NeuroImage 59 (2012) 3881–3888

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

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

Marina Laganaro [a,*], Andrea Valente [a], Cyril Perret [b]

[a] FAPSE, University of Geneva, Switzerland
[b] Université Paris 13, Sorbonne Paris Cité, UTRPP EA 4403, France

ARTICLE INFO

Article history:
Received 15 June 2011
Revised 28 September 2011
Accepted 20 October 2011
Available online 4 November 2011

Keywords:
ERP
Language production
Picture naming
Processing speed
Variability

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 analysed 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.

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 couple of hundreds of milliseconds. Research on speech production has analysed the specific cognitive processes 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 and Laganaro, 2012; 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

* Corresponding author at: FAPSE, University of Geneva, 40, Bd Pont d'Arve, CH-1211 Geneva 4, Switzerland.
E-mail address: marina.laganaro@unige.ch (M. Laganaro).

1053-8119/$ – see front matter
doi:10.1016/j.neuroimage.2011.10.082

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.

Material and methods

Subjects

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 behavioural 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).

Material

120 words and their corresponding pictures were selected from two French databases (Alario and 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).

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∗9.5 cm (approximately 4.52° 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 min 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 behavioural analyses).

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.

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, Frantext, New et al., 2004.