The Speech Focus Position Effect on Jaw–Finger Coordination in a Pointing Task

Amélie Rochet-Capellan
GIPSA Laboratory, Department of Speech and Cognition, Grenoble, France; Centre de National de la Recherche Scientifique (CNRS), Paris, France; and Grenoble University, Grenoble, France

Rafael Laboissière
U864 Espace et Action, INSERM, University Claude Bernard Lyon 1, Bron, France, and Max Planck Institute for Human Cognitive and Brain Sciences, Munich, Germany

Arturo Galván
Max Planck Institute for Human Cognitive and Brain Sciences

Jean-Luc Schwartz
GIPSA Laboratory, Department of Speech and Cognition, Grenoble, France; CNRS; and Grenoble University

Purpose: This article investigates jaw–finger coordination in a task involving pointing to a target while naming it with a 'CVCV (e.g., /'papa/) versus CV'CV (e.g., /pa'pa/) word. According to the authors' working hypothesis, the pointing apex (gesture extremum) would be synchronized with the apex of the jaw-opening gesture corresponding to the stressed syllable.
Method: Jaw and finger motions were recorded using Optotrak (Northern Digital, Waterloo, Ontario, Canada). The effects of stress position on jaw–finger coordination were tested across different target positions (near vs. far) and different consonants in the target word (/t/ vs. /p/). Twenty native Portuguese Brazilian speakers participated in the experiment (all conditions).
Results: Jaw response starts earlier, and finger–target alignment period is longer for CV'CV words than for 'CVCV ones. The apex of the jaw-opening gesture for the stressed syllable appears synchronized with the onset of the finger–target alignment period (corresponding to the pointing apex) for 'CVCV words and with the offset of that period for CV'CV words.
Conclusions: For both stress conditions, the stressed syllable occurs within the finger–target alignment period because of tight finger–jaw coordination. This result is interpreted as evidence for an anchoring of the speech deictic site (part of speech that shows) in the pointing gesture.

KEY WORDS: deixis, pointing, speech–hand coordination, lexical stress

Hand and mouth often work together in human behaviors, mainly in alimentation and communication. This link has motivated a large body of research. For example, Iverson and Thelen (1999) showed that spontaneous co-occurrence of hand and mouth movements appears right after birth. Then, at around 6–8 months, hand and mouth start to mutually entrain each other in rhythmic activities characterized by manual and oral babbling. Gestures and speech are then produced sequentially at around 9–14 months and eventually are synchronized at the age of 16–18 months. Interplay of hand and mouth motor control is also observed in adults' behavior. For example, when speakers open their mouth while grasping an object such as a piece of fruit, the apertures of both the grasp and the mouth are adapted to the size of the object (Gentilucci, Benuzzi, Gangitano, & Grimaldi, 2001). The observation of an action realized by one part of the body (e.g., bringing a fruit to the mouth) also affects the production of an action realized by the other (e.g., uttering a syllable; see Gentilucci, 2003; Gentilucci, Santunione, Roy, & Stefanini, 2004). Hand and mouth are also coupled in adults' rhythmic activities. For example, Kelso, Tuller, and Harris (1981) found a 1:1 ratio between the frequency of the repetition of the word "stack" and the simultaneous repetition of a flexion–extension motion of the index finger. In addition, the co-occurrence of hand and mouth movements is clearly observable in face-to-face communication. The origin of this co-ocurrence seems to be motor rather than purely perceptual, as gestures are produced even in situations in which the interlocutor cannot perceive them, such as in phone calls (Iverson & Goldin-Meadow, 1998).

According to McNeill (2000), a variety of gestures can occur in communication, ranging from *gesticulations,* which are global, nonconventionalized, and speech dependent, to *signs* in signed languages, which are segmented, analytic, conventionalized, and performed without speech. This article focuses on a particular type of gesture that can accompany speech in communication, namely *pointing gestures*. The global aim here is to link *deixis,* the component of language that allows referring to objects, with the capacity of synchronizing gesture and voice to show objects. This synchronization may depend on the properties of the motor coordination between hand and mouth, arising from prelinguistic links between the two motor systems.

Pointing Gestures and Language

Our interest in pointing gestures—and, more particularly, in their coordination with speech—originates mainly from five observations reported in the literature. The first observation is that pointing gestures are the principal medium of shared attention, a basic function required for language acquisition (Tomasello, Carpenter, & Liszkowski, 2007). The second observation is that pointing gestures appear to be universal (Butterworth, 2003), despite variability in the form of the gesture across cultures (Haviland, 2000; Wilkins, 2003). The third observation is that pointing gestures are the first and the dominant communicative actions in infant communication. At 12 months, pointing gestures constitute 60% of infants' manual communicative gestures and are often accompanied by vocalizations (Butterworth, 2003). The fourth observation is that pointing gestures are at the cutting edge of language development. Goldin-Meadow and Butcher (2003) showed that the age at which children associate a pointing gesture with a word having complementary meanings is related to the age of two-word productions (for similar conclusions, see also Pizzuto, Capobianco, & Devescovi, 2005; Volterra, Caselli, Capirci, & Pizzuto, 2005). The fifth and last observation is that pointing gestures have been put forward as the canonical form of language demonstrative words (Diessel, 1999; Haviland, 2000). Drawing evidence from developmental and comparative psychology, Diessel (2006) argues that demonstrative words such as "this" or "that" serve the basic communicative function of joint (or shared) attention rather than a specific grammatical function. He provides evidence for considering demonstrative words as particular linguistic objects, defending their universal character and especially their specific and close link with pointing gestures.

Altogether, this body of research on the relationships between pointing gestures and language in general, and between pointing gestures and deixis in particular, led Abry, Vilain, and Schwartz (2004) to consider the connection between hand and voice in deixis as a crucial step in language emergence. They proposed to derive speech and language from the necessity to localize the objects we need to talk about, which requires the hand and mouth coordination. Hence, the understanding of speech–showing and hand–pointing coordination could be considered a key step to understanding the emergence of language deixis.

In this framework, this article investigates the effect of the position of the emphasized part of speech—namely, *speech focus* (the part of speech that shows)—on jaw–finger coordination in a task involving pointing at a target with the hand/finger while naming it.

Processes Underlying Speech–Pointing Synchronization

At least since McNeill's work (1981), it is well known that speech and hand gestures are coordinated in online face-to-face interactions. This phenomenon has motivated studies about the processes involved in speech–hand coordination around, among others, the question of the interaction versus modularity of the two systems. Most often, these studies used a dual-task paradigm: The participant provided both a verbal and a gestural response to a stimulus. The hand dynamics in this dual task are compared to hand dynamics in a gesture-only task, and the speech dynamics in this dual task are compared to the speech dynamics in a speech-only task. For example, in Holender (1980), the task was to name a letter that appeared on a screen and press a key, whereas in Castiello, Paulignan, and Jeannerod (1991), it was to pronounce "tah" in response to a visual stimulus that indicated an object to grasp. More in line with our concerns, Levelt, Richardson, and Heij (1985)—and, later, Feyereisen (1997)—used the dual-task paradigm in order to study pointing gestures. The dual task was to point at an object with the hand while verbally designating it using a "that object" or "this object" utterance (e.g., "this lamp"). According to Levelt et al. (1985), pointing gestures present a double interest for the study of speech and hand synchronization: First, they are strictly dependent on the message being expressed, and, second, the moment at which they reach their target (now referred to as the *pointing apex*) can be easily detected.

Among others, Levelt et al.'s (1985) results showed that for utterances such as "this lamp," the voice onset tends to be synchronized with the pointing apex. Hence, putting the target further from the participant delays both the pointing apex and the voice onset. The voice onset also occurs later in the dual task (when it is accompanied by the pointing gesture) than in the speech-only task. Alternatively, the timing of the pointing apex is essentially the same in both gesture-only and dual tasks. The authors interpret these results as evidence for an adaptation of speech commands to brachiomanual commands rather than the reverse. A delayed verbal response in a dual task as regards a speech-only task is also put forward in Castiello et al. (1991), Feyereisen (1997), and Holender (1980).

However, all of these studies measured the verbal response delay using the acoustic signal only, without considering the speech articulators. As discussed in the next section, the processes of speech–hand coordination might be better described and understood through the dynamic interplay between the orofacial articulation and the hand/finger systems.

Jaw–Hand Coordination Rather Than Voice–Hand Coordination

The motivation for investigating the articulatory motions in speech–hand coordination stems from two kinds of arguments. First, at a methodological level, speech is also a gestural system much like pointing. Following Stetson (1951), a great number of studies have focused on the articulators' motions, characterizing speech as the outcome of a motor system. As suspected by Castiello et al. (1991) and Holender (1980), some motor events might happen before the voice onset. Hence, it is legitimate to investigate when articulators start to move relative to the pointing gesture. In addition, at a theoretical level, speech–pointing coordination has been assumed to emerge in the course of ontogeny from a developmental meeting between the jaw and arm/hand motor control (Ducey-Kaufmann, 2007). According to MacNeilage and Davis's (2000) frame-then-content scenario of speech development, speech motor control begins in young babies with the mastering of the opening/closing oscillations of the jaw, which provides the *speech frame* (MacNeilage & Davis, 2000). The independent and coordinated control of the tongue and the lips (the *content*) would be mastered later. In this frame-then-content sequence experimentally observed in the course of ontogeny (Green, Moore, & Reilly, 2002; Munhall & Jones, 1998), the jaw is considered the carrier of speech gestures. Yet, MacNeilage and Davis did not consider the role of manual gestures in speech acquisition. Different studies put forward a link between the motor control development of brachiomanual and orofacial gestures. Supporting evidence for this link comes from the relationship between the frequencies of hand and jaw oscillations in babbling (Ducey-Kaufmann, 2007; Iverson & Thelen, 1999; Petitto, Holowka, Sergio, & Ostry, 2001)—what Ducey-Kaufmann (2007) referred to as the *sign frame* and the *speech frame,* respectively. According to them, the relationship between the frequencies of the two systems would evolve toward a developmental meeting point between the speech frame and the sign frame. This meeting point is suggested to be the basis of speech–hand coordination and the background for the production of the first words. These two sets of methodological and theoretical arguments lead us to propose a jaw–hand rather than a voice–hand investigation framework for studying speech and manual pointing coordination.

An Attraction Between the Speech Focus and the Hand Focus

The question of interest in the present study concerns the candidate sites for the speech and pointing gesture coordination: Which part of the hand gesture is synchronized with which part of the speech utterance? According to McNeill (1992), in speech–hand coordination, the hand gesture stroke is executed in synchrony with the semantically co-expressive word. Moreover, verbal deixis can be prosodic as well as grammatical (e.g., Lœvenbruck, Baciu, Segebarth, & Abry, 2005). When considering the communicative aim of speech–hand association in deixis, it seems reasonable to assume that the part of the discourse that shows should occur synchronously with the part of the gesture that shows. Thus, synchronization of speech and hand pointing in face-to-face communication could result in an attraction between the *speech focus* (the indexical word and/or the stressed part of the utterance) and the *pointing focus* (the moment at which the arm/hand/finger system is aligned with the target). This hypothesis is compatible with Levelt et al.'s (1985) results, which show a tendency toward synchrony between voice onset corresponding to the demonstrative word ("this" or "that") and the hand–pointing apex. Nevertheless, Levelt et al. did not vary the position of the speech deictic site, which was systematically at the beginning of the utterance (e.g., "this lamp" vs. "that lamp"). In this article, we propose to vary the position of speech focus in a simple way by varying the stressed syllable in CVCV utterances. Our aim is to study how this variation influences the jaw–hand coordination in a task consisting of pointing to a target while naming it with a CVCV word. Our main hypothesis is that the hand–pointing apex should be synchronized with the extremum (or apex) of the jaw-opening gesture corresponding to the stressed syllable, either the first syllable in ˈCVCV utterances (e.g., /ˈpapa/) or the second syllable in CVˈCV utterances (e.g., /pa'pa/). This alignment could be reached either through adaptation of the jaw movement to a constant hand movement in both 'CVCV and CV'CV sequences or through a mutual adaptation involving a modification of both jaw and hand motions across word stress conditions.

Method

Participants and Language

Brazilian Portuguese was chosen because it is one of the languages in which it is possible to find pairs of words that differ only by stress position (e.g., 'CVCV vs. CV'CV). The participants were 20 native Brazilian Portuguese speakers (4 men, 16 women) aged 18–37 years (M = 28.3, SD = 5.3). They were paid 8 euros per hour for their participation. The participants were all right-handed, had no reported history of speech or hearing pathology, and were unaware of the purpose of the experiment.

Experimental Design

The experiment involved a hand-pointing task associated to the utterance of a CVCV disyllable. The main factor was the stress position in the CVCV dissyllable: stress on the first versus the second syllable (e.g., /'papa/ vs. /pa'pa/). The consonant was either /p/ or /t/. The vowel /a/ was selected because it requires a large jaw-opening gesture. Moreover, two spatial targets were used for the pointing gesture (near vs. far). The variation of both the consonant and the target position contributed to focus participants' attention on the task. Hence, the experimental design consisted of three within-subjects two-level crossed factors: stress position (first vs. second), consonant (/t/ vs. /p/), and target position (near vs. far).

Procedure

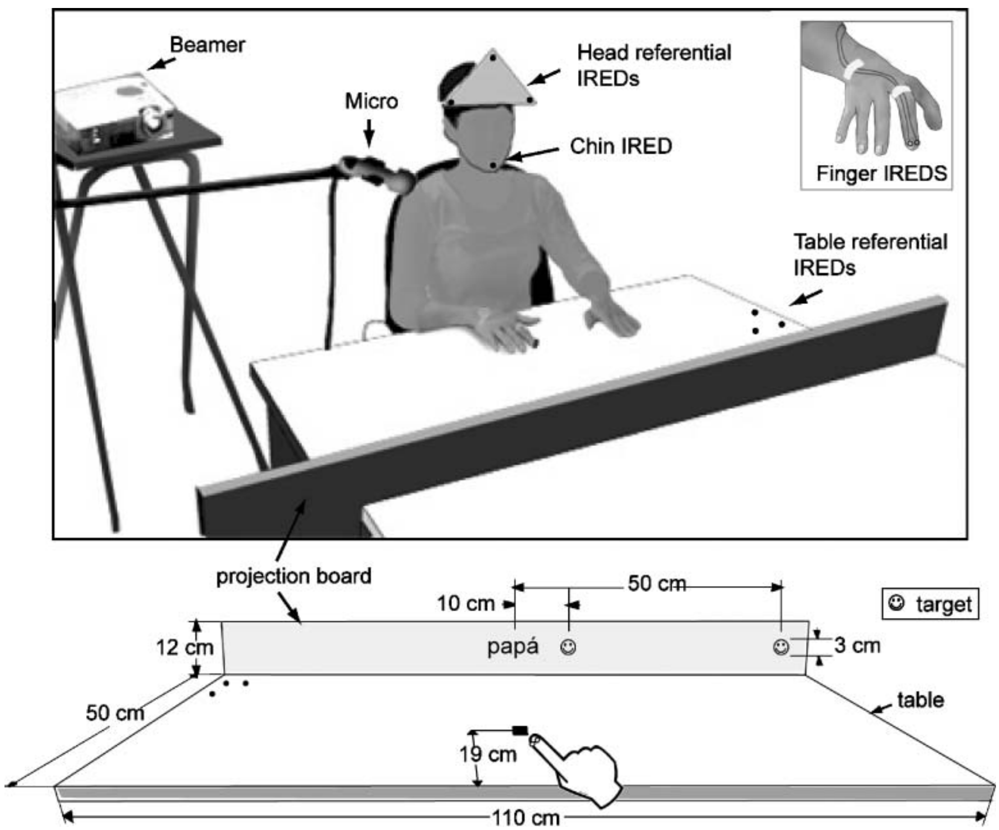
The participants were seated at a table. The targets to point at and the item to pronounce were projected simultaneously on a white screen in front of them using a projector (see Figure 1, top). A black square pasted on the midline of the table, close to the participant's sagittal plane, indicated the finger resting position. The participants were informed that a word and a red smiley sign (the target) would appear on the screen. The target appeared to the participant's right (see Figure 1, bottom) either near (10 cm from midline) or far (50 cm from midline). In order to make the joint gesture/pronunciation task more natural, participants were instructed to use the word displayed as the name of the person represented by the smiley target. Participants were instructed to simultaneously point with the index finger at and name the target as soon as the color of the smiley sign changed from red to green. Prior to the experiment, participants were briefly trained to become familiar with the task: They were asked to simultaneously point at and name objects in the room. They also practiced reading CVCV sequences aloud in order to make sure that they understood the stress instruction properly. The experiment was divided into four blocks. One block contained 4 practice trials followed by 40 experimental trials, 5 for each combination of stress position, consonant, and target position. The order of the trials was randomized for each block and each participant. Blocks were separated by 30-s rest periods. To reduce anticipatory responses to the go signal (smiley target becoming green), the red smiley duration was varied from trial to trial (M = 2.5 s, SD = 0.15 s, normally distributed). The green smiley target lasted on the screen for 1 s in each trial.

Data Recording and Postprocessing

Finger and jaw movements were recorded using Optotrak (Northern Digital, Waterloo, Ontario, Canada), an optoelectronic position measurement system that tracks the three-dimensional motion of infrared-emitting diodes (IREDs). The positions were sampled at 100 Hz. IRED locations are illustrated in Figure 1 (top). Two IREDs were pasted onto the tip of each participant's right forefinger: one on the middle of the nail and the other on the medial side next to the nail. In so doing, at least one of the IREDs was visible by the cameras during the pointing movement, even when participants supinated their hands at the motion apex. A third IRED was attached to the participant's chin. It tracked a flesh point rather than the jaw itself. However, considering the phonetic material in question (stop consonants associated with an open vowel), the motion of this flesh point is a relevant indicator of jaw motion. Head motion was measured by three IREDs attached to a plastic triangle, which was fixed by a strap around the participant's head. The coordinates of the moving IREDs were projected into a fixed referential, defined by three IREDs pasted on the table. Jaw position was then computed in the head-moving reference frame. Principal component analysis (PCA) was applied separately to each of the three 3D trajectories of the two finger IREDs and the jaw IRED. The first principal component explained most of the variance for each IRED and for all participants: 98.8% (SE = 0.2%) and 98.3% (SE = 0.3%) for the two finger IREDs and 95.6% (SE = 0.5%) for the jaw IRED. This component was chosen to represent finger and jaw movements. Signals were low-pass filtered at 15 Hz with a non-phase-distorting Butterworth filter. The sound was simultaneously recorded and sampled at 16 kHz.

The recorded utterances were checked against the correct phonetic and stress instructions. Trials with speech production errors were excluded from the dataset (on

Figure 1. Experimental setup (top) and projection dimensions (bottom). The participants pointed at and named a smiley sign projected on a screen in front of them. Finger and jaw movements were captured using Optotrak (in front of the participant) with two infrared-emitting diodes (IREDs) for the finger, one for the jaw (chin IRED), three for the head (IRED head referential), and three on the table (referential for all moving IREDs). The target name appeared in the midline of the visual field (e.g., /pa'pa/), and the pointing target appeared on the right side, either at a near (10-cm) or a far (50-cm) position.



average, 4.1 trials per participant, with a maximum of 11 errors for 1 participant; see Table 1, first row). Correct trials for which jaw or finger gestures were initiated before the go signal were also discarded. This mainly concerned 1 participant, who made 39 initiation errors (see Table 1, second row). For the remaining data, trials for which the two finger IREDs were partially hidden from the Optotrak cameras were not considered for the analysis (8.2 trials, on average; see Table 1, third row). By default, the middle IRED was chosen for the analysis. When this IRED was masked and the left one was visible, the left IRED was taken for the analysis (the mean correlation between the two finger IREDs is above .99 for each of the three coordinates x, y, and z). Similarly, trials for which the jaw IRED was partially hidden were discarded from the analysis, which mainly concerned 1 participant for 12 trials (see Table 1, fourth row). A three within-subject factors analysis of variance (ANOVA) shows that stress position, consonant, and target position do not significantly affect either the number of utterance and initiation errors or the number of trials with hidden finger or jaw IREDs.

Table 1. Number of rejected trials over all 20 participants and all experimental conditions.

Problem	*M*	*SE*	Min	Max
Utterance error[a]	4.1	0.86	0	11
Initiation error[b]	2.35	1.94	0	39
Masked finger[c]	8.2	1.50	0	21
Masked jaw[d]	0.75	0.60	0	12

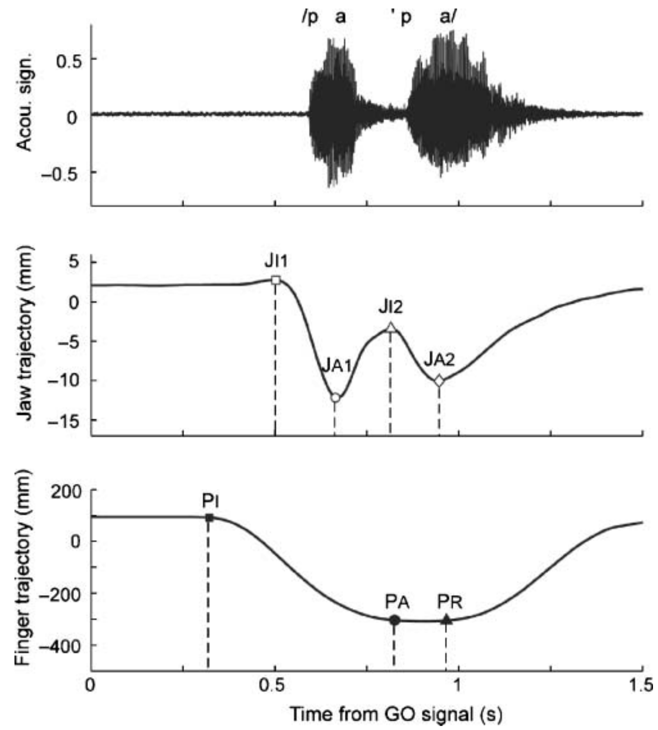
Note. *SE* = standard error; Min = minimum; Max = maximum.

[a]incorrect pronunciation. [b]response before the go signal. [c]Two finger infrared-emitting diodes (IREDs) partially masked. [d]Jaw IRED partially masked.

Labeling and Measurements

Figure 2 displays an example of signals and labeling for a /pa'pa/ trial. Onset and offset events of the finger and jaw movements correspond to 10% of peak velocity.

Figure 2. Example of signals and labeling for a /pa'pa/ trial. Top panel: Acoustic signal (Acou. sign.); middle panel: jaw trajectory; bottom panel: finger trajectory. *Onsets* and *offsets* of strokes are defined as the instant when velocity reaches 10% of the maximum value during the corresponding stroke. J_{I1}, J_{I2}, and J_{A1}, J_{A2}, respectively, correspond to the onset and apex (offset) times of the first and second jaw-opening gestures. P_I and P_A correspond to the onset and apex (offset) times of the pointing forward stroke. P_R is the onset time of the finger return movement. Note that in this example, even though stress position affects the second syllable, the amplitude of the jaw movement is smaller for the second syllable than for the first one. This corresponds to the general pattern observed in this experiment (see jaw-opening motion panels in Figure 5).



All times correspond to the elapsed times from the go signal to the event.

For the jaw (see Figure 2, middle panel), the analysis was focused on the two opening strokes for the /a/ vowels. J_{I1} and J_{A1} are the initiation and apex times for the first opening stroke (stroke onset and offset, respectively). Similarly, J_{I2} and J_{A2} are the initiation and apex times for the second opening stroke. For the finger, the trajectory can be split into three parts: the forward stroke, the pointing plateau, and the return stroke (see Figure 2, bottom panel). The *forward stroke* corresponds to the pointing gesture toward the target. It starts at P_I (initiation time of pointing gesture) and ends at P_A (apex time of pointing gesture). The *pointing plateau* is the amount of time during which the finger remains pointed at the target. It starts at P_A and ends at P_R (onset time of the return stroke). The *return stroke* corresponds to the movement of the finger back to its rest position. This stroke was not considered here because it is not really part of the pointing task. Note that Levelt et al. (1985) only investigated the pointing forward stroke. Yet the observation of a plateau in the apex position shows that the finger-pointing task does not end with the alignment of the finger and the target but rather with the onset of the return stroke. In our study, for the finger motion analysis we considered both the forward stroke and the plateau.

Hypotheses

Under the assumption that there is a tight temporal coordination of jaw- and finger-pointing gestures, the apex of the jaw-opening gesture corresponding to the stressed syllable should be synchronized with the pointing apex. This might induce a significant effect of the stress position on the delay between the apex of the first jaw-opening gesture, J_{A1}, and the apex of the pointing gesture, P_A, that should be shorter in the first-syllable stress condition than in the second-syllable stress condition. The increase in the delay between J_{A1} and P_A in the second-syllable stress condition would result in the apex of the second jaw-opening gesture, J_{A2}, being closer to P_A. This should occur regardless of the target position and of the utterance consonant. Hence, the effect of stress position on the timing of the pointing apex relative to the two jaw-opening gestures' apices should be similar for the two target positions and the two consonants. Finally, the effect of stress position on jaw–finger coordination could result from either an adaptation of jaw movement to hand movement (or the converse) or a mutual adaptation between the two motor systems. This can be evaluated through the analysis of the effect of stress position on timing, duration, and amplitude of finger and jaw gestures.

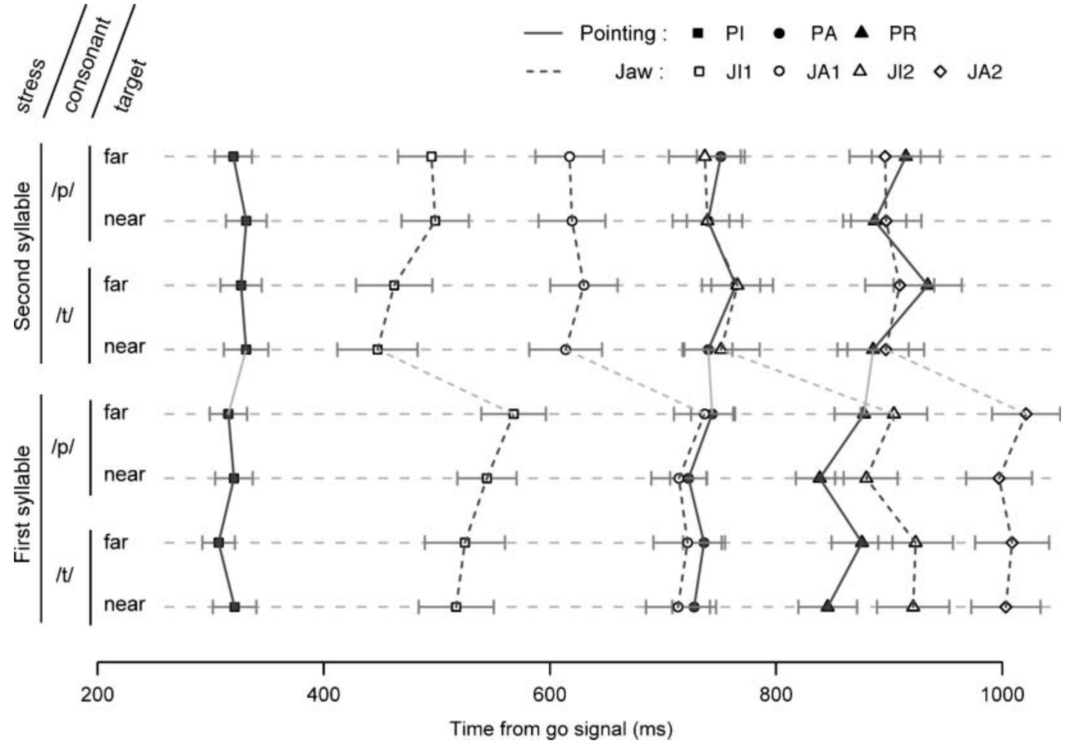
Results

Figure 3 shows the mean temporal position of jaw and finger events, computed for the 20 participants, depending on stress position, consonant, and target position. The effects of the experimental variations on each measured variable were tested using three-way (stress position, consonant, and target position) within-subject ANOVAs. The *p* value for the significance level was fixed at .05.

Absolute and Relative Timing of Finger and Jaw Motions

Initiation times. The pointing motion starts, on average, 322 ms after the go signal (see P_I values in Figure 3). There was no significant main effect of any of the three factors on P_I. In contrast, the initiation of the first jaw stroke (J_{I1}) occurs significantly earlier in the second-syllable stress condition (476 ms) than in the first-syllable stress condition (539 ms), $F(1, 19) = 15.9$, $p < .001$, and for /t/ (488 ms) as compared with /p/ (527 ms), $F(1, 19) = 13.5$, $p < .01$. Neither the Stress Point × Consonant interaction nor the Stress Position × Target Position interaction was significant. Hence, the onset of the jaw movement (J_{I1}) depended on stress position and occurred later than the onset of the finger movement (P_I). Analysis of $J_{I1} - P_I$ delay shows that J_{I1} was closer to P_I in the second-syllable stress condition (149 ms) than in the first-syllable stress condition (223 ms), $F(1, 19) = 14.1$, $p < .01$, and for /t/ (166 ms) as compared with /p/ (206 ms), $F(1, 19) = 13.3$, $p < .01$. Furthermore, $J_{I1} - P_I$ was significantly shorter

Figure 3. Means of the elapsed time from the go signal to the pointing and jaw events relative to stress position (first vs. second syllable), consonant (/t/ vs. /p/), and target position (near vs. far). P_I, P_A, and P_R, respectively, correspond to the onset, the apex, and the onset of the return of the finger-pointing movement. J_{I1}, J_{A1}, and J_{I2}, $J_{A2,}$ respectively, correspond to the onset and the apex of the first and the second jaw-opening strokes.



for the near (176 ms) than for the far (195 ms) target position, $F(1, 19) = 10.9$, $p < .01$. Interaction effects on $J_{I1} - P_I$ were not significant. Hence, the jaw started to move after the finger but significantly closer to the beginning of the finger movement when stress was on the second rather than on the first syllable. Moreover, the effect of stress position on the delay between P_I and J_{I1} resulted from an effect on J_{I1}. The stress position had no significant effect on P_I.

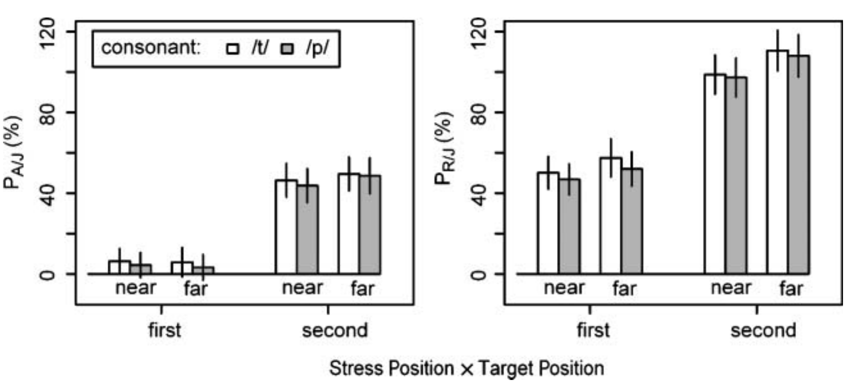
Apex times. The pointing apex (P_A) occurred 16 ms later in the far (749 ms) than in the near (733 ms) target condition. The effect of target position was significant, $F(1, 19) = 9.2$, $p < .01$. Stress position and consonant type did not have any significant effect on P_A. By contrast, the stress position significantly influenced the apex times of the jaw-opening gestures. The apex of the first jaw-opening gesture (J_{A1}) occurred 101 ms earlier in the second-syllable stress condition (620 ms) than in the first-syllable stress condition (721 ms), $F(1, 19) = 54.0$, $p < .0001$. A similar delay across stress conditions was observed for the apex of the second jaw-opening gesture (J_{A2}), which occurred, on average, 108 ms earlier in the second-syllable stress condition (900 ms) than in the first-syllable stress condition (1,008 ms), $F(1, 19) = 54.8$, $p < .0001$. Target position effect was also significant for both J_{A1}, $F(1, 19) = 7.1$, $p < .05$, and J_{A2}, $F(1, 19) = 5.9$, $p < .05$. The apices of both jaw-opening gestures occurred earlier in the near (665 ms for J_{A1} and 949 ms for J_{A2}) than in the far (677 ms for J_{A1} and 959 ms for J_{A2}) target condition. The increase of the target distance thus delayed the apices of the two jaw-opening gestures by about 10 ms. This delay was equivalent to the nonsignificant delay observed for the onset of the jaw movement, J_{I1}, and seemed to partly compensate for the 16-ms delay observed for the pointing apex, P_A. Finally, the consonant had no significant effect on the apex times of the two jaw-opening gestures.

In addition, the delay observed for J_{A1} in the first-syllable stress condition compared with the second-syllable condition resulted in J_{A1} being closer to P_A in the first case. The study of $J_{A1} - P_A$ showed that P_A occurred about 11 ms after J_{A1} in the first-syllable stress condition and about 129 ms after J_{A1} in the second-syllable stress condition. This effect of stress position was significant, $F(1, 19) = 52.3$, $p < .0001$. Alternatively, neither the target position nor the consonant had a significant effect on $J_{A1} - P_A$. To better characterize the temporal relationships between finger and jaw apices, we computed the position of pointing apex, P_A, relatively to the apices of the two jaw-opening gestures, J_{A1} and J_{A2}:

$$P_{A/J} = \frac{P_A - J_{A1}}{J_{A2} - J_{A1}}, \text{in percent.} \qquad (1)$$

A value of 0% for $P_{A/J}$ corresponded to a case for which P_A occurs at the same time as J_{A1}. A value of 100% corresponded to a case for which P_A is synchronized with J_{A2}. Figure 4 (left) shows $P_{A/J}$ means and standard errors against stress conditions. It appeared that $P_{A/J}$ mean was greater when the second syllable was stressed (47%) rather than the first one (5%), $F(1, 19) = 66.0$, $p < .0001$. The effects of target position and consonant type as well as the interactions were not significant. In line with the jaw–finger synchronization hypothesis, stress position influenced the relative times of the apices of the jaw-opening gestures and of the pointing gesture. Although the pointing apex was very close in time to the apex of the first jaw-opening gesture when the first syllable was stressed, it occurred at about an equal distance in time from the apices of the two jaw-opening gestures when the second syllable was stressed. Strikingly, in this last condition, the pointing apex seemed closely synchronized with the initiation of the second opening gesture of the jaw, J_{I2}. On average, P_A occurred just 1 ms after J_{I2} when the second syllable was stressed and 175 ms before J_{I2}

Figure 4. Means and standard errors of the time of the pointing apex ($P_{A/J}$, left) and of the time of the onset of the pointing return ($P_{R/J}$, right) relative to the apices of the two jaw-opening gestures, depending on stress position (first vs. second syllable), target position (near vs. far), and consonant (/t/ vs. /p/).



when the first syllable was stressed (see Figure 3). Moreover, in the second-syllable stress condition, the apex of the second jaw-opening gesture was very close to P_R, the onset of the pointing gesture return stroke (that is, the offset of the pointing plateau).

Offset of the pointing plateau (P_R) relative to the jaw apices. P_R occurred later when the second syllable was stressed (906 ms) rather than the first one (860 ms), $F(1, 19) = 16.0, p < .001$, as well as when the target was far (901 ms) as compared to when it was near (864 ms), $F(1, 19) = 28.0, p < .0001$. The consonant factor, as well as all of the interactions, was not significant. The pointing plateau duration ($P_R - P_A$) was significantly greater when the second syllable was stressed (157 ms) rather than the first one (127 ms), $F(1, 19) = 14.1, p < .01$. It was also significantly greater when the target was far (152 ms) rather than near (132 ms), $F(1, 19) = 13.9, p < .01$. As for P_R, neither the consonant factor nor the interactions were significant. Analysis of the $J_{A2} - P_R$ interval showed that P_R occurred 5 ms after J_{A2} in the second-syllable stress condition and 148 ms before J_{A2} in the first-syllable stress condition. This effect of stress position on $J_{A2} - P_R$ was significant, $F(1, 19) = 89.3, p < .0001$. The delay between J_{A2} and P_R was also greater when the target was near (85 ms) rather than far (58 ms), $F(1, 19) = 16.9, p < .001$, but did not depend on the consonant. Interactions on $J_{A2} - P_R$ were not significant. Similar to what we did for the pointing apex, we computed the position of P_R relative to the apices of the two jaw-opening gestures:

$$P_{R/J} = \frac{P_R - J_{A1}}{J_{A2} - J_{A1}}, \text{in percent.} \qquad (2)$$

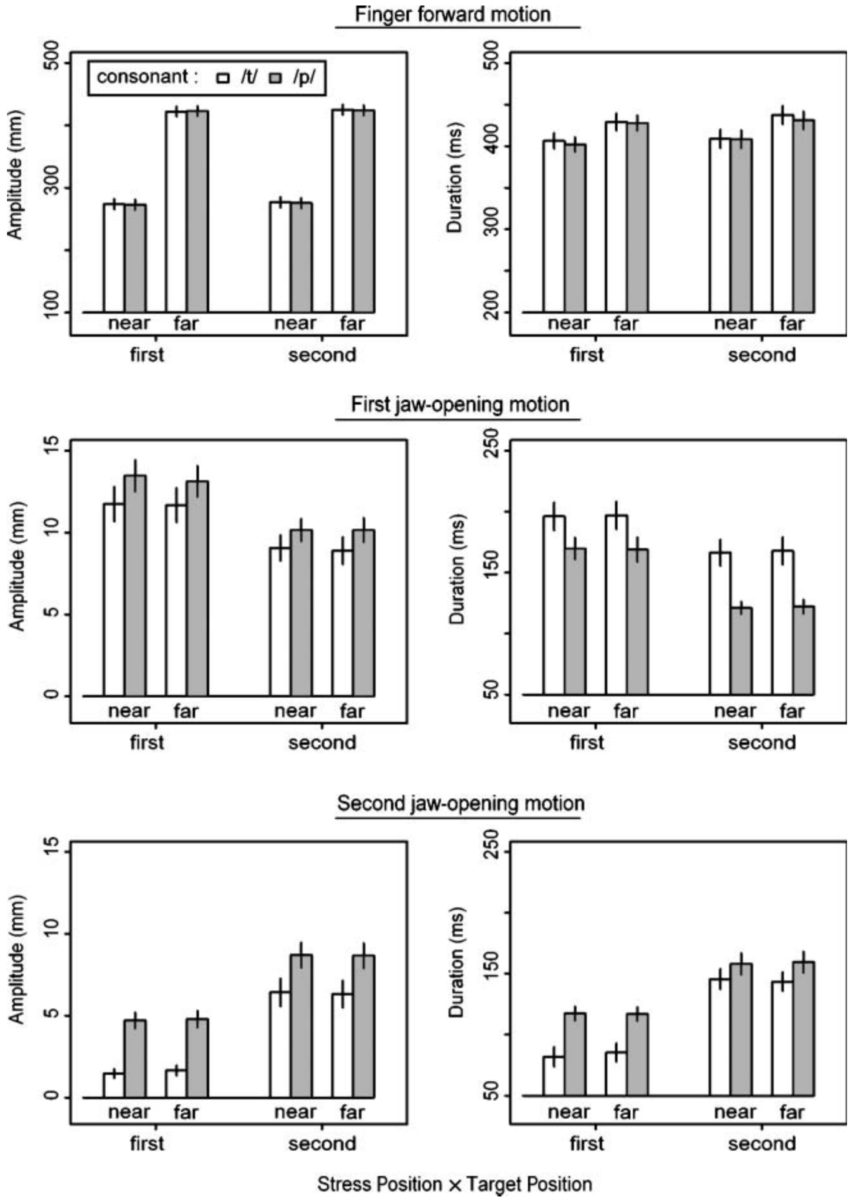
Figure 4 (right) shows that $P_{R/J}$ was close to 100% for the second-syllable stress condition (104%), whereas it was close to 50% for the first-syllable stress condition (51%). $P_{R/J}$ was also greater when the target was far (82%) compared to when it was near (73%) and for /t/ (79%) as compared to /p/ (76%). The three main effects were significant—stress position, $F(1, 19) = 102.9, p < .0001$; target position, $F(1, 19) = 12.4, p < .01$; consonant, $F(1, 19) = 4.9, p < .05$—but none of the interactions were. When the first syllable was stressed, the finger left its apex position before the apex of the second jaw-opening gesture, at about an equal distance in time from the apices of the two jaw-opening gestures. Actually, as mentioned in the previous section, in the first-syllable stress condition, the finger even left its apex position 175 ms before the apex of the second jaw-opening stroke. By contrast, in the second-syllable stress condition, the finger left its apex position just after the apex of the second jaw-opening gesture was reached.

In summary, the pointing apex, P_A, was close to the apex of the first jaw-opening gesture, J_{A1}, when the first syllable was stressed, whereas it occurred at about the midpoint between the apices of the two jaw-opening gestures, J_{A1} and J_{A2}, when the second syllable was stressed. In the second-syllable stress condition, the jaw events (J_{I1}, J_{A1}, J_{I2}, and J_{A2}) occurred earlier than in the first-syllable stress condition, and the onset of the pointing gesture return (P_R) was delayed. These two facts resulted in J_{A2} being close to P_R. By contrast, the timing of the pointing-forward stroke did not depend on stress position. The increase of the target distance delayed the pointing apex but did not significantly influence its time position relative to the apices of the jaw-opening motions: $P_{A/J}$ was approximately the same in the two target conditions. This originates from the fact that the apices of the jaw-opening gestures were delayed when the target was far compared to when it was near. Regardless of the experimental conditions, the stress in speech always occurred sometime during the pointing plateau, either at the plateau onset for 'CVCV words or at the plateau offset for CV'CV words. This resulted in the entire pointing plateau occurring before the onset of the second jaw-opening stroke, when the first syllable was stressed, and after it, when the second syllable was stressed. More precisely, the pointing plateau occurred during the jaw-closing stroke after the stressed vowel in the first-stress syllable condition and during the jaw-opening stroke toward the stressed vowel when the second syllable was stressed. In the section that follows, we investigate the finger and jaw motion in detail to further characterize the impacts of stress, target, and consonant on each system separately.

Detailed Description of Finger-Forward and Jaw-Opening Strokes

Finger-forward stroke. The previous analysis showed that the pointing onset, P_I, did not depend on the experimental condition, whereas the pointing apex, P_A, occurred significantly later when the target was far than when it was near. This timing pattern implied a greater duration of the forward stroke (computed as $P_A - P_I$) when the target was far (431 ms) rather than near (406 ms), $F(1, 19) = 26.3, p < .0001$ (see Figure 5, first row, right). Neither the effect of stress position nor the effect of consonant on the duration of the pointing-forward stroke was significant. The amplitude of this stroke (computed as the distance between the pointing spatial positions at P_A and P_I; see Figure 5, first row, left) was also greater when the target was far (423 mm) compared to when it was near (275 mm), $F(1, 19) = 2222.0, p < .0001$. The 2-mm difference observed in the first-syllable stress condition (348 mm) compared to the second-syllable stress condition (350 mm) was also significant, $F(1, 19) = 8.3, p < .01$, but it did not significantly interact with target position. The consonant did

Figure 5. Means and standard errors of the amplitude and the duration of the finger forward stroke (first row) and of the two jaw-opening strokes (second and third rows), depending on stress position (first vs. second syllable), target position (near vs. far), and consonant (/t/ vs. /p/).



not have a significant effect on the amplitude of the pointing-forward stroke. Hence, the pointing-forward stroke was mainly influenced by target position: The alignment of the finger with the target required greater amplitude and duration when the target was far compared to when it was near. This leads one to consider that the pointing-forward stroke is "target-driven": Its main objective is the alignment of the finger with the target. Consequently, the speech system may have to adapt in order to achieve the relative timing pattern between the apices of finger- and jaw-opening gestures, as previously observed.

Jaw opening strokes. As described previously, the initiations and apex events of the jaw-opening strokes clearly depended on the experimental conditions. The duration of the first opening stroke (computed as $J_{A1} - J_{I1}$; see Figure 5, second row, right) was 39 ms greater when the first syllable was stressed (183 ms) rather than when the second one was stressed (144 ms), $F(1, 19) = 26.2$, $p < .0001$. It was also 37 ms greater for /t/ (182 ms) compared to /p/ (145 ms), $F(1, 19) = 14.5, p < .01$. The Stress Position × Consonant interaction was also significant: The effect of stress position was larger for /p/ (first syllable: 168 ms, second syllable: 122 ms) than for /t/ (first syllable: 197 ms, second syllable: 167 ms), $F(1, 19) = 5.5$, $p < .05$. However, target position did not significantly influence the duration of the first opening stroke. Similarly, the amplitude of the first jaw stroke (distance between the jaw spatial positions at J_{I1} and J_{A1}; see Figure 5, second row, left) was greater when the first syllable was stressed (12.5 mm) than when the second one was stressed (9.6 mm), $F(1, 19) = 21.1, p < .001$. Consonant also had a significant effect: The mean amplitude of the first opening gesture was greater for /p/ (11.7 mm) than for /t/ (10.3 mm), $F(1, 19) = 9.7$, $p < .01$. Neither the target position nor the interactions was significant. Contrary to the first opening stroke, the duration of the second opening stroke ($J_{A2} - J_{A12}$; see Figure 5, third row, right) was greater when the second syllable was stressed (152 ms) rather than the first one (100 ms), $F(1, 19) = 42.1$, $p < .0001$, and for /p/ (138 ms) as compared to /t/ (114 ms), $F(1, 19) = 47.1$, $p < .0001$. There was a significant Stress Position × Consonant interaction, $F(1, 19) = 8.2$, $p = .05$: The effect of stress position was greater for /t/ (first syllable: 84 ms, second syllable: 145 ms) than for /p/ (first syllable: 117 ms, second syllable: 159 ms). Target position did not significantly influence $J_{A2} - J_{A12}$. A similar pattern was observed for the amplitude of the second stroke (difference between jaw spatial positions at J_{I2} and J_{A2}; see Figure 5, third row, left), which was greater when the second syllable was stressed (7.5 mm) rather than the first one (3.2 mm), $F(1, 19) = 40.3$, $p < .0001$, and for /p/ (6.7 mm) compared to /t/ (4.0 mm), $F(1, 19) = 45.8$, $p < .0001$. There was also a significant Stress Position × Consonant interaction, $F(1, 19) = 5.7$, $p < .05$: The effect of stress position was greater for /t/ (first: 1.6 mm, second: 6.4 mm) than for /p/ (first: 4.8 mm, second: 8.7 mm). There was no significant effect of target position. These results show that in order to produce stress as instructed, speakers increase both the duration and the amplitude of the jaw-opening stroke corresponding to the stressed syllable. The specific articulatory configurations of /p/ and /t/ resulted in significant effects of the consonant factor on the amplitude and the duration of the jaw-opening strokes. However, these effects of consonant appeared to have no influence on the timing of the apices of the jaw-opening gesture. Finally, the target position did not have a significant effect on the amplitude and the duration of the jaw-opening strokes: Increase in distance to the target induced a delay of approximately 10 ms for all jaw events studied here. These results support the view that the timing of the pointing apex relative to the timing of the apices of the jaw-opening gestures mainly originates from an adaptation of the jaw.

In summary, the analyses showed that jaw–finger coordination basically consists of (a) synchronizing the pointing apex with the apex of the first jaw-opening gesture when the first syllable is stressed and (b) synchronizing the onset of the finger return gesture with the apex of the second jaw-opening gesture when the second syllable is stressed. This pattern seems to rely on four types of adaptation. The first is an adaptation of the finger pointing-forward stroke to the spatial target: Amplitudes and durations increase when the target is far rather than near. The second is an adaptation of the jaw to the phonetic goal: Specific durations and amplitudes correspond to /p/ versus /t/, and amplitudes and durations are greater for stressed versus unstressed syllables. The third is an adaptation of the jaw to the pointing-forward stroke: When the first syllable is stressed, whichever the spatial target is, the apex of the jaw-opening gesture corresponding to the stressed syllable is synchronized with the pointing apex. When the second syllable is stressed, the delay between the apices of the two jaw-opening gestures decreases, possibly allowing synchronization between the pointing apex and the onset of the jaw-opening gesture corresponding to the second syllable. The fourth is an adaptation of the duration of the pointing plateau, possibly compensating for the incomplete finger–jaw apex synchrony when the second syllable is stressed. In this condition, the increase of the duration of the pointing plateau allows for the apex of the jaw-opening stroke that corresponds to the stressed syllable to occur within the finger-to-target alignment period.

These results are now discussed in light of previous observations on speech–hand coordination and within the framework of the jaw–hand/finger coordination hypothesis.

Discussion

The change in stress position from a 'CVCV to a CV'CV clearly influences the jaw–finger coordination. However, it does so in a different way than we had predicted. There are bidirectional adaptations from the jaw to the hand and from the hand to the jaw, which are superimposed on the intrinsic target-driven behaviors corresponding to each system. We discuss these three components (intrinsic behavior, jaw-to-hand adaptions, and hand-to-jaw adaptations) separately, and we conclude with a discussion about the way the two systems could be coupled in a dynamic systems approach.

A Pointing Task and a Speech Task

Obviously, the first point that must be made in this discussion is that, apart from any link between speech and hand gestures, the hand and speech systems reach their respective targets regardless of the individual requirements on each system. In the pointing task, the target must be designated and the movement must be adapted to the target position, which involves larger and longer movements when the target is far compared to when it is near. In the speech task, a speaker must utter the correct consonants with the correct stress pattern, which involves an increase of the amplitude and/or the duration of the opening gesture for the stressed syllable. The phonetic identity of the consonant (labial vs. coronal) influences the initiation times of the jaw-opening strokes but not their apex times. It also affects the amplitude and duration of the jaw-opening strokes. Asymmetry in the jaw cycles for labial and coronal constrictions in CV syllables is a phenomenon also observed in the repetition of CVCV words when speech rate is increased. One possible explanation for this finding is that, for anatomical reasons, the position of the jaw has a stronger influence on the position of the lower lip than on the position of the tongue (Rochet-Capellan & Schwartz, 2007).

The Jaw Meets the Hand

With regard to the first and major connection between the two systems, it appears that speech adapts to the hand rather than the reverse and that this adaptation is related to the jaw movements. Our prediction was that the synchronization between the apex of the jaw-opening stroke corresponding to the stressed syllable and the apex of the pointing gesture would be a stable attraction point for jaw–finger coordination. This appears to be the case only when the first syllable is stressed. In this case, the observed large delay between the beginning of the pointing gesture and the beginning of the first jaw-opening gesture is in accordance with previous results obtained using the dual-task paradigm (see introduction). When the second syllable is stressed, the delay between the onsets of the pointing gesture and the jaw movement is reduced because of an early onset of the jaw movement. However, this is not enough for the apex of the pointing gesture to be synchronized with the apex of the second jaw-opening gesture. It could seem surprising that when the second syllable is stressed, the onset of the jaw movement occurs 148 ms (on average) after the onset of the hand movement, whereas the apex of the second jaw-opening stroke occurs 151 ms after the apex of the pointing gesture. Initiating the first jaw-opening gesture at the same time as the pointing gesture could be sufficient for the apex of the second jaw-opening gesture to be synchronized with the apex of the pointing gesture. The hand then compensates this failed synchronization during its return to rest phase, as discussed later. However, in terms of jaw adaptation, the synchronization seems partly unachieved. This does not mean that there is no meeting point between the jaw and the hand when the second syllable is stressed. In this condition, the jaw produces the first cycle early enough to entirely incorporate the pointing plateau inside the second cycle. Therefore, the general working hypothesis—that the part of the discourse that shows is related to the part of the gesture that shows—is verified. The movement of the jaw is organized such that the pointing plateau unambiguously occurs within the appropriate temporal domain of the spoken utterance that is the jaw cycle corresponding to the stressed syllable. This specifies the "deictic sites" that we were seeking.

There seems to exist another meeting point that we had not predicted: When the second syllable is stressed, the finger reaches its apex more or less exactly at the same time as the jaw reaches its highest position, which is likely to closely correspond to the consonantal contact time for the second syllable (between the lips for /p/, or between the tongue and the palate for /t/). Interestingly, this is reminiscent of the synchrony put forward by Attina, Cathiard, and Beautemps (2006) between the hand target and the achievement of the consonantal closure in cued speech. This manually augmented speech communication system involves hand positions and finger configurations added to the speech flow at a syllabic rhythm, providing complementary information for people who are deaf to almost perfectly understand oral communication. In their study of cued speech production by French speakers, Attina and colleagues found a clear synchrony between (a) the apices of hand movements achieving their adequate position for a given syllable and (b) the consonantal closure times in the corresponding speech.

In summary, "the jaw meets the hand" in the sense that the jaw adapts the onset of its movement in order for the pointing plateau to occur within the jaw cycle for the stressed syllable. This adaptation depends on the specific pointing and speech requirements. It occurs within a temporal window providing two possible meeting points for the pointing gesture apex: one corresponding to the apex of the first jaw-opening gesture (when the first syllable is stressed) and another corresponding to the onset of the second jaw-opening gesture (when the second syllable is stressed). The adaptation is achieved through the adequate delay of the onset of the jaw movement. If the target to point at is far, the pointing stroke is longer by about 16 ms, and the onset of the jaw movement is delayed by about 10 ms.

The Hand Meets the Jaw

There is also a clear hand/finger adaptation to the movement of the jaw. It appears that the hand waits for the apex of the jaw-opening stroke corresponding to the stressed vowel to occur before initiating its return stroke.

This is illustrated by the observed significant effect of stress position on the pointing plateau offset. To our knowledge, this is the first experimental evidence of such adaptation of the pointing gesture to speech. This could be the way the hand "corrects" for the lack of jaw–finger synchrony at the apex time—that is, by providing another meeting point thanks to the expansion of the plateau duration. This ensures that the jaw-opening gesture apex corresponding to the stressed syllable does occur within the pointing plateau, in all cases.

Two Coupled Dynamic Systems With Synchronization Sites

Generally, there appears to be a bidirectional link between the hand/finger and the jaw. This could be implemented by a series of local adaptations from one to the other, driven by specific temporal commands. The fact that there are simultaneous specific requirements for both systems rather suggests a kind of coupling between two dynamic systems, with the stress command specifying the adequate sites of synchronization. Actually, in the debate about modularity versus interactivity of the motor control of speech and hand-pointing systems, it has been assumed that when the speech and the hand systems move at the same time, the brain coordinates them as a single structure. In this framework, Kelso et al. (1981) showed that rhythmic tasks involving speech and hand movements displayed preferential phasing relationships between the articulators and the hand, as two coupled oscillator systems. Hence, according to these authors, speech–hand coordination should be described inside a coupled-oscillators modeling framework, such as the Haken-Kelso-Bunz (HKB) model introduced by Haken, Kelso, and Bunz (1985) for bimanual coordination. Other contributions to the dynamic systems approach of speech–hand coordination showed that the link between the two systems is not absolute (Smith, McFarland, & Weber, 1986) and also depends on high-level cognitive factors (see Treffner & Peter, 2002, for an improvement of the HKB model using intentional and attentional factors). However, all of these contributions support the view that the two systems could be linked through their relative oscillatory frequencies. In addition, developmental studies suggested a ratio of 2:1 between the preferential oscillatory frequencies of the jaw and the hand performing pointing gestures (Ducey-Kaufmann, 2007). This implies that, at most, two syllables could be achieved within one pointing gesture without influencing the pointing gesture. In a preliminary study (Rochet-Capellan, Schwartz, Laboissière, & Galván, 2007), we showed that for a task consisting of pointing and naming a target with a 1-, 2-, 3- versus 4-CV syllable word twice in rapid succession (e.g., /pa/+gesture and /pa/+gesture again), the delay between the apices of the two pointing gestures is the same for 1- and 2-CV words while it significantly increases from 2- to 3-CV words. This preliminary result provides some confirmation of the 2:1 preferential ratio.

Finally, the systematic delay of the speech gesture relative to the pointing gesture, together with the large adaptation of speech to the hand gesture, suggest that the speech response could be anchored in the pointing gesture through the online monitoring of the hand movement. This could explain why Levelt et al. (1985) observed an effect of pointing perturbation on speech only if the perturbation occurred at the beginning of the hand gesture. Thus, the speech response would be carried by the pointing gesture, in agreement with the proposal that the manual activity drives the coordination with the oral system (Iverson & Thelen, 1999). This hypothesis would need further investigation, especially about the question of the information used by the speech system to initiate its response.

Conclusion

Taken together, the results agree with the idea that speech focus is anchored in the pointing gesture for deictic expressions. This anchoring seems to be supported by a synchronization of the speech frame (the jaw cycle) with the sign frame (the pointing gesture) as suggested in the "vocalize to localize" framework (Abry et al., 2004) for the origins of language. This results from two independent levels of speech adaptation. First, the effects of stress position and consonant order show an "internal" adaptation of the jaw-opening gestures to the phonetic goal. In addition, the effect of spatial target position suggests an "external" adaptation for the synchronization with the pointing gesture within the adequate deictic site.

Further investigation will be needed to better understand the coordination of the jaw and hand in deictic tasks. But the interesting point about this study is that there is, indeed, a coordination, implemented in a more or less complex way in the speaker's brain, and that this coordination can be put forward and studied by analyzing online behavior in various linguistic tasks involving pointing using the voice and the hand.

Acknowledgments

This work is part of the "Patipapa" project funded by the French Ministry of Research (Action Concertée Incitative "Systèmes Complexes en Sciences Humaines et Sociales"). Rafael Laboissière now works as a researcher at Centre National de la Recherche Scientifique. We also wish to thank Christian Abry and Coriandre Vilain for insightful discussions and Marion Dohen for proofreading this article.

References

Abry, C., Vilain, A., & Schwartz, J. (2004). Introduction: Vocalize to localize? A call for better crosstalk between auditory and visual communication systems researchers: From meerkats to humans. *Interaction Studies, 5*(3), 313–325.

Attina, V., Cathiard, M.-A., & Beautemps, D. (2006). Temporal measures of hand and speech coordination during French cued speech production. In S. Gibet, N. Courty, & J.-F. Kamp (Eds.), *GW 2005, Lecture Notes in Artificial Intelligence, 3881* (pp. 13–24). New York: Springer-Verlag.

Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 9–33). Mahwah, NJ: Erlbaum.

Castiello, U., Paulignan, Y., & Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness: A study in normal subjects. *Brain, 114,* 2639–2655.

Diessel, H. (1999). *Demonstratives: Form, function, and grammaticalization*. Amsterdam: John Benjamins.

Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics, 17,* 463–489.

Ducey-Kaufmann, V. (2007). *Le cadre de la parole et le cadre du signe: un rendez-vous développementale* [The speech frame and the sign frame: A developmental rendezvous]. Unpublished doctoral dissertation, Université Stendhal.

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

Gentilucci, M. (2003). Grasp observation influences speech production. *European Journal of Neuroscience, 17,* 179–184.

Gentilucci, M., Benuzzi, F., Gangitano, M., & Grimaldi, S. (2001). Grasp with hand and mouth: A kinematic study on healthy subjects. *Journal of Neurophysiology, 86,* 1685–1699.

Gentilucci, M., Santunione, P., Roy, A. C., & Stefanini, S. (2004). Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. *European Journal of Neuroscience, 19,* 190–202.

Goldin-Meadow, S., & Butcher, C. (2003). Pointing toward two-word speech in young children. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 85–107). Mahwah, NJ: Erlbaum.

Green, J., Moore, C., & Reilly, K. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research, 45,* 66–79.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetic, 51*(5), 347–356.

Haviland, J. B. (2000). Pointing, gesture spaces, and mental maps. In D. McNeill (Ed.), *Language and gesture* (pp. 13–46). Cambridge, United Kingdom: Cambridge University Press.

Holender, D. (1980). Interference between a vocal and a manual response to the same stimulus. In G. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 421–432). Amsterdam: North Holland Publishing Company.

Iverson, J., & Goldin-Meadow, S. (1998, November 19). Why people gesture as they speak. *Nature, 396,* 228.

Iverson, J., & Thelen, E. (1999). Hand, mouth, and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies, 6,* 19–40.

Kelso, J. A. D., Tuller, B., & Harris, K. S. (1981). A "dynamic pattern" perspective on the control and coordination of movement. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 137–173). New York: Springer-Verlag.

Levelt, W. J. M., Richardson, G., & Heij, W. L. (1985). Pointing and voicing in deictic expressions. *Journal of Memory and Language, 24,* 133–164.

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. *Journal of Neurolinguistics, 18,* 237–258.

MacNeilage, P., & Davis, B. (2000, April 21). On the origins of internal structure of word forms. *Science, 288,* 527–531.

McNeill, D. (1981). Action, thought, and language. *Cognition, 10,* 201–208.

McNeill, D. (1992). *Hand and mind*. Chicago: University of Chicago Press.

McNeill, D. (2000). Introduction. In D. McNeill (Ed.), *Language and gesture* (pp. 1–10). Cambridge, United Kingdom: Cambridge University Press.

Munhall, K., & Jones, J. (1998). Articulatory evidence for syllabic structure. *Behavioral and Brain Science, 21,* 524–525.

Petitto, L., Holowka, S., Sergio, L., & Ostry, D. (2001, September 6). Language rhythms in baby hand movement. *Nature, 413,* 35–36.

Pizzuto, E., Capobianco, M., & Devescovi, A. (2005). Gestural-vocal deixis and representational skills in early language development. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 6*(2), 223–252.

Rochet-Capellan, A., & Schwartz, J.-L. (2007). An articulatory basis for the labial-to-coronal effect: /pata/ seems a more stable articulatory pattern than /tapa/. *The Journal of the Acoustical Society of America, 121,* 3740–3754.

Rochet-Capellan, A., Schwartz, J.-L., Laboissière, R., & Galván, A. (2007). Two CV syllables for one pointing gesture as an optimal ratio for jaw-arm coordination in a deictic task: A preliminary study. In D. Kayser & S. Vosniadu (Eds.), *Proceedings of the European Cognitive Science Conference* (EuroCogSci07; pp. 608–613). Delphi, Greece: Routledge.

Smith, A., McFarland, D. H., & Weber, C. M. (1986). Interactions between speech and finger movements: An exploration on the dynamic pattern perspective. *Journal of Speech and Hearing Research, 29,* 471–480.

Stetson, R. H. (1951). *Motor phonetics* (2nd ed.). Amsterdam: North Holland Publishing Company.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78,* 705–722.

Treffner, P., & Peter, M. (2002). Intentional and attentional dynamics of speech-hand coordination. *Human Movement Science, 21*(5–6), 641–697.

Volterra, V., Caselli, M. C., Capirci, O., & Pizzuto, E. (2005). Gesture and the emergence and development of language. In M. Tomasello & D. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates* (pp. 3–40). Mahwah, NJ: Erlbaum.

Wilkins, D. (2003). Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 171–215). Mahwah, NJ: Erlbaum.

Received July 24, 2007

Accepted April 7, 2008

DOI: 10.1044/1092-4388(2008/07-0173)

Contact author: Amélie Rochet-Capellan, who is now a postdoctoral fellow with the Motor Control Laboratory, Department of Psychology, McGill University, 1205 Dr. Penfield Ave., Montreal, Quebec H3A 1B1, Canada. E-mail: amelie@motion.psych.mcgill.ca.