. Hormis la segmentation de la manualité, la question du geste principal dans un « BH » peut être pertinente à se poser : ne serait-il pas intéressant de pouvoir segmenter à l'intérieur d'une même « handedness » pour marquer avec quelle main le stroke se fait ? Dans le corpus proposé, le locuteur B utilise principalement sa main droite. Or, dans l'exemple de la Figure 9, alors que le geste est effectué avec les deux mains, la main droite donne le repère alors que la main gauche donne l'information principale et effectue le stroke : « longer ». Dans quelle mesure ce changement de main principale est-il significatif ? Une segmentation à l'intérieur d'une manualité pour indiquer avec quelle main le stroke s'effectue permettrait donc les repérages de ce genre de récurrences et leur analyse.



(00 : 34 : 500) [et en fait vous avez longé l'hippodrome]

FIGURE 9 Main dominante sur le stroke

3.2 Dimensions du geste

La question des dimensions du geste est une question complexe à l'heure actuelle en études de la gestuelle. Depuis que ce domaine existe, les chercheurs ont tenté de concevoir des typologies d'annotation du geste (Efron, 1941 ; Ekman et Friesen, 1969 ; McNeill, 1992, etc.). Ces typologies ne sont pas faciles à utiliser car elles inscrivent les gestes dans des catégories dont les frontières sont parfois très poreuses et mal définies. La typologie la plus utilisée ces dernières années est celle de McNeill (1992) même si elle demeure controversée car parfois difficile à manipuler. McNeill lui-même (2005) a suggéré de considérer ces catégories davantage comme des dimensions que comme des types restrictifs (un même geste pouvant avoir plusieurs dimensions). Pour le présent corpus, nous avons utilisé une typologie enrichie (voir 1.2.3). L'intérêt pour nous d'annoter les dimensions du geste est double. Premièrement, les hypothèses de nos études impliquent souvent les dimensions du geste (on peut faire l'hypothèse qu'une condition va éliciter plus de gestes iconiques que de gestes métaphoriques, par exemple). Deuxièmement, une grande partie des études en gestuelle actuellement utilisent la typologie de McNeill ce qui nous permet d'effectuer des comparaisons entre différents travaux et de reproduire des expérimentations en comparant les résultats sur des bases similaires.

Dimensions	Primaire	Secondaire
Butterworth	2	0
Aborted	4	0
Beat	2	10
Deictic	21	8
Emblem	3	0
Iconic	20	7
Interactive	2	1
Metaphoric	6	1

TABLE 5 Statistiques des annotations des dimensions

La Table 5 présente les statistiques des annotations pour les dimensions primaire et secondaire. On peut constater tout d'abord que les battements (beats) sont en général identifiés sur la piste [secondary_dimension] car ils sont, la plupart du temps, superposés sur d'autres gestes. Ensuite, on remarque que les dimensions les plus fréquentes sont les iconiques et les déictiques ce qui n'est pas surprenant avec ce genre de tâche langagière. Plusieurs gestes déictiques ont une dimension secondaire iconique et inversement. La plus grande difficulté pour notre équipe a été de définir quelle dimension primait sur l'autre (Figures 10 à 13). On voit dans ces discussions que l'interprétation repose à la fois sur la signification du geste (illustration d'un item/d'une action vs pointage d'un lieu ou d'une direction) et sur l'accompagnement verbal. En effet, selon le cadre théorique de McNeill (1992) dans lequel les membres de notre équipe s'inscrivent, le geste et la parole forment un seul et même système et doivent être analysés simultanément.



(00 : 30 : 570) [droite c'est plein de restaurants euh ça s'appelle]

FIGURE 10 Exemple de déictique à dimension secondaire iconique

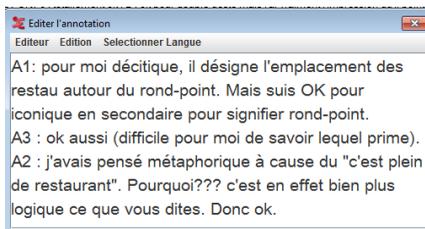


FIGURE 11 Dialogue entre les annotateurs sur l'exemple de la Figure 10



(00 : 22 : 409) [statut du David vous allez vers les Goudes]

FIGURE 12 Exemple d'iconique à dimension secondaire déictique

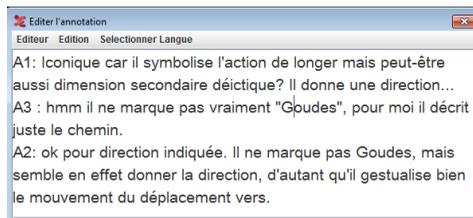


FIGURE 13 Dialogue entre les annotateurs sur l'exemple de la Figure 12

Dans l'exemple 12, le geste décrit l'action de longer plus qu'il ne donne une direction car il ne marque pas le point d'arrivée dans l'espace gestuel. On considère alors ici que le geste est d'abord un iconique avec une dimension secondaire déictique.

3.3 Phases du geste

Ceci est sans conteste l'aspect le plus complexe que nous ayons eu à annoter. Connaître les phases du geste peut être pertinent lorsque l'on travaille par exemple sur la synchronisation entre geste et verbal ou entre geste et prosodie. Il est alors important de déterminer où se trouve le pic du geste (le stroke) et de l'associer à un mot, une syllabe ou un phénomène acoustique. Annoter les phases du geste amène nécessairement à s'interroger sur les frontières temporelles du geste : où commence-t-il ? Où se termine-t-il ? Quelle est sa partie la plus signifiante ? Les gestualistes proposent différentes phases. La Table 6 synthétise les phases que nous avons utilisées et les définitions données par les auteurs (nous avons gardé la version originale).

Preparation	(optional) : «the limb moves away from its rest position to a position in gesture space where the stroke begins. The preparation phase typically anticipates the linguistic segments that are coexpressive with the gesture's meaning.» (McNeill, 1992 : 83)
Stroke	« (obligatory in the sense that absent a stroke, a gesture is not said to occur)[...] The stroke is the gesture phase with meaning; it is also the phase with effort [...]. In a large sample of gestures, the stroke is synchronous with co-expressive speech about 90 percent of the time.» (McNeill, 2005 : 31) / It is the phase of the excursion in which the movement dynamics of 'effort' and 'shape' are manifested with greatest clarity. » (Kendon, 2004 :112)
Hold	« in general is any temporary cessation of movement with leaving the gesture hierarchy (in contrast to a rest, which means exiting the hierarchy). » (McNeill, 1992 : 83)
Retraction	(optional) «The hands return to rest (not always the same position as at the start). There may not be a retraction phase if the speaker immediately moves into a new stroke. » (McNeill, 2005 : 31) / « when the hand (or other body part) relaxes and is returned to some position of rest [...]» (Kendon, 2004 : 112)
Partial retraction	« After the stroke, if the hand approaches a resting position but shifts to a preparation before reaching it, the interrupted retraction phase is called a <i>partial retraction</i> . » (Kipp, 2005 : 52)

TABLE 6 Les définitions des phases du geste

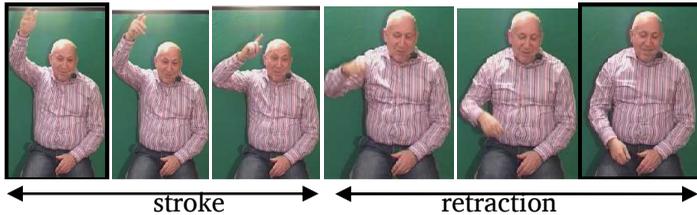
La catégorie *Hold* a été beaucoup découpée en micro catégories. Ainsi, Kita (1993) parle de *prestroke hold* et *poststroke hold* et McNeill (2005) distingue également des *stroke holds* et des *independent holds*. Ces phases ajoutant de la complexité à un découpage déjà ardu, nous avons fait le choix de ne pas les utiliser.

La *préparation* du geste est donc l'étape du début du geste jusqu'au *stroke*. Le *stroke* ou pic gestuel est le point le plus signifiant du geste. Il est, la plupart du temps, accompagné du mot clé de la phrase. Ce que l'on peut remarquer dans ce corpus, c'est que lors de la préparation du geste, la main produit souvent un mouvement dans le sens contraire du sens dans lequel le *stroke* s'effectuera. Par exemple, lorsque la main droite va effectuer un mouvement de droite à gauche, elle commence par reculer de gauche à droite, comme pour armer le geste. Dans ces cas là, la distinction des phases est plus aisée puisque le premier mouvement dans une direction constitue la préparation et lorsque la main amorce le mouvement dans la direction inverse, on est en présence du *stroke* (Figure 14).



(00 :39 :680) [Escale Borély vous remontez]

FIGURE 14 Exemple de distinction preparation/stroke

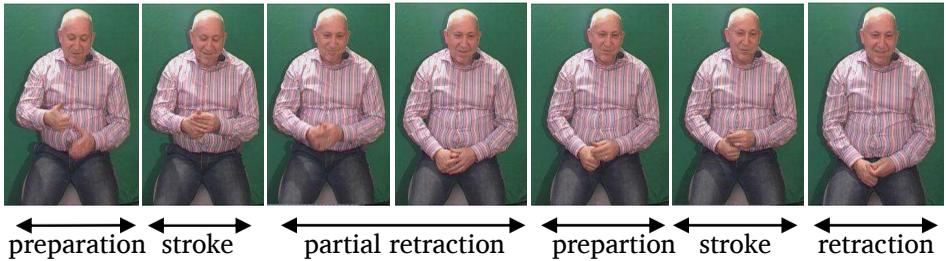


(01 :10 :747) [T/\$ la prison la fameuse prison des \$Baumettes]

FIGURE 15 Exemple de rétraction

La rétraction du geste, quant à elle, est déterminée au moment où le geste est moins clair, moins précis et que la main retourne à sa position de repos comme on peut le voir dans la Figure 15.

Nous avons eu des divergences parfois importantes lors de cette segmentation. Ainsi, le segment « enfin sa son moulage de pouce là » (Figure 16) a été problématique, mais en cela intéressant : y a-t-il un geste ou deux ?



(00 : 59 : 890) [enfin sa son moulage de pouce là]

FIGURE 16 Le moulage du pouce

La main droite prend la forme du pouce et la main gauche vient se placer autour de la main droite, l'ensemble descend alors jusqu'en position de repos. C'est comme s'il posait ses mains pour mieux réfléchir au mot qu'il cherche, une sorte de « silence gestuel », avant de remonter les deux mains jointes jusqu'au niveau de la poitrine. Pendant tout le mouvement les mains ne se sont pas lâchées, pourtant elles sont retournées en position de repos. Si on considère la position de repos comme la fin d'un geste, il y a alors deux gestes similaires. Cependant, si on considère la position des mains, en l'occurrence la même configuration que le stroke, alors il n'y a qu'un seul geste et il y a deux strokes avec une rétraction partielle entre les deux.

4 Conclusion

Ce travail d'annotation à trois gestualistes a été à la fois frustrant (principalement à cause du peu de temps disponible et de l'éloignement des annotateurs) mais a également été extrêmement enrichissant. Le fait d'avoir à comparer et justifier son travail avec d'autres partenaires oblige nécessairement à repenser sa propre méthode. Ce travail nous a, par ailleurs, incité à relire les ouvrages fondamentaux en études de la gestuelle afin d'éclaircir certaines définitions. Nous constatons, comme c'est souvent le cas dans un domaine scientifique jeune, que la méthodologie ainsi que les notions essentielles sont loin de faire consensus et manquent parfois de clarté. Il faut aussi tenir compte du fait que les études sur les coverbaux réalisées il y a 10 ou 20 ans ne bénéficiaient pas des moyens technologiques actuels, notamment des logiciels qui permettent une précision nouvelle dans l'analyse du geste. Ces outils ouvrent de nouvelles voies mais requièrent également de nouvelles balises méthodologiques.

Remerciements

Merci à Stéphane Rauzy (LPL) pour les calculs du kappa.

Références

- BAVELAS, J., CHOVIL, N., COATES, L. & ROE, L. (1995). Gestures specialized for dialogue. In *Personality and social psychology bulletin*, 21, pages 394-405.
- BERTRAND R., BLACHE P., ESPESSER R., FERRE G., MEUNIER C., PRIEGO-VALVERDE B., RAUZY S. (2008). Le CID - Corpus of Interactional Data - Annotation et Exploitation Multimodale de Parole Conversationnelle. *Traitement automatique des langues (TAL)* 49 (3), 105-134.
- EFRON, D. (1941) *Gesture 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*. The Hague ; Paris : Mouton.
- EKMAN, P. & FRIESEN W. V. (1969). The Repertoire of Nonverbal Behavior : Categories, Origins, Usage, and Coding. In *Semiotica*, 1, The Hague : Mouton Publishers, pp. 49-97.
- KENDON, A. (2004). *Gesture : Visible Action as Utterance*. Cambridge : Cambridge University Press.
- KIPP, M. (2005). *Gesture Generation By Imitation: From Human Behavior To Computer Character Animation*. Universal-Publishers.
- KITA, S. (1993). *Language and thought interface: A study of spontaneous gestures and Japanese mimetics*. Unpublished PhD Diss. Chicago: University of Chicago.
- McNEILL, D. (1992). *Hand and Mind : What gestures reveal about thought*. Chicago: The University of Chicago Press.
- McNEILL, D. (2005). *Gesture & thought*. Chicago: The University of Chicago Press.
- SANTOS, F. (2010). *Le kappa de Cohen : un outil de mesure de l'accord inter-juges sur des caractères qualitatifs*. Publication en ligne. [consultée le 28/03/2012]. http://frederic.santos.perso.sfr.fr/pdf/stat/Kappa_Cohen.pdf.
- TELLIER, M. et STAM, G. (2010). Découvrir le pouvoir de ses mains : La gestuelle des futurs enseignants de langue. In *Actes du Colloque « Spécificités et diversité des interactions didactiques : disciplines, finalités, contextes »*, Lyon.
- TELLIER, M., GUARDIOLA, M., BIGI, B. (2011). Types de gestes et utilisation de l'espace gestuel dans une description spatiale : méthodologie de l'annotation. In *Actes de l'Atelier DEGELS, 18èmes conférence annuelle Traitement Automatique des Langues Naturelles (TALN)* (2011 juin 27-juillet 1 : Montpellier, FRANCE). Montpellier: Université de Montpellier II. 2011, pages 45-56.