Article scientifique | Article | 2012 | Accepted version | Open Access

- -

Time course of word production in fast and slow speakers: A high density ERP topographic study

- -

Laganaro, Marina; Valente, Andrea; Perret, Cyril

This publication URL: https://archive-ouverte.unige.ch//unige:27439
Publication DOI: 10.1016/j.neuroimage.2011.10.082

Time course of word production in fast and slow speakers: a high density ERP topographic study

Marina Laganaro[1], Andrea Valente[1], Cyril Perret[2]

1. FAPSE, University of Geneva, Switzerland

2. Université Paris 13, Sorbonne Paris Cité, UTRPP EA 4403

Corresponding Author:

Marina Laganaro

FAPSE

University of Geneva

40, Bd Pont d'Arve,

CH-1211 Geneva 4

E-mail: marina.laganaro@unige.ch

Tel. + 41 22 3798223

Tables: 2

Figures: 3

N. of words: 5123

Abstract

The transformation of an abstract concept into an articulated word is achieved through a series of encoding processes, which time course has been repeatedly investigated in the psycholinguistic and neuroimaging literature on single word production. The estimates of the time course issued from previous investigations represent the timing of process duration for mean processing speed: as production speed varies significantly across speakers, a crucial question is how the timing of encoding processing varies with speed. Here we investigated whether between-subjects variability in the speed of speech production is distributed along all encoding processes or if it is accounted for by a specific processing stage. We analyzed event-related electroencephalographical (ERP) correlates during overt picture naming in 45 subjects divided into three speed subgroups according to their production latencies. Production speed modulated waveform amplitudes in the time window ranging from about 200 to 350 ms after picture presentation and the duration of a stable electrophysiological spatial configuration in the same time period. The remaining time windows from picture onset to 200 ms before articulation were unaffected by speed. By contrast, the manipulation of a psycholinguistic variable, word age-of-acquisition, modulated ERPs in all speed subgroups in a different and later time period, starting at around 400 ms after picture presentation, associated with phonological encoding processes. These results indicate that the between-subject variability in the speed of single word production is principally accounted for by the timing of a stable electrophysiological activity in the 200-350 ms time period, presumably associated with lexical selection.

1. Introduction

Speakers produce two to three words per second in connected speech, with some variability due to individual speech rate (Miller et al, 1984). Between-subjects variability in speech rate involves differences in the articulation rate and in the number and duration of pauses. Even at slow speech rates, speakers transform an abstract idea into the articulation of physical speech sounds corresponding to a single word in a few hundreds of milliseconds. Research on speech production has analysed the mental operations involved in the transformation of an idea into an articulatory plan (Garrett, 1975; Levelt, 1989). There is a general agreement between different models of speech production that the speaker encodes a pre-linguistic concept into a lexical-semantic representation leading to the selection of the appropriate word (lexical selection); then the phonological makeup of the sentence (the word form) is encoded (phonological encoding), which drives the selection of the appropriate muscle commands to start articulating. Psycholinguistic experimental investigations coupled with neuroimaging studies allowing high temporal resolution (electroencephalography, EEG and magnetoencephalography, MEG) have provided accurate estimates of the time course of these different encoding processes from concept to articulation (Indefrey and Levelt, 2004). The time course of single word production has particularly been investigated using picture naming tasks, in which speakers have to produce a word corresponding to a concept represented by a picture. In this kind of speech production task, visual and conceptual processes are estimated to take place from 0 to about 150-175 ms after picture presentation, followed by lexical-semantic (lexical selection) processes until about 275 ms. The encoding of the phonological form is thought to occur between 275 and 400-450 ms after picture onset, followed by phonetic encoding and motor execution. The timing of single word encoding has been

repeatedly confirmed in recent ERP studies, in particular regarding lexical selection and phonological encoding processes (Costa, et al., 2009; Laganaro, et al., 2009; Maess et al., 2002; Perret & Laganaro, in press; Strijkers, et al., 2010; Vihla, et al., 2006). These estimates represent an average timing across different words and different speakers. However, specific linguistic properties of the words, such as their frequency of use (Oldfield and Wingfield, 1965) or their age-of-acquisition (Morrison, et al., 1992), are known to affect the speed of word production. More importantly for our purpose here, it is also widely known that the overall processing time for identical words varies across speakers. For instance, in simple picture naming tasks production latencies can vary by a factor of two, even among subjects from a homogeneous population (i.e. undergraduate students).

Given the between-subject variability in processing speed, the estimates of the time course of encoding processes presented above represent an average including both, slow and fast speakers. Therefore, we may wonder whether differences in processing speed are distributed across all encoding processes or if only certain specific cognitive processes vary according to the speed of speech production. In other words, the question is whether all encoding processes from concept to articulation are stretched in slow speakers relative to fast speakers, or if processing speed is associated with variable encoding times for a particular process. Schuhmann et al. (2009) had to deal with the interpretation of which encoding process was associated with a specific time period in subjects with very short production latencies (during the production of a limited number of monosyllabic words): They hypothesized that speed affects all the processes involved in speech production equally. Alternatively, one may hypothesize that production speed depends on a specific encoding process, either at pre-linguistic levels (conceptualisation) or during word encoding (lexical selection, phonological or phonetic encoding). To our knowledge this question has never been addressed directly.

Here we investigated the variability in processing speed during speech production by comparing event-related electroencephalographical (ERP) correlates during picture naming in fast and slow subjects. Taking advantage of topographic (spatio-temporal) ERP analyses (Murray, et al., 2008; Michel et al., 2009), we examined the duration of specific electrophysiological patterns (functional microstates, Lehmann, 1987; Michel et al., 2009) across slow and fast speakers and their correlation with production latencies. If speed of word production is distributed along all the speech encoding processes as hypothesised by Schuhmann et al. (2009), then differences between slow and fast speakers should be observed in several time windows from the moment a picture appears on the screen to articulation. On the other hand, if differences in processing speed are linked to a specific encoding process, then ERP divergences between slow and fast speakers should be limited to a given time window, which can be associated to a specific encoding process. As an additional comparison point to index specific encoding processes we manipulated a psycholinguistic variable known to reliably affect production latencies, namely word age-of-acquisition (AoA hereafter). Effects of AoA on production latencies have been repeatedly reported in picture naming paradigms independently of other psycholinguistic variables (Alario et al., 2004; Bonin, et al., 2002; Chalard, et al., 2003; Cuetos et al., 1999; Morrison and Ellis, 1995) and of production speed (Morrison et al., 2002). In addition, there is converging evidence from psycholinguistic (Chalard and Bonin, 2006; Morrison, et al., 1992), neuropsychological (Kittredge, et al., 2008) and ERP investigation (Laganaro and Perret, 2011) in favour of a lexical-phonological locus of the AoA effect. The double comparison of the time period modulated by speed with (1) the estimates of timing of speech encoding processes issued from previous studies, and (2) the time window affected by AoA, will enable us to conclude as to whether a specific encoding process accounts for the differences in production speed, or if variations in processing speed are distributed along several/all word encoding processes.

2. Material and methods

<u>2.1 Subjects</u>

45 undergraduate students (8 men) participated in the study. They were all native French speakers, aged 18-35 (mean= 24.06). All were right-handed as determined by the Edinburgh Handedness Scale (Oldfield 1971). The participants gave their informed consent and were paid for their participation.

The 45 subjects were divided in three subgroups of 15 subjects each, according to their mean production latencies (*slow*-, *mean*- and *fast*-speed subgroups, see behavioral results). There was no significant difference in age between the three subgroups ($F<1$): the *slow* subgroup (N=15, 4 men) had a mean age of 23.3 (s.d.=3.52), the *mean*-speed subgroup (N=15, 2 men) had a mean age of 25.6 (s.d.=5,18), and the *fast* subgroup (N=15, 2 men) had a mean age of 24.1 (s.d.=4,39)..

<u>2.2 Material</u>

120 words and their corresponding pictures were selected from two French databases (Alario & Ferrand, 1999; Bonin et al, 2003). All picture-words had a high name agreement. 60 stimuli were early-acquired words (EAW) and the other half were late-acquired words (LAW). Early- and late-acquired words were matched on the first phoneme (Kessler et al., 2002) and on length. In addition, the following psycholinguistic variables were balanced across AoA conditions (see Table 1): name agreement, image agreement, conceptual familiarity, visual complexity (from the mentioned databases), lexical frequency and syllable frequency (from Frantext, New et al., 2004).

[Table 1 about here]

2.3 Procedure

Participants were tested individually in a soundproof dark room. They sat 60 cm in front of the screen. The presentation of trials was controlled by the E-Prime software (E-Studio). Pictures were presented in constant size of 9.5*9.5cm (approximately 4.52 degree of visual angle) on a grey screen.

Before the experiment, participants were familiarized with all the pictures and their corresponding names on a paper sheet. An experimental trial had the following structure: first, a "+" sign was presented for 500 ms. Then, the picture appeared on the screen for 2000 ms. The participants had to produce overtly the word corresponding to the picture. A blank screen lasting 2000 ms was displayed before the next trial. All items were presented in a pseudo-random order, preceded by 4 warming-up filler trials. The experiment lasted about 15 minutes with a break after 60 stimuli. Production latencies were measured by means of a voice key and productions were digitized for further systematic latency and accuracy check with a speech analysis software (see behavioral analyses).

2.4 EEG acquisition and pre-analyses

EEG was recorded continuously using the Active-Two Biosemi EEG system (Biosemi V.O.F. Amsterdam, Netherlands) with 128 channels covering the entire scalp. Signals were sampled at 512 Hz with band-pass filters set between 0.16 and 100 Hz.

Stimulus-aligned (forward) epochs of 450 ms and response-aligned (backward) epochs of 450 ms were averaged across conditions. Response-aligned epochs lasted from 550 to 100 ms before the production latency of each individual trial; stimulus-aligned epochs started at the moment the picture appeared on the screen.

In addition to an automated selection criterion rejecting epochs with amplitudes reaching ±100 µV, each trial was visually inspected, and epochs contaminated by eye blinking, movements or other noise were rejected and excluded from averaging. In addition, only trials with both response-aligned and its corresponding stimulus-aligned uncontaminated epochs were retained, i.e. a minimum of 76 averaged trials per subject.

ERPs were then bandpass-filtered to 0.2–30 Hz and recalculated against the average reference.

For the spatio-temporal segmentation analysis (see below) the stimulus-aligned and response-aligned data from each subject were merged according to each individual subject's RT for the actual averaged trials. In other words, the overlapping signal between the forward stimulus-aligned and the backward response-aligned ERP data was removed. The combination of stimulus- and response-aligned data was introduced by Laganaro and Perret (2011): it allows the individual averaged data (and the group grand-average) to cover the actual time form picture onset to 100 ms before articulation.

2.5 Behavioural analyses

After elimination of errors, production latencies (RT - reaction time - hereafter, i.e. ms separating the onset of the picture and articulation onset) were systematically checked with a speech analysis software (Praat, Boersma and Weenik, 2000), thanks to an inaudible acoustic click at the onset of the picture recorded on the second track of the recording system. Participants were then assigned to three subgroups of 15 subjects each, according to their mean RT (*fast*, *mean* and *slow* subgroups).

RT data were fitted with a linear regression mixed model (Baayen, et al, 2008) with the R-software (R-project, R-development core team 2005; Bates and Sarkar, 2007). Production speed (*fast*, *mean* and *slow*) and AoA (early- and late-acquired words) conditions were

included in the mixed model as fixed effect variables and participants and items as random effect variables.

2.6 ERP analyses

The ERPs were first subjected to waveform analysis to determine the time periods of differences in amplitudes between groups. Then spatio-temporal segmentation analyses were performed on the grand-averages from each condition and statistically tested in the single subjects' data as described below.

2.6.1 Waveform analyses

Waveform analysis was carried out in the following way: ANOVAs were computed on amplitudes of the evoked potentials at each electrode and time point (every 2 ms) over the whole period with the AoA condition as a within subjects factor and speed as between-subjects factor. To correct for multiple comparisons, only differences over at least 5 electrodes from the same region out of 6 scalp regions (left and right anterior, central, posterior) extending over at least 10 ms were retained with a conservative alpha criterion of 0.01.

2.6.2 Topographic pattern analysis (spatio-temporal segmentation)

The second analysis was a topographic (map) pattern analysis. This method is independent of the reference electrode (Michel et al., 2001, 2004) and insensitive to pure amplitude modulations across conditions (topographies of normalized maps are compared). A modified hierarchical clustering analysis (Michel et al., 2001; Pascual-Marqui, et al., 1995), the agglomerative hierarchical clustering (Murray, et al., 2008) was used to determine the most dominant configurations of the electric field at the scalp (topographic maps). A modified version of the cross validation criterion combining a cross-validation criterion and the

Krzanovski-Lai criterion (see Murray et al., 2008) was used to determine the optimal number of maps that explained the best the grand-average data sets across conditions. Statistical smoothing was used to eliminate temporally isolated topographic maps with low strength. This procedure is described in detail in Pascual-Marqui et al. (1995, see also Brunet et al., 2011; Murray et al., 2008). Additionally, a given topography had to be present for at least 10 time frames (20 ms). We first applied a spatio-temporal segmentation on the six grand-average data from each subgroup (*fast*, *mean*, *slow*) and each AoA condition (early- and late-acquired words). Then, the pattern of map templates observed in the averaged data was statistically tested by comparing each of these map templates with the moment-by-moment scalp topography of individual subjects' ERPs. Each time point was labelled according to the map with which it best correlated spatially, yielding a measure of map presence. This procedure referred to as 'fitting' allowed to establish how well a cluster map explained individual patterns of activity (GEV: Global Explained Variance) and its duration. These analyses were performed using the Cartool software (Brunet et al., 2011).

In order to analyse whether one map is more representative of one group or whether it lasts longer in one condition, GEV and durational measures observed in each subject's data were used for statistical analysis. Analyses of variance were applied to these measures with subjects as random variable and conditions (production speed and AoA) as fixed factors.

3. Results

3.1 <u>Behavioural Results (RTs)</u>

Production latencies above 1400 ms and below 400 ms as well as errors were removed from the analysis (7.6% of the data). Mean RT on the whole group of 45 subjects was 818 ms (SD = 95.67 ms). The 45 subjects were divided into 3 subgroups of 15 subjects each according to their mean production speed (see Table 2).

Early-acquired words were produced 59 ms faster than late-acquired words (F(1, 5145) = 23.23, MS= 452425, *p* <.0001) without interaction with speed subgroups F(1, 5145) = 3.23, MS= 48699, *p* =.145). Differences in RT were significant across all speed subgroups (*mean vs. fast*: t(4975)= -4.04, p<.0001; *mean* vs. *slow*: t(4975)=6.93, p<.0001).

3.2 ERPs

Figure 1 displays the statistical comparison of amplitudes between conditions. Amplitudes differed between *speed* subgroups on a very few electrodes around 100 ms and consistently in the time period ranging from around 200-220 ms to 350 ms on posterior left and anterior left and right electrodes (Figure 1.A). Amplitudes diverged between AoA conditions in a very different time period, from 380-400 ms after picture presentation to 200 ms before articulation (Figure 1B.). No interaction appeared between speed subgroups and AoA, except on very sparse electrodes in the time period preceding articulation (Figure 1C.).
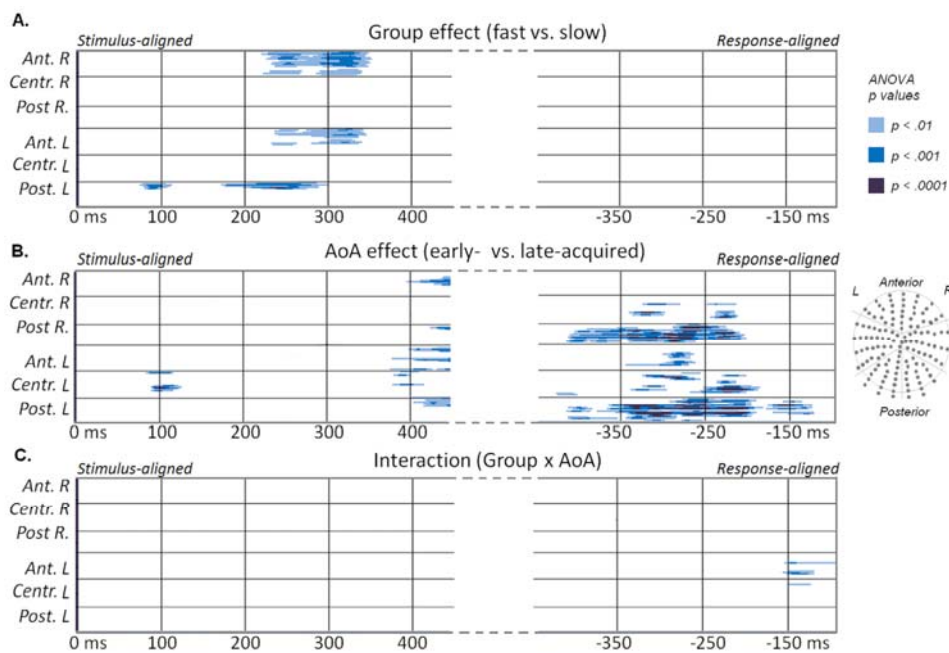
The spatio-temporal segmentation applied to the 6 grand-averages from 50 ms after picture presentation to 100 ms before articulation revealed 6 different topographies accounting for 92.41% of the variance (see Figure 2). The following fitting periods were applied to verify presence and duration of maps in the individual data: from 50 to 200 ms, from 200 to the 450 ms and from 450 to 100 ms before each individual RT in each condition. To account for between-subjects variability, map templates crossing the fitting borders were also entered in each fitting period (maps "A", "B", "C" in the first fitting period, "B", "C", "D", "E" in the second period and "D", "E", "F" in the last period).

Maps labelled "A" and "B" did not differ between conditions (all Fs<1). The stable topographic map "C" differed between speed subgroups on duration ($F_{(2, 42)}$ = 13.65, p<.0001) and on GEV($F_{(2, 42)}$ = 7.24, p<.01); no difference appeared between AoA conditions and no interaction between speed subgroups and AoA (both Fs<1). The mean duration of the stable electrophysiological activity "C" was 74 ms for the *fast*, 85 ms for the *mean*, and 157 for the *slow* subgroups. Planned comparisons indicated significant differences between the *slow* subgroup and each of the two subgroups (p<.0001 on duration and p<.01 on GEV), and no difference between *fast* and *mean*-speed subgroups (p=.55). The longer duration of map "C" was confirmed by a difference in the onset of the next stable spatial configuration: the onset of Map "D" (the first time frame of presence of map "D") significantly differed across speed groups ($F_{(2, 42)}$ = 13.80, p<.0001), not across AoA conditions ($F_{(1, 44)}$ = 1.4, p=.24; interaction: F<1). The stable topographic map "D", had similar duration between speed subgroups ($F_{(2,42)}$=1.7, p=.19), but displayed significantly longer duration for *late-acquired* relative to *early-acquired* words ($F_{(1, 44)}$=3.7, p=.05), with no interaction between AoA and speed (F<1) and no significant differences on GEV (Fs<1.2,

ps>.29). The mean difference in duration of map "D" between early- and late-acquired words was 41 ms.

Finally, there was no significant difference on maps "E" (all F<1) and "F" (speed subgroups: F(2,42)=2.3, p=.1; AoA: F(1,44)=1.35, p=.25; interaction: F<1).

[Fig. 2 about here]
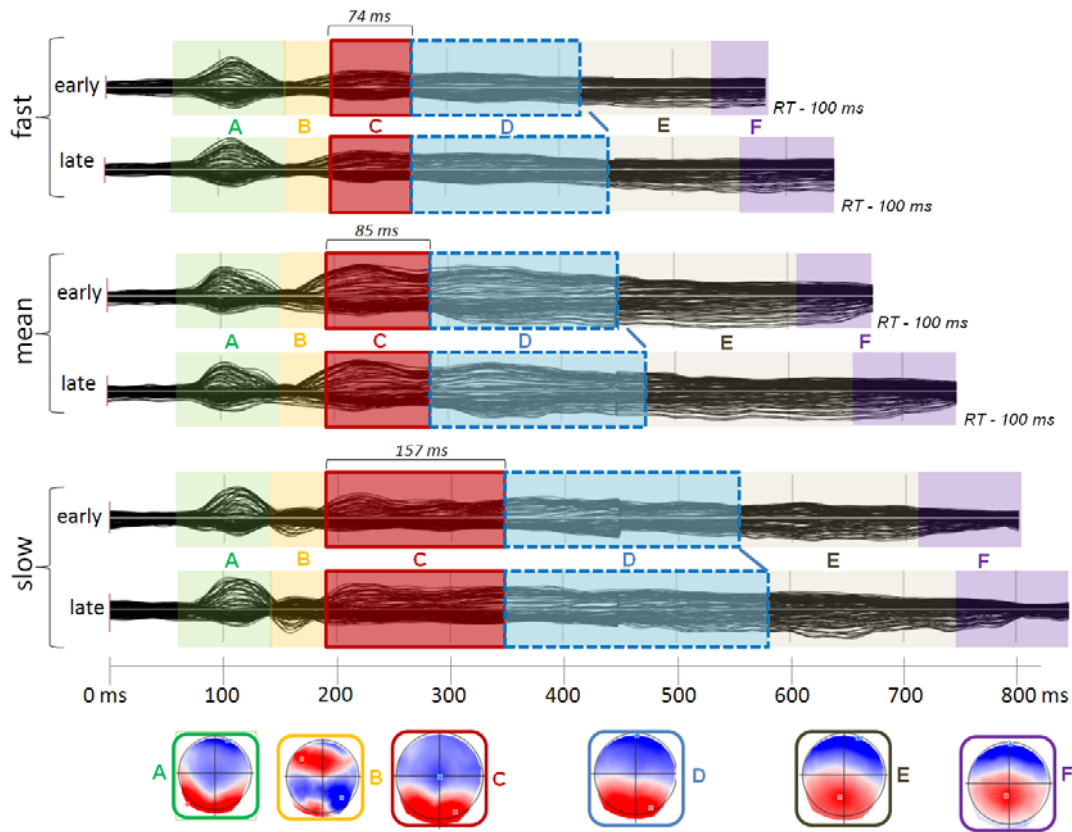
3.3 Analyses limited to two groups of fast and slow subjects

In order to analyze with increased power the effect of speed and its distribution across different periods of stable electrophysiological patterns, we restricted the analyses to two data sets by comparing only two groups of subjects across early and late acquired words. The subgroups included the 20 fastest (RT= 730 ms, SD=42.54) and the 20 slowest subjects (RT=907 ms, SD=54.62). The waveform analysis confirmed a significant effect of speed on amplitudes in the time window from 200 to 330 ms after picture onset on anterior right and

13

central-posterior left sites (Figure 3A.). This time-window corresponds to a N2-N3 like peak on electrodes from frontal regions at scalp (Fz in Figure 3A) and to a P2 like peak on posterior electrodes (POz in Figure 3A). An additional time period of diverging amplitudes between the *slowest* and the *fastest* subjects appeared in the last 80 ms. The spatio-temporal cluster analysis applied to the two grand-average data from 50 ms after picture onset to 100 ms before RT disclosed 6 different topographic maps accounting for 95.04% of the data. The duration of each stable map template in the two subgroups is displayed in Figure 3B. The fitting in the individuals revealed significant differences between the fastest and slowest speakers on duration and GEV for the scalp map "C" (duration: $t(38) = 3.78$, $p<.001$; GEV: $t(38) = 3.14$, $p<.01$) and on duration of map "F" (duration: $t(38)=2.19$, $p<.05$; GEV: $t<1$). All other differences on map duration and GEV were not significant (map "B": $t(38) = 1.75$, $p =.09$ on duration; all other $ts<1$).

Correlations calculated between individual RTs and duration (number of time frames) of stable scalp topographies in the subjects' data indicated a significant positive correlation between production latencies and duration of the scalp map "C" ($r(40) = .446$, $p<.01$) and a marginal correlation with scalp map "F" ($r(40) = .289$, $p = .066$). All other correlations between map duration and RTs were not significant ($r< .238$, $ps > .13$).

[Figure 3 about here]

A.

Stimulus-aligned — RT-aligned

B.
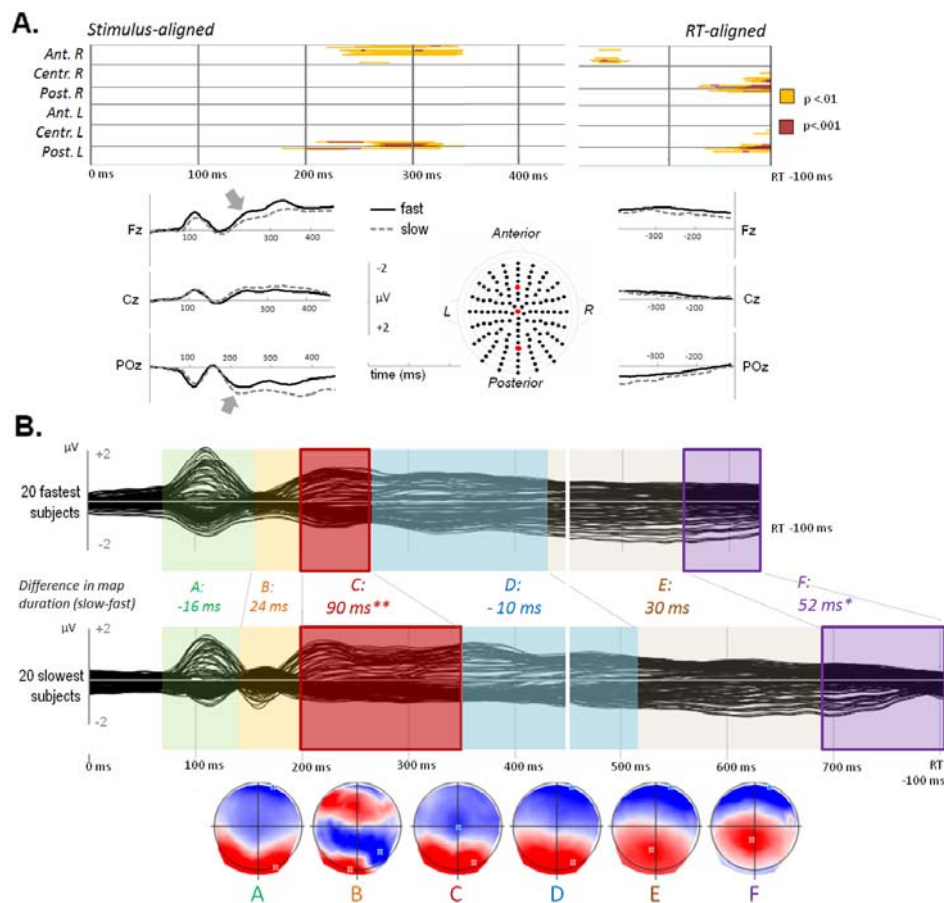
4. Discussion

We investigated whether the between-subject variability in processing speed is distributed along all encoding processes or if it is due to specific processing stages involved in speech production. Converging results from the waveform analysis and the topographic pattern analysis point to an effect of production speed in the time window ranging from 200 to 350 ms after picture presentation. By contrast, the manipulated psycholinguistic word property (AoA) modulated ERPs in a different and later time window. No further significant differences were observed across speed subgroups from picture onset to 200 ms before articulation: only the very last analysed period (from about 200 ms to 100 ms before articulation) was also modulated by speed. Also importantly, the only significant correlation between duration of stable electrophysiological activities and individual RTs appeared on duration of the scalp map starting around 190 ms.

15

The main consequence of these results is that differences across slow and fast speakers are not equally distributed along all encoding processes. The principal effect of speed falls within a time period which has been associated with conceptually driven lexical selection in previous research. Investigations on the time course of processes underlying single word production have estimated the beginning of lexical selection at around 180-200 ms (Indefrey and Levelt, 2004, see the Introduction). Although this timing is founded on rather indirect evidence from behavioural studies and from ERP paradigms using mainly metalinguistic tasks, it has been repeatedly validated in recent ERP studies using overt production paradigms. In particular, several ERP investigations using different approaches (semantic interference effects: Aristei et al., 2011; Costa et al., 2009; Maess et al., 2009; lexical frequency and between-languages cognates effects: Strijkers et al., 2010) have converged in suggesting that lexical selection is engaged around 200 ms after picture presentation (although other studies suggested that conceptual-semantic processes can extend this time period, Schendan & Maher , 2006; Schendan & Kutas, 2003; Sitikiva et al., 2006). The effects attributed to lexical selection were observed in a P200 peak range on posterior electrodes (Aristei et al., 2011, with average reference; Costa et al., 2009 and Strijkers et al., 2010 with nose reference) and on a N300 like peak on central and anterior sites (Strijkers et al., 2010). The exploration of waveforms on single electrodes in the present study (Figure 3) indicated that the effect of speed fall within P2 like and N2 like peaks at posterior and anterior scalp sites respectively.

Crucially, production speed affected ERPs in a different and earlier time window relative to the effect of AoA. Behavioural studies analysing production latencies in healthy subjects (Chalard and Bonin, 2006; Morrison et al., 1992) and on accuracy in aphasic patients (Kittredge et al., 2008) pointed to an effect of word AoA arising during phonological

encoding (but see Belke et al., 2005; Johnson and Barry, 2005 for alternative results favouring a lexical-semantic locus of AoA effects). The time window displaying the AoA effect has been associated with the encoding of word forms (phonological encoding, from approximately 280 to 450 ms) by Indefrey and Levelt (2004). The approximate onset of phonological encoding processes has also received support from ERP investigations using different paradigms. Vihla et al. (2006) showed that ERPs differed between tasks involving phonological encoding and semantic categorisation tasks at around 300 ms after picture presentation. Laganaro et al. (2009; 2011) showed that ERPs diverged in brain-damaged (aphasic) speakers with impaired phonological encoding relative to healthy controls at about 300 ms after picture presentation. In a comparison between spoken and written picture naming Perret and Laganaro (in press) observed ERP divergences on amplitudes and modality specific topographical configurations from around 260 ms. Although direct verification of the time window associated with phonological encoding and its duration has not been provided so far, these investigations converge in suggesting that phonological encoding is engaged around 260-300 ms after picture presentation. Also, the time window displaying AoA effects in the present study is congruent with the effect reported in a previous ERP investigation using different stimuli and a smaller sample (Laganaro and Perret, 2011). Taken together, they reliably indicate that longer encoding times for late-acquired words fall within the time window associated with the encoding of word forms.

More importantly for the aim of the present study, both AoA and production speed had an independent effect on RTs and modulated amplitudes and duration of stable electrophysiological patterns in two different time windows. The time-window of the main effect of speed was observed on a scalp topography that preceded the one which was modulated by AoA. We can therefore conclude that the time period preceding phonological

encoding, presumably engaging lexical selection, yields the largest difference between *fast* and slow speakers.

The very last analysed period (extending from about 200 to 100 ms before articulation) also yielded a significant portion (50 ms) of the difference between fast and slow speakers. Although there is not enough insight from previous studies using high temporal resolution paradigms to infer about the processes involved during the very last hundreds of milliseconds, pre-motor control processes are certainly engaged during the period preceding the vocal onset. Electromyographic (EMG) studies measuring the neuromuscular activity during speech production have reported inter-individual variability on different parameters of pre-vocal EMG (Gracco, 1988). For instance, laryngeal EMG onset varied from 185 to 55 ms before vocal onset in a simple vocal reaction time study in slow versus fast subjects (Shipp, Izdebski & Morrissey, 1984). Therefore, the difference in duration of the stable electrophysiological map "F" might be due to different motor control strategies between fast and slow speakers, but speech motor planning (phonetic encoding) may also be involved during this very last period.

Taken together, the 90 ms difference in duration of the stable brain activity starting around 200 ms (map "C" in Figure 3) and the difference observed on the last stable brain activity (52 ms for map "F") match the largest part of the 170 ms difference in RTs between fast and slow speakers. Only the duration of map "C" had a significant but partial correlation with RTs. This suggests that the largest part of variation in production speed is accounted for by processes taking place in the 200-350 ms time period and to a lesser extent in the last 200 ms preceding articulation. The association of longer lasting stable brain activity in the 200-300 ms time window with longer lasting pre-articulatory activity might indicate some dependence

between lexical selection and motor control/programming: further studies will need to verify and investigate the nature of this association.

The remaining residual difference between speed subgroups is distributed among other time periods, without reaching significance in the present ERP analyses. One possible reason why these smaller differences are not captured lies in the statistical smoothing applied on time periods of 20 ms to eliminate periods of topographical instability in the transition between two stable electrophysiological activities (Lehmann, 1987).

The question then is why is the between-subject variability not equally distributed along all encoding processes?

A careful comparison of previously published results indicate that there is virtually no variation in the time period yielding the effect of the manipulated variables thought to affect conceptually driven lexical selection across a number of studies, despite variations in mean RT, number of items and amount of item repetition. For instance, mean RTs were around 820 ms for the fastest condition in the study by Costa et al. (2009), 750 ms in Aristei et al. (2011) and 700 ms in Strijkers et al. (2010), but all these investigations reported effects attributed to lexical selection at around 200 ms after picture onset.

There is much less evidence about the duration of lexical selection. Among the previously mentioned studies, only Costa et al. (2009) assessed the duration of the time-period affected by the manipulation aimed at indexing lexical selection (semantic interference). Their estimate (around 180 ms) exceeded the 75 ms suggested by Indefrey and Levelt (2004). The authors proposed two alternative interpretations to this longer timing: the semantic interference effect might exceed lexical selection processes and affect by cascading subsequent encoding processes; alternatively lexical selection may last longer in their study due to an increased competition between lexical representations. According to this second

interpretation, lexical selection can be stretched under certain circumstances, which may also be the case for *slow* speakers in the present study.

Finally, a crucial question is why did production speed not modulate the time window associated with phonological encoding, particularly in view of the fact that word properties (AoA) did affect ERPs in this later time period. We can attempt an interpretation in terms of the automaticity of encoding processes. Although all processes involved in picture naming are highly automatic, word form encoding is probably more automatic than lexical selection. Ferreira and Pashler (2002) provided experimental evidence in favor of this hypothesis: they showed that the stage of lexical selection interferes with a concurrent dual-task, whereas phonological encoding does not affect it. A similar interpretation in terms of cognitive control during the stage of lexical selection has been suggested by Hartsuiker and Notebaert (2010) for the number of pauses observed during a modified multiple pictures naming paradigm. In other words, lexical selection may require more attentional control than other word encoding processes and therefore be more dependent on between-subjects variability. By contrast, word-form encoding speed is related to the strength of connections between lexical and phonological codes, which is more dependent on word properties such as age-of-acquisition.

Under the attentional control hypothesis, *slow* speakers are those who undergo a higher cognitive cost in the control necessary for lexical selection. However, we cannot exclude that other encoding processes are also engaged during the main time-period affected by speed. It is widely acknowledged that speech encoding processing involves some degree of interactivity, either in terms of cascading activation or feed-back activation between phonological and lexical representations for instance (Dell, 1986; see Goldrick, 2006 for a review). The observation that distinct periods of stable electrophysiological spatial configuration are affected by different variables (here production speed and word AoA) suggests that these effects arise while different underlying brain generators are engaged, it does not exclude per

se some influence of other processes during these time periods. So far these accounts are compatible with both, serial models of language encoding (Levelt et at., 1999) and connectionist models accounting for sequential lexical and phonological processing (e.g. Foygel and Dell, 2000). By contrast, the present data are clearly not compatible with interpretation of timing obtained by applying proportional differences to all encoding processes relative to the mean estimates in studies involving shorter or longer production latencies. Further research will need to address wether the present results are specific to the context of picture naming and to variability within a homogeneous population or if they generalise to other populations and to other language production tasks.

Conclusion

The present investigation showed that production speed and word age of acquisition modulated ERPs in two distinct time windows. The largest part of between-subject variability in the speed of speech production is accounted for by the timing of a stable electrophysiological activity observed in the time period presumably associated with lexical selection. The duration of the stable electrophysiological activity starting around 200 ms varied with speed, shifting the onset of the following stable brain activities. As a consequence, any interpretation of RT data that diverges significantly from mean speed estimates cannot rely on to the timing of lexical selection and of the following encoding processes made in current models of speech production.

References

Alario, FX. Ferrand, L. 1999. A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. Behav Res Methods Instruments Computers 31, 531-552.

Alario, FX., Ferrand, L., Laganaro, M., New, B., Frauenfelder, UH. Segui, J. 2004. Predictors of picture naming speed. Behav Res Methods Instruments Computers 36,140-155.

Aristei, S., Melinger, A., Abdel Rahman, R. 2011. Electrophysiological chronometry of semantic context effects in language production. J Cogn Neurosci 23, 1567–1586.

Baayen, RH., Davidson, DJ., Bates, DM. 2008. Mixed-effects modeling with crossed random effects for subjects and items. J Mem Lang 59, 390-412.

Bates, DM. Sarkar, D. 2007. Lmer4: Linear mixed-effects models using S4 classes, R package version 0.99875-6.

Belke, E., Brysbaert, M., Meyer, AS. Ghyselinck, M. 2005. Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. Cognition 96, B45-B54.

Boersma, P. Weenik, D. 2007. Praat: doing phonetics by computer. Institute of Phonetic Sciences of the University of Amsterdam. http://www.praat.org.

Bonin, P., Chalard, M., Méot, A. Fayol, M. 2002. The determinants of spoken and written picture naming latencies. Br J Psychol 93, 89-114.

Bonin, P., Peerman, R., Malardier, N., Méot, A. Chalard, M. 2003. A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. Behav Res Methods Instruments Computers 35, 158-157.

Brunet, D., Murray, MM. Michel, CM. 2011. Spatio-temporal analysis of multichannel EEG: CARTOOL. Computational Intelligence and Neuroscience, in press.

Chalard, M., Bonin, P., Méot, A, Boyer, B. Fayol, M. 2003. Objective age-of-acquisition AoA norms for a set of 230 objects names in French: Relationships with psycholinguistic variables, the English data from Morrison et al. 1997, and naming latencies. European J Cogn Psychol 15, 209-245.

Chalard, M., Bonin, P. 2006. Age-of-acquisition effects in picture naming: Are they structural and/or semantic in nature? Visual Cognition 137-8, 864-883

Costa, A., Strijkers, K., Martin, C. Thierry, G. 2009. The time course of word retrieval revealed by event-related brain potentials during overt speech. Proc Natl Acad Sci, 106, 21442-6.

Cuetos, F., Ellis, AW., Alvarez, B. 1999. Naming times for the Snodgrass and Vanderwart pictures in Spanish. Behav Res Methods Instruments Computers 31, 650-658.

Dell, G.S. 1986. A spreading activation theory of retrieval in sentence production. Psych Rev 93, 283-321.

Ferreira, V. S., Pashler, H. 2002. Central bottleneck influences on the processing stages of word production. J Exp Psychol Learn Mem Cognition 28, 1187-1199.

Foygel, D.. Dell, G.S. 2000. Models of impaired lexical access in speech production. J Mem Lang 43, 182–216.

Garrett, M.F. 1975. The analysis of sentence production. In G.H. Bower Ed., The psychology of learning and motivation, Vol. 9. New York: Academic Press.

Goldrick, M. 2006. Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. Lang Cogn Proc 21, 817-855

Gracco, VL 1988. Timing factors in the coordination of speech movements. J Neurosci 8, 4628-4639.

Hartsuiker, R.J., Notebaert, L. 2010. Lexical access problems lead to disfluencies in speech. Exp Psychol. 57, 169-177.

Indefrey, P., Levelt, W. 2004. The spatial and temporal signatures of word production components. Cognition 92, 101-144

Johnston, R.A., Barry, C. 2005. Age of Acquisition effects in the semantic processing of pictures. Mem Cognition 33, 905-912.

Kessler, B., Treiman, R., & Mullenix, J. 2002. Phonetic biases in voice key response time measurements. J Mem Lang 47, 145-171.

Kittredge, AK., Dell, GS., Verkuilen, J. Schwartz, MF. 2008. Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. Cogn Neuropsychology 25, 463-492.

Laganaro, M., Morand, S. Schnider, A. 2009. Time course of evoked-potential changes in different forms of anomia in aphasia. J Cogn Neurosci 21, 1499-1510.

Laganaro, M. Perret, C. 2011. Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. Brain Topogr 24, 19-29.

Lehmann, D. 1987. Principles of spatial analysis. In: AS. Gevins, A. Remond eds Handbook of electroencephalography and clinical Neurophysiology. Vol. 1: Methods of Analysis of Brain Electrical and Magnetic Signals. Elsevier, Amsterdam. 309-354.

Levelt, W. 1989. Speaking: from intention to articulation. Cambridge, Mass: MIT Press.

Levelt, W., Roelofs, A., Meyer, A.S. 1999. A theory of lexical access in speech production. Behav Brain Sci 22, 1-75.

Maess, B., Friederici, A.D., Damian, M., Meyer, A.S. Levelt, W.J.M. 2002. Semantic Category Interference in Overt Picture Naming: Sharpening Current Density Localization by PCA. J Cogn Neurosci 14, 455-462.

Michel, C.M., Thut, G., Morand, S., Khateb, A., Pegna, A.J., Grave de Peralta, R., 2001. Electric source imaging of human brain functions. Brain Res Review. 36, 108–118

Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R. 2004. EEG source imaging. Clin Neurophys 115, 2195-2222.

Michel, CM., Koenig, T., Brandeis, D., Gianotti, L.R.R. 2009. Electric neuroimaging. Cambridge: Cambridge University Press.

Miller, J. L., Grosjean, F., Lomanto, C. 1984. Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. Phonetica 41, 215–225.

Morrison, CM., Ellis, AW. Quinlam, PT. 1992. Age of acquisition, not word frequency, affects object naming, not object recognition. Mem Cognition 20, 705-714.

Morrison, CM. Ellis, AW. 1995. Roles of word frequency and age of acquisition in word naming and lexical decision. J Exp Psychol: Learn Mem Cognition 21,116-133

Morrison, C.M., Hirsh, K.W., Chappell, T.D., Ellis, A.W. 2002. Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. The European J Cogn Psychol 14, 435-459.

Murray, MM., Brunet, D. Michel, C. 2008. Topographic ERP analyses, a step-by-step tutorial review. Brain Topogr 20, 249-269.

New, B., Pallier, C., Brysbaert, M. Ferrand, L. 2004. Lexique2: A new French lexical database. Behav Res Methods Instruments Computers 36, 516-524.

Oldfield, RC. 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neurospychologia 9, 97-113.

Oldfield, RC. Wingfield, A. 1965. Response latencies in naming objects. Quarterly J Exp Psychol 17, 273-281.

Pascual-Marqui, RD., Michel, CM. Lehmann D 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Transactions on Biomedical Engineering 42 658–665.

Perret. C., Laganaro, M. Comparison of electrophysiological correlates of writing and speaking: a topographic ERP analysis. Brain Topogr in press

Schendan HE, Kutas M. 2003. Time course of processes and representations supporting visual object identification and memory. J Cogn Neurosci 15, 111-135.

Schendan, H.E., Maher, S.M. 2006. Object knowledge during entry-level categorization is activated and modified by implicit memory after 200 ms. NeuroImage, 44, 1423-38

Sitnikova, T., West, W.C., Kuperberg, G.R., Holcomb, P.J. 2006. The neural organization of semantic memory: Electrophysiological activity suggests feature-based segregation. Biol Psychol 71, 326-340

Shipp, T., Izdebski, K., Morrissey, P. 1984. Physiological stages of vocal reaction time. J Speech Hear Res 27, 173-178.

Schuhmann, T., Schiller, NO., Goebel, R. Sack, AT. 2009. The temporal characteristics of functional activation in Broca's area during overt picture naming. Cortex 45, 1111-6.

Strijkers, K., Costa, A. Thierry, G. 2010. Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. Cereb Cortex 20, 912-928.

Vihla, M., Laine, M., Salmelin, R. 2006. Cortical dynamics of visual/semantic vs. Phonological analysis in picture naming. NeuroImage 33, 732-738.

TABLES

Table 1. Properties of the 120 words and corresponding pictures

	AoA	*h*NA	IA	Fam	VC	LexFreq	SyllFreq
Early-Acquired (EAW)	**1.86**	.15	3.54	3.02	2.94	21.14	2531.69
Late-Acquired (LAW)	**2.67**	.19	3.73	2.96	3.08	14.78	2201.22
p-value	***<.0001***	*.3796*	*.1401*	*.6993*	*.3946*	*.2503*	*.6539*

AoA: adult rated Age of Acquisition measures on a 5-points scale (1=learned at 0-3 years, 5=learned after 12); *h*NA: Name Agreement h-statistic; IA: Image Agreement; Fam: Conceptual Familiarity; VC: Visual Complexity; LexFreq.: Lexical frequency; SyllFreq: mean syllable frequency.

Table 2. Behavioural results

	fast	*mean*	*slow*
N	15	15	15
RT	714	815	926
EA words	*683*	*780*	*902*
LA words	*744*	*851*	*950*

Figures

Figure 1. Significant differences (ANOVA's p values) on ERP waveform amplitude on each electrode (Y axes) and time point (X axes) A. between speed subgroups (*slow* and *fast*), B. between AoA conditions, C. in the interaction between speed and AoA, with the arrangement of the 128 electrodes.

Figure 2. Grand-average ERPs (128 electrodes) from each speed subgroup and each AoA condition from onset to 100 ms before RT and temporal distribution of the topographic maps revealed by the spatio-temporal segmentation analysis in each data. Bottom: map templates for the six stable topographies observed from 50 ms after picture presentation to 100 ms before articulation (positive values in red and negative values in blue with display of maximal and minimal scalp field potentials).

Figure 3. A. Significant differences (p values) on ERP waveform amplitude on each electrode (Y axes) and time point (X axes) between the two speed subgroups and averaged ERP waveforms for the fast and the slow subgroups with the arrangement and electrode position of the displayed waveforms (Cz, Fz and POz).
B. Grand-average ERPs (128 electrodes, from onset to 100 ms before RT) from the 20 fastest and 20 slowest subjects and temporal distribution of the stable scalp topographies revealed by the spatio-temporal segmentation analysis. Bottom: map templates for the six stable topographies observed from 50 ms after picture presentation to 100 ms before articulation (positive values in red and negative values in blue with display of maximal and minimal scalp field potentials).