Why does picture naming take longer than word reading? 
Contribution of articulatory processes

Stéphanie Riès\textsuperscript{1,2}, Thierry Legou\textsuperscript{1}, Boris Burle\textsuperscript{1}, F.-Xavier Alario\textsuperscript{1}, & Nicole Malfait\textsuperscript{1}
\textsuperscript{1}CNRS & Aix-Marseille Université
\textsuperscript{2}Helen Wills Neuroscience Institute, University of California at Berkeley

Corresponding author:
Nicole Malfait
Institut de Neurosciences de la Timone (INT)
UMR 7289, CNRS - Aix Marseille Université
Campus santé Timone
27, boulevard Jean Moulin
13385 Marseille cedex 5
nmalfait@gmail.com

Abstract: 202 words
Pages: 17
Words: 4189
Figures: 3

Keywords: speech production, naming, reading, verbal response-times, articulation, electromyography

Acknowledgment: This research was supported by grant no. ANR-07-JCJC-0074 from the Agence Nationale de la Recherche (France), by European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 241077 & 263575), by an MNRT grant from the Ministère de l'Enseignement et de la Recherche (France) awarded to Stéphanie Riès and by IFR 131 Sciences du Cerveau et de la Cognition (France). We thank Licia Damiani for collecting the data.
ABSTRACT

Since the 19th Century, it has been known that response latencies are longer for naming pictures than for reading words aloud. While several interpretations have been proposed, a common general assumption is that this difference stems from cognitive word-selection processes and not from articulatory processes. Here we show that, contrary to this widely accepted view, articulatory processes are also affected by the task performed. To demonstrate this we used a procedure that has never been used in research on language processing: response-latency fractionating. Along with vocal onsets, we recorded the electromyographic (EMG) activity of facial muscles while participants named pictures or read words aloud. On the basis of these measures, we were able to fractionate the verbal response latencies into two types of time intervals: pre-motor times (from stimulus presentation to EMG onset), mostly reflecting cognitive processes, and motor times (from EMG onset to vocal onset), related to motor execution processes. We show that pre-motor and motor times are both longer in picture naming than in reading, although articulation is already initiated in the latter measure. Future studies based on this new approach should bring valuable clues for a better understanding of the relation between cognitive and motor processes involved in speech production.
Response latencies are longer for naming pictures than for reading words aloud. This established fact has been a focus of interest in psycholinguistics for over a century (Cattell, 1885; Ferrand, 1999). It has been interpreted in different ways. Some studies favor the so-called “semantic hypothesis”, according to which time-consuming access to semantic information is only required in picture naming (Theios and Amrhein, 1989). According to the “uncertainty hypothesis”, the effect is due to a more equivocal stimulus-response association in picture naming compared to word reading. A picture, but not a printed word, can be described with different labels (Fraisse, 1969; Ferrand, 1999). Underlying these hypotheses is the idea that performance differences (e.g. in verbal response times) is entirely due to central cognitive stages, such as differences in word-selection processes, while the articulatory processes would be insensitive to the nature of the task (e.g. Roelofs, 2004; Levelt, Roelofs, Meyer, 1999).

Kello, Plaut and MacWhinney (2000) and Damian (2003) investigated the relationship between the cognitive and motor processes involved in speech production using a Stroop task (Stroop, 1935). In addition to verbal response onsets, they measured verbal response durations. When high time-pressure was imposed, a Stroop effect was observed on verbal response durations. According to Kello et al. (2000), this demonstrates that task demands can flexibly change speech production processes from “serial”, where each processing stage must be over before the next one starts, to “cascaded”, where articulatory processes can be initiated before semantic retrieval and phonological access are complete. In the latter case, how cognitive access occurs may influence the way articulation is performed. Hence this account questioned the dissociation between cognitive and motor processes that is generally taken for granted. However, Damian (2003) failed to replicate the critical effect reported by these authors, and argued on the contrary that in “conceptually driven” tasks cognitive processes cannot cascade down to the articulatory stage.

Close consideration of word-form encoding processes suggests that words and pictures
Articulatory processes in naming and reading

may differ in how they trigger the verbal response. In word reading, but not in picture naming ("conceptually driven" responses), partial sub-lexical information is afforded to permeate through to output before pre-production processes are complete. Consistent with this idea, Hennessey and Kirsner (1999) showed that articulatory durations of the same word were longer for printed words than for pictures. In their view, this difference (observed for low frequency items only) indicates that response execution can be stretched out on-line to compensate for an early start on the basis of partial information. In contrast to this view, however, there is also evidence supporting the idea that reading a word aloud starts only when the computation of the phonology of the entire verbal response is complete, at least with monosyllabic words (Rastle, Harrington, Coltheart, Palethorpe, 2000).

Despite apparent inconsistencies, these studies are valuable as they address an issue that is seldom investigated in language processing research, namely the relationship between the cognitive and motor processes involved in speech production. We submit, however, that the measures most commonly used may have been suboptimal. On the one hand, verbal response times (RTs) reflect the sum of processes occurring from stimulus presentation to vocal onset, without distinguishing between cognitive and motor processes. Indeed, verbal RTs include, not only the time allocated to word selection, but also the time needed to plan and initiate the articulatory motor processes preceding vocal onset. On the other hand, verbal response durations are measured from vocal onset, and thus do not include the early cognitive phase.

In the frame of chronometric-analysis studies using manual responses, RT fractionating procedures have been developed, intended to help disentangle time-delays reflecting response-selection from delays related to motor response execution (e.g., Botwinick and Thompson, 1966; Burle, Possamaï, Vidal, Bonnet, Hasbroucq, 2002). The idea is to break the RTs into two time-intervals on the basis of the electromyographic (EMG) activity that can be recorded preceding the overt manual response (e.g. from the flexor pollicis brevis muscle for thumb button-press responses.
Articulatory processes in naming and reading

The first delay, between the presentation of the stimulus and the onset of EMG activity, would essentially reflect stimulus encoding, response selection and motor planning, and has been referred to as “pre-motor time” (pre-MT). The second time-interval, that separates EMG onset from button-press, would characterize motor response execution processes, and has been referred to as “motor time” (MT). Applying this procedure, these studies could demonstrate that task manipulations (e.g. the number of possible responses) initially thought to influence pre-motor processes only also affect MTs (Possamaï, Burle, Osman, Hasbroucq, 2002).

To our knowledge, RT fractionating procedures have never been applied in the field of language production research (although see discussion of a “not [...] observable” execution-acoustic interval in Rastle, Croot, Harrington, Coltheart, 2005). Here we exploited this method as a means to test whether motor response execution processes engaged in articulation are affected by task demands. Along with vocal onsets, we measured electromyographic (EMG) activity of lip muscles as participants named pictures or read words aloud. On the basis of this latter measure, we divided verbal response-times into two time intervals: pre-MTs and MTs (Figure 1). There is, however, a clear difference between the way this method was originally applied to simple manual responses and the way we exploited it here. In typical thumb button-press responses there is a direct causal link between flexor pollicis brevis EMG and thumb flexion. In contrast, EMG activities recorded from facial muscles and voice onsets correspond to different effectors. (Therefore, trials in which the EMG activities follow the vocal onset can also be observed, for details see Supplementary Materials). In addition, EMG activities trigger the action of effector muscles, whereas vocal onset is the effect of an effector. Thus fractionating verbal RT allowed us to assess whether task demands induce systematic variations in the time (phase) relationship between the actions of two vocal-tract effectors (protrusion and/or aperture of the lips and glottal aperture).

Under the hypothesis that articulatory motor execution is insensitive to the origin of the speech to be produced, MTs should not be affected by the task performed. If, on the contrary, MTs
differ across reading and naming, it would directly challenge this widely-held idea. This hypothesis was tested in two similar experiments, both contrasting reading and naming, which differed as follows. In the first experiment, the two tasks were performed in different blocks involving different but carefully matched materials, whereas in the second experiment the words to be read corresponded to the names of the pictures, and the two types of trials were intermixed within the testing blocks. Other differences are detailed below.

**Figure 1: Verbal RT fractionating; the standard naming latency is divided into a pre-motor time (pre-MT; between stimulus and EMG onsets) and a motor time (MT; between EMG and vocal onsets).**

**EXPERIMENT 1**

**Methods**

**Participants**

Ten native French-speakers with normal or corrected to normal vision participated in the experiment (mean age: 21, σ=3 years). The data from one participant were removed from the analysis due to voice-key problems.

**Materials and Design**

The stimuli were 40 line-drawings of common objects selected from published
collections (e.g., Alario and Ferrand, 1999) or designed for the experiment (mean name-agreement: 92%, σ=12%) and 40 French written words. The stimuli were presented on a CRT screen positioned 150 cm from the participant. The pictures were presented in black on an 11x11 cm white square and the words were written in white on a black background, covering on average a comparable visual angle of 4°. The picture names and words' initial phonological properties (1st-phoneme: /b/, /f/, /m/, /p/, /v/; and type of second phoneme: open-palatal, closed-palatal and closed-velar), lexical frequency and number of syllables were matched (see Supplementary Materials).

Each experimental run comprised 40 trials, which were either the 40 pictures or the 40 written words presented pseudo-randomly. Each participant was tested on 8 runs. The order of runs was counterbalanced across participants.

Procedure

Each trial consisted of the following sequence: (1) a fixation point ("+" sign) for 500 ms, (2) the picture or word presented until the participant responded or a 1500 ms deadline was reached, (3) a blank screen for 2000 ms, controlled by Eprime 2.0 (Psychology Software Tools, Inc.). In a preliminary familiarization phase, the 40 words and 40 drawings that would be used in the test phase were each presented once in two separate runs, with counterbalanced orders across participants. The instruction was to read or name each of the items. For the pictures, participants were corrected when producing an unexpected name. The actual test phase lasted about 30 minutes. Participants were asked to name the pictures and read the words aloud as fast and as accurately as possible.

Vocal responses were recorded with a piezzo-electric microphone, sampled at 22050 Hz. EMG activity was recorded monopolarly with pre-amplified surface electrodes Ag/AgCl (BioSemi, Inc.) from three oro-facial muscles: orbicularis oris, risorius and mentalis; sampling rate was 2048 Hz (filters: DC to 268 Hz, 3 db/octave). A passive reference electrode was placed on the earlobe, and two active reference electrodes positioned on the arm-wrist.
Data processing

The EMG data were filtered off-line (high-pass=10 Hz). Because of poor signal-to-noise ratio, EMG recorded from the \textit{mentalis} was discarded from the analyses. For off-line detection of the vocal and EMG onsets the same algorithm was used, based on the Teager-Kaiser operator that permits the detection of abrupt onsets of energy in the signal (Li, Zhou, Aruin, 2007). Processed by this operator, the signal was then filtered (moving average window) to discard detections of unexpected isolated pulses (e.g. due to lip opening). For each trial, the detected onset was then visually checked and readjusted when needed. This was performed blind to the task (reading or naming). Trials where the baseline on the EMG channels was too noisy, as well as those where recording failed (e.g. faulty voice-key triggering or no verbal response within 1500ms) were excluded from further analysis. Trials were coded as errors in case of partial or complete production of incorrect or unexpected words, verbal disfluencies (stuttering, utterance repairs, etc.).

Mixed-effects analysis of variance (ANOVA) were run on the three different measures (verbal RTs, pre-MTs and MTs) with the factors Task (naming vs reading), Muscle (\textit{orbicularis} or \textit{risoris}) and 1st-phoneme (/b/, /f/, /m/, /p/ or /v/) as fixed effects, and with participants (F1) or linguistic items (F2) as random variables. For the linguistic item analysis, Task and 1st-phoneme were between-items factors, and Muscle a within-item variable. Student two-tailed t-tests were used for post-hoc pairwise comparisons. For all tests, a significance level of 0.05 was used. P-values were adjusted for multiple comparisons using the Bonferroni correction.

Results

An average of 4.5% of trials (σ=3.1%) per participant were rejected due to poor EMG signal-to-noise ratio. Results are summarized in Figure 2. Verbal RTs were longer for pictures than for words (Δ=185 ms, σ=49 ms; $F_1(1,8)=139.16$, $p<0.001$, $F_2(1,70)=484.15$, $p<0.001$). The error rate was higher for pictures (3.56%, σ=3.73%) than for words (0.65%, σ=0.98%; $t_1(8)=2.90$, $p<0.05$; $t_2(52.45)=3.49$, $p<0.01$). There was a significant main effect of 1st-phoneme ($F_1(4,32)$
Articulatory processes in naming and reading

=4.66; \( p<0.01 \); \( F_2(4,70)=4.31, p<0.01 \), due to utterances starting with /p/ associated with longer RTs, although none of the pairwise comparisons reached significance (/p/ vs. /v/: \( t_1(8)=2.07, p=0.29 \); all other comparisons: \( t_1's(8)<1 \) and \( t_2's<1^1 \)). The interaction between Task and 1st-phoneme was not significant (\( F_1(4,32)=1.24, p=0.31 \); \( F_2(4,70)<1 \)).

Pre-MTs were longer for pictures than for words (\( \Delta=164 \text{ ms}, \sigma=45 \text{ ms}; F_1(1,8)=115.88; p<0.001; F_2(1,70)=549.82, p<0.001 \)). There was no effect of Muscle (\( F_1(1,8)=2.76, p=0.10 \)), and no interaction between Task and Muscle (\( F_1(1,8)=1.45, p=0.26; F_2(1,70)=2.23, p=0.14 \)). A main effect of 1st-phoneme (\( F_1(4,32)=3.18; p<0.05; F_2(4,70)=4.94, p<0.01 \)) was observed. However, while /f/ and /v/ tended to produce longer responses, pairwise differences did not reach significance (all \( t_1's(8)<1.21 \) and all \( t_2's<1^1 \)). There were no interactions between 1st-phoneme and Task (\( F_1(4,32)=1.10, p=0.37; F_2(4,70)<1 \)) or between 1st-phoneme and Muscle (\( F_1(4,32)=1.82, p=0.15 \); although it was significant by items: \( F_2(4,70)=5.83, p<0.001 \), and no three-way interaction (\( F_1(4,32)<1; F_2(4,70)=1.36, p=0.26 \)).

Figure 2: For experiment 1, pre-motor and motor times across 1st-phonemes in the reading and naming tasks averaged over muscles. As expected, pre-motor times are longer for naming than for reading (left). Importantly, motor times are also longer for naming than for reading (right; note that motor times are depicted time-locked to speech onset).

Critically, MTs were longer in naming picture than in reading words (\( \Delta=23 \text{ ms}, \sigma=30 \text{ ms};^1 \)

---

^1 The degrees of freedom for the by-item t-tests varied from 22.99 to 29.96
Articulatory processes in naming and reading

$F_1(1,8)=5.10; \ p=0.05; \ F_2(1,70)=26.58, \ p<0.001)$. There was no main effect of Muscle$^2$ ($F_1(1,8)<1;\ F_2(1,70)=2.76, \ p=0.10$), and no interaction between Muscle and Task ($F_1(1,8)=1.45, \ p=0.26; \ F_2(1,70)=2.23, \ p=0.14$), but a significant main effect of 1st-phoneme ($F_1(4,32)=11.46; \ p<0.001; \ F_2(4,70)=27.71, \ p<0.001$). Longer MTs were observed for utterances starting with /p/ (/p/ vs /b/: $t_1(8)=5.75, \ p<0.01; \ t_2(29.85)=6.79, \ p<0.001$; other $t_1's(8)<1$ and $t_2's<1.30^3$). There were no interaction between 1st-phoneme and Task ($F_1(4,32)<1; \ F_2(4,70)<1$), 1st-phoneme and Muscle ($F_1(4,32)=1.82, \ p=0.15$; although significant by items: $F_2(4,70)=5.83, \ p<0.001$), and no three-way interaction ($F_1(4,32)<1; \ F_2(4,70)=1.36, \ p=0.26$). Further details about inter-participant and inter-trial variability are provided as Supplementary Materials.

EXPERIMENT 2

Methods

Participants

Eighteen native French-speakers with normal or corrected to normal vision participated in the experiment (mean age: 20.6, $\sigma=1.5$ years). The data of 6 participants were excluded from the analysis due to over-noisy EMG recordings.

Materials & Design

The stimuli were 12 line-drawings of common objects (mean name agreement: 96%, $\sigma=6$%) and 12 French written words corresponding to the names of the pictures (see Supplementary Materials). The 1st phonemes were: /b/, /m/, or /p/. Items beginning with /f/ or /v/ were not used given the variability they introduced in Experiment 1. The type of second phoneme (open-palatal, closed-palatal and closed-velar) was controlled as in the previous experiment.

The familiarization phase (see Experiment 1) was followed by 3 testing runs in which

---

$^2$ The estimates for the effects involving the Muscle factor are not expected to differ between PMT and MT analyses. This is because the performance for all the different muscles are estimated within trials, with the same RT for all muscles. In other words, PMT + MT is constant across muscles. As a consequence, the variance associated to this factor is the same for PMTs and for MTs. Importantly, this is not the case for the factors Task or 1st-phoneme, for which the estimates involve different trials.

$^3$
all 12 pictures and 12 words were presented each 4 times (96 trials in total) intermixed in a pseudo-random order (stimuli involving the same response were separated by at least two other items).

**Procedure**

The procedure was the same as in Experiment 1 except for the following modifications. The duration of the fixation point varied randomly between 500 and 1000 ms. The voice and EMG were recorded by the same device (Keithley Instruments, Inc.). Bipolar montages of 6 mm-diameter Ag/AgCl surface electrodes (Grass Technologies, Inc.) were used to record EMG activity from 4 facial muscles: *levator labii superioris*, *risorius*, *orbicularis oris*, and *depressor labii inferioris*. (Recording of the *mentalis* muscle activity was discarded as only poor signal-to-noise ratio was obtained in Experiment 1.) Sampling rate was 2000 Hz and the reference electrode was placed over the left collarbone. The acoustic signal was recorded at 28000 Hz. Stimuli-presentation durations were kept constant (1500ms) across trials, and both pictures and words were presented in black on a white background.

**Data processing**

The acoustic and EMG data were analyzed as in Experiment 1 except that, for the linguistic item analysis, 1st-phoneme was a between-items factor, while both Task and Muscle were within-item variables.

**Results**

An average of 4.3% of trials (σ=4.3%) were rejected due to poor EMG signal-to-noise ratio. Figure 3 summarizes the results. Again, Verbal RTs were longer for pictures than for words (Δ=90 ms, σ=38 ms; $F_1(1,11)=69.19$, $p<0.001$; $F_2(1,9)=208.86$, $p<0.001$). Error rates were very low for both pictures and words (0.44%, σ=1.32% and 0.18%, σ=0.33% respectively; pictures vs. words: $t_1(11)<1$; $t_2(11)=-1.49$, $p=0.17$). There was an effect of 1st-phoneme ($F_1(2,22)=221.93$, $p<0.001$; $F_2(2,9)=37.86$, $p<0.001$), due to items starting with /p/ (/p/ vs. /m/: $t_1(11)=17.91$, $p<0.001$; $t_2(11)=7.29$, $p<0.05$; /m/ vs. /b/: $t_1(11)<1$, $t_2(11)<1$). The interaction between Task and 1st-phoneme was
not significant ($F_1(2,22) = 1.43$, $p = 0.26$; $F_2(2,9) < 1$).

Pre-MTs were longer for pictures than for words ($\Delta = 74$ ms, $\sigma = 29$ ms; $F_1(1,11) = 77.46$, $p < 0.001$; $F_2(1,9) = 143.84$, $p < 0.001$). There was a main effect of Muscle ($F_1(3,33) = 15.66$, $p < 0.001$; $F_2(3,27) = 314.43$, $p < 0.001$). Pre-MTs were longer for *risorius* than for *orbicularis oris* ($t_1(11) = 3.99$, $p < 0.01$; $t_2(11) = 15.30$, $p < 0.001$; *orbicularis oris* vs. *depressor labii inferioris*: $t_1(11) = 2.28$, $p = 0.13$; $t_2(11) = 10.55$, $p < 0.001$; *levator labii superioris*: $t_1(11) < 1$; $t_2(11) = 3.33$, $p < 0.05$). There was no main effect of 1st-phoneme ($F_1(2,22) < 1$; $F_2(2,9) < 1$). A significant interaction was observed between Task and Muscle, with a larger effect of Task for muscles for which pre-MTs were larger ($F_1(3,33) = 3.77$, $p < 0.05$; $F_2(3,27) = 5.73$, $p < 0.01$; pairwise comparisons for *orbicularis oris* vs. *depressor labii inferioris*: $t_1(11) = 3.49$, $p < 0.05$; $t_2(11) = 3.47$, $p < 0.05$; other $t_1's(11) < 1$ and $t_2's(11) < 1$). The interaction between Task and 1st-phoneme was marginal by subject ($F_1(2,22) = 2.92$, $p = 0.07$; $F_2(2,9) < 1$; all pairwise comparisons: $t_1's(11) < 1$; $t_2's(11) < 1.25$). There was no interaction between 1st-phoneme and Muscle ($F_1(6,66) < 1$; $F_2(6,27) < 1$) and no three-way interaction ($F_1(6,66) = 1.69$, $p = 0.14$; $F_2(6,27) = 1.61$, $p = 0.18$).

Confirming the results reported in Experiment 1, MTs were longer for pictures than for words ($\Delta = 16$ ms, $\sigma = 22$ ms; $F_1(1,11) = 6.38$, $p < 0.05$, $F_2(1,9) = 17.40$, $p < 0.01$). There was a main effect of Muscle ($F_1(3,33) = 15.67$, $p < 0.001$; $F_2(3,27) = 316.10$, $p < 0.001$). MTs were longer for *orbicularis oris* than for *risorius* ($t_1(11) = 3.99$, $p < 0.01$; $t_2(11) = 15.29$, $p < 0.001$; *orbicularis oris* vs. *depressor labii inferioris*: $t_1(11) = 2.28$, $p = 0.13$; $t_2(11) = 10.49$, $p < 0.001$; *risorius* vs. *levator labii superioris*: $t_1(11) < 1$; $t_2(11) = 3.32$, $p < 0.01$). A significant interaction between Task and Muscle indicated that the effect of Task was larger for muscles for which MTs were larger ($F_1(3,33) = 3.78$, $p < 0.05$, $F_2(3,27) = 5.72$, $p < 0.01$; pairwise comparisons *depressor labii inferioris* vs. *orbicularis oris*: $t_1(11) = 3.45$, $p < 0.05$, $t_2(11) = 3.45$, $p < 0.01$; other $t_1's(11) < 1$ and $t_2's(11) < 1$). MTs were also affected by 1st-phoneme ($F_1(2,22) = 127.69$, $p < 0.001$; $F_2(2,9) = 31.03$, $p < 0.001$; items starting with /p/ being slower (/p/ vs. /b/: $t_1(11) = 17.04$, $p < 0.001$; $t_2(11) = 9.94$, $p < 0.01$; /b/ vs. /m/: $t_1(11) < 1$; $t_2(11) < 1$). The effect
Articulatory processes in naming and reading

of 1st-phoneme did not interact with Task ($F_1(2,22)<1; F_2(2,9)<1$) nor with Muscle ($F_1(6,66)<1; F_2(6,27)<1$). There was no three-way interaction ($F_1(6,66)=1.69, p=0.14; F_2(6,27)=1.60, p=0.18$).

Figure 3: Pre-motor and motor times broken down by task, muscle and 1st-phoneme in Experiment 2. Here too, both pre-motor and motor times are affected by the task.

DISCUSSION

Our main finding is that the well-known difference between naming and reading in verbal RTs is not due solely to increased pre-MTs but also to a significant lengthening of MTs for naming compared to reading. Contrary to a widely-held hypothesis, motor processes involved in the execution of the verbal response are also affected by the task being performed. Dividing the verbal RTs on the basis of EMG activities recorded from facial muscles allowed us to demonstrate that the nature of the stimulus can have an impact on the coordination of the articulatory “gestures” of the different effectors (lip and glottis) (Browman and Goldstein, 1992).

One may argue that these results simply reflect differences in strategies. In Experiment 1, where word and picture naming were presented in separate blocks, participants may have set a priori different response thresholds for the two tasks. However, the same task effect was observed in Experiment 2 where words and pictures were intermixed. Thus, we suggest as more likely an interpretation in which central, non-motor, effects would spread-out into the motor articulatory processes, in agreement with a cascaded model of information processing (McClelland, 1979). In
Articulatory processes in naming and reading

this view, information is continuously accumulated through different processing stages, and a response is produced when the last stage has reached a predefined level of information accumulation. The critical point is that, as soon as the first level starts accumulating information, it transmits downstream its partial output. If one level is affected by an experimental manipulation, this will also impact downstream levels. In the present context, a task processing difference at a cognitive level will percolate down to affect motor process.

Such a cascading hypothesis implies the concurrent engagement of cognitive processes and motor output, and a dynamic flow of information between them. (For example, segments produced in errors show articulatory features that are biased towards the non-produced target segment; thus information about the target segment or close alternatives is thought to have cascaded to articulatory processes: Goldrick and Blumstein, 2006; McMillan and Corley, 2010; see also Kello et al., 2001). The available evidence comprises two features that allow us to better specify some aspects of these dynamics. First, in our data, the task effect was larger the earlier the muscle was activated. If the task effect stems from cognitive processes that start before the onset of articulation, there would be a shorter delay between their initiation and the EMG onset for the muscles activated first. The cascading of the task effect would therefore be more perceptible (thus possibly larger) for those early muscles than for muscles activated later.

Second, there is a contrast between the direction of the task effects reported here (shorter MTs for word than pictures) and those reported by Hennessey and Kirsner (1999; for low frequency items, durations were shorter for pictures than for words). In their account, word duration is longer because articulatory execution is stretched out to accommodate the fact that it was triggered on the basis of partial information (e.g. word beginning) that needs to be completed online. To provide a unitary account of both findings, one may speculate post-hoc that the articulation of a smaller chunk (e.g. a word onset in reading) will be swifter and more efficient than the articulation of a complete phonetic program (e.g., a whole word in picture naming), but that subsequent execution is stretched
to compensate for a speedy start. This interpretation raises the intriguing possibility that online adjustments of articulation processes do occur and are measurable, even in tasks as simple as single word production (see also Kawamoto, Liu, Mura, Sanchez, 2008). We note however that the “sub-lexical” hypothesis has been partially challenged on different grounds (e.g. Rastle et al., 2000).

As noted earlier, influential current explanations for the naming-reading difference in verbal-RTs are diverse, but all are tied to processes occurring before the initiation of articulation: semantic access (Theios and Amrhein, 1989), response selection as indexed by stimulus-response compatibility (Fraisse, 1969; Ferrand, 1999), or partial phonological or phonetic access (i.e., the “sub-lexical hypothesis”, see previous paragraph). Adjudicating between the “semantic”, the “uncertainty”, and the “sub-lexical” hypothesis on the basis of the current data is not straightforward. It is worth noting however that the “uncertainty hypothesis” (Fraisse, 1969) finds support in some converging independent non-linguistic evidence, as response uncertainty has been shown to affect MTs in manual responses tasks (Possamaï et al., 2002).

In conclusion, we show that the well-known difference in verbal-RT performance between naming and reading is not due to the lengthening of cognitive processes solely, as often taken for granted. Our findings demonstrate that motor articulatory processes are also affected by the speech-production conditions, so the duration of the articulatory stage contributes in its own right to the age-old difference between latencies in word reading and picture naming. In addition, our study presents a procedure entirely new to research on language processing—verbal RT-fractionating—that permits finer grained analyses than commonly performed. Future studies based on this new approach intended to capture finer grained effects should bring valuable clues for a better understanding of the relation between cognitive and motor processes.
REFERENCES


Cattell, J.M. (1885) Über die zeit der erkennung und benennung von schriftzeichen, bildern und farben [The time it takes to recognize and name letters, pictures, and colors]. Philisophische Studien, 2, 635-650.


Articulatory processes in naming and reading


