Development of a speech-discriminating electromyogram system for routine ambulatory recordings for the low-level masseter muscle activity

Short title: Speech-discriminating ambulatory electromyogram system

Yohei KUMAZAKI, Mami NAITO, Shigehisa KAWAKAMI, Atsutoshi HIRATA, Kazuhiro OKI, & Shogo MINAGI

Department of Occlusal and Oral Functional Rehabilitation, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University

Corresponding author:
Shigehisa KAWAKAMI
Department of Occlusal and Oral Functional Rehabilitation, Okayama University, 2-5-1 Shikata-cho, Kita-ku, Okayama 700-8525, Japan
TEL: +81-86-6687
FAX: +81-86-6689
e-mail address: s-kawa-10@t.okadai.jp
Abstract

Previous work suggests a relationship between sustained low-level tooth clenching and the etiology of myogenous temporomandibular disorder (TMD) pain. This study aimed to establish a reliable system with which to evaluate low-level electromyographic (EMG) activity related to low-level tooth clenching while discriminating speech activity, which is one of the most common facial muscle activities to be discriminated from low-level clenching. This device should facilitate the clinical evaluation of awake muscle activity in TMD patients.

Eight female and eight male subjects (38.9 ± 11.3 years) participated in the study to evaluate the validity of estimation of speech duration. Actual speech duration was defined by one examiner by pointing out the timing of beginning and end point of each speech on wave-editing software. Speech duration, as detected by a voice sensor system, which was activated by a voice loudness of 54.71 ± 5.00 dB, was significantly correlated with the above actual speech duration (p < 0.01, R² = 0.9935).

An actual recording with the system was carried out in one TMD patient and one healthy volunteer and revealed that the duration of diurnal EMG activity higher than 5% MVC was 1649.16 s and 95.99 s, respectively. As the voice sensor system adopted in this study could define the exact onset and offset of each segment of speech, EMG activity during speech could be precisely discriminated. The results of this study demonstrate that the EMG system with voice sensor system would be an effective tool for the evaluation of low-level masticatory muscle activity.
Introduction

Awake bruxism is mainly characterized by a clenching-type activity and is associated with psychosocial factors, e.g., anxiety, stress sensitivity, depression, and personality characteristics, and a number of psychopathological symptoms, e.g., state anxiety and trait anxiety (1, 2). Epidemiological studies have demonstrated that tooth-clenching is a risk factor of masticatory muscle pain (3-5). Sato et al. (6) reported that 52.4% of patients suffering from Temporomandibular Disorders (TMD) with muscular pain had a habit of diurnal tooth contacting. Svensson et al. suggested the possible importance of low-level clenching in the etiology of TMD (7). In a systematic review made by Manfredini et al. (8), they stated “Experimental sustained jaw clenching may provoke acute muscle tenderness, but it is not analogous to myogenous TMD pain; therefore, such studies may not help clarify the clinical relationship between bruxism and TMD”. However, Farella et al. (9) reported that sustained low-level tooth clenching in healthy young women induced a delayed soreness in the jaw elevator muscles, which might simulate myogenous TMD pain. The results of these studies indicate the importance of revealing the relationship between sustained low-level tooth clenching and myogenous TMD pain in TMD patients.

Many lines of research have been taken to measure diurnal electromyographic (EMG) activity employing ambulatory EMG recording systems, and the accuracy of these systems has been investigated (10-12). In the measurement of low-level masseter muscle EMG activity, the overlap between facial and masticatory muscle surface EMG signals has been regarded as a significant obstacle for the evaluation of the actual activity occurring within the masseter muscle. For example, with regard to sleep bruxism, it has been reported to be difficult to distinguish between bruxism events and other muscle activity; e.g., during chewing, talking, and grimacing. The distinction will require polysomnographic recordings including video and sound registrations (13). Among these facial activities, speech is especially known to result in sustained EMG activities that are often a sustained part of daily social activities (14-16) and which could potentially contaminate the EMG signal that is meant to be recording activity from the jaw muscles. Considering the importance of the duration of low-level EMG activity, speech would be one of the most common facial muscle activities that should be discriminated from low-level clenching. Watanabe et al. combined a self-reported event diary and voice recording system to determine the threshold EMG level for the post analysis of recorded temporalis EMG (12). However, as they accordingly used only the EMG activity threshold for the discrimination of speech, low-level clenching could also be regarded as EMG activity during speech. Therefore, a reliable EMG recording system, which could clearly identify sustained EMG activity associated
with speech from other low-level EMG activities, is necessary to progress research in this field.

This study establishes a reliable system with which to evaluate low-level EMG activity while discriminating speech activity.

Materials and methods

Evaluation of voice detecting system

Sixteen subjects without any symptom of pronunciation functions, eight females and eight males (mean age ± standard deviation of 38.9 ± 11.3 years) participated in the session. For the detection of voice, a voice operated trigger switch (VOX) was used with a condenser microphone attached to the neck skin adjacent to the larynx (Fig. 1). VOX circuit used in this study is shown in Appendix 1. Another condenser microphone was attached to the neck skin on the opposite side of the larynx using transparent adhesive film (Cathereep, Nichiban Co., Ltd., Tokyo, Japan). The VOX generated a signal voltage for each phonation, which was recorded as a negative electrical trigger signal at the beginning of the phonation and a positive electrical trigger signal at the end of the phonation. As shown in Fig 1a, the VOX trigger signal was recorded using one channel of a two-channel digital recorder (ICR-PS004M, SANYO Electric Co., Ltd., Osaka, Japan) and voice recordings were made using the other channel of the recorder in MPEG audio layer 3 (mp3) format with 64 kbps. The timing of the triggers was objectively identified using a threshold for the digitized trigger signals. The onset and offset of speech was defined by auditory evaluation made by one examiner YK. The onset of speech was defined as the beginning of the audible sound production, and the offset was defined as the end of audible sound production. As the sound pressure is irregular in the conversation depending on the pronounced words, several VOX trigger pairs of onset and offset could be observed even in one identical pronounced sentence. Therefore, if the interval from one offset trigger to the next onset trigger was less than 13.8 s, the conversation was regarded to be continued to the next sentence. The duration of 13.8 s was defined from preliminary experiment on the actual analysis of daily conversations as the longest quiet duration from offset to onset. Subject was instructed to make a quiescent interval of a minimum of 3 seconds between each sentence during the speech, which was confirmed by the waveform on the wave-editing software (Sound Engine, Coderium Co., Ltd., Sapporo, Japan).

The loudness of the voice necessary to generate a VOX trigger was evaluated using a sound level meter (LA-1210, Ono Sokki Co., Ltd., Yokohama, Japan). The sound level meter was located 20 cm from the mouth. Five measurements were made for each subject during the phonation of /a/ with gradually increasing loudness from 0 to 65 dB over a 2 s period in a sound
EMG and VOX recording of TMD and non-TMD subjects

Ambulatory EMG recordings were made for two subjects, one a patient (68-year-old male) who sought TMD treatment at the TMD clinic of Okayama University Hospital and the other a healthy volunteer (26-year-old male) who showed no signs or symptoms of TMD which was confirmed according to the Research Diagnostic Criteria for Temporomandibular Disorders. Subjects were instructed to gently and gradually bite an occlusal force detector until the bite force reached 500-gf on the left first molar, which was ipsilateral with the EMG electrode placement to produce the basis for the EMG standardization. The occlusal force detector was designed to generate a trigger signal with the bite force of 500-gf. Each subject carried out this low-level biting five times. The second channel of the recorder was firstly connected to the occlusal force detector and then reconnected to the VOX after the five low-level bitings. Our preliminary experiment revealed that 500-gf was generally above the lower border of the linear relationship between the occlusal force and the recorded EMG amplitude in this EMG recording system (Appendix 2) (Naito et al., submitted).

The EMG was recorded using differential surface electrodes on the left masseter muscle. A simultaneous recording of VOX signals was also made. The recordings took place on a weekday from 9:00 am for approximately 14 hours until the subject fell asleep. The subjects were instructed to write down daily activities, such as eating meals, reading books and so on in a diary. The EMG of the left masseter muscle was recorded by means of three disposable Ag/AgCl surface electrodes (Vitrode F-150S, Nihon Kohden Corp., Tokyo, Japan) with an active diameter of 17 mm and center-to-center distances of 25 mm. The line connecting the centers of the signal electrodes was parallel to the main direction of the muscle. The electrodes and cables were secured to the buccal skin with thin biocompatible adhesive tape. Before the electrodes were fixed, the skin was vigorously rubbed with a pad soaked in 70% ethyl alcohol. Subjects were instructed to perform maximal voluntary clenching (MVC) three times for 2 s at intervals of 2 s. Subjects were encouraged by the examiner to clench as hard as possible. For EMG standardization, the EMG signal level of 20% MVC, 5% MVC and 500-gf bite were used as the thresholds in the signal analysis (12, 17).

The recording hardware for EMG consisted of an analog signal processing and differential amplification integrated hybrid Circuit (NB-6201HS, Nabtesco Co., Kobe, Japan), which included a high-pass filter (10 Hz) and low-pass filter (1000 Hz), and a two-channel digital
The recording was made as a linear recording with sampling frequency of 22.05 kHz and 16-bit representation, and was stored in mp3 format with 64 kbps. This system could record for 29 hours with a battery, and the weight of the whole system was 156 g (Fig. 1c, 1d). Figure 3 shows that the recorded EMG had an excellent signal-to-noise (S/N) ratio throughout the recording period with low baseline noise.

The protocol used in this study was approved by the ethics committee of Okayama University (No.1124). Informed consent was obtained from all participants.

**Data analysis**

The recorded mp3 data were processed off line, converted into RIFF waveform audio format (WAV), filtered with a notch filter (60 Hz), and then downsampl ed to 100 Hz. The converted WAV file size was approximately 5 Gbytes for 14 hours’ recording. EMG thresholds of 5% and 20% MVC and an EMG level of 500-gf biting on the first left molar were used. The duration of EMG activity above these thresholds was divided into 1) speech, 2) mastication, and 3) any other EMG activities other than 1) or 2). The EMG activities during mastication were distinguished with the diary and EMG signal pattern by one examiner, YK. The EMG signal for mastication was discriminated as it appeared as a typical series of the repetition of short bursts as shown in Fig. 3b.

**Results**

**Evaluation of voice detecting system**

The loudness (mean ± standard deviation) of the subjects’ voices needed to generate the VOX trigger was 54.71 ± 5.00 dB. This can be regarded as the typical loudness of daily conversation (18). Therefore, the VOX ON trigger was generated with almost all of the daily conversation except only for whispered talk. As shown in Fig. 4, the wave of the recorded speech was observed as clusters of large and small waves of different length, which generated several sets of triggers in one speech period. Analysis of the relationship between the recorded onset of the speech and the onset of the VOX trigger revealed that the mean time lag, which was caused by the property of the analogue circuit of VOX, from the start of the speech to the first VOX onset was 0.23 ± 0.06 s. The relationship between the recorded end of the speech and the offset of the VOX trigger revealed that the mean time lag from the offset of the last VOX trigger to the end of the speech was 0.17 ± 0.07 s. The duration of speech heard by the examiner, YK, and the duration of speech calculated according to the VOX trigger had significant concordance (p < 0.01,
EMG and VOX recording of TMD and non-TMD subjects

The whole recording time of EMG and VOX was 14 hours for the healthy volunteer and 13.5 hours for the TMD patient, respectively. The MVC values of the TMD patient and the healthy volunteer at the beginning and at the end of the recording sessions were 18832.0 and 14853.0, 8378.1 and 8952.7, respectively. The baseline values of the TMD patient and the healthy volunteer at the beginning and at the end of the recording sessions were 20.7±25.1 and 36.9±22.3, 49.3±10.4 and 64.6±26.8. Figure 6 shows typical example of the recorded EMG of the subject with VOX trigger. It should be noted that some part of EMG signals exceed 500gf level during speech as shown in Fig. 6a. Long-lasting muscle activities of various duration of a level similar to that of speech could be observed (Fig. 6b). Therefore, the EMG data during speech were discriminated and summarized separately. ‘>500gf’ in Table 1 stands for ‘the duration of speech accompanied with EMG activity >500gf’. ‘>5%’ in Table 1 stands for ‘the duration of speech accompanied with EMG activity >5%. ‘>20%’ in Table 1 stands for ‘the duration of speech accompanied with EMG activity >20%. As shown in Table 1, the calculated total speech duration for the day was 1419.09 s for the healthy volunteer and 1945.48 s for the patient. In these speech durations, the EMG activity was higher than 500-gf for 370.86 s in the healthy volunteer and 1679.91 s in the patient, the EMG activity was higher than 5% MVC for 95.99 s in the healthy volunteer and 1649.16 s in the patient, and the EMG activity was higher than 20% MVC for 1.22 s in the healthy volunteer and 373.35 s in the patient. In the healthy volunteer, EMG activity was observed for 7647.53 s (>500-gf), 1997.62 s (>5% MVC) and 81.94 s (>20% MVC) except for speech and mastication. In contrast, in the TMD patient, EMG activity was observed for 26342.37 s (>500-gf), 22067.4 s (>5% MVC) and 6007.05 s (>20% MVC) except for speech and mastication. For mastication in the healthy subject, the sum of the duration of EMG activity was 1809.34 s (>500-gf), 1520.73 s (>5% MVC) and 1206.76 s (>20% MVC). For mastication in the TMD patient, the sum of the duration of EMG activity was 2481.01 s (>500-gf), 2376.43 s (>5% MVC) and 1505.96 s (>20% MVC).

Discussion

Recent progress in the development of medical electronic devices has markedly altered the limit of the traditional modality in the evaluation of physiological function. Concomitantly with evolutional trends in the field of medical and welfare robotics, the recording of EMG activity with an extremely high S/N ratio has become possible. In the latest review article published in 2012 (19), it was concluded that there is no evidence to support the use of EMG for
the diagnosis of TMD. The use of EMG in that review was mainly chair-side. Obviously, the usefulness of the electromyogram in a study depends on the purpose of its use and how the electromyogram is used. As the recorded EMG activity has become much more precise as subsequently described, the observation and application of electromyography in understanding the function of the stomatognathic system seems to meet the new generation of medical electronic devices. As shown in the present study, the latest EMG system, having a differential amplifier using low-impedance adhesive electrodes, can record EMG activity with a high S/N ratio even when surrounded by electrical equipment that might generate electric noise. As the recording system could eliminate artifacts possibly arising from the motion of the electrical cables, precise recording with a high S/N ratio could be maintained all day long. The EMG recording system reported in this study would be useful especially for the evaluation of long sustained low-level clenching.

As was mentioned in the review by Al-Saleh (19), the standardization of the recorded data would be one of the most important issues for the evaluation of low-level clenching. Although percent expression relative to the EMG signal during maximal voluntary clenching (%MVC) has been adopted as one of the most common standardization methods, it might not be suitable for the evaluation of low-level clenching for TMD patients. As the low-level clenching that could be measured by the present EMG recording system was less than 5% MVC, the standardized value using %MVC could easily be affected by possible fluctuation of the real strength of muscular activity during so called maximum voluntary clench. The strength of the voluntary clenching might be affected by the pain condition in TMD patient. Also, in denture wearing patient, for example, the strength of voluntary clenching could be affected by the quality of denture prosthesis and pain in denture supporting tissues, thus resulting in possible incorrect evaluation of maximum muscle contraction. Therefore, in addition to %MVC, EMG activity during 500-gf biting at the left first molar, which was ipsilateral with the EMG electrode placement, was adopted as the basis for the EMG standardization in this study, which might be used as a way of standardization of EMG in patient who has any difficulty in exerting real maximum contraction. Our preliminary experiment revealed that 500-gf was generally slightly above the lower border of the linear relationship between the occlusal force and the recorded EMG amplitude in this EMG recording system (data example shown in Appendix 2). Considering the diversity of the subject population of low-level EMG measurement in future studies, EMG standardization should be capable of handling partial or full edentulous subjects, whose MVC could be apparently smaller than those of full dentition subjects.

In the present study, although EMG data were presented for only one patient example, the observed duration of diurnal low-level clenching in the TMD patient was approximately 300 minutes longer than the duration in a normal example, thus suggesting the possible importance
of diurnal sustained low-level clenching as shown in a previous research (9). To evaluate the long-lasting low-level masticatory muscle activity, it is necessary to distinguish orofacial normal functional muscle activity from parafunctional masticatory muscle activity. From the standpoint of the prolongation of activity, masticatory muscle and facial muscle activity during speech would be one of the most frequently observed low-intensity phenomena. In contrast, EMG activity resulting from simple facial expressions would not be supposed to last long. Although EMG activities from facial musculatures during conversation would be discriminated using VOX system, smiling of extreme strength without speaking could not be discriminated by this system. However, considering the facial expression in daily life, it might not be natural to assume that extreme smile in non face-to-face condition would last too long and too often to affect the results of the EMG analysis, except for some occupational situations. Therefore, it might be important to clarify the occupational condition of subjects in future studies.

Several studies have attempted to distinguish speech activity in diurnal ambulatory recordings of a masticatory muscle EMG. Watanabe et al. (12) reported one method for the estimation of an EMG threshold for eliminating speech activity from the surface EMG of temporal muscle that was simultaneously recorded with the voice. However, defining one EMG threshold implies the high possibility of eliminating parafunctional muscle activity that is equivalent to or lower than the speech level. In fact, in the TMD patient in the present study, long-lasting muscle activities of various durations of a level similar to that of speech were observed. As the patient was a lecturer at a private cram school, the long speech duration should relate to his lecturing. However, in addition to his actual speech, the EMG data suggested that the patient has the habit of diurnal low-level clenching. As the VOX system adopted in this study could define the exact onset and offset of each speech segment, EMG activity lower than that during speech could be precisely evaluated. In addition, as the VOX system does not need to record the speech, neither the subjects nor the evaluator needs to consider the privacy of speech contents.

The results obtained in this study revealed that the VOX system with a high S/N EMG system would be an effective tool for the evaluation of low-level masticatory muscle activity while the subject is awake.

Acknowledgement
Authors would thank Professor Greg Murray (University of Sydney) for his valuable scientific advice. This research was supported by MEXT KAKENHI (grant number 24390433).
Conflict of interest
The authors declare no conflict of interest.

Appendix 1
The electrical circuit of VOX used in the present study.

Appendix 2
An example of preliminary data showing the relationship between the amplitude of the masseter muscle EMG and the ipsilateral occlusal bite force. Although occlusal bite force larger than 400-gf resulted in quite linear relationship ($R^2=0.9589$) with EMG amplitude recorded in this EMG recording system, the bite force less than 400-gf, drawn with open dots, tended to show gradual decrease of the regression inclination.
Table 1
Breakdown EMG durations in healthy volunteer and TMD patient
The accumulated duration of EMG activity above thresholds of 5% and 20% MVC and an EMG level of 500-gf was divided into speech, mastication, and the duration except for speech and mastication.

<table>
<thead>
<tr>
<th></th>
<th>accumulated duration (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
</tr>
<tr>
<td>Healthy volunteer</td>
<td></td>
</tr>
<tr>
<td>speech</td>
<td>1419.09</td>
</tr>
<tr>
<td>mastication</td>
<td>1809.34</td>
</tr>
<tr>
<td>except for speech and mastication</td>
<td>7647.53</td>
</tr>
<tr>
<td>Patient</td>
<td></td>
</tr>
<tr>
<td>speech</td>
<td>1945.48</td>
</tr>
<tr>
<td>mastication</td>
<td>2481.01</td>
</tr>
<tr>
<td>except for speech and mastication</td>
<td>26342.37</td>
</tr>
</tbody>
</table>
Figure Legends

Fig. 1  Experimental systems with VOX
a, Experimental setting for the simultaneous recording of the VOX trigger and speech. Voice input from one microphone, higher than an initially set threshold, generates a trigger signal. Another microphone was used to record the raw voice in the recorder.  b, Recording setup of the VOX trigger and EMG. One microphone was used to generate trigger signal during speech. EMG was recorded to another channel of the recorder.  c, Wearing the EMG and VOX system.  d, The EMG and VOX system.

Fig. 2  Raw waveform for the evaluation of VOX trigger threshold level
The shaded arrow shows the trigger onset associated with the gradual increase in voice loudness, and the open arrow shows the trigger offset. Dotted line shows termination of voice. Note that the trigger offset has a slight time lag after the termination of the voice as shown by the time duration between the dotted line and open arrow.

Fig. 3  Example of EMG recording
a, MVC recordings at the beginning and end of the 14-hour recording session. Note that the baseline of the signal has an excellent S/N ratio even at the end of the full daytime ambulant recording session.  b, Typical EMG pattern for mastication.

Fig. 4  Raw data example of a speech recording with VOX trigger
ON–OFF triggers of 22 sets were observed during this recorded speech of 24 seconds.

Fig. 5  Relationship between actual speech duration and estimated speech duration calculated from the VOX trigger
Significant correlation (p < 0.01, $R^2 = 0.9935$) was observed between these durations.

Fig. 6  Example of EMG with VOX recording
a, Recorded data example during speech.  b, Recorded data example not during speech. Note some EMG bursts during speech surpassed 500-gf threshold.
References


Fig. 1a
Fig. 1c

- Surface electrodes
- Microphone
- Waist bag
Fig. 1d

- microphone
- amplifier
- battery
- surface electrodes
- VOX circuit
- digital recorder
Fig. 2

Voice waveform amplitude

VOX trigger amplitude

onset offset
Fig. 3a

EMG amplitude

MVC at the beginning of the recording

MVC at the end of 14 hours' recording
Fig. 3b
Fig. 4

Voice waveform amplitude

VOX trigger amplitude

dB
-6.0
-12.0
-Inf
-12.0
-6.0
-2.5
-0.0
-2.5
-6.0
-12.0
-Inf
-12.0
-6.0
-2.5

1s.
Fig. 5

$y = 1.0186x - 0.1509$

$R^2 = 0.9935$

Speech duration calculated from VOX triggers (S) vs. Actual speech duration (S)
Fig. 6

(a) EMG amplitude vs. dB

(b) VOX trigger amplitude with threshold at 500gf

Speech
Appendix 1
Appendix 2

![Graph showing the relationship between EMG amplitude and bite force (gf). The graph displays a linear trend with error bars indicating variability.](image-url)