



De l'influence du lien communicatif entre geste manuel et parole dans leur coordination

Benjamin Roustan, Marion Dohen

► To cite this version:

Benjamin Roustan, Marion Dohen. De l'influence du lien communicatif entre geste manuel et parole dans leur coordination. RJCP 2011 - 9ème Rencontres des Jeunes Chercheurs en Parole, May 2011, Grenoble, France. pp.n.c. hal-00642471

HAL Id: hal-00642471

<https://hal.science/hal-00642471>

Submitted on 18 Nov 2011

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.

De l'influence du lien communicatif entre geste manuel et parole dans leur coordination

Benjamin Roustan & Marion Dohen

GIPSA-Lab, UMR 5216 CNRS/Grenoble INP/UJF/U. Stendhal

Département Parole et Cognition, gipsa-lab

961 rue de la Houille Blanche - Domaine universitaire - BP 46

38402 Saint Martin d'Hères CEDEX. FRANCE

{benjamin.roustan, marion.dohen}@gipsa-lab.grenoble-inp.fr

ABSTRACT

Manual gestures are often produced in spoken communication. It however remains unclear how they are coordinated to speech and the role the communicative link between the manual gesture and speech plays in this coordination. This study analyzes this issue in a designation task in which participants produced prosodic focus as well as three gesture types (pointing, beat and control gesture). In a previous experiment we showed that when pointing and prosodic focus show exactly the same thing, they are synchronized. In this study, prosodic focus and pointing did not show exactly the same thing (modulation of the communicative link between the manual gesture and speech) and we analyzed the effect this had on their coordination. The results show that the communicative relationship between gesture and speech has an influence on their coordination. The results also highlight inter-individual differences.

Keywords: multimodality, manual gesture, coordination, prosodic focus, pointing

1. Introduction

Dans un contexte de communication naturelle, la production de parole est très souvent accompagnée de gestes manuels (*e.g.* McNeill, 1992; Kendon, 2004). On ne sait cependant pas très bien comment ces gestes sont coordonnés à la parole. Certains auteurs proposent que la prosodie jouerait un rôle important dans cette coordination (*e.g.* Bolinger, 1983; McClave, 1998). Ce rôle serait particulièrement important dans le cas du geste de *battement* (geste rapide de la main de haut en bas) qui serait notamment lié à l'expression de l'emphase et de la focalisation prosodiques (*e.g.* McNeill, 1992; McClave, 1998). L'étude de ce geste et de sa coordination avec la parole pourrait donc se révéler particulièrement intéressante.

Un cadre communicatif est également particulièrement intéressant pour l'étude de la coordination entre parole et gestes manuels de part sa nature intrinsèquement multimodale : la désignation. On désigne avec la main en pointant c'est-à-dire en montrant un objet dans l'espace qui devient centre de l'attention partagée. Mais on désigne aussi en utilisant la parole et notamment la focalisation prosodique qui permet de faire émerger un mot ou un groupe de mots pour le dé-

signer comme étant particulièrement informatif. C'est ainsi, que le pointage et la focalisation prosodique sont liés (voir Løevenbruck et al., 2009, pour plus de précisions). Quelques études préliminaires ont déjà exploré la coordination temporelle entre pointage et focalisation. En particulier, de Ruiter (1998) trouve que le début du mouvement de pointer dépend de l'emplacement de la focalisation au sein d'un syntagme (groupe nominal) de type adjectif + nom. Rochet-Capellan et al. (2008) trouvent que l'exécution du geste est déplacée vers la syllabe accentuée d'un non-mot de 2 syllabes. Dans une étude précédente (Roustan & Dohen, 2010, nommée ci-après *exp1*) nous avons utilisé la capture de mouvements pour étudier la coordination temporelle entre la focalisation contrastive prosodique et différents types de gestes (pointage, battement et un geste de contrôle *i.e.* appui sur un bouton). Des phrases simples (de type Sujet-Verbe-Objet) étaient produites avec focalisation soit sur le sujet soit sur l'objet. Les participants devaient produire un geste manuel en même temps qu'ils parlaient. Les résultats montrent que la focalisation prosodique « attire » le geste et ce, quelque soit son type (les apex de ceux-ci étant réalisés en même temps que l'élément focalisé). La coordination la plus fine observée était entre le geste de pointage et les cibles articulatoires de la parole. C'est pour le geste de pointage que l'attraction était la plus forte et que la coordination observée était la plus fine et la plus consistante (intra- et inter-sujets).

Les résultats quantitatifs sur la coordination gestes manuels/parole sont très variables. Cette variabilité pourrait-être due à la nature des relations communicatives qu'entretiennent geste manuel et parole –relations toujours différentes– dans les différentes études. C'est précisément l'influence du lien communicatif entre geste manuel et parole sur leur coordination que cette étude vise à explorer. Dans *exp1*, le pointage manuel et la focalisation prosodique portaient précisément sur le même élément (ex : focalisation de « BALAI » dans l'énoncé « Baba vend le balai » *et* pointage vers une image de balai). Dans cette étude, une manipulation expérimentale sera mise en place pour que pointage et focalisation ne désignent plus tout à fait la même chose. Dans l'étude de de Ruiter, l'objet désigné manuellement était le syntagme complet alors que la focalisation désignait seulement un élément de ce syntagme (ex : focalisation de « VERT » dans l'énoncé « Le crocodile vert » *et* pointage vers une image de crocodile vert). Or l'auteur

trouve une dépendance de la coordination à la focalisation. Ces résultats suggèrent donc que, même si le pointage désigne une unité de sens plus large (objet+couleur) que ce qui est focalisé prosodiquement (objet ou couleur), son exécution est modifiée par l'emplacement de la focalisation prosodique. Une des limites de cette étude est qu'elle n'a analysé l'influence de la focalisation que pour l'initiation du geste et pas pour la partie du geste qui montre *i.e.* l'apex.

2. Matériels et Méthodes

2.1. Participants

Les mêmes 10 participants que ceux de *exp1* ont passé cette expérience (l'ordre de passage des deux expériences était randomisé). Tous (8 femmes, 2 hommes) étaient droitiers de langue maternelle française et âgés de 30,2 ans en moyenne.

2.2. Matériel

Les phrases utilisées étaient toutes de type Sujet-Verbe-Objet (*e.g.* « Le bonbon est rouge »). La structure syllabique était fixée à $S=1+2\text{syll}$ (article défini + nom d'objet) ; $V=1\text{syll}$ (verbe être, présent simple) ; $O=1\text{syll}$ (couleur). Toutes les syllabes étaient de type CV sauf pour O pour lequel le type était CVC.

2.3. Plan d'expérience

Nous avons étudié quatre conditions *geste* : pas de geste (parole seule), geste de pointage (communicatif et deictique), geste de battement (communicatif mais non deictique), geste de contrôle (appui sur un bouton – non communicatif, non deictique). Deux conditions de *focalisation* ont été analysées : focalisation contrastive prosodique sur le sujet (SF) ou sur l'objet (OF).

Les participants devaient effectuer une tâche de correction qui élicitait naturellement la production de focalisation contrastive prosodique. Ils entendaient un prompt audio (proposition déclarative), deux images apparaissaient ensuite sur un écran. Ces images induisaient naturellement une correction sur le prompt entendu, comme illustré dans l'exemple suivant :

	SF	OF
Prompt	- Le bonbon est vert.	- Le balai est rouge.
Images	bonbon rouge & ballon vert	bonbon vert & balai jaune
Participant	- le BALLON est vert	- le balai est JAUNE

L'instruction était simplement de produire le geste en parallèle de la correction. Dans la condition de *geste* de pointage, les participants devaient pointer vers l'image correspondante à la correction. Ils pointaient donc vers l'objet tout entier (*e.g.* un balai vert) alors que la focalisation se faisait seulement sur une propriété de celui-ci (*e.g.* le fait qu'il soit vert). Dans la condition *geste* de battement, ils devaient produire un mouvement rapide vertical de la main. Dans la condition *geste* de contrôle, ils devaient appuyer sur un bouton situé à côté d'eux. Aucune autre indication ne leur était donnée.

L'expérience était découpée en quatre blocs (un pour chaque condition *geste*). Chaque bloc était composé de 24 essais (6 phrases $\times$ 2 conditions *focalisation* $\times$ 2 répétitions). L'ordre de passage des blocs ainsi que celui des essais au sein de ceux-ci étaient aléatoires et

différents pour tous les participants.

2.4. Dispositif expérimental

Les participants étaient assis devant un écran semi-transparent sur lequel apparaissaient des images. Une position de repos était marquée sur une table située sur leur droite. Le bouton utilisé pour la condition *geste* de contrôle était situé sur cette même table, en avant. Les participants devaient placer leur index droit sur la position de repos avant et après la production d'un geste manuel. Un système de capture de mouvements 3D par infra-rouges (Optotrak NDI) a permis de suivre les mouvements articulatoires (4 marqueurs autour de la bouche) et manuels (marqueur sur l'ongle de l'index droit). Un microphone était utilisé pour enregistrer les productions vocales.

2.5. Mesures

Deux juges indépendants ont vérifié la conformité des productions gestuelles (type de geste correct) et acoustiques (phrase correcte, production de focalisation sur le bon élément). Toutes les erreurs de productions ont été écartées des analyses qui sont présentées dans ce texte.

Praat a été utilisé pour étiqueter les limites des syllabes. Les mouvements manuels ont été caractérisés par deux points : l'apex (P_A) et le début du mouvement de retour (P_R). On appelle tenue du geste le segment $[P_A; P_R]$. Pour le pointage, l'apex correspond à l'extension maximale du mouvement, il en est de même pour le contrôle (appui sur le bouton). Pour le geste de battement, le point le plus bas du battement vertical a été annoté comme apex. Une normalisation temporelle selon la durée de la production acoustique a été réalisée pour éviter des effets du contenu sémantique de la phrase ainsi que du temps de réponse. À l'issue de cette normalisation, le temps « 0 » correspond au début de l'énoncé et le temps « 1 » à sa fin.

3. Prédictions

Le but de cette étude est de comparer la coordination parole / geste manuel obtenu à ceux de *exp1*. Dans *exp1*, pointage et focalisation désignaient la même chose. Dans cette étude, la focalisation désigne une sous-partie de l'objet pointé (soit sa nature, soit sa couleur). Rappelons que dans *exp1*, pointage et focalisation étaient synchronisés. Ici, si on considère l'énoncé « Le ballon est rouge », on s'attend à observer la même coordination parole / pointage que l'élément focalisé soit « ballon » ou « rouge » puisque dans les deux cas, le pointage manuel désigne l'image d'un ballon rouge. Se pose alors la question du patron de coordination. Plusieurs possibilités sont envisageables : ① le pointage englobe la globalité de l'énoncé (à la fois les éléments « ballon » et « rouge ») ② le pointage englobe une partie des deux éléments ③ le pointage est synchronisé avec l'un des deux éléments. Concernant les gestes de battements, s'ils sont coordonnés avec des indices prosodiques, on s'attend à n'observer aucune différence par rapport à *exp1*. Enfin, il en est de même pour le geste de contrôle qui n'est pas lié à la parole en terme d'intention communicative.

4. Résultats

Les variables dépendantes (instants de réalisation de $P_A : t_{P_A}$ et $P_R : t_{P_R}$) ont été testées par des ANOVAs à deux facteurs : *focalisation* (2 niveaux : SF et OF) et *geste* (3 niveaux avec geste : pointage, battement et contrôle). Les résultats de ces ANOVAs sont présentées dans la Table 1. On note un effet significatif de la condition focalisation à la fois sur l'instant de réalisation de l'apex et sur celui du retour : comme dans *exp1* le geste est produit plus tard quand la focalisation touche la fin de l'énoncé. On observe de plus un effet significatif de la condition *geste* mais seulement pour l'instant de réalisation de l'apex. Ceci traduit le fait que les différents gestes ne sont pas réalisés de la même manière ce qui est trivial puisqu'ils sont différents. L'interaction n'est jamais significative ce qui montre que la focalisation impacte tous les gestes de manière similaire. Ces résultats sont très similaires à ceux de *exp1*. Ils sont en accord avec nos prédictions sauf pour le geste de pointage.

Table 1: Résultats des ANOVAs sur t_{P_A} et t_{P_R}

	t_{P_A}	t_{P_R}
Focalisation	$F(1, 9) = 11.49, p < .01$	$F(1, 9) = 14.78, p < .01$
Geste	$F(2, 18) = 25.77, p < .001$	$F(2, 18) = 1.49, p = .25$
Focalisation $\times$ Geste	$F(2, 18) = 0.34, p = .71$	$F(2, 18) = 2.13, p = .15$

Une analyse plus fine des données montre qu'on peut cependant classer les participants en deux sous-groupes qui utilisent des stratégies de coordination geste/parole très différentes (en particulier pour le geste de pointage). Ceux du groupe 1 (G1 : 4 participants) utilisent la même stratégie de coordination que dans *exp1*. Ceux du groupe 2 (G2 : 6 participants) adoptent une stratégie différente. Des ANOVAs similaires à celles de la Table 1 peuvent être réalisées sur ces deux sous-groupes et sont présentées en Table 2. L'effet de la focalisation est alors significatif seulement pour le Groupe 1. Ceci suggère que les participants du groupe 2 coordonnent geste et parole de la même façon quelque soit l'élément focalisé (ex : « ballon » ou « rouge »).

Table 2: Résultats des ANOVAs sur t_{P_A} et t_{P_R} pour Groupe 1 (même stratégie que dans *exp1*) et Groupe 2 (stratégie différente)

	t_{P_A}	t_{P_R}
Focalisation	$F(1, 3) = 53.36, p < .01$	$F(1, 3) = 94.43, p < .01$
Geste	$F(2, 6) = 13.83, p < .001$	$F(2, 6) = 2.23, p = .19$
Focalisation $\times$ Geste	$F(2, 6) = 0.73, p = .52$	$F(2, 6) = 6.7, p < .05$
Focalisation	$F(1, 5) = 2.25, p = .19$	$F(1, 5) = 3.85, p = .11$
Geste	$F(2, 10) = 12.09, p < .01$	$F(2, 10) = 0.19, p = .83$
Focalisation $\times$ Geste	$F(2, 10) = 0.48, p = .63$	$F(2, 10) = .07, p = .93$

Afin d'affiner cette analyse, nous avons tracé les emplacements des gestes de pointage relativement à la parole pour chaque participant et dans chaque expérience (voir Figure 1). De plus, la Figure 2 montre la différence de réalisation temporelle pour l'apex et le début acoustique de la focalisation relativement au début de l'énoncé entre *exp1* et cette expérience (une valeur négative correspond à un événement intervenant plus tôt dans cette expérience que dans *exp1*).

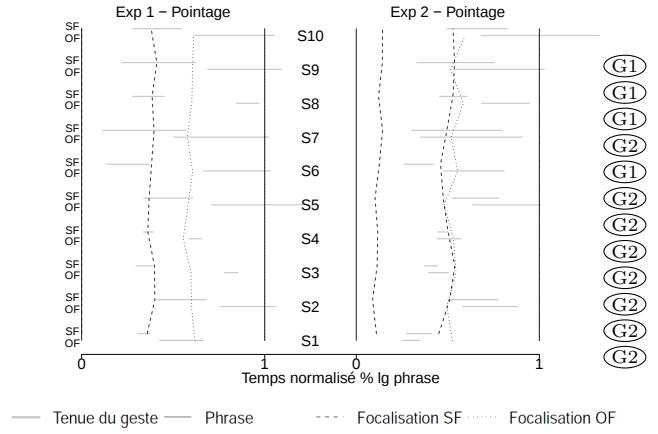


Figure 1: Organisation temporelle de la parole et du geste de pointage pour chaque participant

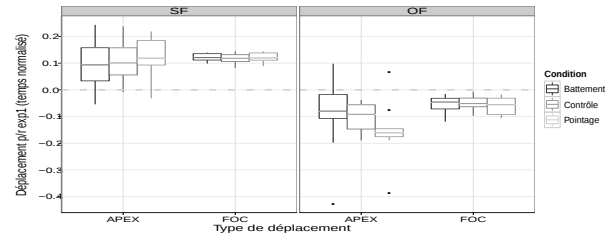


Figure 2: Déplacement temporels pour l'apex et la partie focalisée entre *exp1* et l'expérience courante

La Figure 2 montre qu'en SF l'élément focalisés est systématiquement produit plus tard dans cette expérience par rapport à *exp1*. Ceci est simplement dû au fait que dans cette expérience, il y a un déterminant avant le nom focalisé, ce qui n'était pas le cas dans *exp1*. Le décalage des gestes et de la parole étant identique, ce résultat suggère une coordination similaire à *exp1* en condition SF. Cependant, en OF, les résultats diffèrent et dépendent du type de geste.

Pointage : Pour le Groupe 2 (sauf S5), la Figure 1 montre qu'en OF, le pointage braccchio-manuel est largement déporté depuis l'objet (*exp1*) vers le sujet (l'expérience courante). Par ailleurs, la Figure 2 confirme cette tendance et permet de voir que ce déplacement de l'exécution du geste ne peut pas être dû à un simple déplacement de l'objet (déplacement de l'objet entre *exp1* et cette expérience : -0.061 , déplacement de l'apex : -0.155 qui est deux fois plus grand que le déplacement de l'objet). Ceci est confirmé par des tests t de Welch (déplacement focalisation *vs* déplacement apex : $t(10.76) = 2.56, p = 0.03$).

Geste de battement : Dans *exp1*, nous n'avons pas trouvé de synchronisation spécifique entre battements et la focalisation prosodique même si ces gestes étaient aussi attirés par la focalisation. Ce manque de coordination fine avait été partiellement expliqué par le fait que, d'après les participants, ces gestes étaient difficiles à réaliser à la demande et de manière non spon-tanée. Ici les résultats sont semblables à *exp1* : le geste de battement est attiré par la focalisation mais on ne trouve pas de synchronisation spécifique. La Figure 2 montre que la différence entre les deux études est beaucoup moins prononcée que pour le pointage et que le décalage temporel de l'apex en OF entre *exp1* et cette étude n'est pas statistiquement différent du

déplacement de l'objet : la coordination est similaire à ce qu'on observait dans *exp1*.

Geste de contrôle : Dans *exp1*, le schéma de coordination observé pour le geste de contrôle était très proche de celui observé pour le geste de pointage bien que la coordination soit beaucoup moins fine et plus variable. Dans cette expérience, le schéma de coordination est différent de celui observé pour *exp1* mais reste proche du schéma du geste de pointage : le geste arrive plus tôt qu'en *exp1* dans la condition OF.

5. Conclusion et discussion

Le but de cette expérience était d'étudier l'influence du lien communicatif entre geste manuel et parole sur leur coordination dans le cadre d'une tâche de désignation. Plus précisément, le pointage manuel désignait à la fois la nature de l'objet et sa couleur (*e.g.* un ballon rouge) alors qu'une seule de ces deux informations était focalisée dans la parole (*e.g.* « le ballon est ROUGE »). Cette étude utilisait le même protocole qu'une étude précédente (Roustan & Dohen, 2010, *exp1*) dans laquelle il y avait correspondance exacte entre ce qui était désigné par le pointage et par la focalisation en parole (ex : « Baba vend le BALAI » avec pointage vers l'image d'un balai). Ainsi, le but de cette expérience était surtout de comparer les stratégies de coordination des participants dans les deux expériences. Les productions de dix participants ont été acquises (capture de mouvements et enregistrement acoustique) pour deux conditions de focalisation (sujet vs objet) et quatre conditions de geste (parole seule vs parole + geste de pointage vs parole + geste de battement vs parole + geste de contrôle).

Une analyse générale des résultats a d'abord suggéré des patrons de coordination similaires à ceux observés dans *exp1* : la focalisation « attire » le geste. Ceci correspondait à nos prédictions pour les gestes de battement et de contrôle mais pas pour le geste de pointage. Cependant, une analyse plus détaillée des données a permis de montrer que les participants peuvent en fait être divisés en deux groupes : certains utilisent le même patron coordinatif que dans *exp1* mais pas les autres (surtout pour le geste de pointage). Pour ce dernier groupe et pour le geste de pointage, le patron coordinatif était similaire à celui de *exp1* dans la condition focalisation SF. Cependant, ce n'était pas le cas en OF où une tendance à décaler la production du geste vers l'objet était toujours observée, mais avec une amplitude beaucoup moins grande, ce décalage (trop important) ne pouvant être expliqué par un décalage temporel de l'objet entre les deux expériences. Au lieu d'avoir synchronisation systématique entre pointage et focalisation comme dans *exp1*, on n'observe ici qu'un léger décalage du pointage vers la focalisation. Comme dans *exp1*, le geste de contrôle semble obéir à un patron coordinatif similaire à celui du geste de pointage. Ceci était expliqué dans *exp1* par le fait que ce geste était trop proche du geste de pointage (extension du bras pour appuyer sur un bouton). Cependant, les résultats ici observés sont plus proches de *exp1* que pour le geste de pointage. Enfin, pour le geste de battement, bien qu'il existe une légère différence entre les deux expériences, la coordi-

nation geste manuel / parole reste très proche de celle observée pour *exp1*.

La coordination entre gestes manuels et parole est donc modulée par le lien communicationnel entre ces deux modalités. Dans *exp1*, le pointage était synchronisé avec la focalisation parce qu'ils désignaient strictement la même chose. Cette synchronisation n'est cependant plus observée quand la focalisation désigne une sous-partie de l'énoncé alors que le pointage désigne toute l'information contenue dans l'énoncé (cas de la présente étude). Bien que le geste de battement ne soit pas synchronisé très finement avec la focalisation, il est « attiré » par celle-ci de la même manière dans les deux expériences puisque cette fois, le lien communicationnel parole / geste manuel ne change pas.

Notons que la deixis n'est pas multimodale uniquement par la combinaison parole / geste manuel. En parole spontanée, elle peut également être accompagnée de mouvements de la tête et de pointage oculaire par exemple. Nous n'avons pas exploré ces gestes dans cette étude. Il serait néanmoins intéressant d'étudier comment toutes ces modalités se coordonnent entre elles et si l'usage d'une modalité additionnelle peut avoir un effet sur la coordination parole / geste manuel.

6. Remerciements

Nous remercions tous nos participants ainsi que Coriandre Vilain pour son aide technique et Jean-Luc Schwartz pour ses conseils avisés.

Références

- Bolinger, D. (1983). Intonation and gesture. *American Speech*, 58, 156-174.
- de Ruiter, J. (1998). *Gesture and speech production*. Thèse de doctorat non publiée, Catholic University of Nijmegen, Netherlands.
- Kendon, A. (2004). *Gesture : Visible action as utterance*. Cambridge University Press. Paperback.
- Lœvenbruck, H., Dohen, M., & Vilain, C. (2009). Pointing is special. In S. Fuchs, P. Perrier, & H. Lœvenbruck (Eds.), *Some aspects of speech and the brain* (p. 211-258). Peter Lang Verlag.
- McClave, E. Z. (1998). Pitch and manual gestures. *Journal of Psycholinguistic Research*, 27(1), 69-89.
- McNeill, D. (1992). *Hand and mind : What gestures reveal about thought*. University Of Chicago Press. Hardcover.
- Rochet-Capellan, A., Laboissière, R., Galvan, A., & Schwartz, J.-L. (2008, December). The speech focus position effect on jaw-finger coordination in a pointing task. *Journal of Speech, Language, and Hearing Research*, 51, 1507-1521.
- Roustan, B., & Dohen, M. (2010, May). Co-production of contrastive prosodic focus and manual gestures : Temporal coordination and effects on the acoustic and articulatory correlates of focus. In *Speech prosody 2010 conference*. Chicago, USA.