Does the Number of Syllables Affect the Finger Pointing Movement in a Pointing-naming Task?

Amélie Rochet-Capellan[2,1], Coriandre Vilain[1], Marion Dohen[1], Rafael Laboissière[3,4] and Jean-Luc Schwartz[1]

[1]Gipsa-Lab, Speech and Cognition Department, CNRS- Grenoble University, France
[2]Motor Control Lab, McGill University, Montréal, Canada
[3] INSERM, U864, Espace et Action, Bron, France
[4]Max Planck Institute for Human Cognitive and Brain Sciences, Munich, Germany

E-mail: amelie@motion.mcgill.ca, coriandre.vilain@gipsa-lab.inpg.fr, marion.dohen@gipsa-lab.inpg.fr, rafael.laboissiere@inserm.fr, jean-luc.schwartz@gipsa-lab.inpg.fr,

Abstract

This study analyses the finger-pointing gesture in a pointing-naming (PN) task for a group of native speakers of Brazilian and a group of native speakers of French. It shows that the increase of the number of syllables (n-syl.) in the target name slightly delays the timing of the finger gesture toward the target. The duration of that gesture also tends to be shorter for 1- and 2- syllable(s) in the PN task rather than in a pointing-alone (P) task. Finally, the finger-target alignment is globally longer for the P task than for the PN task. This duration then gradually increases with the increase of n-syl. in the PN task. These results reproduce the main published results showing that speech/pointing coordination is mainly achieved by 'on line' adaptation of speech and 'off-line' adaptation of the finger. Yet, the results also suggest that speech could 'calibrate' the finger-pointing gesture. Hence, in the simultaneous designation of a target, speech and hand should be considered as two 'collaborative' rather than two 'competitive' systems.

1 Introduction

A number of descriptive studies investigated the hand pointing gesture and its relationship with speech in ecological situations [1-3]. Yet, few experimental works analyzed the pointing/speech coordination. Using an experimental approach (pointing at a target saying 'that lamp'), Levelt et al. [4] showed that the voice/finger coordination mainly consisted in the adaptation of the voice onset to the pointing gesture. By contrast, the pointing gesture was weakly affected by the speech presence (see also [5]).

Using a similar approach, we showed that the jaw/finger coordination in a pointing-naming (PN) task adapted to the stress position within the target name (/pápa/ *vs.* /papá/) [6]. This adaptation occurred in two ways: (1) the jaw adapted 'on line' its onset in the course of the pointing gesture toward the target; and (2) the finger adapted 'off line' the duration of the finger-target alignment phase. This two-dimensional adaptation ensured that *the part of the discourse that shows* occurred during *the part of the gesture that shows*.

We then attempted to extend this conclusion to the variation of the number of syllables (*n-syl.*) in the naming utterance. The study involved two groups of speakers of languages with different stress patterns: native speakers of Brazilian-Portuguese (hereafter, 'B-group') and native speakers of French (hereafter, 'F-group'). This paper presents a first subset of results from this study: the effect of *n-syl.* on the timing and the kinematics of the pointing gesture.

2 Method

Participants and task

The 'B-group', 3 males and 7 females, performed the PN task as a pilot study always after the experiment reported in [6]. The 'F-group', 8 males and 7 females performed three tasks in counterbalanced orders: a

PN task; a Pointing-alone (P) task and a Naming-alone task (not reported here). All the participants were right-handed.

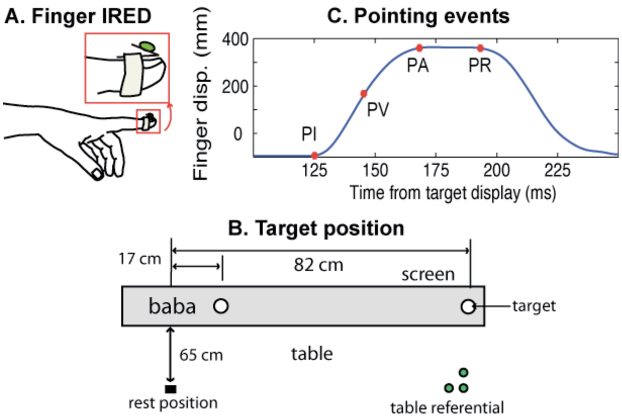

Figure 1. (A) Optotrak IREDs position, (B) Target position and (C) Optotrak signal labeling, finger displacement against time (PI: Pointing Initialization, PA: Apex, PR: Return, see text for details).

Procedure and data processing

For both studies, the participant was seated at a table. In the PN task, the target (a smiley symbol) was projected to the participant's right on a screen in front of him/her together with a word (see figure 1.B). The task was to point the target with the right forefinger and name it with the word at the same time, as soon as the target color changed (GO-signal). A mark on the table showed the rest position for the forefinger. The distance from the rest position to the screen differed in the pilot and the main study but the angle from the rest position to the target were comparable (see figure 1.B for the 'F-group', [6] for 'B-group'). The target name was /pá/ vs. /papá/ vs. /papapá/ for the 'B-group', and /ba/ vs. /baba/ vs. /bababa/ for the 'F-group'. The /ba/ syllable was used with the French language in order to avoid /papa/ which is a frequent word of the language. The 'F-group' performed 14 trials for each of the 6 experimental conditions [target position (2) × *n-syl.* (3)], randomly organized in two blocks. The 'B-group' performed 20 trials per condition randomly organized into four blocks. A same procedure was used for the P task achieved by the 'F-group', despite the fact that no target-name was projected. The task consisted in pointing the target, it included 14 trials for each target position organized in a single block.

The movements of the finger were tracked with an Optotrak. The positions of the IREDs (Infrared-emitting diodes) were sampled at 100 Hz for the 'F-group' and 200 Hz for the 'B-group'. Due to masking issues of the finger IRED observed in the procedure used for the 'B-group' (see [6]), for the 'F-group', the finger IRED was mounted on a piece of metal attached to the tip of the forefinger (Figure 1.A). The data processing was exactly the same as in [6]. The events labeled for the pointing gesture were PV, the velocity peak, PI and PA, respectively the initialization and the apex (at 10% of PV) and PR, the onset of return to the rest position as 10% of the velocity peak of the return motion (see figure 1 C).

3 Results

For the PN task, two-factors ANOVAs were conducted [target position (near vs. far) × *n-syl.* (1- *vs.* 2- *vs.* 3- *syl.*)]. T-tests with Bonferroni corrections were used as post-hoc comparisons between levels of the target name factor. Effects are considered as significant for $p<.05$.

Finger timing in the PN task

Figure 2 displays the mean and the standard errors of the pointing events from the GO-signal.

Effect of the target position. For the 'F-group', the increase in the target distance delays the pointing gesture by about 24 ms for PV ($p<.001$), 31 ms for PA ($p<.001$), and 45 ms for PR ($p<.01$). This effect of the target position does not depend on the target name (1-, 2- or 3-*syl.*) and it does not appear for PI. For the 'B-group', the delay of the finger events when the target is far *vs.* near is about 9 ms for PV, 7 ms for PA and 27 ms for PR. By contrast, PI occurs 20 ms earlier when the target is far rather than near. The global effect of the target position is significant only for PI ($p<.05$) and PR ($p<.01$). Yet, unlike for the 'F-group', it tends to interact with the effect of *n-syl.*

Effect of *n-syl*. For both groups, the global effect of *n-syl.* is significant for the occurrences of PI, PV, PA and PR. For the 'F-group', PI is delayed by about 16 ms from 1- to 2-*syl.* ($p_{bf}<.01$) and 4 ms from 2- to 3-*syl.* (NS). For PV, these delays are, respectively 13 ms ($p_{bf}<.05$) and 8 ms (NS). For PA, the delays are 20 ms ($p_{bf}<.01$) and 9 ms (NS) and for PR, 74 ms ($p_{bf}<.0001$) and 106 ms ($p_{bf}<.0001$). For the 'B-group', PI is delayed by about 26 ms from 1- to 2-*syl.* (p_{bf}<.01) while the difference between 2- and 3-*syl.* is only -2 ms (NS). These differences are +26 ms (p_{bf}<.01) and -1 ms (NS) for PV, 36 ms (p_{bf}<.001) and -3 ms (NS) for PA and 79 ms (p_{bf}<.0001) and 85 ms (p_{bf}<.0001) for PR.

Table I. Means of pointing gesture parameters according to the language, the target position (Near, N. vs. Far, F) and to the number of syllables in the target name (P=P-task), see text for details.

Param.	Targe	Brazilian				French				
		Number of syllables								
		1	2	3	m.	P	1	2	3	m.
P_{dur}	N.	400	403	402	*402*	452	420	426	427	*431*
(ms)	F.	419	433	433	*423*	473	452	454	463	*460*
	m.	*410*	*418*	*418*	*415*	*462*	*436*	*440*	*445*	*446*
P_{amp}	N.	256	258	259	*258*	291	284	283	284	*285*
(mm)	F.	409	411	409	*409*	448	440	438	439	*441*
	m.	*332*	*335*	*334*	*334*	*370*	*362*	*360*	*362*	*363*
P_{ali}	N.	112	163	240	*172*	280	150	206	308	*236*
ms	F.	136	172	270	*193*	315	172	225	316	*257*
	m.	*124*	*168*	*255*	*182*	*298*	*161*	*216*	*312*	*247*
P_{vel}	N.	1083	1099	1097	*1093*	1098	1120	1108	1124	*1117*
mm/s	F.	1690	1642	1653	*1662*	1626	1669	1666	1633	*1656*
	m.	*1387*	*1370*	*1375*	*1378*	*1362*	*1394*	*1387*	*1378*	*1374*

Hence, for both groups, the tendency is to delay PI, PV and PA from 1- to 2-*syl.* while the timing remains quite equivalent from 2- to 3-*syl.* For PR, the delay occurs both from 1- to 2-*syl.* and from 2- to 3-*syl.* The effect of *n-syl.* on the timing of the pointing events does not significantly interact with the effect of target position for the 'F-groups'. For the 'B-group', the lead of PI in the far target condition is larger in the 2-*syl.* condition than in the 1- and 3-*syl.* ones while for PV, PA and PR, the target position has an effect for 1- and 3-*syl.* but not for 2-*syl.*

Parameters of movement in the PN task

Table I summarizes the means and the standard errors of the movement parameters. P_{dur} and P_{amp} are respectively the duration and the amplitude from PI to PA. P_{vel} is the amplitude of peak velocity and P_{ali} is the duration of the finger-target alignment period, from PA to PR. For the 'F-group', the effect of *n-syl.* on P_{amp}, P_{dur} and P_{vel} is not significant. These parameters clearly increase from the near to the far target position: +156 mm for P_{amp} (p<.0001); +32 ms for P_{dur} (p<.001) and +539 mm/s for P_{vel} (p<.0001). By contrast, P_{ali} only depends on *n-syl.* (p<.0001): +54 ms from 1- to 2-*syl.* (p_{bf}<.0001) and +97 ms from 2- to -3 *syl.* (p_{bf}<.0001). Results of the 'B-group' are similar: P_{amp} increases as the distance to the target increases (+152 mm, p<.01). P_{dur} (+26 ms, p<.01) and P_{vel} as well (+569 mm/s, p<.0001), while *n-syl.* has no effect on these parameters. Then, P_{ali} is influenced both by *n-syl* (p<.0001), [+44 ms from 1- to 2-*syl.* (p_{bf}<.0001); +87 ms from 2- to 3-*syl.* (p_{bf}<.0001)] and by the target position [+11 ms from near to far, p<.05].

PN vs. P task

For the 'F-group', the effect of the target position on the pointing gesture is globally similar in the P and the PN task. Then, PI and PV are delayed in the PN task as compared to the P task for 2-*syl.* (PI: +30 ms, PV: +24 ms, p_{bf}<0.001) and 3-*syl.* (+33 ms for PI and PA, p_{bf}<.0001) but not for 1-*syl.* By contrast, time of PR occurrence, P_{dur} and P_{ali} do not significantly differ between the P task and the 3-*syl.* condition of the PN task, while they are greater for the P task than for 1 *syl.* [+155 ms for PR; +28 ms for P_{dur}; +140 ms for P_{ali}, p<.001], and 2-*syl.* [+77 ms for PR; +23 ms for P_{dur}: and +82 ms for P_{ali}, p<.05. No significant difference appears between P and PN for PA, P_{amp} and P_{vel}.

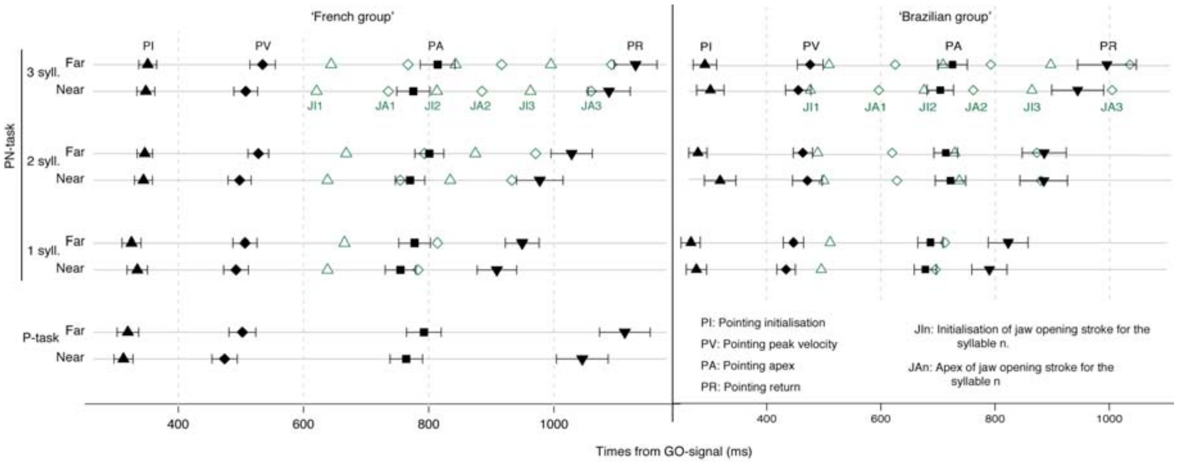

Figure 2. Occurrences of pointing events according to the number of syllables, the target position and the group of participants. The analysis of jaw events are not detailed in the present paper.

4 Discussion and conclusions

This analysis of the pointing gesture first showed that the kinematics of the pointing gesture is quite similar for the 'F-group' and the 'B-group'. The main difference concerned the 2-*syl.* condition. A possible explanation is that, before achieving the task reported here, the 'B-group' participated in our study on stress position in which the target name could be /pápa/ vs. /papá/ [6]. Achieving this task may have affected the present results. Analyses of between-subjects' variability should help to understand this difference.

A strong characteristic of the pointing gesture, also observed in [6-7], is the tendency towards isochrony. Despite a strong increase in the amplitude of the gesture with the increase of the target distance, the parallel increase in velocity results in quite comparable durations, even though the gesture is longer for the far *vs.* near target.

Then, the comparison of the P and PN task for the 'F-group' does not perfectly replicate previous results showing no difference in pointing duration between a P task and a Pointing+Speech task [3]. The present results show that pointing duration tends to be shorter when the gesture comes with a 1- or 2-*syl.* utterance rather than when it is realized alone. This difference disappears when *n-syl*. increases to 3-*syl.* Moreover, in the PN task, the timing of pointing is slightly delayed with the increase of *n-syl*. Yet, a first analysis of jaw timing (figure 2) seems to agree with previous results [3-6] showing that jaw/pointing coordination within the finger movement towards the target mainly results from an adaptation of speech onset: the jaw anticipates more and more with *n-syl.* increases.

The tendency to delay pointing events with the increase of *n-syl*. could result from an interaction between the two systems [3, 4]. However, this hypothesis does not explain the decrease in pointing duration from the P task to the PN task. An alternative hypothesis could be that the speech utterance 'calibrates' the pointing gesture. Even if the global effect of *n-syl*. on pointing duration is not significant in the PN task, differences between the P and PN tasks also seem to be linked to *n-syl*.

But the main effect of *n-syl*. on the pointing gesture lies in the 'off-line' adaptation of the duration of finger-target alignment that increases with *n-syl*. This result agrees with our previous results showing a strong adaptation of the duration of the finger-target alignment to the stress position in 2-*syl.* utterances [6-7]. The interesting point here is that, for the P task, the finger-target alignment lasts longer than the PN task. It therefore seems that here, again, speech could 'calibrates' the finger pointing action.

These data are quite coherent with our previous results. They also suggest that speech/pointing co-occurrence should be envisioned as the product of two collaborating systems working to achieve a coherence in multimodal pointing rather than an autonomous manual system driving speech in an asymmetric way as proposed in [3]. The detailed analysis of speech events and their relationships with pointing is the imminent next step. It should contribute to better understanding the present results.

Acknowledgement

The first author thanks the Fyssen Foundation for its financial support.

References

[1] S. Kita (Ed.), *Pointing: Where language culture and cognition meet*. Lawrence Erlbaum Associates, 2003.

[2] C. Abry, A. Vilain, J.-L. Schwartz (eds), Vocalize to Localize II: Special issues of *Interaction Studies,* 6(2), 2005.

[3] J.B. Haviland, Pointing, gesture spaces, and mental maps. In D. McNeill (Ed.), *Language and gesture*, pp. 13–46, 2000.

[4] W. J. M Levelt, G. Richardson, and W. L. Heij, Pointing and voicing in deictic expressions. *Journal of Memory and Language*, 24: 133–164, 1985.

[5] P. Feyereisen. The competition between gesture and speech production in dual-task paradigms. *Journal of Memory and Language*, 36(1): 13–33, 1997.

[6] A. Rochet-Capellan, J-L. Schwartz, R. Laboissière and A. Galván. Pointing a target while naming it with a CVCV sequence: Speech focus position effect on finger and jaw coordination, *Journal of Speech Language and Hearing Research*, 2008, in press.

[7] A. Rochet-Capellan, J.-L. Schwartz, R. Laboissière and A. Galván. Pointing to a target while naming it with labial-coronal or coronal-labial CVCV: the effect of consonants order and stress position on jaw-finger coordination. *In Proceedings of InterSpeech* 2007.