CEREBRAL CORRELATES OF MULTIMODAL POINTING: AN FMRI STUDY OF PROSODIC FOCUS, SYNTACTIC EXTRACTION, DIGITAL- AND OCULAR- POINTING

Hélène Lœvenbruck[1], Coriandre Vilain[1], Francesca Carota[1,2], Monica Baciu[3], Christian Abry[1], Laurent Lamalle[4], Cédric Pichat[3], Christoph Segebarth[5]

[1] ICP / Speech and Cognition Department, GIPSA-lab, UMR CNRS 5216, Grenoble, France
[2] Institut des Sciences Cognitives, UMR 5015 CNRS UCB Lyon 1, Bron, France
[3] Laboratoire de Psychologie et NeuroCognition, UMR CNRS 5105, UPMF, Grenoble, France
[4] INSERM IFR n° 1, RMN biomédicale, Unité IRM 3T, CHU de Grenoble, France
[5] Unité Mixte INSERM / Univ. J. Fourier, U594, Grenoble, France
{helene.loevenbruck; Coriandre.Vilain}@gipsa-lab.inpg.fr; francesca.carota@isc.cnrs.fr; christian.abry@u-grenoble3.fr; {Monica.Baciu; cedric.pichat}@upmf-grenoble.fr; {Laurent.Lamalle; Christoph.Segebarth}@ujf-grenoble.fr

ABSTRACT

Deixis or pointing plays a crucial role in language acquisition and speech communication and can be conveyed in several modalities. The aim of this paper is to explore the cerebral substrate of multimodal pointing actions. We present an fMRI study of pointing including: 1) index finger pointing, 2) eye pointing, 3) prosodic focus production, 4) syntactic extraction (during speech production). Fifteen subjects were examined while they gave digital, ocular and oral responses inside the 3T imager. Results of a random effect group analysis show that digital and prosodic pointings recruit the parietal lobe bilaterally, while ocular and syntactic pointings do not. A grammaticalization process is suggested to explain the lack of parietal activation in the syntactic condition. Further analyses are carried out on the link between digital and prosodic parietal activations.

Keywords: Prosody, syntax, speech production, manual and ocular pointing, fMRI.

1. INTRODUCTION

Deixis, or the ability to draw the listener's attention to an object – gradually acquired by children, first by pointing with the eyes, then the finger, then with intonation and finally with syntax–, is crucial in speech communication [1, 2]. All pointing actions share the same purpose of bringing relevant information to the interlocutors' attention. Pointing could exhibit common cerebral correlates for its several modalities. The existence of a specific cerebral network shared by different pointing modalities is suggested by previous works. Digital and ocular pointing would recruit a network including the left posterior parietal and frontal cortex [3]. Verbal pointing, such as prosodic pointing (i.e. focus) and syntactic pointing (i.e. syntactic extraction) would share a common neural network. Prosodic pointing recruits a left temporo-parieto-frontal network, including Wernicke's area, the supramarginal gyrus and Broca's region [4, but 5]. Syntactic pointing only involves Broca's region [4]. Activation of the inferior parietal lobule during prosodic pointing suggests the existence of a functional continuity with digital (and ocular) pointing; the activation of Broca's area in both syntactic and prosodic conditions evokes a second common region.

The aim of this fMRI study was to investigate the cerebral correlates of pointing. To test the hypothesis that multimodal pointing could involve a continuum of cerebral regions, we designed an fMRI paradigm including a continuum of conditions: 1) index finger pointing, 2) eye pointing, 3) prosodic focus, 4) syntactic extraction.

2. METHODS

2.1. Subjects

Sixteen healthy right-handed volunteers, aged 18-55 years, native speakers of French were examined. The study was performed in accordance with the institutional review board regulations. One subject did not perform one of the tasks and had to be excluded from the analysis. The results are reported for the fifteen remaining subjects.

2.2. Stimuli

Stimuli consisted of two types (1 and 2) of images illustrating a girl (Lise) and a boy (Jules) alternatively placed one on the right and the other on the left side of the screen. Type 1 images showed the girl holding a book, while the boy did not; Type 2 images showed the reverse. In the middle of all images, a red or green fixation cross was displayed. A blank screen with a mid-centered black cross preceded each stimulus.

2.3. Task and procedures

The tasks consisted of verbal and non-verbal pointing to a character. Control conditions were included. The same question: “Est-ce que Lise tient le livre?” (Does Lise hold the book?) was used for all tasks. Subjects were instructed: (a) to confirm the question when type 1 images were presented (*control*), (b) to point to Jules when type 2 images were presented (*contrastive pointing*).

Each of these tasks consisted of two phases: a *preparation* phase, indicated to subjects by the red fixation cross, and an *execution* phase indicated by the green cross. The separation into these two phases was motivated by a study on ocular and digital pointing where preparation recruits many of the areas involved in pointing execution (see [4]).

Tasks: During prosodic pointing, subjects uttered « JULES tient le livre » (Jules holds the book), with a contrastive focus on “Jules”. During syntactic pointing, they uttered « C’est Jules qui tient le livre » (It is Jules who holds the book), with syntactic extraction of “Jules”. During digital pointing, subjects pointed with the right index finger to Jules. During ocular pointing, subjects looked in the direction of the target character.

Controls: In prosodic and syntactic controls, subjects neutrally uttered: « Lise tient le livre » (Lise holds the book). In digital control, they performed a downward finger movement. In ocular control, they made a downward eye movement.

2.4. fMRI paradigm

A pseudorandom, event-related fMRI paradigm was employed. Four functional scans were acquired, one for each type of pointing.

Each functional scan included a sequence of the following four conditions: Preparation of the control task (Pc), Preparation + Execution of the control task (PEc), Preparation of the pointing task (Pp), Preparation + Execution of the pointing task (PEp). A null event (NE, black fixation cross) was added to the four conditions. Five conditions (Pc, PEc, Pp, PEp, NE) were thus included.

The five conditions were alternated between scans and between subjects. 24 repetitions of each condition were presented. Trials were presented as events lasting 4.5 s. The duration of the PEp and PEc conditions was of 0.5 s for the initial fixation cross + 2 s for the preparation phase + 2 s for the execution phase. The Pp and Pc conditions consisted of: 0.5 s for the fixation cross + 4 s for the preparation phase alone. The NE condition also lasted 4.5 s. The trial sequences were presented following a pseudo-random order. Total duration of a scan was approximately 9 mn (4.5 s × 5 conditions × 24 trials).

2.5. Apparatus

Stimuli were presented using the Presentation software [6]. They were viewed through a mirror attached in front of the subject’s eyes. Ear plugs and anti-noise headphones protected subjects against the scanner noise. Three types of behavioural responses were recorded: vocal production, eye movement and index finger movement. Verbal responses were recorded using an fMRI-compatible microphone. To minimize the amount of noise recorded, the microphone was positioned out of the scanner, at one extremity of a wave guide consisting of a soft plastic tube. The other extremity of the tube was connected with a mask placed over the subjects’ mouth. This apparatus reinforced the signal-to-noise ratio. Eye position was monitored using an ASL 504 eye-tracker (*Applied Science Laboratories*, Bedford, MA) coupled with the scanner. Right index finger responses were recorded using a digital camera placed out of the fMRI room, behind the window.

2.6. fMRI data acquisition and processing

A whole-body 3 Tesla MRI imager (Bruker) with gradient echo acquisition was used to measure blood oxygenation level-dependent contrast over the whole brain (repetition time: 2.5s; echo time: 30 ms; field of view: 216 x 216 mm, acquisition matrix: 72x72; reconstruction matrix: 128 x 128; 7 dummy scans). Forty-one 3.5 mm axial interleaved slices were imaged adjacent and parallel to the bi-commissural plane, encompassing the whole brain and the cerebellum. A high-resolution 3D anatomical scan was obtained. Anatomical images were acquired using a sagittal

MPRAGE sequence (inversion time: 900ms, volume: 176 x 224 x 256 mm; resolution: 1.375 x 1.750 x 1.33 mm; acquisition matrix: 128 x 128 x 192; reconstruction matrix: 256 x 128 x 128). A B0 fieldmap was acquired twice. Functional data analysis was performed using SPM2 software (Wellcome Department of Cognitive Neurology, London). Functional data were realigned to correct for head motion using a rigid body transformation. A spatial normalisation was applied. The functional images were spatially smoothed.

To examine cerebral activation during the crucial part of the trials, the haemodynamic response to the onset of each event was modelled with a delayed haemodynamic response function, shifted to onset 3.5 s later. Contrasts between conditions were determined voxelwise using the General Linear Model. To perform a random effect analysis, the contrast images (pointing *vs.* control) were calculated for each subject individually and were then entered into a one-sample t test with a significance threshold of p=0.02 uncorrected.

3. RESULTS

3.1. Behavioural results

Monitoring the subjects' utterances included word accuracy and correct prosodic contour productions. Word recognition was possible in the audio signals from all subjects but 2. F0 examination was possible on 9 out of 16 subjects. In the other cases, technical problems with the microphone or the mask drastically degraded the signal. Word accuracy and F0 patterns were correct in all available prosodic and syntactic conditions (and their controls). Eye-tracking data were analyzed with the eye-tracker software. Horizontal positions of the eyes were checked using a Matlab script. Preliminary results suggest that subjects behaved according to the instructions. An overview of the video data on finger movements suggests that all the subjects performed adequately.

3.2. FMRI results

Patterns of activation were examined for 15 subjects for each of the 4 following contrasts:

- digital pointing tasks *vs.* digital control
- ocular pointing tasks *vs.*ocular control
- prosodic pointing tasks *vs.*prosodic control
- syntactic pointing tasks *vs.*syntactic control

We first examined the activations for the preparation phases, as done in [3]. The random effect group analysis did not reveal any significant activation in any of the contrasts. A possible explanation is that the preparation task might not have been well explained to the subjects. Some subjects reported that they had been executing the pointing or control task covertly, while others reported that they were trying to prepare for the following task, as instructed. We therefore decided to examine the execution phase only, for which monitoring ensured correct behaviour.

Table 1 represents the peaks of activations and their corresponding Talairach coordinates provided by the random effect group analysis.

Table 1: Talairach coordinates and Z-scores of activated regions in the pointing *vs.* control conditions

Region	Talairach coord. in mm (x, y, z)	Z-scores
Digital pointing vs. control		
R. Precentral Gyrus, BA 4	34 -20 59	3,63
R. Sup. Parietal Lobule, BA 7	16 -57 64	3,24
L. Postcentral Gyrus, BA 2	-59 -27 40	3,10
L. Supramarginal Gyrus, BA 40	-42 -45 38	3,06
R. Supramarginal Gyrus, BA 40	55 -35 32	2,98
Ocular pointing vs. control		
L. Mid. Occipital Gyrus, BA 18	-32 -83 2	3,57
L. Precentral Gyrus, BA 4	-16 -30 68	3,41
R. Parahippoc. Gyrus, BA 28	18 -22 -17	3,39
R. Lingual Gyrus, BA 17	22 -89 -1	3,34
L.Middle Frontal Gyrus, BA 8	-20 23 40	3,29
Prosodic pointing vs. control		
R. Sup. Parietal Lobule, BA 7	28 -47 60	3,76
R. insula, BA 13	30 20 10	3,46
L. Cuneus, BA 18	-16 -71 17	3,17
L. Precuneus, BA 7	-20 -64 31	3,13
L. Inferior Frontal Gyrus, BA 47	-46 13 -6	3,13
L. Sup. Temporal Gyrus, BA 42	-61 -28 15	2,91
L. Precuneus, BA 7	-12 -46 58	2,86
R. Sup.Temporal Gyrus, BA 22	61 -56 14	2,75
R. Precentral Gyrus, BA 44	51 10 13	2,75
R. Lingual Gyrus, BA 19	16 -60 3	2,65
L. Inf. Parietal Lobule, BA 40	-61 -33 35	2,59
Syntactic pointing vs. control		
R. Precentral Gyrus, BA 4	53 -12 37	4,02
L. Postcentral Gyrus, BA 43	-53 -7 22	3,44
L. Middle Frontal Gyrus, BA 47	-38 39 -7	3,01
R. Middle Frontal Gyrus, BA 46	42 16 19	2,99
R. Mid. Temporal Gyrus, BA 21	59 -16 -9	2,98
L. Mid. Temporal Gyrus, BA 21	-57 -18 -7	2,61

Figs. 1 and 2 represent functional activations during each pointing condition *vs.* its control.

Activation for the *digital pointing* was slightly superior in the right hemisphere (RH) compared to the left (LH), contrary to expectations. Most of the activation lies bilaterally within the parietal lobe, in the superior parietal lobule (BA 7) and the supramarginal gyrus (BA 40).

The *ocular condition* activated the left hemisphere predominantly: in the middle occipital gyrus (BA 18) and in the Frontal Eye Field (BA 8).

The *prosodic condition* activated left perisylvian regions (BA 47, 42) and supramarginal gyrus as well as right perisylvian regions (BA 13, 22) and superior parietal lobule (BA 7).

The *syntactic condition* activated left perisylvian regions (BA 43, 47, 21) as well as right precentral and middle temporal gyri (BA 4, 21).

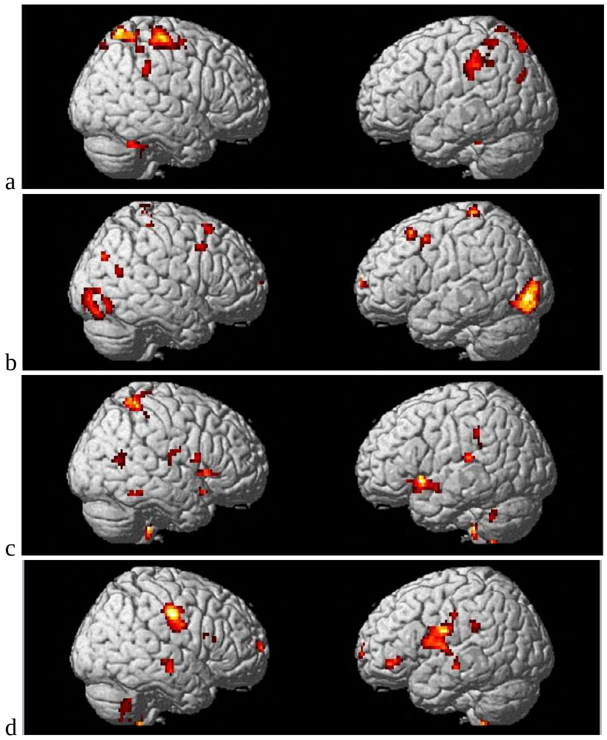

Figure 1: Activations during pointing *vs.* control conditions rendered on a sagittal template, in **a.** digital- **b.** ocular- **c.** prosodic- **d.** syntactic- modes. The left hemisphere is on the right, the right one on the left.

4. DISCUSSION AND CONCLUSION

A specially designed apparatus allowed us to monitor multimodal pointing during fMRI sessions. Crucially, overt speech responses were recorded during noisy 3T fMRI session. The fMRI results obtained for 15 subjects suggest the activation of the superior and inferior parietal lobes in manual (bilaterally) and prosodic pointing (right and left respectively). In accordance with [4], parietal activation did not reach statistical significance in the syntactic condition, which suggests that when grammaticalization occurs, parietal lobe activation becomes unnecessary and is replaced by perisylvian activation. Although there was no significant parietal activation during ocular pointing, this condition activated an occipito-frontal network, including the FEF, as reported by studies on ocular saccades [3]. The lack of parietal activation in the ocular condition needs further investigation. More analyses of the data are currently carried out to further explore the possible link between parietal activations related to digital and prosodic pointing.

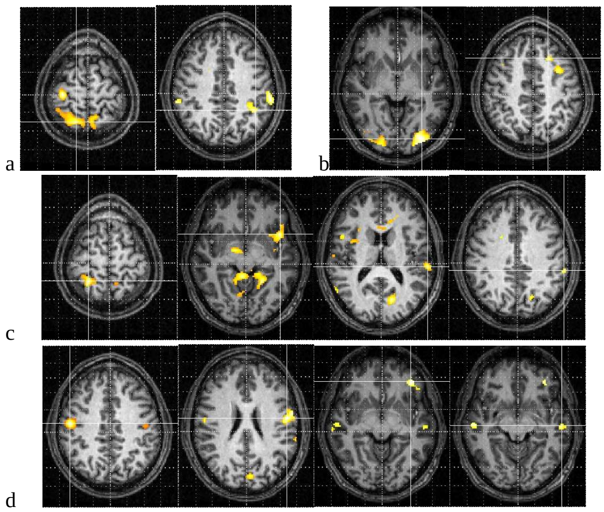

Figure 2: Activations during pointing *vs.* control conditions in **a.** digital- (BA 7 in RH and BA 40 in LH) **b.** ocular- (BA 18 and BA 8 in LH) **c.** prosodic- (BA 7 in RH, BA 47, 42, 40 in LH) **d.** syntactic- (BA 4 in RH, BA 43, 47, 21 in LH) modes. The LH is on the right, the RH on the left.

5. REFERENCES

[1] Kita, S., 2003. Pointing: where language, culture, and cognition meet. Mahwah, NJ: Lawrence Erlbaum.

[2] Goldin-Meadow, S., 1999. The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, 3, 419-429.

[3] Astafiev S. V., Shulman, G. L., Stanley, C. M., Snyder, A. Z., Van Essen, D. C., Corbetta M., 2003. Functional Organization of Human Intraparietal and Frontal Cortex for Attending, Looking, and Pointing. *Journal of Neuroscience*, 23, 4689-4699.

[4] Lœvenbruck H., Baciu M., Segebarth C., Abry C., 2005. The left inferior frontal gyrus under focus: an fMRI study of the production of deixis via syntactic extraction and prosodic focus. *J. Neurolinguistics*, 18, 237-258.

[5] Dogil G., Ackermann H., Grodd W., Haider H., Kamp H., Mayer J., Riecker A., Wildgruber D., 2002. The speaking brain: a tutorial introduction to fMRI experiments in the production of speech, prosody and syntax. *J. Neurolinguistics*, 15, 1, 59-90.

[6] http://nbs.neurobs.com/presentation/