The role of pre-SMA for time-critical speech perception – a transcranial magnetic stimulation (TMS) study

Susanne Dietrich, Florian Müller-Dahlhaus, Ulf Ziemann, Hermann Ackermann, Ingo Hertrich

Department of General Neurology, Center for Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

ABSTRACT

The pre-supplementary motor area (pre-SMA) is engaged in speech comprehension under difficult circumstances such as poor acoustic signal quality, cognitive load, or time-critical conditions. Previous studies found left pre-SMA activated when subjects listen to accelerated speech. Here we tested the functional role of the pre-SMA for accelerated speech comprehension by inducing a transient “virtual lesion” using continuous theta-burst stimulation (cTBS). Participants were tested (1) prior to (pre-baseline), (2) 10 min after (test condition for the cTBS effect), and (3) 60 min after stimulation (post-baseline) using a sentence repetition task (formant-synthesized at rates of 8, 10, 12, 14, and 16 syllables/s). Speech comprehension was quantified by the percentage of syllables in correctly reproduced words. For high speech rates, subjects showed decreased performance in the test conditions as compared to the baselines. This transient suppression of speech comprehension indicates that pre-SMA contributes to time-critical encoding of phonetic-linguistic information.

Keywords: cognitive control, motor system, virtual lesion, prediction, inner speech generation

1. INTRODUCTION

The supplementary motor area (SMA) can be subdivided into SMA proper and a more anterior part, i.e., the pre-SMA [1]. The former region seems to be primarily involved in motor control tasks as an interface for movement initiation and temporal triggering in case of, e.g., syllable repetitions [2]. By contrast, the pre-SMA is assumed to be associated with cognitive control functions beyond the motor domain [3]. For instance, pre-SMA was found to be involved in task switching [4], managing inhibitory mechanisms, e.g., in stop-signal tasks [5, 6], response selection processes [7, 8], and complex sequencing, e.g., coordination of prosody and syntax [9]. Regarding speech and language processing, pre-SMA was found to be relevant especially under conditions with high demands, e.g., under time-critical circumstances. A recent functional magnetic resonance imaging (fMRI) study suggested that the “bottleneck” for understanding accelerated speech is limited by frontal cortex functions rather than auditory processing, as indicated by activation of pre-SMA and inferior frontal gyrus (IFG) when speech rate reaches the limits of intelligibility [10]. Accordingly, a further fMRI study found increased left pre-SMA activation in individuals trained to comprehend ultra-fast speech at rates of 16-18 syllables per second (syl/s) [11]. Thereby, the authors suggested that pre-SMA is involved in the coordination of phonological-phonetic representations (left hemisphere) with syllable-prosodic event timing (right hemisphere), adjusting inner speech components to the timing of the incoming speech signal [12].

The hypothesized functional role of pre-SMA was tested in healthy subjects by means of continuous theta-burst transcranial magnetic stimulation (cTBS). This application has been introduced as a method to induce a transient "virtual lesion" for ca. 30 min, changing motor cortex excitability in an inhibitory way [13]. The stimulation site was determined on the basis of previous fMRI data [11]. Speech comprehension was measured with a sentence repetitions task (behavioural response) comprising formant-synthesized sentences at distinct syllable rates ranging from moderately fast (8 syl/s) to ultra-fast speech (16 syl/s). Suppression of pre-SMA was hypothesized to influence only speech comprehension at high speech rates.

2. METHODS

2.1. Participants

Thirty-eight adult volunteers participated in the study. Half of them underwent cTBS (mean age = 29.63, SD = 8.56), the remaining subjects served as a control group (no cTBS, age = 30.79, SD = 7.90). All participants were male, right-handed (Edinburgh handedness inventory), native German speakers, without hearing deficits as determined by means of an audiogram. The experimental procedures were approved by the ethics committee of the University of Tübingen.

2.2. Stimuli and procedure

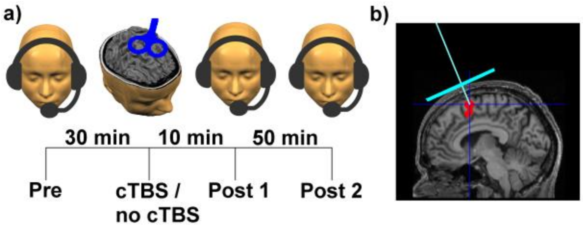
Participants performed sentence repetition tasks encompassing a set of 45 sentences of a length of 18 syllables each (approximately 10 words in order to limit memory load). The stimuli were based on newspaper materials and converted to speech by formant synthesis (“eloquence” implemented in the screen-reader software JAWS) at five distinct speech rates: 8, 10, 12, 14, and 16 syl/s (three subtests for baseline and test conditions, 9 items per rate condition within each subset). The sentences were played via headphones within a sound attenuated room, subjects being asked to repeat them “as good as possible” and “as fast as possible” after sentence offset, even when they failed to grasp all words. The subjects’ repetitions were digitally recorded and underwent subsequent quantitative evaluation of speech comprehension (percentage of syllables within correctly reproduced words, ignoring minor errors such deviant as plural endings). Participants performed the repetition task prior to cTBS (pre-stimulation baseline = pre), 10 min after cTBS (assumed suppression to the cTBS effect = post1), and 60 min after cTBS application (post-stimulation baseline = post2) (Fig. 1a). Prior to the experimental session, a set of 18 practice sentences was presented to the subjects to get acquainted with the test situation and the sound of the speech synthesizer. The three subsets of stimuli were rotated across participants with regard to the baseline (pre, post2) and test runs (post1). The time intervals between pre, post1, and post2 testing were equal for the experimental group (cTBS) and the control group (no cTBS).

2.3. cTBS stimulation

Application of cTBS was done with a conventional monophasic stimulation coil (Magstim company, UK). Similarly as reported by Huang et al. [13], 600 pulses were applied in a theta burst-pattern (burst of three pulses at 50 Hz, repeated every 200 ms for a duration of 40 s). Intensity was set at 120% resting motor threshold (RMT) for the abductor pollicus brevis muscle of the left hand. RMT was determined as the minimum stimulator output required to produce a motor-evoked potential (MEP) > 50 µV peak amplitude in at least 5 out of 10 consecutive trials.

Using a neuro-navigation system (Localite GmbH), the stimulation site (pre-SMA) was determined on the basis of fMRI data [11] (MNI coordinates -6, 9, 60) projected onto an individual anatomical MRI dataset for each subject. The coil was placed tangentially to the scalp (Fig. 1b).

Figure 1: (a) Experimental design, (b) stimulation site used for neuro-navigated cTBS (MNI group coordinate x = -6, y = 9, z = 60). The activation cluster resulted from a previous fMRI study [11].



2.4. Data analyses

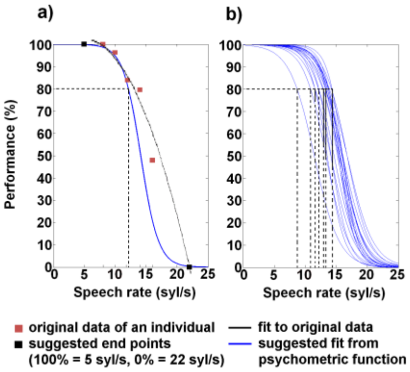
2.4.1. Descriptive overview

Descriptive statistics were used to show repetition performance and speech reaction time across syllable rates (8, 10, 12, 14, 16 syl/s). Speech reaction time was measured from acoustic stimulus offset to the acoustic onset of the subjects' repetitions.

2.4.2. Psychometric function

In order to obtain a single estimate of performance for each testing, the percentage of correct material was plotted against syllable rate, and a psychometric function [14, 15] was fitted to the data (Fig. 2a). Subsequently, the individual syllable rate at which subjects’ performance of speech comprehension amounts to 80% was determined from this function (Fig. 2b).

Figure 2: (a) Extraction of an individual’s psychometric function using five original data points (8, 10, 12, 14, 16 syl/s). (b) By means of the psychometric function, the individual syllable rate at which subjects’ performance of speech comprehension amounts 80% was calculated.



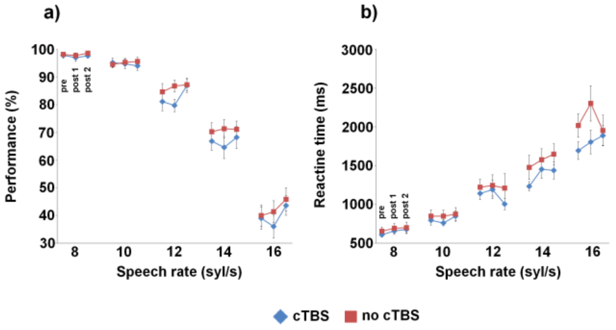
The 80% value was chosen because at this point speech comprehension is still present, but under time-critical circumstances, requiring the hypothesized function of pre-SMA. The transient cTBS effect, in comparison to the control group (no cTBS), was tested after the data were normalized to the pre-baseline performance. Statistics were conducted by means of a repeated measurements ANOVA with *Time* (pre/post1/post2) as a within-subject factor and *Condition* (cTBS/no cTBS) as a between subject factor. Post hoc, the cTBS effect (post1 testing) was tested by subtracting the post1-recordings from the mean of the pre- and post2-baselines (one-tailed *T*-test, hypothesizing impaired speech comprehension 10 min after cTBS.

3. RESULTS

3.1. Descriptive data on the effects of speech rate and cTBS

Speech comprehension declined at high syllable rates (Fig. 3a), concomitant with an increase of speech reaction time (Fig. 3b). Under the cTBS condition, a transient reduction of performance at syllable rates of 12 syl/s or faster occurs as a “dip” (blue lines in Fig. 3a), which is absent in the control group (red lines). Although less consistent, a transient increase of reaction time (peak) can be shown as well (blue lines in Fig. 3b).

Figure 3: (a) Performance of speech comprehension and (b) reaction time across the five syllable rates under the cTBS and control sessions during pre-, post1-, and post2 measurements.

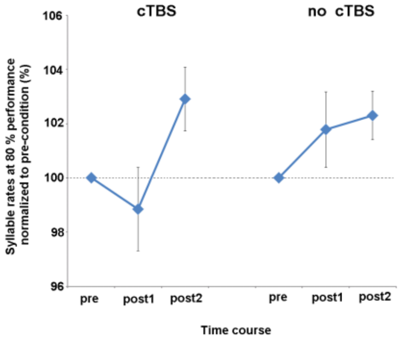


3.2. cTBS effect on speech comprehension under time-critical conditions

Regarding the experimental group, mean syllable rate at which subjects showed 80% correct sentence repetition was 12.79 syl/s (*SE* = 0.32) for the pre-stimulation baseline, 12.59 (0.29) for the post-1 test condition, and 13.14 (0.33) for the post2-baseline. The respective values for the control group were pre: 13.09, (0.32), post1: 13.27 (0.27), and post2: 13.37, SE 0.30).

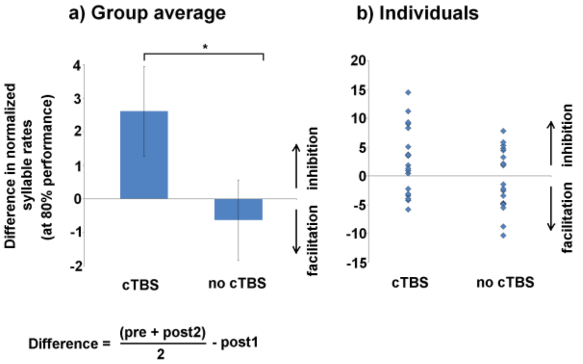
Repeated measures ANOVA yielded a significant main effect of the within-subject factor *Time* (pre/post1/post2; *F* [2, 35] = 5.006, *p* < .012, GG-corr.) Furthermore, the between-subjects factor *Condition* (cTBS versus no cTBS group) tentatively interacted with the quadratic trend of *Time*, (*F* [1, 36] = 3.274, *p* = 0.079), as indicated by a “dip” in Figure 4 for the cTBS group. Post hoc analysis (two-tailed paired *T* test) showed significant pre-post2 (*p* < .001) and post1-post2 differences (*p* < .019).

Figure 4: Normalized performance during pre-, post1-, and post2 measurements for the cTBS group and the control group.



Post hoc analysis of the two-way *Time* × *Condition* interaction revealed a significant group effect (*T* = 1.809, *p* < .039) on the difference measure (mean of pre and post2 minus post1) as illustrated in Figure 5a (individual data points are plotted in Fig. 5b).

Figure 5: Transient inhibition of pre-SMA by cTBS reduces speech comprehension. (a) Bar chart of the quadratic trend across the three time points, quantifying the “dip” in speech comprehension for the cTBS group (subtraction of post1 data from the mean of pre- and post2-baselines). (b) Individual subject data for the quadratic trend.



3. DISCUSSION

The aim of the present study was to test the functional role of the pre-SMA for speech comprehension under time-critical circumstances. As hypothesized, continuous theta-burst stimulation (cTBS), inducing a transient "virtual lesion" in the pre-SMA, resulted in reduced sentence repetition performance for fast syllable rates. Although single subjects yielded no or only weak effects after cTBS application, the group average clearly showed an impairment of speech perception. In previous studies, cTBS has mostly been applied to research on motor cortex, which can easily be tested by means of electromyography [13]. By contrast, the functional relevance of pre-SMA for speech comprehension can only be tested by a behavioral task, which might be more difficult. Nevertheless, the present data indicate that it is possible and that TMS methods can be applied to test higher cognitive functions.

In the present data, transient suppression of speech comprehension was only observed for high speech rates of 12 syl/s or faster. Accordingly, pre-SMA was found to be stronger activated during presentation of ultra-fast as compared to moderately fast speech [11, 16] and to be particularly active near the limit of intelligibility [10].

In addition to the cTBS effect in the post1 test, the present results showed an increase in correct sentence repetitions for the post2-baseline. Since this effect was present in both subject groups, it may indicate a training effect during the entire experiment rather than, e.g., representing a delayed facilitatory effect of cTBS.

Although some functional role of pre-SMA seems evident from the present data as well as previous fMRI data, its differential contribution to the entire process of speech processing should be considered more in detail. Presumably, time-critical speech perception can not totally be performed in a bottom-up mode. Utilizing the general redundancy in speech, the speech generation mechanism can make predictions for upcoming speech material in order to save time during lexical access. To these ends, the data stream of predictions has to be synchronized with the incoming signal. This synchronization can best be performed on the basis of the prosodic structure that is predominantly represented in the right hemisphere. Regarding the timing of motor events, the right hemisphere has an inhibitory control function on left-dominant forward action control, working as a kind of "brake" [17]. Similar control mechanisms may be present for predictive inner speech generation without any overt motor activity. Evidence for the involvement of SMA and pre-SMA in predictive language mechanisms has been provided in a review paper emphasizing the role of these areas as output regions from cerebellar-thalamic and basal ganglia-thalamic circuits [18].

Thus, the present study, showing that a virtual lesion of pre-SMA impairs speech comprehension under time-critical conditions, fits into recent theories considering the pre-SMA as a superordinate structure of cognitive control with regard to top-down aspects of language reception and speech processing.

4. REFERENCES

[1] Nachev, P., Kennard, C., & Husain, M. 2008. Functional role of the supplementary and pre-supplementary motor areas. *Nat Rev Neurosci 9*, 856-869.

[2] Brendel, B., Hertrich, I., Erb, M., Lindner, A., Riecker, A., Grodd, W., & Ackermann, H. 2010. The contribution of mesiofrontal cortex to the preparation and execution of repetitive syllable productions: An fMRI study. *NeuroImage 50*, 1219-1230.

[3] Chouinard, P. A., & Paus, T. 2010. What have we learned from 'perturbing' the human cortical motor system with transcranial magnetic stimulation? *Frontiers in Human Neuroscience 4*, 173.

[4] Kennerley, S. W., Sakai K., & Rushworth M. F. 2004. Organization of Action Sequences and the Role of the Pre-SMA. *Journal of Neurophysiology 91*, 978-993.

[5] Chen, C.-Y., Muggleton, N. G., Tzeng, O. J. L., Hung, D. L., & Juan, C.-H. 2009. Control of prepotent responses by the superior medial frontal cortex. *NeuroImage 44*, 537-545.

[6] Kwon, Y. H., & Kwon, J. W. (2013). Response Inhibition Induced in the Stop-signal Task by Transcranial Direct Current Stimulation of the Pre-supplementary Motor Area and Primary Sensoriomotor Cortex. *J. Physical Therapy Sci. 25*, 1083-1086.

[7] Tremblay, P., & Gracco, V. L. 2006. Contribution of the frontal lobe to externally and internally specified verbal responses: fMRI evidence. *NeuroImage 33*, 947-957.

[8] Tremblay, P., & Gracco, V. L. 2010. On the selection of words and oral motor responses: Evidence of a response-independent fronto-parietal network. *Cortex 46*, 15-28.

[9] Kotz, S. A., Schwartze, M. 2010. Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences* 14, 392-399.

[10] Vagharchakian, L., Dehaene-Lambertz, G., Pallier, C., Dehaene, S. 2012. A Temporal Bottleneck in the Language Comprehension Network. *Journal of Neuroscience* 32, 9089-9102.

[11] Dietrich, S., Hertrich, I., Ackermann, H. 2013. Training of ultra-fast speech comprehension induces functional reorganization of the central-visual

system in late-blind humans. *Front Hum Neurosci* 7, 701.

[12] Hertrich, I., Dietrich, S., Ackermann, H. 2013. How can audiovisual pathways enhance the temporal resolution of time-compressed speech in blind subjects? *Front Psychol* 4, 530.

[13] Huang, Y., Edwards, M. J., Rounis, E., Bhatia, K. P., Rothwell, J. C. 2005. Theta burst Stimulation of the human motor cortex. *Neuron* 45, 201–6.

[14] Wichmann, F. A., Hill, N. J. 2001. The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics* 63, 1293-1313.

[15] Wichmann, F. A., Hill, N. J. 2001. The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics* 63, 1314-1329.

[16] Dietrich, S., Hertrich, I., Ackermann, H. 2013. Ultra-fast speech comprehension in blind subjects engages primary visual cortex, fusiform gyrus, and pulvinar – a functional magnetic resonance imaging (fMRI) study. *BMC Neurosci* 14, 74.

[17] Aron, A. R., Robbins, T. W., & Poldrack, R. A. 2014. Inhibition and the right inferior frontal cortex: one decade on. *Trends in Cognitive Sciences 18*, 177-185.

[18] Schwartze, M., Tavano, A., Schröger, E., & Kotz, S. A. 2012. Temporal aspects of prediction in audition: Cortical and subcortical neural mechanisms. *International Journal of Psychophysiology, 83* 200-207.