Dynamics of phonological–phonetic encoding in word production: Evidence from diverging ERPs between stroke patients and controls

Laganaro, Marina; Python, Grégoire; Toepel, Ulrike

How to cite

LAGANARO, Marina, PYTHON, Grégoire, TOEPEL, Ulrike. Dynamics of phonological–phonetic encoding in word production: Evidence from diverging ERPs between stroke patients and controls. In: Brain and language, 2013, vol. 126, n° 2, p. 123–132. doi: 10.1016/j.bandl.2013.03.004

This publication URL: https://archive-ouverte.unige.ch/unige:89189
Publication DOI: 10.1016/j.bandl.2013.03.004

Laganaro, M., Python, G., & Toepel, U. (2013). Dynamics of phonological-phonetic encoding in word production: Evidence from diverging ERPs between stroke patients and controls. *Brain and Language, 126,* 123-132

Dynamics of phonological-phonetic encoding in word production: Evidence from diverging ERPs between stroke patients and controls

Marina Laganaro[1], Grégoire Python[2], Ulrike Toepel[2]

1. FAPSE, University of Geneva, Switzerland

2. Neurorehabilitation Unit, Department of Clinical Neurosciences, Vaudois University Hospital Center and University of Lausanne, Lausanne, Switzerland

Corresponding Author:

Marina Laganaro
FAPSE, University of Geneva
40, Bd Pont d'Arve,
CH-1211 Geneva 4
E-mail: marina.laganaro@unige.ch

Keywords: ERP, speech production, picture naming, phonological-phonetic, aphasia

Tables: 1

Figures: 4

N. of words: 8'001

Abstract

While the dynamics of lexical-semantic and lexical-phonological encoding in word production have been investigated in several event-related potential (ERP) studies, the estimated time course of phonological-phonetic encoding is the result of rather indirect evidence. We investigated the dynamics of phonological-phonetic encoding combining ERP analyses covering the entire encoding process in picture naming and word reading tasks by comparing ERP modulations in eight brain-damaged speakers presenting impaired phonological-phonetic encoding relative to 16 healthy controls. ERPs diverged between groups in terms of local waveform amplitude and global topography at ~400ms after stimulus onset in the picture naming task and at ~320-350ms in word reading and sustained until 100ms before articulation onset. These divergences appeared in later time windows than those found in patients with underlying lexical-semantic and lexical-phonological impairment in previous studies, providing evidence that phonological-phonetic encoding is engaged around 400ms in picture naming and around 330ms in word reading.

Introduction

Expressing an abstract idea through articulated speech involves the retrieval, preparation and articulation of several syllables per second. Theories of speech production (Levelt, 1989; Levelt et al., 1999; Dell, 1988) agree on the main processing stages involved in the transformation of a concept into articulated speech sounds, despite differing opinions on the independence and processing flow between the different encoding processes. The main processing stages involve semantic, lexical, phonological and phonetic encoding. Unlike all other encoding levels, aspects of phonological-phonetic encoding (the transformation of an abstract phonological code into specified phonetic plans) have not been implemented to the same extent in all speech production models. With respect to the goal of the present study, the lack of specification in phonological-phonetic encoding particularly relates to its time course, i.e. the period in time between conceptualization and articulation, during which an abstract phonological form is transformed into motor plans.

The dynamics of speech encoding processes have been intensively investigated with behavioral and electrophysiological techniques during single word production, leading to an estimate of the relative time course of the main underlying encoding processes (Indefrey and Levelt 2004; Indefrey, 2011). However, although the time course of lexical-semantic and lexical-phonological encoding during single word production has been directly investigated and confirmed by several event-related potential (ERP) studies, the estimate of phonological-phonetic encoding processes is based upon rather indirect evidence. Our aim here is to provide evidence for the time course of phonological-phonetic encoding, taking advantage of the combination of ERP analyses that cover the entire encoding process (from

concept to the onset of articulation) and the analysis of ERP modulations in brain-damaged speakers presenting impaired phonological-phonetic encoding.

In the following, we will first review the main encoding processes involved in single word production and the electrophysiological evidence on their relative time course in both healthy subjects and speakers who are language-impaired as a result of stroke. Then, we will present ERP analysis methods that allow one to track late- encoding processes, i.e. processes close to actual speech articulation.

Time course of speech encoding processes

Psycholinguistic experimental investigations coupled with neuroimaging studies allowing high temporal resolution (electroencephalography, EEG and magnetoencephalography, MEG) have modelled the mental operations involved in the planning of speech production (Levelt et al. 1999; Dell, 1986) and their time course (Indefrey and Levelt, 2004; Indefrey, 2011). This field of research has largely been dominated by investigations of behavioural and neural responses in single word production since, on the one hand, words constitute the building blocks for sentence production, and, on the other hand, experimental and theoretical constraints have given priority to single word paradigms. The timing of mental operations involved in speech production has been estimated in a meta-analysis by Indefrey and Levelt (2004, updated in Indefrey, 2011), in particular regarding picture naming tasks, i.e. the transformation of a concept into single word articulation. In picture naming tasks, it is estimated that visual and conceptual processes take place at 0 to ~200ms after picture presentation, followed by lexical-semantic (i.e. lexical selection) processes up to ~275ms. Encoding the phonological word form is then thought to occur between 275-450ms after picture onset, followed by phonetic encoding and motor execution. The time window

between 275-450ms has further been functionally subdivided: while the first interval (from ~275 to ~355ms) is supposed to correspond to the retrieval of phonological codes (i.e. lexical-phonological encoding), the processes occurring between 355-455ms have been associated with post-lexical phonological processes (i.e. syllabification or the application of phonological rules to the retrieved segmental content).

Notably, timing estimates reported in the first meta-analysis (Indefrey & Levelt, 2004) were based primarily on studies using metalinguistic tasks (e.g. Schmitt, Münte & Kutas, 2000; Van Turennout, Hagoort & Brown, 1998), while in a recent update (Indefrey, 2011), these temporal dynamics also incorporated the results of recent ERP studies that used overt speech production tasks, specifically examining lexical selection (i.e. lexical-semantic processes) and phonological encoding (i.e. lexical-phonological processes). These recent ERP studies investigated overt speech production via semantic interference paradigms (Aristei et al., 2011; Costa et al., 2009; Maess et al., 2009), the manipulation of psycholinguistic variables supposed to affect specific encoding stages such as lexical frequency (Strijkers et al., 2010) and word age of acquisition (Laganaro & Perret, 2011), or compared behavioural tasks entailing different encoding processes (Vihla et al., 2006). Commonly, these studies have reported that lexical selection is engaged ~200ms after picture presentation, while lexical-phonological encoding processes take place ~260-300ms after picture onset. Finally, converging results have also been reported by ERP studies investigating word production in brain-damaged language impaired (aphasic) speakers. Laganaro et al. (2009) reported ERP differences in a picture naming task between aphasic patients with impairments predominantly in lexical-semantic processes and control subjects within 100-250ms after picture onset. Later ERP divergences, i.e. starting ~300ms after picture presentation, were observed, on the other hand, in aphasic speakers with predominantly lexical-phonological

encoding impairments (Laganaro et al., 2009, 2011). Similar results were reported in a MEG study on recovery from anomia in patients with lexical-phonological impairments (Cornelissen et al., 2003). Here, MEG changes between pre- and post-treatment assessments were observed starting at 300ms after picture presentation. In contrast to the aforementioned findings, little evidence is so far available on the time window associated with the transition from phonological to phonetic encoding. These estimates are largely based on a MEG study by Salmelin et al. (1994) showing neural activations in Broca's area at 400-500ms during picture naming. Since Broca's area is supposed to underlie phonetic encoding, this finding provides evidence on the time interval dedicated to phonetic encoding in picture naming (but see Schumann et al. 2009, 2011, for earlier involvement of Broca's area during picture naming).

Post-lexical phonological-phonetic processes are engaged in all language production tasks independent of the specific stimulus type, meaning that also reading aloud entails such processes. In contrast to picture naming, the retrieval of the phonological code in word reading is mediated by orthography, but post-lexical phonological-phonetic processes are supposed to be largely common between the two tasks (Roelofs, 2004). To our knowledge the time course of phonological-phonetic encoding while reading aloud has not been addressed directly, especially since ERP studies on word reading have mainly used lexical decision tasks. However, given that visual-orthographic processes are engaged until ~200ms after word presentation (Appelbaum et al., 2009) and lexical access is estimated to occur around 200-250ms (see Hauk et al, 2012 for a review), post-lexical encoding processes should be effective after these time intervals. In these veins, Möller et al. (2007) asked healthy participants to overtly read aloud sequences of word pairs giving rise to phonological-phonetic errors in the form of spoonerisms (SLIP task). They found that ERPs to

speech errors differed from correct productions between 350-600ms after the onset of the target word pairs. This finding thus provides evidence as to the time window that specifically corresponds to phonological-phonetic encoding in reading aloud word pairs.

An approach to capture phonological-phonetic encoding processes with ERPs

One of the reasons why phonological-phonetic encoding has so far been less investigated in ERP paradigms is most probably due to the proximity of these processes to articulation. Phonological-phonetic processes refer to the transformation of an abstract phonological representation into speech motor plans that guide articulation. Psycholinguistic models of speech production claim that an abstract linguistic phonological form is planned before a more specified motor pattern is encoded (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). Based on this abstract phonological code, phonetic encoding processes specify articulation plans that serve as a basis for motor commands. These phonetic plans are not built on-line each time a speaker plans an utterance, but rather stored gestural scores are retrieved from an abstract phonological plan. These stored phonetic plans, which are considered to be syllable-sized (Crompton, 1981, 1982; Levelt et al., 1999), code invariant aspects of motor execution. Context-dependent parameters and the regulation of neuromuscular programs are then only defined at a later encoding stage (Levelt, 1989). Thus, phonological-phonetic encoding is thought to immediately precede motor execution.

Yet, phonological-phonetic encoding processes are probably not tackled in ERP investigations that use metalinguistic tasks or in delayed production paradigms. Moreover, in designs employing overt production, these encoding processes are unlikely to be tracked when fixed epochs of stimulus-aligned ERPs are analyzed. However, the most frequent

strategy to avoid ERP artefacts due to motor execution in overt immediate (i.e. non-delayed) speech production tasks consists of examining stimulus-aligned fixed ERP epochs that do not extend beyond the shortest response latency (Maess et al., 2002; Koester & Schiller, 2008). Such analysis approach yet leaves a core question unanswered, i.e. to what extent late-latency encoding processes (i.e. phonological and phonetic encoding) are captured. For instance, when fixed stimulus-aligned ERP epochs of 500ms (Maess et al., 2002) and 700ms (Koester & Schiller, 2008) are analysed, the entire encoding processes for fast trials and subjects (i.e. when production latencies do not extend beyond the analysed epoch) might well be captured. In contrast, for "slower" trials and subjects (i.e. when production latencies extend beyond the analysed epochs), the capture of later encoding processes might be truncated.

In our approach, we take advantage of an ERP analysis method that combines stimulus-aligned and response-aligned ERPs according to each subject's mean production latency, in order to encompass the entire time window of speech encoding processes. This method, which was introduced by Laganaro and Perret (2011) and applied in previous speech production studies (Laganaro et al., 2012; Perret & Laganaro, 2012), allows one to capture encoding processes that are closer to articulation (phonological-phonetic encoding) despite of variations in production speed. Further, by coupling electrode-wise ERP analyses covering the entire time window of speech encoding with global topographic ERP analyses (by means of a *spatio-temporal segmentation,* see below), periods of stable electrophysiological activity can be identified (i.e. stable topographic maps or functional microstates, see below). These topographies can be shown to differ qualitatively (different ERP topographies explain participants' responses) or quantitatively (similar topographies explain participants' responses at varying latencies) across experimental conditions from the onset of a stimulus

to the onset of speech articulation. In extension, the occurrence of qualitatively different response topographies points to (at least partially) differing neural generators over the respective time period (Murray et al., 2008; Michel & Murray, 2012).

In order to investigate the time interval associated with phonological-phonetic encoding, we analyzed ERP divergences between stroke patients with phonological-phonetic impairments and healthy controls. Our hypothesis is that patients present ERPs diverging from those of healthy controls in the time window associated with the impaired encoding process. The approach adopted for analysis has been applied in previous studies on different forms of anomia (Laganaro, 2009) and confirmed in a single case study comparing pre- to post-stroke ERPs, proving its methodological reliability (Laganaro et al., 2011). We predict that, in the current study, no differences should be observed between the ERP patterns in patients and controls for the time windows previously associated with lexical-semantic and lexical-phonological processes and impairments. ERP divergences should, on the other hand, occur in later time windows than those found in patients with underlying lexical-semantic and lexical-phonological impairment (Laganaro et al., 2009; 2011), i.e. after ~300ms in the picture naming task. In order to allow for comparisons with the results of previous studies, we employed a picture naming task. Since patients are expected to display phonological-phonetic impairments in all speech production tasks - these not being limited to the lexical-semantically driven picture naming task alone - a word reading task was additionally employed in the study design. Although so far there is only very limited data on the latency of phonological encoding processes involved in word reading aloud, phonological-phonetic encoding and impairment is thought to be common to both picture naming and word reading tasks (see above). That is, we predict that divergent ERPs between patients and healthy controls should be observed in both tasks alike.

Methods

Participants

Eight patients with phonological-phonetic impairments following a left-hemispheric stroke participated in the study. All were right-handed native French speakers and were selected according to the following criteria. They presented phonological and/or phonetic errors in all speech production tasks (spontaneous, naming, reading, repetition) with no other language impairment at the moment of the study according to a standard aphasia assessment battery (Montreal-Toulouse 86 Aphasia Battery, Nespoulous et al., 1992). Crucially, in order to avoid any influence of lexical-semantic or lexical-phonological impairments, word finding difficulties or the production of semantic paraphasias constituted exclusion criteria.

[Table 1 about here]

All patients included in the present study: (i) did not suffer from word finding impairments in picture naming at the moment of the study, except some delayed responses in two patients (see Table 1); (ii) did not produce semantic paraphasias in any language production task; (iii) had scores within normal range at the Pyramids- and Palm-Tree-Test (Howard & Patterson, 1992); (iv) produced phonemic-phonetic errors in naming, reading aloud and repetition tasks. Patients' phonemic-phonetic errors were mainly single phoneme errors (phoneme substitution, omission or addition, or phonetic distortion). The rate of phonetic distortion errors was quantified in a sentence repetition task composed of 96 two-word adjectival noun-phrases. Four patients (P1, P2, P3 and P4, i.e. "phonetic-dominant" sub-group) had a rate of phonetic distortions exceeding 5%, which indicates the presence of apraxia of speech (AoS) according to standard criteria (McNeil, Pratt, & Fossett 2004; Romani & Galluzzi, 2005;

Ziegler, 2009). The other four patients had a diagnosis of mild conduction aphasia and produced mainly well-formed phonemic paraphasias (i.e. "phonological-dominant" sub-group).

Healthy control subjects

The control subjects were 16 healthy volunteers. Half of them were individually matched by age, gender and education to the patients ("matched controls" sub-group, mean age: 48.8, 4 men); the other 8 controls were aged from 32 to 73 ("other controls" sub-group, mean age: 47.4, 4 men). All participants gave a written informed consent to participate in the study, which was approved by the local ethics committee.

Materials

Picture naming task

120 words and their corresponding pictures (line drawings on white squares of 280 x 280 pixels) were selected from two French databases (Alario & Ferrand, 1999; Bonin et al, 2003). All pictures and their corresponding words had a high name agreement (over 75%, mean name agreement = 92.5%) to ensure that participants give the same name for a same picture (see Alario et al., 2004). The stimuli were monosyllabic (N=40), bisyllabic (N=60) and trisyllabic (N=20) words of lexical frequency varying from 0.13 to 227 occurrences per million words (mean = 17.3) in the French database Lexique (New et al., 2004).

Word reading task

The stimuli were the same 120 words as for the picture naming task and were printed in black on white rectangles of 300 x 70 pixels.

Procedure

Participants were tested individually in a sound-proof dimly lit room. They sat approximately 60cm in front of a PC screen. The presentation of the trials was controlled by the E-Prime software (E-Studio). Before the experiment, participants were familiarized with all pictures and their corresponding names on a paper sheet. All participants first underwent the picture naming task, followed by the word reading task. An experimental trial had the following structure: first, a "+" sign was presented for 500ms on screen. Then, the picture (naming task) or word (reading task) appeared on the screen for 3000ms. Participants had to overtly produce the word corresponding to the picture or to the written word. A blank screen lasting 2000ms was displayed before the next trial for control subjects while each stimulus was launched manually for patients. All items were presented in pseudo-random order, preceded by 4 warm-up filler trials. In both tasks, a break was introduced after 60 stimuli. Pictures were presented in constant size on a grey screen. Word productions were digitized and production latencies (hereafter: reaction times [RT] in ms, i.e. the time separating the onset of the picture or of the word and the articulation onset) were systematically checked with a speech analysis software (Check-Vocal, Protopapas, 2007).

EEG acquisition and pre-processing

EEG was recorded continuously using the Active-Two Biosemi EEG system (Biosemi V.O.F. Amsterdam, Netherlands) from 128 channels covering the entire scalp. Signals were sampled at 512Hz with an online band-pass filter set to 0.16-100Hz. The custom online reference of the system is the common mode sense (CMS active electrode) – driven right leg (CMS-DRL) which drives the average potentials as close as possible to the amplifier zero (details of this setup can be found on the Biosemi website: http://www.biosemi.com). Offline, ERPs were then band- pass filtered to 0.2–30Hz and recalculated against the average reference. ERP epochs with amplitudes exceeding ±100µV were automatically rejected. Only trials with

correct productions and valid RTs were retained. Further, each trial was visually inspected, and epochs contaminated by eye blinks, movements or other noise artifacts were rejected and excluded from averaging. This procedure was applied separately to stimulus- and response-aligned epochs. In addition, only trials with both response-aligned and its corresponding stimulus-aligned uncontaminated epochs were retained. As a result, a minimum of 63 averaged trials per subject entered the ERP analyses (patients: 63-97 epochs [mean= 82.25] in naming and 82-109 epochs [mean = 105.5] in the reading task; controls: 78-118 epochs [mean= 100.5] in naming and 66-115 epochs [mean= 98] in reading). Artifact electrodes were interpolated using 3-D splines interpolation (Perrin et al., 1987).

Stimulus-aligned (i.e. forward) epochs and response-aligned (i.e. backward) epochs were first averaged for each single subject and then averaged per participant sub-group. The duration of each response-aligned and stimulus-aligned ERP epoch was 450ms (=230 sampling points at 512Hz recording rate) for the picture naming task and 390ms for the reading task. For the common topographic analyses of stimulus- and response-aligned ERPs (see below), the respective epochs were averaged for each individual and group. To assure that there was no temporal overlap between stimulus- and response-aligned data (i.e. in cases of RTs shorter than 900ms in picture naming and shorter than 780ms in reading), time periods of overlap were removed from response-aligned ERPs. As a result, the combination of stimulus- and response-aligned individual and group averaged ERPs covered the exact time interval from picture or word onset, respectively, to 100ms before articulation. In those patients whose mean production latencies exceeded 1000ms, preparation for the common topographic analyses required the extraction of longer stimulus- and response-aligned ERP epochs (up to 586ms) to fully cover the time interval from picture or word onset to 100ms before articulation.

ERP analyses

The ERPs obtained for the picture naming and the word reading task were analysed separately. The ERPs were first subjected to a sampling point-wise ERP waveform analysis to determine the time periods presenting local differences in ERP amplitudes between patients and controls. Then, a spatio-temporal segmentation was performed on the group-averaged ERPs of each sub-group to determine topographic differences across groups and statistically validate them in the responses of single subjects as described below.

ERP waveform and global field power analyses

For both production tasks, electrode-wise and sampling point-wise unpaired t-tests served to compare the local ERP amplitudes in control subjects and patients over the entire analysed time period. Therein, the responses of patients were once compared to the entire control group and once only to the sub-group of age-matched controls. An additional analysis compared the responses of the two control sub-groups. To correct for multiple comparisons, only ERP differences present at ≥5 electrodes from the same region out of 6 scalp regions (left and right anterior, central, posterior) and lasting ≥30ms were retained with a conservative alpha criterion of 0.01. For differences in global field power (GFP) unpaired t-test were computed on the GFP between groups at each time-frame, with an alpha criterion of 0.05 and a time-window of 20ms of consecutive significant difference.

Global topographic ERP pattern analysis (spatio-temporal segmentation)

The second analysis conducted was a global topographic ERP (map) pattern analysis. This analysis entails a spatio-temporal segmentation of the ERPs over periods of electrophysiological stability (i.e. topographic maps or ERP microstates). The spatio-temporal

segmentation procedure compresses the variability of ERPs in a series of template maps which summarize the data and serve to determine which topographic template best explains participants' ERP responses to each experimental condition. In extension, it provides insights into how groups or conditions differ in terms of likely underlying neurophysiologic mechanisms (Murray et al., 2008; Michel & Murray, 2012), in addition to the temporal information about ERP differences. This method is independent of the reference electrode (Michel et al., 2001, 2004) and insensitive to pure amplitude modulations across conditions (since topographies of normalized maps are compared). The spatio-temporal segmentation was applied to the group-averaged data (from picture onset to 100ms before articulation) of the four sub-groups: "matched" controls, "other" controls, "phonological-dominant" patient and "phonetic-dominant" patient sub-groups. We used a modified hierarchical clustering algorithm (Michel et al., 2001; Pascual-Marqui et al., 1995), the agglomerative hierarchical clustering, to determine the most dominant electric field configurations at the scalp (topographic ERP maps). To determine the optimal number of ERP maps that best explain the group-average data across sub-groups a combination of a cross-validation and the Krzanovski-Lai criterion was used (see Murray et al., 2008). Statistical smoothing was applied to eliminate temporally isolated topographic maps with low explanatory power. This procedure is described in detail in Pascual-Marqui et al. (1995, see also Brunet et al., 2011; Murray et al., 2008). In accordance with the criteria for the local ERP waveform analyses, a given ERP topography had to be present ≥30ms. The pattern of topographic map templates observed in the group-averaged data was statistically tested by comparing each of these map templates with the moment-by-moment scalp topography of individual subjects' ERPs. Each data sampling point was labelled according to the template map with which it best correlated spatially, yielding a measure of map presence in milliseconds. This procedure

referred to as 'fitting' allowed to establish how well a topographic template map explained single subject responses in each experimental task.

Results

Picture naming

Behavioral data

Production accuracy was above 97% in controls and varied from 66% to 99% in patients (for error distribution in patients see Table 1). Patients were on average 200ms slower than controls (patients: 1054ms [SD=146ms], matched controls: 823ms [SD=89ms], other controls: 878ms [SD=62ms]). Log transformed RTs were used in the statistical analyses to approximate a normal distribution (Shapiro-Wilk normality test: $W= 0.97$, $p = .61$). The statistical analysis revealed main effects of group on production latencies ($F(2,21)=8.38$, $p<.01$). Patients' production latencies differed from both control sub-groups (Fisher $ps<.01$), but the production latencies of both control sub-groups did not ($p=.18$).

ERPs

In the ERP waveform analysis, the comparison between patients and all controls indicated amplitude differences starting around 370ms after picture onset until the end of the analyzed period (i.e. 100ms before articulation). Similar differences were observed when only comparing the responses of patients to age-matched controls (see Figure 1). On the other hand, virtually no differences were apparent when comparing the responses of both control sub-groups. Global field power also differed between patients and controls from

about 410ms in the stimulus-aligned ERPs and between 330 and 180ms before articulation onset in the response-aligned data.

[Figure 1 about here]

The spatio-temporal segmentation was applied to the group-averaged ERPs of each sub-group (two control sub-groups and two patient sub-groups) from picture onset to articulation onset. Eight topographic map templates were identified accounting for 87.4 % of the variance in the data. Three fitting periods were established on the basis of this result (see Figure 2): from 50-300ms (encompassing the ERP template maps A, B, C), from 300-450ms (map templates C, D, E) and from 450ms to the end of the individual ERP epochs (100ms before articulation).

[Figure 2 about here]

While the same sequence of three first stable electrophysiological patterns was observed in all sub-groups (template maps A, B, C from 50 to approximately 390-410ms in Figure 2), different distributions of topographic maps appeared starting around 390ms in patients compared to controls. Template map D appeared in the averaged ERPs of both control sub-groups while template map E characterized patients' ERPs over approximately the same time-period. The presence of these two stable electrophysiological patterns in controls and patients was further assessed by fitting the template maps obtained at group level to single subject responses over the 400-500ms time window. While template map D was observed in 13 control subjects and 3 patients (χ^2=4.5, p<.05), template map E characterized ERPs in all patients, but only appeared in 2 out of 16 controls (χ^2=10.15, p<.01). The topographic map F in Figure 1b characterized the last 150 ms in both control sub-groups and also

appeared from about 580 to 800ms in 5 patients (including all the patients from the "phonetic" sub-group). Template map H characterized nearly exclusively the responses in patients over the last 100ms of the analyzed period; it was observed in the last fitting period in 6 out of 8 patients, but only in 3 out of 16 controls (χ^2=7.3, p < .01). Interestingly, the template map G appeared in the group-average data of the "phonological" patients sub-group only, yet driven by the responses of two patients only (P4 and P5), who did not display a specific behavioral pattern relative to the other patients in this sub-group. In control subjects, this particular topography was not observed at all.

Reading

Due to technical problems in data collection, the reading task data of one patient (P1) was not available for analysis. Therefore, the matched control subject was also removed from the analyses as well as a random subject from the "other" control group.

Behavioral data

Reading accuracy was above 99% in controls and varied from 67% to 100% correct in patients (see Table 1). Production latencies were on average 220ms slower in patients (790ms, SD=125ms) than in controls (respectively 567ms [SD= 52ms] and 578ms [SD=84ms] in "matched and "other" controls). Log transformed RTs were used in the statistical analyses to approximate a normal distribution (Shapiro-Wilk normality test: W = 0.95, p = 0.39). There was a significant effect of group (F(2, 18) = 12.9, p < .001) and patients were slower than both control sub-groups (Fisher *ps* < .001), whereas control sub-groups did not differ significantly (p = .85).

ERP analyses

ERP amplitudes diverged between patients and the combined control group from ~350ms after stimulus onset to 150ms before word articulation (Figure 3). Similar differences were observed when only comparing the responses of patients to matched controls. No differences were apparent in the comparison between the two control sub-groups. GFP differences between controls and patients were observed from 330ms after word onset to 180ms before articulation onset in addition to a short 30ms period of different GFP at 230-260ms.

[Figure 3 about here]

The spatio-temporal segmentation applied to the group-averaged ERPs of the two control sub-groups and the two patient sub-groups from word onset to the end of the response-aligned ERPs identified 6 topographic ERP template maps accounting for 79% of the overall variance in the data. Three fitting periods were established on the basis of this result (see Figure 4): from 80-160ms (coinciding with the template maps A and B), from 160-300ms (in accordance with the template maps B, C, D) and from 300ms to the end of each individual ERP epochs (i.e. 100ms before articulation onset; encompassing the template maps D, E, F).

[Figure 4 about here]

The same sequence of four stable ERP topographies was present in all sub-groups until ~300ms (template maps A, B, C, D). In succession, the template map E characterized the ERP responses of both control sub-groups from ~300-320ms until 150ms before articulation, while template map D still sustained over this time window in patients. The fitting from group-level to individuals' responses confirmed that map E was present in 13 out of 14

controls, but only in 4 out of 7 patients ($\chi 2=3.86$, $p<.05$). Of these patients, three had been classified as presenting phonological-dominant and one as showing phonetic-dominant deficits. The final ERP template map F, on the other hand, better characterized the responses of patients than those of controls. The fitting of this group template map to individuals' responses proved its presence in 6 out of 7 patients, but only in 5 out of 14 controls ($\chi 2=4.68$, $p<.05$) .

Discussion

The current study investigated the dynamics of phonological-phonetic encoding during single word production by comparing ERPs in stroke patients with impaired phonological-phonetic encoding and healthy controls. ERPs were found to diverge between groups in terms of local waveform amplitude and global topography, from ~380-400ms after stimulus onset in the picture naming task and from ~320-350ms in the word reading task. In both tasks, differences between controls and patients sustained until the end of the analyzed individual ERP periods, i.e. until 100-150ms before articulation onset.

As previous studies have mainly contributed to our understanding of the dynamics of encoding processes in picture naming tasks, the interpretation of our present findings will primarily be embedded in this literature. Inferences on the outcomes in the word reading aloud task will then be drawn on the basis of previous results.

As clarified in the introduction, phonological-phonetic processes refer to the application of phonological rules to the retrieved phonological content and to the encoding of articulatory plans. We therefore applied strict selection criteria to include only patients who produced

phonological-phonetic errors without word finding difficulties or semantic errors (i.e. without lexical-phonological impairment and without lexical-semantic impairment). Our study thus provides evidence that local and global ERP patterns in patients with phonological-phonetic impairments (vs. controls) diverged at a later state than previously reported in patients with lexical-semantic anomia (i.e. between 100-250ms after stimulus onset) or lexical-phonological impairments (i.e. from 300ms after stimulus onset; Cornelissen et al., 2003; Laganaro et al., 2008; 2009; 2011). In previous studies, the atypical ERPs observed in sub-groups of anomic patients converged with the dynamics of encoding processes estimated in healthy subjects (Indefrey & Levelt, 2004; Indefrey, 2011). The ERPs of patients did not differ from controls during the first 400ms in the present study, suggesting unimpaired visual, lexical-semantic and lexical-phonological encoding processes. Approximately 400ms after stimulus onset patients displayed ERP waveforms that differed from healthy controls with both different and longer lasting periods of stable global electrophysiological patterns (topographic maps) in addition to longer RTs. These differences are probably due to more laborious encoding processes in the patients, but also to an increased necessity to detect and correct errors (monitoring processes). In other words, encoding processes differed after 400ms in the patient group both quantitatively and qualitatively (longer phonological-phonetic encoding and possibly increased monitoring processes than in non-brain damaged controls), but the onset of differences between groups pinpoints the time course of phonological-phonetic encoding in healthy controls.

A crucial comparison here concerns the time window of atypical ERPs between patients deemed to have lexical-phonological impairment in previous studies and those in the present study. Indeed, patients classified as impaired at the lexical-phonological level in former studies or at the phonological-phonetic level in the present study all produced

phonemic paraphasias in picture naming tasks. However, formerly investigated patients also presented high rates of non-responses (up to 50% in Cornelissen et al., 2003; up to 79% in Laganaro et al., 2009), indicating underlying word-finding impairments (i.e. lexical-phonological impairments). By contrast, word-finding impairments were absent in the eight patients of the present study, i.e. they only produced single phoneme errors in all production tasks. This pattern of errors had previously been attributed to underlying "post-lexical" impairment (Kohn & Smith 1994; Goldrick & Rapp, 2007). Although the exact definition of "post-lexical" impairment varies between authors, it globally refers to patterns in which the lexeme has been correctly retrieved, but further application of phonological rules and phonetic specification fails, hence resulting in phonological-phonetic impairments. Thus, ERPs diverged from controls around 330ms after picture presentation in patients with underlying lexical-phonological impairment in previous findings and about 100ms later in patients with phonological-phonetic impairment in the present data. The periods of divergence between patient sub-groups then seem to overlap from around 400ms. Unfortunately, further comparisons with previous studies are limited due to methodological differences. While in the study of Cornelissen et al. (2003) a pre-post treatment comparison was used, Laganaro et al. (2009) employed a delayed production task whose results are comparable to immediate production only during the first 350ms (Laganaro & Perret, 2011).

Two further points should be considered. First, in the time course of encoding processes proposed by Indefrey and Levelt (2004), post-lexical phonological processes (syllabification of the retrieved phonological code) are engaged from about 355-455ms after picture onset. However, the mean production latencies in the control subjects of the present study are longer than the 600ms estimated by Indefrey and Levelt (2004), based mainly on studies

using much smaller sets of stimuli with several repetitions (see recent picture naming ERP studies by Costa et al. 2009; and Aristei et al., 2011 for mean RTs around 800ms as in the present study). That is, encoding processes in the present study appear to be shifted temporally (see Laganaro et al., 2012, for a comparison between fast and slow subjects). The onset of ERP differences between controls and patients suggests that in the control group post-lexical phonological processes are engaged around 400ms after picture onset for production latencies of about 850ms.

Second, one might wish to distinguish between the dynamics of phonological and phonetic encoding. Our eight patients were classified according to the dominance of phonological (e.g. well-formed phoneme substitution) or phonetic errors (e.g. phoneme distortion). However, the time interval of diverging ERPs relative to controls was identical in the two sub-groups of patients. There are two possible explanations for this lack of difference between patients with phonological- and phonetic-dominant impairments. First, patients did not display pure phonetic nor phonological impairments as a group. The fact that these two patterns of impairment often overlap and are difficult to distinguish has been repeatedly acknowledged in the literature (McNeil, Robin & Schmidt, 2008; Duffy, 2005) and an increasing number of studies propose an overlap between phonological and phonetic encoding processes and impairments (see Laganaro, 2012). In addition, although the meta-analysis of the dynamics of word production reviewed in the Introduction was based on a model involving serial encoding processes, Indefrey and Levelt (2004) also suggested that phonetic encoding may start as soon as the first phonological syllable has been encoded, therefore assuming an overlap between phonological and phonetic processes for polysyllabic words.

In summary, the onset of diverging ERPs in patients with impaired phonological-phonetic encoding provides evidence for the time interval during which phonological-phonetic processes are engaged in healthy controls when naming pictures and reading words aloud. In particular, divergences were apparent until the end of the analyzed period, i.e. 100ms before articulation in both production tasks. Along with previous assumptions from the psycholinguistic literature (Roelofs, 2004), the present results suggest that similar phonological-phonetic encoding processes are engaged in both tasks. The observed time window of diverging ERPs between patients and controls suggests that phonological encoding may be engaged after 300ms in reading. This result is in line with the estimation that lexical access occurs in an earlier time window (around 200-250ms), and also broadly corresponds with the time course of ERP modulations on error trials in the SLIP task starting around 350ms after stimulus presentation (Möller et al., 2007), although production was elicited with word pairs in this latter study.

It should be pointed out that the time interval of diverging ERPs between participant groups was found to be much longer in naming than in reading. This might be attributed to several reasons. Participants performed the reading task always after completing the picture naming task, so that phonological-phonetic encoding during picture naming might have been eased by the previous task. On the other hand, reading is supposed to entail faster motor preparation (Riès et al., 2012). Most relevant for the current report, however, is that outcomes generally converged between tasks, thus confirming diverging ERPs between patients and controls over the time window immediately preceding articulation onset, starting respectively at ~400ms after picture presentation in the naming task and at ~330ms after word onset in the reading task. That is, the respective time intervals most likely are representative of the dynamics of phonological-phonetic encoding in healthy controls.

Acknowledgements This research was supported by Swiss National Science Foundation grants no. PP001-118969 and PP00P1_140796 to Marina Laganaro.

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. *Behavior Research Methods, Instruments, and Computers, 31*, 531-552.

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

Appelbaum, L.G., Liotti, M., Perez, R., Fox, S.P. and Woldorff, M.G. (2009). The temporal dynamics of implicit processing of non-letter, letter, and word-forms in the human visual cortex. *Frontiers Human Neuroscience, 3*:56.

Aristei, S., Melinger, A., Abdel Rahman, R. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience, 23,* 1567–1586.

Bonin, P., Peerman, R., Malardier, N., Méot, A. & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studiespp. French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, and Computers, 35*, 158-157.

Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience, 2011, Article ID 813870. doi:10.1155/2011/813870*

Cornelissen, K., Laine, M., Tarkiainen, A., Jarvensivu, T., Martin, N. & Salmelin, R. (2003). Adult Brain Plasticity Elicited by Anomia Treatment. *Journal of Cognitive Neuroscience, 15,* 444–461.

Costa, A., Strijkers, K., Martin, C. Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences, 106*, 21442-6.

Crompton, A. (1981). Syllables and segments in speech production. *Linguistics, 19*, 663-716.

Crompton, A. (1982). Syllables and segments in speech production. In A. Cutler (Ed.). Slips of the tongue and language production (109-162). Berlin: Mouton.

Dell, G. S. (1986). A spreading activation theory of retrieval in language production. *Psychological Review, 93,* 283–321.

Duffy JR. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management.* St. Louis, MO: Mosby-Year Book.

Goldrick, M. & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production*. Cognition, 102*, 219–260.

Hauk, O., Coutout, C., Holden, A., Chen, Y. (2012). The time-course of single-word reading: Evidence from fast behavioral and brain responses. *Neuroimage, 60,* 462-1477

Howard, D. & Patterson, K. (1992). *The Pyramid and Palm Trees Test*. Thames Valley Test Compagny, Bury St Edmund.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology 2:255*.

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

Koester, D., Schiller, NO. (2008). Morphological priming in overt language production: electrophysiological evidence from Dutch. *Neuroimage. 42*, 1622-1630.

Kohn, S.E. & Smith, K.L. (1994). Distinction between two phonological output deficits*. Applied psycholinguistics*, *15*, 75-95.

Laganaro, M. (2012). Patterns of impairments in AoS and mechanisms of interaction between phonological and phonetic encoding. *Journal of Speech, Language, Hearing Research, 55,* S1535-43.

Laganaro, M. (2012). Patterns of impairments in AoS and mechanisms of interaction between phonological and phonetic encoding*. Journal of Speech, Language and Hearing Research, 55,* 1535-1543.

Laganaro, M., Morand, S. Michel, CM, Spinelli, L., Schnider, A. (2011). ERP correlates of word production before and after stroke in an aphasic patient*. Journal of Cognitive Neuroscience, 23*, 374-381.

Laganaro, M., Morand, S., Schnider, A. (2009). Time course of evoked-potential changes in different forms of anomia in aphasia. *J. of Cognitive Neuroscience*, *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 Topography, 24,* 19-29.

Laganaro, M., Valente, A., Perret, C. (2012). Time course of word production in fast and slow speakers: a high density ERP topographic study. *NeuroImage, 59*, 3881-3888.

Levelt, W.J.M. (1989). *Speaking: from intention to articulation*. Cambridge, Mass: MIT Press.

Levelt, W.J.M., Roelofs, A. & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 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 Cognitive Neuroscience, 14*, 455-462.

McNeil MR, Robin DA, Schmidt RA. (2008). Apraxia of Speech: Definition and Differential Diagnosis. In: MR McNeil, *Clinical Management of Sensorimotor Speech Disorders*. 2nd Ed. New York: Thieme Medical Publishers.

McNeil, MR., Pratt, SR. & Fossett, TRD. (2004). The differential diagnosis of apraxia of speech. In: B. Maassen (ed.). *Speech motor control in normal and disordered speech*. New York: Oxford University Press.

Michel, C.M., Murray, M.M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage, 61,* 371-385.

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, 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.

Möller, J., Jansma, B.M., Rodriguez-Fornells, A., Münte, T.F. (2007). What the brain does before the tongue slips. *Cerebral Cortex, 17(5)*, 1173-8.

Murray, MM., Brunet, D. & Michel, C. (2008). Topographic ERP analyses, a step-by-step tutorial review. *Brain Topography, 20,* 249-269.

Nespoulous, J. L., Lecours, A. R., Lafond, D., Lemay, A., Puel, M., Joannette, Y., Cot, F. & Rascol, A. (1992). *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie (MT86)*. Isbergues, France.: L'Ortho-Edition.

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

Perrin, F., Pernier, J., Bertrand, O., Giard, M.H.,& Echallier, J.F., (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr. Clin. Neurophysiol. 66*, 75-81.

Pascual-Marqui, R.D. , Michel, C.M. & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans Biomed Eng*., 42, 658–665.

Perret, C., Laganaro, M. (2012). Comparison of electrophysiological correlates of writing and speaking: a topographic ERP analysis. *Brain Topography, 25*, 64-72.

Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*, 859–862.

Riès, S., Legou, T., Burle, B., Alario, FX. & Malfait, N. (2012). Why does picture naming take longer than word reading? Contribution of articulatory processes*. Psychonomic Bulletin & Review*, 19(5), 955-961.

Roelofs, A. (2004). Seriality of phonological encoding in naming objects and reading their names. *Memory & Cognition*, *32*, 212-222.

Romani, C. & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory disorders and phonological difficulties. *Cognitive Neuropsychology, 27*, 817-850.

Salmelin, R., Hari, R., Lounasmaa, O. V., & Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature, 368*, 463–465.

Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology, 37*, 473–484.

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.

Schuhmann, T., Schiller, NO., Goebel, R. & Sack, AT. (2012). Speaking of which: dissecting the neurocognitive network of language production in picture naming. *Cerebral Cortex, 22,* 701-709.

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

Van Turennout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280*, 572-754.

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

Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech *Language and Cognitive Processes, 24,* 631-661.

TABLES

TABLE 1

Table 1. Patient demographic data and rate of phonemic-phonetic errors in picture naming, reading (words and pseudo-words) and of phonetic distortions in a sentence repetition task.

				Phonemic-phonetic errors*			Phonetic distortions
Pt	Age	Gender	Diagnosis	Picture Naming (N=120)	Word Reading (N=120)	Pseudo-word Reading	Sentence Repetition
P1	52	f	AoS	1%	NA	9%	7%
P2	29	m	AoS	4%	4%	14%	19%
P3	51	f	AoS	30%	33%	20%	14%
P4	46	m	CA	8%	10%	15%	2%
P5	62	f	CA	3%**	3%	10%	0%
P6	51	m	CA	1%	3%	10%	2%
P7	51	f	CA	2%	0	13%	1%
P8	51	m	AoS&BA	10%**	3%	20%	6%

* No other errors, no word finding difficulties in picture naming, except long latencies in **(P5: 30% latencies beyond 3000ms; P8: 24% latencies beyond 3000ms).

AoS: apraxia of speech; CA: conduction aphasia; BA: Broca aphasia

FIGURES

Figure 1. a. Significant differences (p values) on ERP waveform amplitude on each electrode (Y axes) and time point (X axes) from picture onset to 100 ms before articulation in the picture naming task: between patients and controls (top), between patients and matched controls (middle) and between matched and other controls (bottom), with the arrangement of the 128 electrodes. b. Mean global field power (GFP) for the patient and control groups with results of statistical analysis (1-p values) and averaged stimulus-aligned and response-aligned ERP waveforms with the arrangement and electrode position of the displayed waveforms (Fz, POz, T7 and T8).

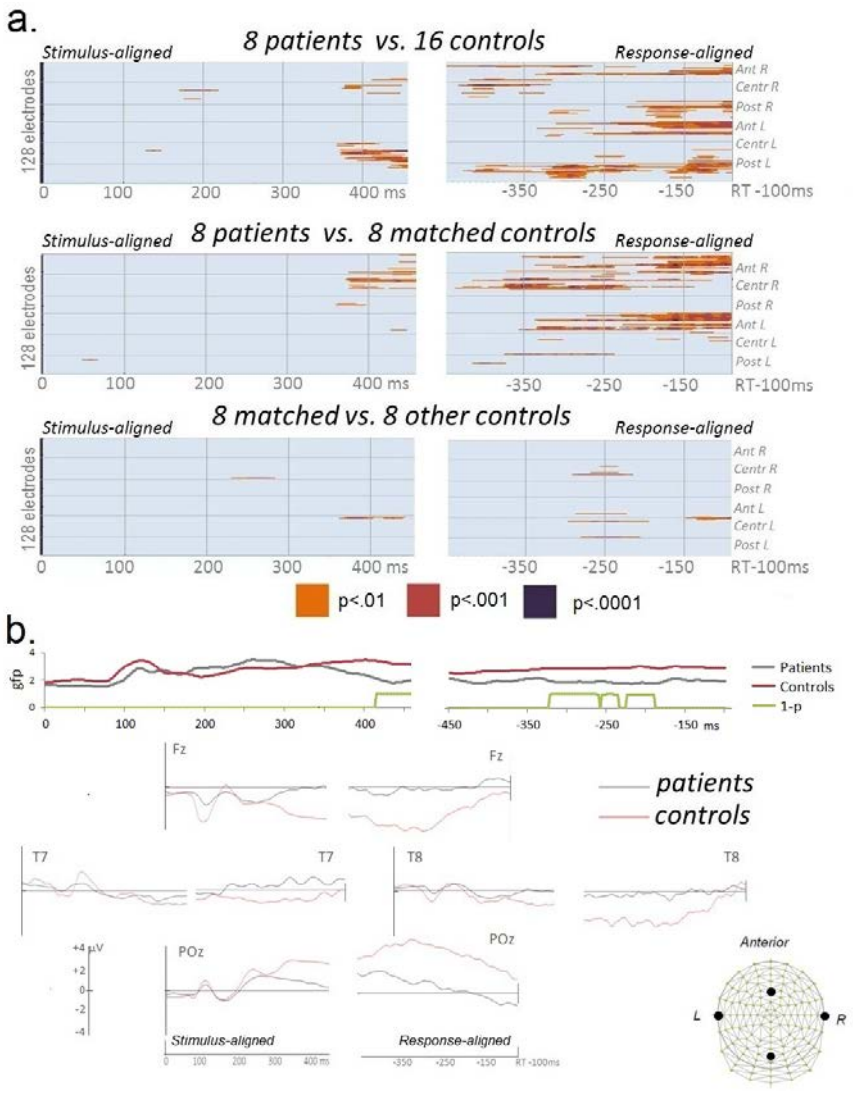

Figure 2. Group-averaged ERPs (128 electrodes) in the patient and control sub-groups and temporal distribution of the topographic maps revealed by the spatio-temporal segmentation analysis in each data set. Template maps for the 8 stable topographies observed from 50ms after picture presentation onset to 100ms before articulation are displayed with positive values in red and negative values in blue and with display of the maxima and minima of scalp field potentials.

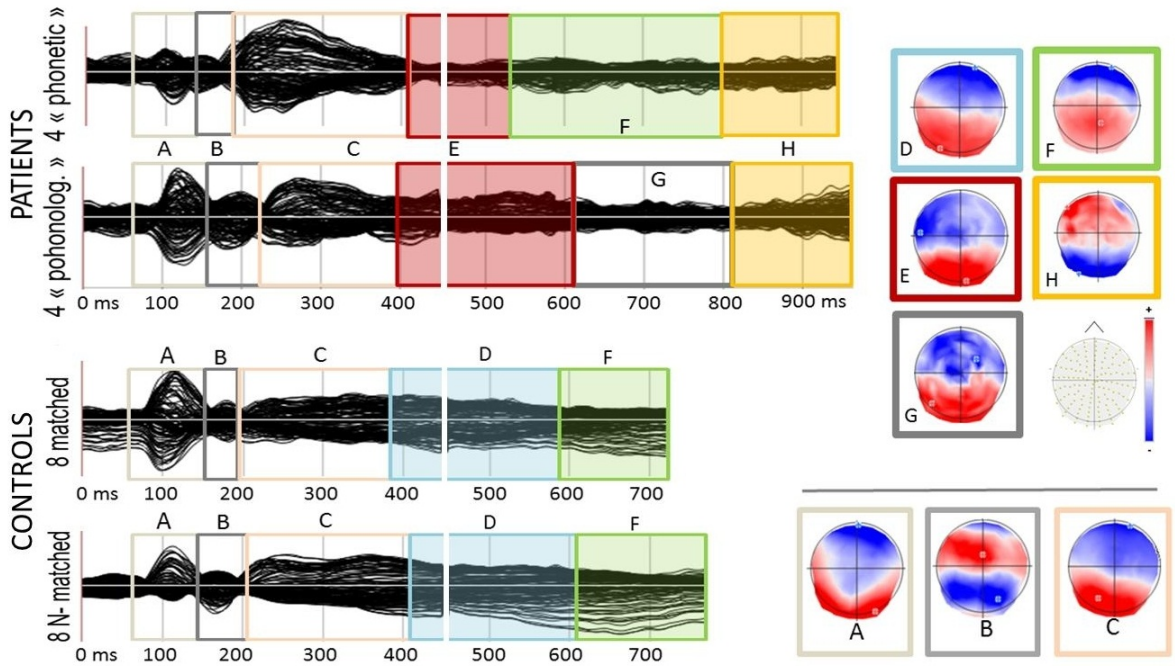

Figure 3. a. Significant differences (p values) on ERP waveform amplitude on each electrode (Y axes) and time point (X axes) from word onset to 100ms before articulation in the word naming task: between patients and controls (top), between patients and matched controls (bottom). **b**.

Mean global field power (GFP) for the patient and control groups with results of statistical analysis (1-p values) and averaged stimulus-aligned and response-aligned ERP waveforms with the arrangement and electrode position of the displayed waveforms (Fz, POz, T7 and T8).

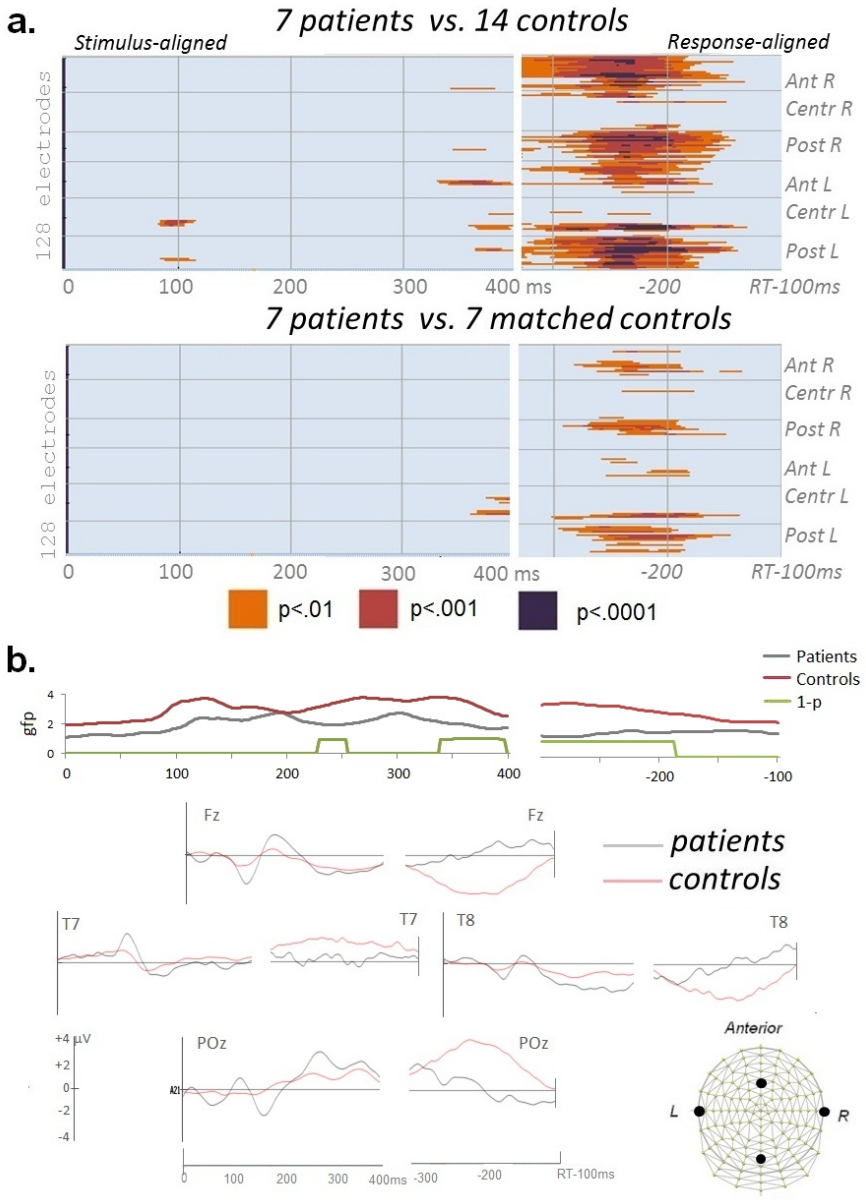

Figure 4. Grand-average ERPs (128 electrodes) from each patient and control sub-group in the word naming task and temporal distribution of the topographic maps revealed by the spatio-temporal segmentation analysis with map templates for the 6 stable topographies observed from 50 ms after picture presentation to 100 ms before articulation are displayed with positive values in red and negative values in blue and with display of maximal and minimal scalp field potentials.