New methods for averaging EMG records

Öhman, S.

journal: STL-QPSR
volume: 7
number: 1
year: 1966
pages: 005-008

http://www.speech.kth.se/qpsr
Electromyographic (EMG) potentials picked up from various speech muscles are usually first integrated by means of low-pass filtering and then the smoothed records from repeated readings of the same utterance are averaged. Thus, if a given utterance has been read \( m \) times and if the \( m \) smoothed EMG records picked up from a given electrode during these utterances are denoted by \( f_1(t), f_2(t), \ldots, f_m(t) \), the average \( f(t) \) based on these utterances is obtained by computing the following formula

\[
f(t) = \frac{1}{m} \sum_{i=1}^{m} f_i(t-t_i)
\]

The time displacements \( t_i \) of Eq. (1) are usually determined on the basis of sound spectrograms or oscillograms. I.e., a certain acoustic event - the burst of a stop consonant, for instance - is localized in the acoustic record of each of the \( m \) utterances and is taken as the time reference for the averaging process.

From the articulatory point of view, however, an utterance \( u \) is a temporal segment of gestures \( g_1, g_2, \ldots, g_k \) which may be represented by

\[
u(t) = \sum_{j=1}^{k} g_j(t-t_j)
\]

In repeated readings of the utterance of Eq. (2) two types of variability will occur; there will be small random variations in the shapes of the gestures \( g_j \) and there will also be small random variations in the times of execution \( t_j \) of the gestures. If the averaging method implied by Eq. (1) is used then both these types of variability will contribute indiscriminately to the variance of the averaged EMG record. Even if the gestures of which the utterance consists are identical in all the \( m \) readings the dispersion about the mean of Eq. (1) may be considerable owing to fluctuations in the relative timing of the gestures. Furthermore, sharp dips and humps of the individual records will in general be smoothed out by the averaging process, especially if they are located far away from the acoustic time reference. To avoid these difficulties the following alternative procedures have been adopted.
Matched averages

The m records $f_1, f_2, \ldots, f_m$ are first inspected visually and prominent humps and dips that recur in all or practically all of the records are marked. Then mark off on the EMG records the time points of all important acoustic segment boundaries as determined from sound spectrograms or oscillograms. The next step is to superpose the m records $f_1, f_2, \ldots, f_m$ on top of each other (by tracing on transparent paper, for instance) so that the first hump or dip is reproduced with minimum dispersion. I.e., the successive records are translated back and forth along the time axis until the first hump of the new record fits the first hump of the growing average optimally with respect to overall shape. This procedure is repeated (on a new transparent paper) for each hump or dip; so that in the end a result similar to that of Fig. I-B-1 is obtained.

In Fig. I-B-1 five smoothed EMG records representing the electrical activity during the utterance /sɛja dɔ:biːdare/ as picked up from a small area of the overlapping parts of M. depressor anguli oris and M. depressor labii inferioris have been superimposed in five different ways. The arrows indicate the humps with respect to which the matching was performed. The solid curve of the bottom graph was obtained by patching together segments of the averages of the five matched curve families displayed above in Fig. I-B-1. These segments were taken from the neighborhood of the humps marked by arrows in the figure.

Evidently, the bottom curve of Fig. I-B-1 has a more sharply defined shape than any average based on Eq. (1) would have. Furthermore, the matching procedure just described furnishes additional information about the affinities between the humps and the segment boundaries. If one keeps track of the segment boundaries associated with the individual records when the latter are matched with respect to a given hump he can subsequently determine which of those boundaries have the smallest dispersion in relation to this hump. Thus, in Fig. I-B-1 it was found that the implosion boundary of the /b/ was minimally spread with respect to the matched average of curve family 3 while the explosion boundary of the /b/ was minimally spread with respect to curve family 4. We can thus correlate the two close lying EMG humps in the vicinity of the /b/ with different phases of the production of that consonant.
Fig. I-B-1. Smoothed EMG records as picked up from a needle electrode that was positioned in the overlapping parts of M. depressor anguli oris and M. depressor labii inferioris. The superposed traces in each part of the figure represent five readings of the same utterance. These traces have been translated in different ways relative to each other in the parts numbered 1 through 5. Curve number 6 is derived from the averages based on 1 through 5.
The matched average method does not distort the time scale of the final representation of the EMG activity. In fact the distances between the successive mean segment boundary time coordinates will be exactly the same as those obtained with Eq. (1). To show this we denote by $b_{ij}$ the time coordinate of the $i$'th acoustic segment boundary of the $j$'th reading of the utterance in question. By $h_{kj}$ we mean the amount by which it is necessary to translate the EMG record associated with the $j$'th reading of the utterance to get an optimal match with respect to the $k$'th hump. Clearly, the mean time coordinate $b_i$ of the $i$'th segment boundary is

$$b_i = \frac{1}{m} \sum_{j=1}^{m} b_{ij} \quad (m \text{ readings}) \quad (3)$$

When the records are translated for match with respect to the $k$'th hump the mean time coordinate $b'_i$ of the $i$'th segment boundary will be displaced somewhat so that

$$b'_i = \frac{1}{m} \sum_{j=1}^{m} (b_{ij} + h_{kj}) = b_i + h_k \quad (4)$$

where

$$h_k = \frac{1}{m} \sum_{j=1}^{m} h_{kj} \quad (5)$$

However

$$b'_{i+1} - b'_i = b'_{i+1} - b_i \quad (6)$$

That is, the distance between any pair of adjacent mean segment boundary time coordinates is independent of the hump with respect to which the matching was performed and is identical with the corresponding distance obtained with the method implied by Eq. (1). Hence, if there are $k$ humps in each record then it is sufficient to mark the mean segment boundary time coordinates each time the $m$ records have been matched with respect to some hump in order to get a common time scale for all $k$ curve families.

All EMG illustrations in the paper by Öhman, Leandorson, and Persson of the present QPSR are matched averages in the sense just described.
Multiple time reference averages

Although the matched average method may be automatized
a simpler method is preferable when the analysis is to be made by
means of a digital computer. Experience from manual work with
matched averages suggests that each EMG hump is always well corre-
lated with some one of the major acoustic segment boundaries. Using
this assumption we are presently developing a computer program for
automatic averaging of the EMG data. This program will store the
values $b_{ij}$ of visually determined segment boundary time coordinates
together with periodically sampled values of the EMG records $f_1$, $f_2$, \ldots, $f_m$ picked up by a fixed electrode during $m$ readings of an
utterance. If there are $n$ segment boundaries in each record the
program will compute $n$ average curves

$$f^*_k(t) = \frac{1}{m} \sum_{i=1}^{m} f_i(t-b_{ik}) \quad k = 1 \ldots n \quad (7)$$

using the successive $n$ segment boundaries as time references. The
final (multiple time reference average) curve will be defined as

$$f_1^*(t); \quad -\infty < t \leq (b_1+b_2)/2$$

$$f(t) = f_k^*(t); \quad (b_{k-1}+b_k)/2 < t \leq (b_k+b_{k+1})/2 \quad 1 < k < n \quad (8)$$

$$f_n^*(t); \quad (b_{n-1}+b_n)/2 < t < \infty$$

where $b_k$ is defined as in Eq. (3). For reasons similar to those
discussed in connection with the matched average method the time
scale will not be distorted by the method of Eqs. (7) and (8). The
difficulties of Eq. (1) are avoided however.