Can recordings of tongue motor-evoked potentials be improved? Comparison of two electrode types and their reliability

Running title: **Tongue MEPs assessed with two electrode types**

William Poncin[1], Simple Futarmal Kothari[2,3], Jin Magara[4,5], Makoto Inoue[5], Eduardo Castrillon[6], Mohit Kothari[2,7]

Affiliations:

1. Pole of Pulmonology, ENT and Dermatology (LUNS), Institute of Experimental and Clinical Research (IREC), Université catholique de Louvain (UCLouvain), Brussels, Belgium.
2. Hammel Neurorehabilitation Centre and University Research Clinic, Department of Clinical Medicine, Aarhus University, Hammel, Denmark
3. Department of Clinical Neurophysiology, Aarhus University Hospital & Department of Clinical Medicine, Aarhus University, Aarhus, Denmark
4. Unit of Dysphagia Rehabilitation, Niigata University Medical & Dental Hospital, Niigata, Japan
5. Division of Dysphagia Rehabilitation, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan
6. Section for Orofacial Pain and Jaw Function, Department of Dentistry and Oral Health, HEALTH, Aarhus University, Denmark.
7. Department of Audiology & Speech Language Pathology, Kasturba Medical College, Mangalore, Manipal Academy of Higher Education, Manipal, India

Corresponding authors

- William Poncin, PT, PhD.
  - Institute of Experimental and Clinical Research (IREC), UCLouvain, Avenue Hippocrate 55, 1200, Brussels, Belgium. Email: william.poncin@uclouvain.be
- Mohit Kothari, BDS, MS, PhD.
  - Hammel Neurorehabilitation Centre and University Research Clinic, Voldbyvej 15, Hammel, Denmark-8450. E-mail: mohkot@rm.dk. Telephone: +45 7841 9544

Acknowledgments

The authors thank the European Respiratory Society who supported this work via a short-term research fellowship. We also thank all the participants for their participation in this study. The authors also wish to express their sincere gratitude to Professor Peter Svensson for his expert consultation and insightful guidance throughout the conduct of this study.

Conflict of interest statement

All authors declared no conflict of interest.

Source of support

This was a researcher-initiated study. We acknowledge the support of the European Respiratory Society (ERS) Fellowship Short-Term Research Fellowship 2023

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Number of Tables & Figures: 6 (2 & 4) / 6

Abstract word count: 250/250 words

Manuscript word count: 3656 words

Number of References: 42/50

ABSTRACT (250/250 words)

Background

In transcranial magnetic studies (TMS), motor-evoked potentials (MEPs) recorded from the tongue provide important insights into corticobulbar excitability and neuromuscular control. However, their clinical and research utility remains limited due to anatomical and technical challenges, particularly those affecting electrode stability and signal quality.

Objective

This study aimed to compare TMS-derived outcomes obtained with traditional surface electrodes and novel suction-based electrodes.

Methods

Nineteen healthy adults participated in a repeated-measures, cross-over design. On Day 1, each participant completed two test sessions per electrode type (intra-session reliability). Thirteen participants returned for a third test session at least one week later (inter-session reliability). Resting motor threshold (rMT) was determined. MEP amplitude (MEP_Amp) and latency (MEP_Lat) were assessed at five single-pulse stimulation intensities (90%, 100%, 120%, 140%, and 160% of rMT). Reliability was analyzed using intraclass correlation coefficients (ICCs, using lower 95% CI bound to gauge reliability strength).

Results

Suction electrodes yielded significantly lower rMT and higher MEP_Amp at 120-160% rMT. For both electrode types, intra-session reliability for rMT was excellent (ICCs >0.90) and poor-to-moderate for MEP_Amp and MEP_Lat (ICCs <0.75). Inter-session reliability was poor (ICCs <0.50) across all outcomes for both electrode types.

Conclusion

Compared to surface electrodes, suction electrodes improved signal quality as indicated by lower rMT and higher MEP_Amp. However, both electrode types exhibited substantial intra- and inter-session variability. Although suction electrodes enhance TMS signal quality, the poor test-retest reliability of tongue MEPs currently limits their utility as biomarkers. Methodological refinements are needed to further improve reliability in TMS-based tongue assessments.

Keywords (max 6): Transcranial magnetic stimulation, tongue, motor-evoked potentials, electrode, reliability, cortical plasticity

BACKGROUND

The human tongue plays a crucial role in essential functions such as speech, mastication, swallowing, and respiration[1,2]. Its functions rely on intricate sensorimotor integration with central neural networks[3], making it a valuable target for investigating neuromotor control and developing rehabilitation strategies for orofacial disorders. Transcranial magnetic stimulation (TMS) has emerged as a widely used, non-invasive tool to assess cortical excitability and neuroplasticity related to tongue function[4]. In TMS studies, motor-evoked potentials (MEPs) can be recorded from tongue muscles following stimulation of the motor cortex, allowing researchers to quantify changes in excitability and monitor training- or pathology-induced neuroplasticity. Notably, studies have shown that the tongue cortical representation can undergo rapid plastic changes in response to training[5-10], and that nociceptive input can interfere with these neuroplastic changes[5,11].

To gain insight into both basic neuroscience and clinical applications, accurate and reliable measurement of TMS-induced MEPs responses is of considerable relevance. High test-retest reliability is mandatory for achieving valid results. However, the literature indicates a high degree of variability in the TMS-induced MEPs[12-14]. This variability could result from technical factors such as participant positioning, as well as the orientation, location, and stability of the TMS coil[15,16]. When controlling for these factors, non-navigated TMS studies reported moderate to good test-retest reliability for resting motor threshold (rMT), MEP peak-to-peak amplitudes (MEP_Amp), or MEP stimulus response (MEP_SR) in different hand muscles or the tibialis anterior[17,18]. However, the reliability of tongue MEPs remains particularly challenging due to the complex anatomical and physiological features of the region, hindering their broader application. Intramuscular electrodes are considered the gold standard due to their signal specificity, but they are invasive, require technical expertise, and may not be tolerated in clinical populations[19]. In contrast, conventional disposable surface electrodes placed on the tongue offer ease of use, yet are uncomfortable for participants and do not allow clear differentiation between TMS-induced MEPs and other electrically evoked signals[20]. Furthermore, inadequate electrode contact, due to involuntary tongue movements or excessive salivation, can cause frequent electrode detachment[21,22], reducing the signal-to-noise ratio and compromising measurement validity.

Recently, a new generation of tongue electrodes has been manufactured. Ita and colleagues developed a suction electrode in which the interior space is maintained at negative pressure[22]. This design keeps the electrodes stable on the tongue surface, potentially improving MEP signal acquisition. Yet, to the best of our knowledge, no study has attempted to verify whether these new suction electrodes positively affect MEP_Amp or the reliability of TMS-derived metrics compared to the standard, more conventional disposable surface electrodes. In this study, we investigated whether there is a potential difference in the MEP_SR curves of the tongue muscle acquired using TMS and test-retest reliability of TMS-derived outcomes (rMT, MEP_Amp and onset latency [MEP_Lat]) between the new suction and traditional electrodes. We hypothesised that, compared to the traditional surface electrodes, the new suction electrodes would lead to more reliable TMS-derived outcomes. In addition, we expected them to produce lower rMT and a steeper MEP_SR curve (i.e., higher MEP_Amp for a given stimulus intensity).

METHODS

Participants

Twenty healthy participants (mean age 41 ± 9 (SD) years, 12 females) were recruited for the study at Hammel Neurorehabilitation Centre and University Research Clinic (HNRC), and at the Department of Dentistry and Oral Health, both affiliated to Aarhus University, Denmark. The Inclusion criteria were: adults ≥18 years of age, to be in a good general and oral health condition, and ability to read and understand the project information. Exclusion criteria were: participants with medical, physical or psychological disorders, allergy to electrode gel, metal implants in head, history of epilepsy, pacemaker/medicinal pumps and pregnancy. All participants gave their written informed consent prior to their participation. The study was approved by the Local Ethics Committee in Central Denmark Region Denmark (Journal number: 1-10-72-9-23 and amendment number: 110938) and was performed in accordance with the Declaration of Helsinki II.

Study design

A repeated-measures cross-over design consisting of two sessions scheduled on Day 1, and a third session scheduled at least one week later on Day 2, was conducted to compare the reliability and performance of two different electrode types for recording tongue MEPs: (1) 15x20mm self-adhesive disposable silver-chloride surface electrode (Cephalon, Nørresundby, Denmark) - hereafter referred to as "A" electrode, and (2) bipolar 6 mm in diameter silver cup electrodes TK221-006 connected with a suction tube through a silicone tube, which can produce negative pressure inside the electrodes (Unique Medical Co., Tokyo, Japan), hereafter referred to as "B" electrode[22] (Figure 1). Between Day 1 and Day 2, there was no carryover of hotspot coordinates or stimulus intensities (*i.e.* new values were established for the second day). All assessments were performed by the same investigator (WP).

Intra-session and inter-session reliability of TMS-derived outcomes were assessed for both electrode types. The intra-session reliability was assessed during the participants' first visit (Day 1). Each participant performed twice the same test procedures with each type of electrode following four consecutive stimulation blocks in either an ABAB or BABA sequence, depending on random assignment. In the ABAB condition, participants were tested using electrode type "A", followed by "B", then A and finally B; the reverse order was applied in the BABA condition. All four blocks (i.e. 2 sessions per electrode type) were completed on the same day, with rest periods between blocks of at least 2 min. Disposable electrodes type A were replaced by new ones whenever they came lose during measurements and the signal quality required it. Inter-session test-retest reliability was assessed by comparing the results obtained at the first test occasion on Day 1 with the results of outcomes acquired during the third session (Day 2). In this follow-up session, each electrode type (A or B) was tested only once, with the order again determined by randomization. Electrodes A were replaced as mentioned above whenever necessary. Condition allocation was based on a computer-generated random sequence, and the experimental procedures and setup were replicated across sessions to ensure methodological consistency.

Recording of motor-evoked potentials

Participants were positioned in a relaxed and reclined position on a bed plinth (Hammel recruitment site) or a dental chair (Aarhus recruitment site) with the backrest tilted at 45°[16]. A snugly fitting flexible head cap was placed over the participant's head in a standardized manner based on anatomical markers and following the International 10–20 Electrode Placement System guidelines[23]. The head was supported by a headrest and was slightly tilted to the right to facilitate optimal placement of the TMS coil and examiner visibility. Electromyographic (EMG) activity was recorded from the right dorsum of the tongue. The ground electrode was placed around the right clavicle bone. Regarding the tongue electrodes, because electrodes A and B differed in size, their placement was standardized using fixed anatomical landmarks rather than edge-to-edge distance. All electrode pairs were positioned longitudinally along the right dorsal tongue surface, with the center of the pair aligned over the midline of the right hemilingual region. The center of the anterior electrode was positioned 10 mm posterior to the tongue tip, and a center-to-center spacing of approximately 15-20 mm (depending on tongue size) was maintained between the anterior and posterior electrodes (Figure 1). Electrodes were connected for bipolar recording of MEPs. EMG signals were amplified, filtered (20 Hz–1 kHz), and stored using Sierra Software (Cephalon, Nørresundby, Denmark).

TMS was applied using a Magstim 200 stimulator (The Magstim Co. Ltd., UK; peak magnetic field = 2T) and a 5-cm angled figure-of-eight coil positioned over the left motor cortex[24,25]. The coil was oriented 45° obliquely to the sagittal midline, generating a current perpendicular to the estimated alignment of the central sulcus. At a constant stimulation intensity, the coil was systematically moved over the left hemisphere to identify the optimal scalp location – or tongue "hotspot" – defined as the site eliciting the largest MEP peak-to-peak response. Following standard procedures, the hotspot search began around anatomical landmarks delineated on the flexible head cap, based on prior studies identifying scalp regions where discernible tongue MEPs are most consistently detected[26,27]. The coil was then moved in all directions to determine the site where tongue MEPs could be elicited at the lowest stimulation intensity. This site was considered the hotspot. The rMT was then defined as the minimum stimulus intensity that produced at least 4 out of 10 discernible MEPs (*i.e.* greater than 50 µV) above the background EMG activity. The minimum stimulus intensity was determined by adjusting the intensity in 5% increments/decrements. The range of stimulation intensities corresponded to percentages of the TMS device's maximum output (from 0% to 100%). The interstimulus interval was between 8 and 12 s.

The MEP_SR curves were assessed by computing the peak-to-peak amplitude at 0.9, 1, 1.2, 1.4 and 1.6 times the rMT at the tongue hotspot (*i.e.* MT90, MT100, MT120, MT140 and MT160, respectively). The average of 8 stimuli applied with an interstimulus interval of 8–12 s was considered the MEP for each intensity of stimulation. A blinded investigator (SFK) retrieved the data stored in Sierra Software and computed the MEP_Lat and MEP_Amp of each individual trace. The averaged MEP_Lat and MEP_Amp were reported.

Statistical analyses

Data distributions were tested for normal distribution using Shapiro Wilk test and checked with histograms, Q-Q plots and boxplots. Depending on their distributions, variables were either expressed as median and interquartile range (IQR) or mean and standard deviation (SD). MEP_Amp data was log-transformed as it was not normally distributed.

Intra-session test-retest reliability was assessed on Day 1 by repeating the same test procedure twice with each electrode type during the initial visit. Reliability was analyzed by comparing the two tests for electrode A (A1 vs A2) and for electrode B (B1 vs B2). The inter-session test-retest reliability was assessed by comparing the first test performed on Day 1 (A1, B1) with the corresponding test on Day 2 (A3, B3), for each electrode type separately (A1 vs A3, B1 vs B3). Two methods were used to explore the variability of the data to gauge reliability. First, the Bland–Altman method was applied to assess the agreement of TMS-derived outcomes (MEP_Amp and MEP_Lat) across different stimulation intensities. Limits of agreement were calculated as the mean difference (bias) ± 1.96 times the standard deviation of the differences. Composite Bland–Altman plots were generated to illustrate how agreement varied with stimulus intensity. Second, intraclass correlation coefficients (ICC) were calculated using a two-way mixed model to verify the intra- and inter-session reliability of the TMS-derived outcomes. The ICC values were expressed with 95% confidence intervals (CI). Based on the lower bound of the 95% CI of the ICC estimate, ICC values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9 are indicative of poor, moderate, good, and excellent reliability, respectively[28].

A paired t-test was used to compare the average rMT values between electrode types (A vs. B), using the mean of the two repeated measurements collected for each electrode on Day 1. Finally, to evaluate the potential difference in MEP SR curves between both electrodes, log-transformed MEP_Amp data collected during the two test sessions at Day 1 were averaged and analyzed using two-way repeated measures ANOVA, with test stimulus intensity (MT90, MT100, MT120, MT140 and MT160) and electrode types (A or B) as within-subject factors. Bonferroni correction was applied for any ANOVA post-hoc comparisons.

Finally, the number of standard self-adhesive electrodes replaced per participant before terminating a test block with acceptable signal quality was counted and reported.

Data were analysed using SPSS 28.0 (IBM software). P values < 0.05 were considered as statistically significant.

RESULTS

Out of 20 enrolled participants on Day 1, data from 19 participants was analyzed, as one participant was excluded from the study due to an excessively high rMT (80% of maximal stimulator output at the tongue hotspot). The demographic characteristics of the 19 remaining participants were as follows: 11 females (58%); mean age: 41 ± 9 (SD) years; mean body mass index: 24.1 ± 3.7 kg/m^2; 17 out of 19 participants (89%) were right hand dominant. Out of the 19 participants that participated on Day 1, 13 participants (68%) also participated on Day 2.

Intra-session reliability

Detailed intra-session reliability data are presented in Table 1 and Figure 2. No significant intra-session test-retest differences were observed for either electrode type across any TMS-derived outcomes, except for MEPs_Lat at MT160 with electrode B (new suction electrode) which was shorter at the second test (B2) (Fig. 2A, 2B). The intra-session test-retest reliability of the rMT, as expressed by the lower bound of the 95% CI of the ICC estimate, was excellent for both electrode types (Table 1). The ICCs for MEPs_Lat indicated poor reliability with both electrodes A and B. The ICCs for MEPs_Amp also indicated poor reliability with both electrodes A and B, except at MT140 where it was good for electrode B.

Inter-session reliability

Detailed inter-session reliability data are presented in Table 2 and Figure 2.

Regarding electrode A (silver-chloride surface electrode), no significant inter-session test-retest differences were observed for rMT (mean difference ± SD: -5 ± 9.1 %, 95% CI [-0.5 to 10.5], p = 0.072). In contrast, most MEP_Lat and MEP_Amp outcomes were significantly different between Day 1 and Day 2 (Fig. 2C, 2D). The lower bound of 95% CI of ICC estimates indicated poor to no reliability for any TMS-derived outcomes (Table 2).

Regarding electrode B (suction electrode), there was a significant difference in rMT between both days, rMT being higher at Day 2 (mean difference ± SD: -6.2 ± 10.0 %, 95% CI [0.1 to 12.2], p = 0.047). MEP_Lat also differed significantly at any stimulator intensity output. In contrast, there was no significant difference in MEP_Amp at any stimulator intensity output (Fig. 2C, 2D). As for electrode A, the lower bound of 95% CI of ICC estimates indicated poor to no reliability for any TMS-derived outcomes collected using electrode B (Table 2).

Comparison of resting Motor Threshold between both electrodes A & B

The average rMT values collected with each electrode type at Day 1 was used for this comparison. There was a significant difference between electrode types, with higher rMT values observed for electrode A (mean difference 2.4 ± 3.1%, P=0.003; Fig. 3).

Comparison of stimulus-response curves between both electrodes A & B

The tongue MEPs were significantly dependent on stimulus intensity (F=147.6; P<0.001) and on type of electrode (F=8.5; P=0.009), with significant interaction between these factors (F=3.5; P=0.021). Post-hoc analyses revealed that this interaction was driven by significantly higher tongue MEPs with electrode B compared to electrode A at higher stimulation intensities, specifically at MT120, MT140, and MT160 (P<0.01; Fig. 4).

Number of self-adhesive electrodes replaced

During the two stimulation blocks using electrodes A on Day 1, a total of 17 electrode pairs were replaced, corresponding to an average of 0.9 ± 1.0 pairs per participant (range: 0-4). On Day 2, 10 electrode pairs were replaced, with a mean of 0.8 ± 1.0 pairs per participant (range: 0-3).

DISCUSSION

This study evaluated the performance and test–retest reliability of two electrode types (traditional disposable silver-chloride electrodes and novel suction-based surface electrodes) for recording tongue MEPs induced by TMS. While the suction electrodes demonstrated some advantages in terms of signal amplitude and rMT, the overall reliability of tongue MEPs remained suboptimal for both electrode types, either intra- or inter-session.

The suction-based electrodes yielded significantly higher MEP_Amp at higher stimulation intensities (120–160% rMT) with no systematic bias between inter-session measurements, as well as lower rMT values compared to traditional surface electrodes, suggesting a potential advantage of the former electrode design in eliciting robust responses at certain intensities[16,29]. These performance gains primarily reflect improved signal quality rather than enhanced measurement reliability, likely resulting from better electrode-tissue contact and reduced signal attenuation due to the negative pressure mechanism stabilizing the electrode on the moist and mobile tongue surface[22]. This is consistent with previous reports highlighting the challenges of using standard self-adhesive electrodes for tongue MEPs, as they are often uncomfortable and poorly tolerated[20]. Saliva accumulation and the need to suppress swallowing to avoid electrode displacement can impair their adherence and compromise signal acquisition[20,21]. Accordingly, on average, one electrode pair per participant had to be replaced to maintain signal quality. However, in some cases, up to four pairs were required, despite instructions for participants to remain relaxed throughout the measurements. Although not formally quantified, this led to a notable increase in preparation time and participant burden, as multiple adjustments were needed to obtain reliable signals with traditional self-adhesive electrodes. In contrast, the suction electrodes demonstrated greater stability, adhering more effectively to the tongue and remaining in place even during swallowing. Experimenters also consistently noted that identifying the tongue hotspot required fewer stimulations and was achieved more rapidly with suction electrodes than with the self-adhesive electrodes. These observational findings should however be confirmed in future studies with quantitative timing data. Taken together, the improved signal quality, faster hotspot localization, and reduced need for readjustment make the suction electrodes more suitable for use in research settings. However, it should be noted that additional comparisons, such as evaluating the signal-to-noise ratio, could have further substantiated the observed benefits of the suction design and provided a more comprehensive assessment of signal quality.

Despite these advantages, both electrode types demonstrated poor intra- and inter-session reliability across all TMS-derived outcomes, with the exception of intra-session reliability for rMT, which was excellent (ICC > 0.9). At near-threshold intensities (90–100% rMT), MEPs amplitudes were typically smaller and closer to background EMG noise, leading to lower signal-to-noise ratios and decreased reliability. In contrast, responses at higher intensities (120–160% rMT) were more robust and consistent, which likely contributed to the improved reliability observed in this range. This pattern aligns with previous evidence showing that MEPs near threshold are inherently more variable due to fluctuating corticospinal excitability[30]. Future work should also aim to refine acquisition protocols around motor threshold intensities, for instance, by omitting the 90% rMT condition and instead increasing the number of stimuli at 100% rMT and/or applying signal averaging methods to enhance waveform quality and reliability in this lower stimulation range.

Similar findings have been reported in studies investigating inter-session reliability of TMS outcomes in more accessible, limb muscles, such as the extensor digitorum communis[31] and the first dorsal interosseous (FDI) muscle[32,33], where MEP amplitude and latency also exhibited poor reliability. However, these challenges appear even more pronounced when assessing the tongue, likely due to its complex anatomy, deep location, and fine motor control. Additionally, like other respiratory muscles, tongue muscles display both tonic and phasic excitability fluctuations during quiet breathing to maintain upper airway patency[34,35]. As respiration was neither controlled nor monitored during TMS delivery, cortical or hypoglossal excitability may have varied depending on the respiratory phase at the time of stimulation. This physiological variability could have introduced significant noise and contributed to signal instability across trials. Future studies should consider synchronizing TMS pulses with respiratory phase or incorporating respiratory monitoring to mitigate this variability. It should also be noted that we adopted a conservative approach in our reliability analysis by reporting the lower bound of the 95% confidence interval for the ICC[36], whereas most TMS reliability studies typically report only the point estimate[31,33,37]. This methodological choice may have contributed to the more cautious interpretation of our results.

The poor inter-session reliability raises concerns about the utility of tongue MEPs as biomarkers for longitudinal studies or therapeutic monitoring. Low ICC values indicate that within-subject variability exceeded between-subject variability[38], suggesting that day-to-day fluctuations in physiological or technical factors may overshadow true neurophysiological changes. From a technical standpoint, the absence of neuro navigation system likely resulted in small variations in coil positioning and angulation between sessions, which are known to significantly affect MEP characteristics [15,39]. In addition, the mean participant age was 41 years (range 26 to 55 years), and age-related changes in corticospinal excitability may have contributed to greater variability in MEP amplitude and latency[39]. Lastly, participants' hydration status has also been shown to influence brain excitability[40]. Given the moist environment of the tongue, this effect may have been amplified, potentially exerting a disproportionate impact on TMS outcomes. While this variability currently limits the clinical applicability of assessing tongue MEPs, they remain a promising non-invasive tool for assessing change in cortical excitability with orofacial muscle training[7,26] and could prove useful for monitoring neuroplasticity associated with aging, stroke or neurodegenerative diseases[41,42], provided that methodological refinements enhance their reliability.

This study has several limitations. First, to reduce testing time, the rMT was determined using 5% increments or decrements in stimulator output. While this approach improved feasibility, it may have limited the precision of rMT estimates and minimized the apparent differences between electrode types. Second, rMT was defined as the minimum stimulus intensity that elicited at least 4 out of 10 MEPs exceeding 50 µV above baseline EMG activity. This 50 µV threshold is higher than that used in some prior studies, which have applied a 5 µV cutoff[7,9,10,16,24,26]. While our more conservative criterion ensures that recorded MEPs reflect robust corticobulbar activation, it also highlights the comparatively low rMT values obtained with the suction electrodes. Third, variability in individual anatomical and physiological factors, such as tongue size, salivation, and participant cooperation, may have influenced electrode positioning and signal acquisition, despite efforts to standardize procedures. Lastly, no a priori sample size calculation was performed, as this was an exploratory, methodological reliability study aimed at assessing the feasibility and variability of TMS-derived tongue outcomes using two electrode configurations. However, the sample size was consistent with previous TMS reliability studies realized on limb muscles[31-33].

CONCLUSION

In conclusion, while suction-based electrodes offer practical advantages in eliciting stronger MEPs and lower rMT, the overall poor test–retest reliability of tongue MEPs, particularly across sessions, limits their current utility as robust biomarkers for longitudinal studies. These findings underscore the need for methodological refinements and standardization in TMS protocols targeting the tongue, including the integration of neuronavigation, respiratory phase monitoring, and hydration control. Addressing these challenges is critical for advancing the clinical and translational use of TMS in orofacial neuroscience and rehabilitation.

REFERENCES

1. Hiiemae KM, Palmer JB. Tongue movements in feeding and speech. *Crit Rev Oral Biol Med*. 2003;14(6):413-29. doi:10.1177/154411130301400604
2. Sawczuk A, Mosier KM. Neural control of tongue movement with respect to respiration and swallowing. *Crit Rev Oral Biol Med*. 2001;12(1):18-37. doi:10.1177/10454411010120010101
3. Maezawa H. Cortical Mechanisms of Tongue Sensorimotor Functions in Humans: A Review of the Magnetoencephalography Approach. *Front Hum Neurosci*. 2017;11:134. doi:10.3389/fnhum.2017.00134
4. Bono D, Belyk M, Longo MR, Dick F. Beyond language: The unspoken sensory-motor representation of the tongue in non-primates, non-human and human primates. *Neurosci Biobehav Rev*. Aug 2022;139:104730. doi:10.1016/j.neubiorev.2022.104730
5. Boudreau S, Romaniello A, Wang K, Svensson P, Sessle BJ, Arendt-Nielsen L. The effects of intra-oral pain on motor cortex neuroplasticity associated with short-term novel tongue-protrusion training in humans. *Pain*. Nov 2007;132(1-2):169-78. doi:10.1016/j.pain.2007.07.019
6. Boudreau SA, Lontis ER, Caltenco H, et al. Features of cortical neuroplasticity associated with multidirectional novel motor skill training: a TMS mapping study. *Exp Brain Res*. Apr 2013;225(4):513-26. doi:10.1007/s00221-012-3391-2
7. Svensson P, Romaniello A, Wang K, Arendt-Nielsen L, Sessle BJ. One hour of tongue-task training is associated with plasticity in corticomotor control of the human tongue musculature. *Exp Brain Res*. Aug 2006;173(1):165-73. doi:10.1007/s00221-006-0380-3
8. Kothari M, Stubbs PW, Figlewski K, et al. Effect of transcranial direct current stimulation on neuroplasticity in corticomotor pathways of the tongue muscles. *J Oral Rehabil*. Sep 2017;44(9):691-701. doi:10.1111/joor.12529
9. Kothari M, Svensson P, Jensen J, et al. Training-induced cortical plasticity compared between three tongue-training paradigms. *Neuroscience*. Aug 29 2013;246:1-12. doi:10.1016/j.neuroscience.2013.04.040
10. Kothari M, Liu X, Baad-Hansen L, Kumar A, Bin G, Svensson P. Influence of visual observational conditions on tongue motor learning. *Eur J Oral Sci*. Dec 2016;124(6):534-539. doi:10.1111/eos.12299
11. Svensson P, Miles TS, McKay D, Ridding MC. Suppression of motor evoked potentials in a hand muscle following prolonged painful stimulation. *Eur J Pain*. 2003;7(1):55-62. doi:10.1016/s1090-3801(02)00050-2
12. Kiers L, Cros D, Chiappa KH, Fang J. Variability of motor potentials evoked by transcranial magnetic stimulation. *Electroencephalogr Clin Neurophysiol*. Dec 1993;89(6):415-23. doi:10.1016/0168-5597(93)90115-6
13. Hashemirad F, Zoghi M, Fitzgerald PB, Jaberzadeh S. Reliability of Motor Evoked Potentials Induced by Transcranial Magnetic Stimulation: The Effects of Initial Motor Evoked Potentials Removal. *Basic Clin Neurosci*. Jan 2017;8(1):43-50. doi:10.15412/j.Bcn.03080106
14. Kothari M, Stubbs PW, Pedersen AR, Jensen J, Nielsen JF. Reliability of surface electromyography measurements from the suprahyoid muscle complex. *J Oral Rehabil*. Sep 2017;44(9):683-690. doi:10.1111/joor.12537
15. Chipchase L, Schabrun S, Cohen L, et al. A checklist for assessing the methodological quality of studies using transcranial magnetic stimulation to study the motor system: an

international consensus study. *Clin Neurophysiol*. Sep 2012;123(9):1698-704. doi:10.1016/j.clinph.2012.05.003

16. Kothari M, Svensson P, Nielsen JF, Baad-Hansen L. Influence of position and stimulation parameters on intracortical inhibition and facilitation in human tongue motor cortex. *Brain Res*. Apr 4 2014;1557:83-9. doi:10.1016/j.brainres.2014.02.017

17. Malcolm MP, Triggs WJ, Light KE, Shechtman O, Khandekar G, Gonzalez Rothi LJ. Reliability of motor cortex transcranial magnetic stimulation in four muscle representations. *Clin Neurophysiol*. May 2006;117(5):1037-46. doi:10.1016/j.clinph.2006.02.005

18. Cacchio A, Paoloni M, Cimini N, et al. Reliability of TMS-related measures of tibialis anterior muscle in patients with chronic stroke and healthy subjects. *J Neurol Sci*. Apr 15 2011;303(1-2):90-4. doi:10.1016/j.jns.2011.01.004

19. Chowdhury RH, Reaz MB, Ali MA, Bakar AA, Chellappan K, Chang TG. Surface electromyography signal processing and classification techniques. *Sensors (Basel)*. Sep 17 2013;13(9):12431-66. doi:10.3390/s130912431

20. Kaczmarczyk I, Rawji V, Rothwell JC, Hodson-Tole E, Sharma N. Comparison between surface electrodes and ultrasound monitoring to measure TMS evoked muscle contraction. *Muscle Nerve*. May 2021;63(5):724-729. doi:10.1002/mus.27192

21. Whillier A, Hommel S, Neef NE, Wolff von Gudenberg A, Paulus W, Sommer M. Adults who stutter lack the specialised pre-speech facilitation found in non-stutterers. *PLoS One*. 2018;13(10):e0202634. doi:10.1371/journal.pone.0202634

22. Ita R, Magara J, Ochiai Y, Tsujimura T, Inoue M. Use of suction electrodes for measurement of intrinsic tongue muscular endurance during lingual pressure generation. *J Oral Rehabil*. Mar 2024;51(3):566-573. doi:10.1111/joor.13612

23. Acharya JN, Hani A, Cheek J, Thirumala P, Tsuchida TN. American Clinical Neurophysiology Society Guideline 2: Guidelines for Standard Electrode Position Nomenclature. *J Clin Neurophysiol*. Aug 2016;33(4):308-11. doi:10.1097/wnp.0000000000000316

24. Komoda Y, Iida T, Kothari M, et al. Repeated tongue lift movement induces neuroplasticity in corticomotor control of tongue and jaw muscles in humans. *Brain Res*. Nov 19 2015;1627:70-9. doi:10.1016/j.brainres.2015.09.016

25. Costa YM, Exposto FG, Kothari M, et al. Masseter corticomotor excitability is decreased after intramuscular administration of nerve growth factor. *Eur J Pain*. Oct 2019;23(9):1619-1630. doi:10.1002/ejp.1438

26. Svensson P, Romaniello A, Arendt-Nielsen L, Sessle BJ. Plasticity in corticomotor control of the human tongue musculature induced by tongue-task training. *Exp Brain Res*. Sep 2003;152(1):42-51. doi:10.1007/s00221-003-1517-2

27. Costa YM, Hayakawa H, Castrillon EE, et al. Impact of a mandibular advancement device on corticomotor plasticity in patients with obstructive sleep apnea. *J Oral Rehabil*. Dec 2024;51(12):2600-2610. doi:10.1111/joor.13846

28. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. Jun 2016;15(2):155-63. doi:10.1016/j.jcm.2016.02.012

29. Kothari M, Baad-Hansen L, Svensson P. Bilateral sensory deprivation of trigeminal afferent fibres on corticomotor control of human tongue musculature: a preliminary study. *J Oral Rehabil*. Sep 2016;43(9):656-61. doi:10.1111/joor.12414

30. Pellegrini M, Zoghi M, Jaberzadeh S. The effect of transcranial magnetic stimulation test intensity on the amplitude, variability and reliability of motor evoked potentials. *Brain Res*. Dec 1 2018;1700:190-198. doi:10.1016/j.brainres.2018.09.002
31. Kraus D, Gharabaghi A. Neuromuscular Plasticity: Disentangling Stable and Variable Motor Maps in the Human Sensorimotor Cortex. *Neural Plast*. 2016;2016:7365609. doi:10.1155/2016/7365609
32. Salihu ATU, Hill KDU, Zoghi MU, Jaberzadeh SU. Examining the Intra-rater Reliability of Transcranial Magnetic Stimulation (TMS)-Induced Motor Evoked Potentials (MEPs) Within and Between Sessions: A Step Towards Ensuring Accuracy of Observed MEP Changes in Repeated Measures Studies conducted by Newly Trained TMS Operators. *Middle East J Rehabil Health Stud*. 2023:e138723. doi:10.5812/mejrh-138723
33. Davila-Pérez P, Jannati A, Fried PJ, Cudeiro Mazaira J, Pascual-Leone A. The Effects of Waveform and Current Direction on the Efficacy and Test-Retest Reliability of Transcranial Magnetic Stimulation. *Neuroscience*. Nov 21 2018;393:97-109. doi:10.1016/j.neuroscience.2018.09.044
34. Fregosi RF. Respiratory related control of hypoglossal motoneurons--knowing what we do not know. *Respir Physiol Neurobiol*. Oct 15 2011;179(1):43-7. doi:10.1016/j.resp.2011.06.023
35. Saboisky JP, Butler JE, Fogel RB, et al. Tonic and phasic respiratory drives to human genioglossus motoneurons during breathing. *J Neurophysiol*. Apr 2006;95(4):2213-21. doi:10.1152/jn.00940.2005
36. Borg DN, Bach AJE, O'Brien JL, Sainani KL. Calculating sample size for reliability studies. *PM R*. Aug 2022;14(8):1018-1025. doi:10.1002/pmrj.12850
37. Weiss C, Nettekoven C, Rehme AK, et al. Mapping the hand, foot and face representations in the primary motor cortex - retest reliability of neuronavigated TMS versus functional MRI. *Neuroimage*. Feb 1 2013;66:531-42. doi:10.1016/j.neuroimage.2012.10.046
38. Müller R, Büttner P. A critical discussion of intraclass correlation coefficients. *Stat Med*. Dec 15-30 1994;13(23-24):2465-76. doi:10.1002/sim.4780132310
39. Forster MT, Limbart M, Seifert V, Senft C. Test-retest reliability of navigated transcranial magnetic stimulation of the motor cortex. *Neurosurgery*. Mar 2014;10 Suppl 1:51-5; discussion 55-6. doi:10.1227/neu.0000000000000075
40. Critzer SS, Bosch TJ, Fercho KA, Scholl JL, Baugh LA. Water and brain function: effects of hydration status on neurostimulation with transcranial magnetic stimulation. *J Neurophysiol*. Sep 1 2024;132(3):791-807. doi:10.1152/jn.00143.2023
41. Cheng IKY, Yiu EML, Chan KMK. Changes in resting motor threshold of the tongue with normal aging and stroke. *Somatosens Mot Res*. Dec 2017;34(4):242-247. doi:10.1080/08990220.2017.1421159
42. Bocci T, Briscese L, Giorli E, Pecori C, Sartucci F. Tongue's motor evoked potentials in the diagnosis of primary lateral sclerosis (PLS): preliminary report. *J Neurol Sci*. May 15 2012;316(1-2):67-71. doi:10.1016/j.jns.2012.01.028

TABLES

Table 1. Intra-session test-retest reliability of TMS-derived outcomes for each type of electrodes.

N=19	Electrodes A (disposable silver-chloride surface electrodes)			Electrodes B (reusable suctioning tongue electrodes)		
	A1	A2	ICC (95% CI)	B1	B2	ICC (95% CI)
rMT (%)	41.1 (10.7)	40.3 (11.4)	0.96 (0.91; 0.99)	38.2 (10.8)	38.4 (9)	0.98 [0.94; 0.99]
MEP_Lat (ms)						
MT90	10.7 (0.8)	10.8 (0.8)	0.78 [0.40 - 0.92]	10.6 (0.9)	10.7 (0.8)	0.67 [0.12 - 0.88]
MT100	10.8 (0.7)	10.8 (0.8)	0.77 [0.39 - 0.91]	10.8 (0.8)	10.8 (0.9)	0.48 [-0.41 - 0.8]
MT120	11.0 (0.8)	10.9 (0.8)	0.79 [0.45 - 0.92]	10.7 (0.6)	10.7 (0.7)	0.85 [0.62 - 0.94]
MT140	10.8 (0.8)	10.8 (0.7)	0.81 [0.50 - 0.93]	10.7 (0.7)	10.7 (0.8)	0.78 [0.42 - 0.92]
MT160	10.8 (0.8)	10.8 (0.7)	0.76 [0.35 - 0.91]	10.9 (0.9)	10.6 (0.7)	0.73 [0.31 - 0.89]
Log MEP_Amp (uV)						
MT90	1.5 [1.4; 1.8]	1.5 [1.4; 1.6]	0.68 [0.14; 0.88]	1.7 (0.2)	1.6 (0.3)	-0.86 [-4.32; 0.32]
MT100	1.7 (0.3)	1.8 (0.2)	0.48 [-0.24; 0.79]	1.8 (0.3)	1.8 (0.2)	0.09 [-1.52; 0.66]
MT120	2.1 (0.3)	2.2 (0.3)	0.55 [-0.07; 0.82]	2.4 (0.3)	2.3 (0.3)	0.44 [-0.47; 0.78]
MT140	2.5 (0.3)	2.5 (0.4)	0.73 [0.27; 0.90]	2.7 (0.4)	2.7 (0.4)	0.90 [0.75; 0.96]
MT160	2.8 (0.4)	2.7 (0.4)	0.77 [0.40; 0.91]	3.0 [2.9; 3.2]	3.1 [2.9; 3.2]	0.81 [0.48; 0.93]

Data are presented as mean (SD) or median [IQR]. Paired t test was used, unless otherwise specified.

A1 and B1 refer to data collected during session 1 on Day 1 using disposable surface electrodes and suction-based electrodes, respectively; A2 and B2 correspond to data from session 2 on Day 1.

Abbreviations: Amp, amplitude; CI, confidence intervals; ICC, intraclass correlation coefficient; Lat, onset latency; MEP, motor-evoked potential; MT, stimulus intensity set at a given power output corresponding to the numbers; rMT, resting motor threshold; TMS, transcranial magnetic stimulation.

Table 2. Inter-session test-retest reliability of TMS-derived outcomes for each type of electrodes.

N=13	Electrodes A (disposable silver-chloride surface electrodes)			Electrodes B (reusable suctioning tongue electrodes)		
	A1	A3	ICC (95% CI)	B1	B3	ICC (95% CI)
rMT (%)	39.2 (11.3)	44.2 (10.4)	0.75 [0.23; 0.92]	36.2 (11.6)	42.3 (10.3)	0.68 [0.05; 0.90]
MEP_Lat (ms)						
MT90	10.7 (0.9)	9.9 (0.7)	-0.56 [-2.16; 0.44]	10.6 (1)	10.0 (0.6)	0.48 [-0.35; 0.83]
MT100	10.9 (0.7)	10.2 (0.8)	-0.47 [-2.22; 0.48]	10.8 (0.9)	10.2 (0.7)	0.27 [-0.71; 0.75]
MT120	11.0 (1.0)	10.1 (0.7)	-0.09 [-1.00; 0.56]	10.6 (0.6)	10.1 (0.5)	0.52 [-0.25; 0.84]
MT140	10.6 (0.8)	10.3 (0.6)	0.23 [-1.22; 0.75]	10.7 (0.8)	10.1 (0.5)	0.50 [-0.26; 0.83]
MT160	10.8 (0.8)	10.1 (0.7)	0.12 [-0.94; 0.69]	11.0 (1.1)	10.0 (0.5)	0.20 [-0.46; 0.68]
Log MEP_Amp (uV)						
MT90	1.7 (0.4)	1.6 (0.2)	-1.50 [-11.8; 0.31]	1.7 (0.2)	1.5 (0.2)	-1.49 [-7.37; 0.26]
MT100	1.7 (0.3)	1.9 (0.3)	0.53 [-0.25; 0.85]	1.8 (0.3)	2.0 (0.4)	0.21 [-1.12; 0.74]
MT120	2.1 (0.3)	2.5 (0.2)	0.57 [-0.26; 0.88]	2.4 (0.3)	2.6 (0.4)	0.11 [-1.20; 0.70]
MT140	2.5 (0.3)	2.7 (0.3)	0.63 [-0.07; 0.88]	2.6 (0.5)	2.9 (0.4)	0.41 [-0.72; 0.82]
MT160	2.7 (0.4)	3.0 (0.3)	0.60 [-0.17; 0.87]	3.0 [2.9 ; 3.2]	3.1 [2.7 ; 3.3]	0.65 [-0.18; 0.90]

Data are presented as mean (SD) or median [IQR]. Paired t test was used, unless otherwise specified.
A1 and B1 refer to data collected during session 1 on Day 1 using disposable surface electrodes and suction-based electrodes, respectively; A3 and B3 correspond to data collected on Day 2 scheduled at least one week after Day 1.
Abbreviations: Amp, amplitude; CI, confidence intervals; ICC, intraclass correlation coefficient; Lat, onset latency; MEP, motor-evoked potential; MT, stimulus intensity set at a given power output corresponding to the numbers; rMT, resting motor threshold; TMS, transcranial magnetic stimulation

FIGURES

Figure 1: Electrodes placement

Figure legend: The figure illustrates the placement of self-adhesive disposable silver-chloride surface electrode (left, referred to as “A” electrode in the main text), and silver cup electrodes connected to a suction tube to create negative pressure inside the electrodes (right, referred to as “B” electrode). Because electrodes A and B differed in size, their placement was standardized using fixed anatomical landmarks rather than edge-to-edge distance. All electrode pairs were positioned longitudinally along the right dorsal tongue surface, with the center of the pair aligned over the midline of the right hemilingual region. The center of the anterior electrode was positioned 10 mm posterior to the tongue tip, and a center-to-center spacing of approximately 15-20 mm (depending on tongue size) was maintained between the anterior and posterior electrodes.

Figure 2: Intra- and inter-session agreement of MEPs latency and amplitude

Figure legend: Composite Bland–Altman plots illustrating how intra- and inter-session agreement of motor-evoked potential (MEP) latency and amplitude vary with transcranial magnetic stimulus output intensity. (A) Intra-session agreement for MEP latency. (B) Intra-session agreement for MEP amplitude. (C) Inter-session agreement for MEP latency. (D) Inter-session agreement for MEP amplitude. Filled symbols represent data obtained with conventional disposable surface electrodes; empty symbols represent data obtained with suction-based electrodes. Circles and error bars indicate the mean difference (bias) and its 95% confidence interval. Lower triangles represent the lower limits of agreement; upper triangles represent the upper limits of agreement.

Figure 3: Comparison of resting motor threshold (rMT) between electrode types

Figure legend: Violin plots illustrating the distribution of resting motor threshold (rMT) values (% of maximum stimulator output) for the two electrode types: conventional disposable surface electrode (A) and suction-based electrode (B). Each participant underwent two transcranial-magnetic stimulation sessions per electrode type at Day 1, and the mean of rMT measurements collected during these two sessions was used for this comparison.

Figure 4. Stimulus–response curves of MEPs amplitude obtained from transcranial magnetic stimulation of the tongue motor cortex.

Figure legend: Log-transformed MEP amplitude data collected during the two test sessions at Day 1 were averaged and analyzed using two-way repeated measures ANOVA, with test stimulus intensities and electrode types as within-subject factors. Data are expressed in means and 95% confidence intervals. Stimulus intensity is expressed in relative percentage of rMT. Grey circles represent MEP amplitudes recorded with disposable surface electrodes, while white squares indicate values obtained with suction-based electrodes. Asterisks denote significantly higher MEP amplitudes for suction-based electrodes at 120%, 140%, and 160% rMT (Bonferroni post hoc, $P < 0.01$).