Brain Stimulation 13 (2020) 267–269

Contents lists available at ScienceDirect

Brain Stimulation

journal homepage: http://www.journals.elsevier.com/brain-stimulation

Automated speech analysis to improve TMS-based language mapping: Algorithm and proof of concept

Transcranial magnetic stimulation (TMS) during confrontational naming can induce anomia or speech arrest and paraphasias and those can be used to locate language areas. This method is qualitative and time-consuming since a trained professional needs to analyze a long video-file of the mapping session. For data analysis we adapted an existing speech analysis method [1] to cope with TMS noise, which exceeds 100 dB SPL with each discharge of the coil. Our algorithm deciphered what was said and the reaction time (RT) of each utterance with a resolution of up to 10 ms, leading to a fast and quantitative analysis.

To develop the noise reduction in combination with the speech analysis algorithms, semi-synthetic data were created, since it is impossible for the human ear to discern speech onset in the loud TMS-noises. During a confrontational naming task (without TMS), voice-recordings were obtained and RTs were manually annotated. Separately, multiple realistic recordings of rTMS noise were recorded, using a figure- 8 coil. Semi-synthetic data were then created by fusion of the rTMS-noises (at the onset) and the voice-recordings with some random delay. Over 1000 recordings were created.

Actual patient data were recorded from three male patients planned for neurosurgery near language areas. Setup was in accordance with published guidelines [2] with two no-TMS recordings as baseline and TMS-mapping of language areas using a 5Hz-1s stimulation at an intensity of 120% of resting motor threshold. Stimulation targets were recorded (Brainsight, Rogue Research, Canada) for offline analysis. All three had direct cortical stimulation (DCS) during awake surgery which we used as gold-standard. The study was approved by the local Ethics Committee of the University Hospitals Leuven. All patients gave written informed consent.

The algorithm contrasted the data likelihood for the model of the expected response with the data likelihood for a generic model of words. An internal penalty parameter avoided the generic model from winning inappropriately. In this task, we adjusted the weight to recognize the correct response in order to maximize the true negatives correctly (but reducing true positives).

Speech enhancement front-end suppressed background noises and rTMS noise [3,4]: recordings were decomposed in a weighted sum of thousands of spectro-temporal exemplars of speech and noise and the noise components were suppressed. The speech recognizer yielded temporal alignments with differing offsets depending on the beginning phoneme of the response: these biases were compensated in the post-processing stage. The algorithm could handle synonyms and new pictures, since the word models were built by joining phoneme models automatically. As output a text file that summarized both accuracy and RT estimates was generated.

In patients, additionally the difference in RT with and without rTMS were computed. As proof of principle, we tested if slowing in reaction time could be used as a quantitative parameter to map language cortex-so trials with correct naming could be used to gauge if the target of stimulation was likely part of language cortex (no-response or incorrect trials were not used since only correctly recognized trials had measured RTs) Fig. 1. Since RTs can inherently vary from trial to trial, differences in RTs were extrapolated over neighboring cortical targets to average out variability, using a spherical area with FWHM of 17mm around each target. This simple averaging did retain information on individual brain anatomy and the intricacies between the tumor and the surrounding brain. The maps were compared with intra-operative mapping and post-operative language function.

The parameters of the automatic routine were set such that in the test data an accuracy of 90.4% with 96% correct detection of no-response events (= specificity) in the presence of TMS noise was reached. In the 3 patients, an overall accuracy of 71% and 96% specificity was obtained. The RT with TMS was on average 13 ms longer than without TMS (paired-t test of 165 observations over 3 patients, $p = 0.0001$).

Comparing DCS with TMS, all DCS-positive points were located in TMS-positive regions of the RT-derived maps. In patient 1 the positive DCS point was located in the zone that showed the largest slowing; also in patient 2 the DCS-positive point was in a region that showed TMS-RT-slowing (largest slowing was seen more anteriorly, not sampled with DCS). In patient 1 and 2, regions with TMS-RT-slowing were not resected, and no clear postoperative deficits were seen. In patient 3, a resection was performed of a DCS-negative region that did show clear slowing using TMS-RT-based mapping: this patient had a severe post-operative language deficit with perseverations, severe naming, reading and repetition problems and moderate comprehension deficits, that largely improved by month three.

We showed that our automated speech analysis method can handle voice recordings from TMS-based mapping and that correctly named words, which are much more common than language errors, could yield information, based on RTs. Our method is the first to our knowledge to yield both accuracy and RTs. A previously published method only identified RTs using an accelerometer over the laryngeal muscles: specificity was 71% for voice onset and non-responses were correctly identified in 88% of events [5]. Our automatic routine yielded an overall accuracy of 90.4% with 96% specificity in test data.

https://doi.org/10.1016/j.brs.2019.10.001
1935-861X/© 2019 Elsevier Inc. All rights reserved.

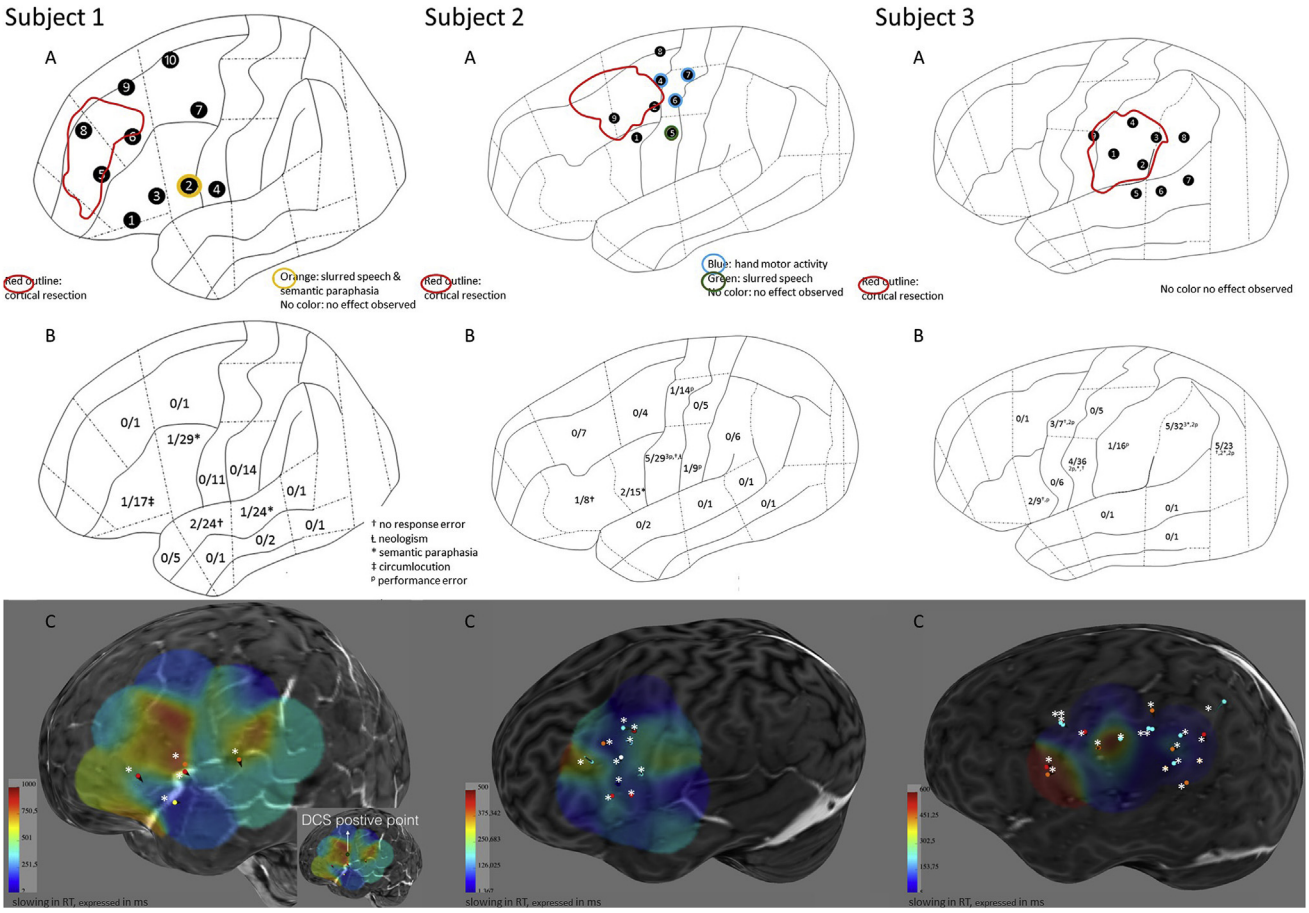

Fig. 1. TMS-based mapping data of the three subjects, compared to intra-operative findings
A. schematic representation of the patients' brain, with location of DCS points superimposed as numbered circles and the resection zone (as derived from post-operative MRI) as red outline. DCS points leading to an observable effect are circled, in blue for motor responses, in green for slurring and in orange for sematic paraphasias.
B. number of trials that were marked as incorrect by a human rater, as a fraction of the total number of trials in different anatomical subregions. No response errors (†), neologisms (Ł), circumlocutions errors (‡), semantic paraphasies (*) and performance errors (P) were observed.
C. TMS-RT-based color-coded map of the three patients in the study, with actual locations of incorrectly named trials-as determined by a human rater-superimposed as arrows (& marked with an extra white star for clarity): red arrow no response, white arrow neologism, yellow arrow circumlocution, blue arrow performance error and orange arrow semantic paraphasia. The maximal slowing in RT differed between patients and the absolute values of the color code (in ms) are given to the right of the respective images. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

As a proof of principle, TMS-RT-derived maps were created, which were in accordance with anatomical knowledge. In addition, all DCS-positive points were located in TMS-positive regions of the TMS-RT-derived maps, and no postoperative deficits were seen after surgeries where no TMS-RT-derived positive regions were resected whereas a severe but transient language deficit was seen after resection of a TMS-RT-derived positive but DCS negative region.

Our results are based on data of only three subjects and these preliminary conclusions need to be corroborated by studies on larger patient numbers. In future experiments, performance could be improved by avoiding naming of words beginning with voiced stop or non-sibilant fricative and to use a microphone with pop-filter.

We can thus conclude that the automated speech recognizer works with data generated during TMS noise. This study is also a proof of principle that RT measurements of correctly named words add relevant data for language mapping.

Acknowledgements

This work was supported by the Institute for innovation by science and technology Flanders (IWT) [TBM grant 09085].

References

[1] Duchateau J, Kong YO, Cleuren L, Latacz L, Roelens J, Samir A, et al. Developing a reading tutor: design and evaluation of dedicated speech recognition and synthesis modules. Speech Commun 2009;51(10):985–94.
[2] Krieg SM, Lioumis P, Mäkelä JP, Wilenius J, Karhu J, Hannula H, et al. Protocol for motor and language mapping by navigated TMS in patients and healthy volunteers; workshop report. Acta Neurochir 2017 Jul [cited 2017 Nov 19];159(7): 1187–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28456870.
[3] Baby D, Virtanen T, Gemmeke JF, van Hamme H. Exemplar-based noise robust automatic speech recognition using modulation spectrogram features. 2014. p. 519–24. December 2014.
[4] Baby D, Van Hamme H. Investigating modulation spectrogram features for deep neural network-based automatic speech recognition. Proc Annu Conf Int Speech Commun Assoc Interspeech 2015;2015-Janua(July):2479–83.
[5] Vitikainen A-M, Mäkelä E, Lioumis P, Jousmäki V, Mäkelä JP. Accelerometer-based automatic voice onset detection in speech mapping with navigated repetitive transcranial magnetic stimulation. J Neurosci Methods 2015. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0165027015001946.

Laura Seynaeve[1,*]
Laboratory for Epilepsy Research, Department of Neurosciences, KU Leuven, Herestraat 49 Box 7003, 3000 Leuven, Belgium

Deepak Baby[2], Hugo Van hamme
Department ESAT-PSI, KU Leuven, Kasteelpark Arenberg 10 Box 2441, 3001 Leuven, Belgium

Steven De Vleeschouwer
Department of Neurosurgery, UZ Leuven and Laboratory for Experimental Neurosurgery, Research Group Experimental Neurosurgery and Neuroanatomy, Department of Neurosciences, UZ Herestraat 49 - box 7003 48, 3000 Leuven, Belgium

Patrick Dupont
Laboratory for Cognitive Neurology, Department of Neurosciences, KU Leuven, O&N II Herestraat 49 - box 1027, 3000 Leuven Belgium

Wim Van Paesschen
Laboratory for Epilepsy Research, Department of Neurosciences, KU Leuven, Herestraat 49 Box 7003, 3000 Leuven, Belgium

Department of Neurology, UZ Leuven, Belgium

[*] Corresponding author.,
E-mail address: laura.seynaeve@uzbrussel.be (L. Seynaeve).

12 September 2019
Available online 11 October 2019

[1] current affiliation: department of neurology, UZ Brussels, Belgium.
[2] current affiliation: IDIAP Research Institute, Martigny, Switzerland.