HAL
open science

Ontogeny of two communicative tools : Distance encoding and multimodality in deictic pointing

Chloe Gonseth, Anne Vilain, Coriandre Emmanuel Vilain

To cite this version:

Chloe Gonseth, Anne Vilain, Coriandre Emmanuel Vilain. Ontogeny of two communicative tools : Distance encoding and multimodality in deictic pointing. Evolang 2012 - 9th International Conference on the Evolution of Language, Mar 2012, Kyoto, Japan. pp.150-157. hal-00704598

HAL Id: hal-00704598

https://hal.science/hal-00704598v1

Submitted on 5 Jun 2012

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.

ONTOGENY OF TWO COMMUNICATIVE TOOLS: DISTANCE ENCODING & MULTIMODALITY IN DEICTIC POINTING

CHLOE GONSETH, ANNE VILAIN, & CORIANDRE VILAIN

GIPSA-Lab, Speech & Cognition Department, Grenoble University
Grenoble, 38000, France

The aim of this paper is to experimentally study the development of speech/gesture relationship within language. The participants, both adults and children, had to designate a target placed at two different distances. They had to perform a deictic gesture accompanied by a deictic word, or to use either of the two modalities independently. Using one vs two pointing modalities allowed us to specify the interaction between the two systems as a useful cooperation, in terms of communicative efficiency. Furthermore, articulatory cues in the vocal modality and kinematic cues in the gestural modality encode the target's distance, and that in a different way in adults as compared with children. The huge variability found in children indicates that they are learning to control the communicative tools they have.

1. Introduction

This study aims at questioning the functional relationship between speech and gesture in the linguistic communication of children and adults, and focusses on the process of spatial deixis. This process, allowing us to attract our interlocutor's attention, is most of the time performed in a multimodal way, as it can be realized both via the vocal and the gestural system, using various pointing gestures. These deictic gestures are universal across cultures; they occupy a particular function in adult language, being ubiquitous in everyday interactions and often accompanied by a deictic word such as "here" or "that" (Kendon, 2004). Data suggest that deictic pointing might be one of the most basic elements of referential communication, and maybe one of the precursors of the phylogenetic development of human language (Arbib, 2005).

In this study dealing with deixis, we investigate (i) the interaction between the vocal and manual systems as a possible cooperation, and (ii) the distance encoding in both vocal and manual pointing as a particular deictic, and therefore communicative function. The distance of a designated object with respect to the speaker and the listener is a spatial feature that we believe can be encoded both

through the kinematic properties of manual pointing gestures and through the lexical and phonetic properties of the deictic terms that are used (see Bonfiglioli et al., 2009, for interaction between lexical encoding of distance and kinematic properties of pointing gestures). Deictic terms encoding distance can be either proximal (e.g. “here”) or distal (e.g. “there”). Typological studies evidence a specific relationship between the encoded distance and the phonological features of the appropriate deictic term. Diessel (1999) notably demonstrated that proximal forms of deictics tend to contain more close vowels than distal forms. Close vowels, contrary to open vowels, are produced with the tongue positioned high in the mouth and with a consequently high jaw position. This is taken as a sound symbolism phenomenon, that is a consistent link beween phonetics and semantics. Our hypothesis is that this typological tendency might be explained by a very basic motor behaviour: speakers tend to open their mouth more when they designate distant objects than when they designate close objects.

In a former work on adult language (Gonseth et al., 2011), we showed that the distance of the target (close vs distant) and the modalities used to designate this target (*speech* vs *speech+gesture*) influenced either acoustical values or articulatory gestures. These findings suggest that distance can be encoded in vocal pointing and that vocal and manual systems cooperate within the particular deictic function. Such a crucial interaction between speech, gesture, and language in adults evidently calls for further studies during the development of linguistic communication. Deictic pointing is indeed crucial when it comes to language ontogenesis. The canonical index finger pointing is the first tool used by children to designate what surrounds them (Kita, 2003). Its use is consistently correlated with important landmarks in the development of verbal production and perception competencies (Özçalışkan & Goldin-Meadow, 2005) and related to gains in language development and future communicative abilities (Butcher & Goldin-Meadow, 2000). Speech/gesture interaction then appears to be a critical factor in the inception of communicative development.

For these reasons, we investigated speech/gesture cooperation and distance encoding in the use of multimodal communication by children, as compared to adults. Though gestures are extremely present in both adult and child communication, their use depends on the speaker’s expertise. Children use the gestural modality as an introduction to the linguistic system, whereas gestures have more complex pragmatic functions in adults (e.g. to emphasize a specific discourse element). We then expect a different cognitive processing, resulting in a different spatial coordination between speech and gesture in children as compared with adults, and an evolution throughout the first years of linguistic development. Distance encoding, as a more subtle strategy, is expected to be

controled later. The present investigation is thus meant to elaborate a developmental schedule of multimodal linguistic pointing.

2. Method

The objective of this study is to replicate the findings of our previous experiment (Gonseth et al. 2011), and to collect new data in 6 to 12 year old children, in order to characterize the evolution of the correlates of distance encoding in both modalities. Are the distance encoding and the speech/gesture cooperation intrinsic features of language use (i.e. observed in both children and adults) or efficient communicative strategies learnt by skilled speakers (i.e. not observed in children, or much more variable than in adult speakers)?

2.1. *Partipants*

Twenty-seven French-speaking adults (ages 18-36 years) participated in the study. Besides, we recorded six French-speaking children (ages 7-12 years, see table 1) as a pilot child sample. All participants were voluntary and naive as to the purpose of the study.

Table 1. Age of each child.

Participant	P1	P2	P3	P4	P5	P6
Age (years)	8	11	11	7	12	8

2.2. *Procedure*

The participants are in a soundproof room in front of two light-emitting diodes (LED), placed at 140 (D1) and 425 cm (D2). Both LEDs are out of arm's reach, but D1 is considerably closer than D2. They are aligned with respect to the subjects' eyes, so that the angle of the arm can not be used to disambiguate which target is pointed at, and so that there has to be a specific gesture for distance encoding. As a reply to the experimenter's question "C'est où?" ("Where is it?"), participants have to designate the turned-on LED, according to three conditions (participants presented 20 trials per condition):

1. Speech: name the turned-on LED using the deictic word "là"/"there", embedded in a carrier sentence (i.e. preceded by "c'est" / "it is").

2. Gesture: point at the target with an index finger pointing.

3. Speech+gesture: simultaneously name and point at the target.

We thus test a cooperation between speech and gesture: vocal and manual pointing are expected to be reinforced in unimodal conditions (*speech* or *gesture*) as compared to the bimodal condition (*speech+gesture*).

3. Recordings & measurements

3.1. *Articulatory and manual data*

Oro-facial and index movements were recorded with the motion capture system Optotrak. Three reference emitters were placed on the forehead, two emitters on the lips, and one emitter on the index. Lip opening was thus measured (lower-lip vertical position minus upper-lip vertical position, in mm), as well as the total duration and the holding duration (s) of the index finger pointing.

3.2. *Acoustical data*

Acoustical features of the vowels were recorded in order to get additional information on articulatory data. Oral productions were recorded via a microphone AKG C1000S. We isolated the vowel in the deictic word “là” ([a]) and measured the first formant (F1, in Hz) in the middle of the vowel.
What we are investigating is the intra-category phonetic variation according to the distance of the target, that is: refering to our previous results, [a] in “là” is expected to be more open when “là” designates a distant target.

4. Results

Failed trials were excluded from the analysis. The median values of acoustical, articulatory, and manual data were used for the statistical analysis, with a significance level fixed at $p < 0.05$, and two fixed factors: (1) three conditions C (*speech*, *gesture*, *speech+gesture*); (2) two distances D (140; 425 cm).
For adults, a repeated-measures analysis of the variance was conducted; for children, an analysis of the variance was conducted for each participant.

4.1. *Articulatory and acoustical data*

4.1.1. *Effect of distance and condition on lip opening*

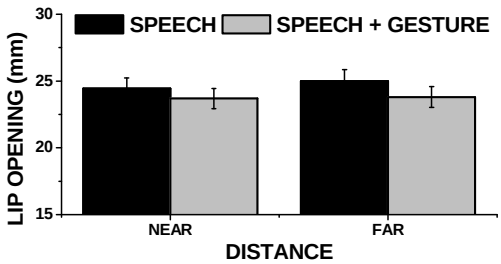

Figure 1. Lip opening values as a function of the condition and the distance for the adults.

Figure 1 illustrates lip opening variations as a function of both distance and condition in adults. Lip opening is influenced by the condition. It is indeed larger in the *speech* condition than in the *speech+gesture* condition ($F(1, 24)=17.06$, $p<0.05$). Furthermore, we find a significant effect of the distance ($F(1, 24)=8.87$, $p<0.05$). Regardless of the condition, lip opening is larger for farther distance.

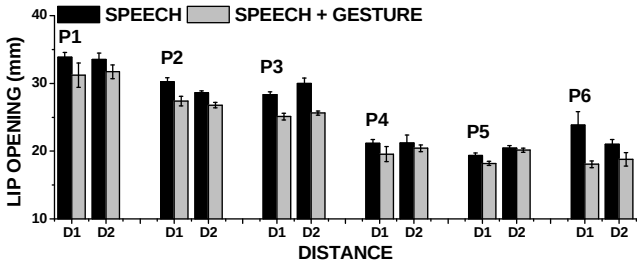

Figure 2. Lip opening values as a function of the condition and the distance for six children.

Figure 2 concerns lip opening variations for six children. For P1 and P4, lip opening does not vary as a function of the distance nor of the condition (P1 D: $F(1, 30)=0.28$, $p=0.60$, ns; C: $F(1,30)=1.99$, $p=0.17$, ns; P4: D: $F(1, 35)=1.17$, $p=0.29$, ns; C: $F(1, 35)=2.41$, $p=0.13$, ns). P2 and P6 show a significant effect of the condition (P2 D: $F(1, 35)=0.99$, $p=0.32$, ns; C: $F(1,35)=13.42$, $p<0.05$; P6 D: $F(1, 21)=1.51$, $p=0.23$, ns; C: $F(1, 21)=5.96$, $p<0.05$), whereas P5 shows a significant effet of the distance (D: $F(1, 32)=8.15$, $p<0.05$; C: $F(1,32)=1.59$, $p=0.21$, ns). Finally, both distance and condition affect the lip opening for P3 (D: $F(1, 35)=4.09$, $p=0.05$; C: $F(1,35)=39.54$, $p<0.05$).

4.1.2. *Effects of distance and condition on F1 values*

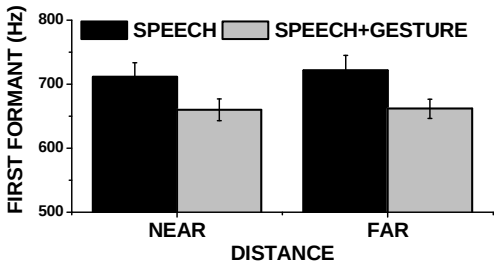

Figure 3. F1 values as a function of the condition and the distance for the adults.

Figure 3 illustrates F1 variations as a function of both distance and condition in adults. A significant effect of the condition is observed ($F(1, 26)=16.67$, $p<0.05$)

but no significant effect of the distance (F(1, 26)=0.84, p=0.36, ns). F1 values are higher in unimodal condition (*speech*) than in bimodal condition (*speech+gesture*) but do not vary as a function of the target's distance.

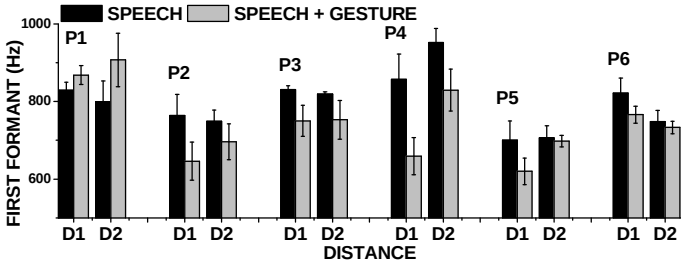

Figure 4. F1 values as a function of the condition and the distance for six children.

Figure 4 concerns F1 variations for six children. Four different patterns are found. Neither the condition nor the distance affect F1 values for P1 (C: F(1, 35)=1.81, p=0.19, ns; D: F(1, 35)=0.45, p=0.50, ns) and P6 (C: F(1, 18)=1.53, p=0.23, ns; D: F(1, 18)=1.80, p=0.19, ns). P2 and P3 show a significant effect of the condition (P2 C: F(1, 35)=8.02, p<0.05; D: F(1,35)=0.45, p=0.51, ns. P3 C: F(1,35)=12.69, p<0.05; D: F(1, 35)=0.17, p=0.68, ns). P5 presents a significant effect of the distance (D: F(1,36)=7.19, p<0.05; C: F(1,36)=0.20, p=0.66, ns). Finally, P4 shows a significant effect of both distance and condition (D: F(1, 37)=3.84, p=0.05; C: F(1, 37)=9.27, p<0.05).

4.2. *Kinematic encoding of the distance*

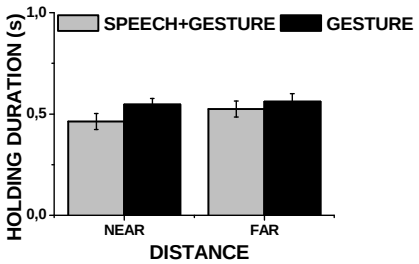

Figure 5. Holding duration as a function of the distance and the condition for the adults.

Figure 6. Total duration as a function of the distance and the condition for the adults.

Figures 5 and 6 illustrate the effect of both distance and condition on gesture durations in adults. We found a significant effect of both distance and condition (HD: D: F(1, 24)=19.84, p<0.05; C: F(1, 24)=32.60, p<0.05; TD: D: F(1,

24)=7.09, p<0.05; C: F(1, 24)=49.47). Adults perform longer manual pointing to designate distant objects and in unimodal condition.

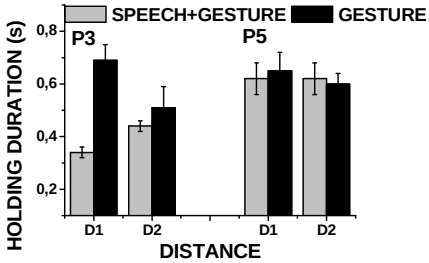

Figure 7. Holding duration as a function of the distance and the condition for two children.

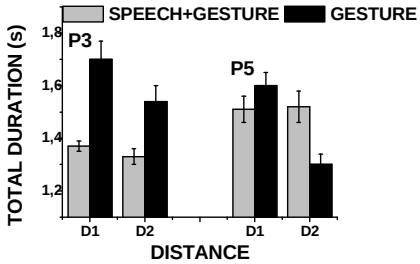

Figure 8. Total duration as a function of the distance and the condition for two children.

Figures 7 and 8 concern gesture durations for two children. Data for four participants are missing due to a technical problem. P3 presents an effect of the condition (HD: F(1, 37)=11.55, p=0.05; TD: F(1, 37)=25.71, p=0.05), but no effect of the distance (TD: F(1, 37)=0.001, p=0.97, ns; HD: F(1, 37)=0.24, p=0.63, ns), whereas no significant effects were found for P5 (C: HD: F(1, 38)=3e-04, p=,0.98 ns; TD: F(1, 38)=4e-04, p=0.98, ns; D: HD: F(1, 38)=0.23, p=0.63, ns; TD: F(1, 38)=0.95, p=0.33, ns).

5. Discussion & conclusion

Our data are, in coherence with our previous study, in favour of a specific distance encoding in both vocal and manual pointing. In adults, articulatory properties of the vocal gesture encode the distance information (i.e. lip opening, which increases with the distance of the target). This articulatory encoding suggests that the tendency of languages worldwide to associate close vowels with proximal deictics and open vowels with distal deictics could originate from a very basic motor correlate. The distance effect is observed here inside vocalic categories, discarding then the lexical origin of the distance encoding in language. For the adult group, gesture durations also depend on the target's distance: the index finger pointing lasts longer when directed to a distant object. Only two children out of six show F1 and/or lip opening increases as a function of the distance. This suggests that the distance encoding is not an intrinsic feature of language use but an efficient communicative strategy, which develops gradually during childhood.

Our study also attests the cooperation between the vocal and manual systems. They are used in a complementary way for the implementation of the

linguistic function of deixis. The information is distributed on the two systems. We indeed observed in adults a significant effect of the mono- vs bimodality of the task on F1 values, lip opening values, and kinematic properties of the manual gesture. Moreover, this effect is significant for four out of six children on one or more parameters. It means that when participants cannot perform any manual pointing, they compensate with a reinforced vocal pointing, and reciprocally. This cooperation is mutual, since the manual pointing also depends on the presence or absence of the vocal pointing. Our hypothesis is that this cooperation might be a very basic feature of language communication, as it is more developped in children than distance encoding, the latter being a more elaborate communicative strategy, depending on more mature pragmatic skills. Further analyses are in progress to precisely determine different developmental profiles. However the large variability found in children indicates that an acquisition process is occuring, leading to the full mastery of these communicative tools.

References

Arbib, M. A. (2005). From Monkey-like Action Recognition to Human Language: An Evolutionary Framework for Neurolinguistics, *Behavioral and Brain Sciences*, 28: 105-167.

Bonfiglioli, C., Finocchiaro, C., Gesierich, B., & Rositani, F. (2009). A kinematic approach to the conceptual representations of this and that, *Cognition*, 111: 270-274.

Butcher, C., & Goldin-Meadow, Susan (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235-257). New York: Cambridge University Press.

Diessel, H. (1999). *Demonstratives: Form, Function, and Grammaticalization*. John Benjamins, Amsterdam.

Gonseth, C., Vilain, A., & Vilain, C. (2011). Deictic pointing: How do speech and gesture cooperate to encode distance information. Submitted.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press, Cambridge.

Kita, S. (Ed) (2003). *Pointing, Where Language, Culture, and Cognition Meet*. Psychology Press.

Özçalışkan, S. & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development, *Cognition*, 96(3), B101-B113.