Variation in EMG activity: a hierarchical approach

Rebecca Z. German,¹,* A. W. Crompton† and A. J. Thexton‡

¹Department of Physical Medicine and Rehabilitation, Johns Hopkins University, School of Medicine, 98 N. Broadway, Suite 409, Baltimore, MD 21231, USA; ‡Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA; †Department of Physiology, King’s College London, London SE1 7EH UK

Synopsis

Recordings of naturally occurring Electromyographic (EMG) signals are variable. One of the first formal and successful attempts to quantify variation in EMG signals was Shaffer and Lauder’s (1985) study examining several levels of variation but not within muscle. The goal of the current study was to quantify the variation that exists at different levels, using more detailed measures of EMG activity than did Shaffer and Lauder (1985). The importance of accounting for different levels of variation in an EMG study is both biological and statistical. Signal variation within the same muscle for a stereotyped action suggests that each recording represents a sample drawn from a pool of a large number of motor units that, while biologically functioning in an integrated fashion, showed statistical variation. Different levels of variation for different muscles could be related to different functions or different tasks of those muscles. The statistical impact of unaccounted or inappropriately analyzed variation can lead to false rejection (type I error) or false acceptance (type II error) of the null hypothesis. Type II errors occur because such variation will accrue to the error, reducing power, and producing an artificially low F-value. Type I errors are associated with pseudoreplication, in which the replicated units are not truly independent, thereby leading to inflated degrees of freedom, and an underestimate of the error mean square. To address these problems, we used a repeated measures, nested multifactor model to measure the relative contribution of different hierarchical levels of variation to the total variation in EMG signals during swallowing. We found that variation at all levels, among electrodes in the same muscle, in sequences of the same animal, and among individuals and between differently named muscles, was significant. These findings suggest that a single intramuscular electrode, recording from a limited sample of the motor units, cannot be relied upon to characterize the activity of an entire muscle. Furthermore, the use of both a repeated-measures model, to avoid pseudoreplication, and a nested model, to account for variation, is critical for a correct testing of biological hypotheses about differences in EMG signals.

Introduction

Recordings of naturally occurring Electromyographic (EMG) signals are variable. As early as 1956, Doty and Bosma documented variability both between successive recordings of swallows obtained from a single electrode and among individuals (Doty and Bosma 1956). One of the first formal and successful attempts to quantify the variation in EMG signals examined the activity associated with capture of aquatic prey by salamanders (Shaffer and Lauder 1985). Shaffer and Lauder examined four levels of variation: (1) among feeding trials, (2) among experimental days, (3) among individuals, and (4) between two morphologically distinct species. They measured 11 EMG response variables (onset and duration of activity in several muscles) with a hierarchical Analysis of Variance (ANOVA) and found that variation among trials accounted for most of the variation, with variation among individuals next. Electrode position did not explain the variation in trials and was a negligible component of overall variation. The variation among the lower levels of the experimental hierarchy was often more significant than was the variation between the two species, which was the level of biological interest.

In view of the limitations of the EMG methods available in 1956, the question arose as to whether the variability in the EMG data reported by Doty and Bosma (1956) was of methodological or biological origin. Thexton et al. (2007) recently re-examined the EMG activity associated with the pharyngeal swallow in unanesthetized decerebrate pigs (Sus scrofa) partly to limit variation caused by studying differently prepared and differently anesthetized animals, as used by Doty and Bosma (1956). In the study by Thexton et al. (2007), significant variation was found at all levels of organization. In particular, unlike Shaffer and Lauder (1986), the swallow-related EMG

From the symposium “Electromyography: Interpretation and Limitations in Functional Analyses of Musculoskeletal Systems” presented at the annual meeting of the Society for Integrative and Comparative Biology, January 2–6, 2008, at San Antonio, Texas.

1E-mail: rz.german@jhmi.edu

Integrative and Comparative Biology, volume 48, number 2, pp. 283–293
doi:10.1093/icb/icn022

Advanced Access publication May 13, 2008
© The Author 2008. Published by Oxford University Press on behalf of the Society for Integrative and Comparative Biology. All rights reserved. For permissions please email: journals.permissions@oxfordjournals.org.
signals differed significantly between electrode pairs in the same muscles, but the study did not include a quantitative comparison among the different hierarchical levels, such as electrodes, sequences, and individuals.

The goal of the present article is to quantify the variation that exists at different levels, using more detailed measures of EMG activity than used by Shaffer and Lauder (1985) and Thexton et al. (2007).

Hierarchical Models, ANOVA, and Pseudoreplication

Nested ANOVA measures the relative contribution of different hierarchical levels of variation to the total variation in a dataset (Neter et al. 1996). Nested designs, as opposed to crossed designs, have one factor where all of its levels, or values, exist within one level, or group, of another factor. The classic example is leaves on a branch. If one factor is branches, and another leaves, the first set of leaves only exist on the first branch, and the second set of leaves only on the second branch. This is in contrast to a crossed design, with for example, the two factors of sex and species, where both sexes exist for both species.

The importance of accounting for different levels of variation in an EMG study is both biological and statistical. First, there is inherent interest in the level at which the signal is variable. Multiple electrodes in the same muscle that give different signals for a stereotyped action suggests that each recording represents a sample drawn from a pool of a large number of motor units that, while biologically functioning in an integrated fashion, show statistical variation. Different levels of variation for different muscles could be related to different functions or different tasks of those muscles. The statistical impact that unaccounted variation can have on results and on their interpretation is false acceptance or false rejection of the null hypothesis (Doncaster and Davey 2007). In a single-factor model, such as one testing for differences among species, additional variation, such as among individuals or trials, will accrue to the mean square error (MSE), and reduce the F-statistic for each of the factors and interactions, so leading to type-II errors (false acceptance of the null hypothesis). Nesting lower factors within higher ones, such as individuals within species, will reduce the error variation, and increase the likelihood of seeing a true species factor effect. Such lower factors are often called “nuisance factors” or account for “nuisance variation” i.e., variation that is not of primary interest for the biological question being asked, but, nevertheless, obscures the true relationship between factor and response (Doncaster and Davey 2007).

Pseudoreplication is a significant problem in biological datasets that have natural hierarchies. It occurs when there is repeated sampling of the same unit of analysis, but the variation is attributed to the factor level (Hurlbert 1984). For example, in a study that only compared species but with individual cycles of muscle activity as the unit of analysis, there would be pseudoreplication if some individuals contributed multiple cycles to the sample. If each individual was represented by one cycle, then the unit of analysis would be either individual or cycle, as they would be interchangeable in that design. However, ANOVA requires independence among units of analysis, and repeated sampling from an unrepresented factor (individual in this case) violates that rule. The consequence is that the number of degrees of freedom for the error increases because there appears to be a larger sample of independent cycles. This lowers MSE, which is the denominator of the F-value used to determine significance (the P-value). The end result is that a type-I error, false rejection of the null hypothesis, or finding an important difference when none exists is more likely.

In EMG studies, there are always many hierarchical levels that could be treated as statistical factors. Beneath the factors of interest, such as differences among sexes, ages, species, or ecological groups, there are locations from which specimen have been collected, breeds of animals, individuals, recording days, multiple recording sessions in one day, cycles within a recording session, and different muscles and multiple electrodes in a single muscle. The tendency to group across these levels stems from the difficulty of collecting data and the desire for a sufficient dataset to ask the biologically interesting questions. Finally, species are not evolutionarily independent, and comparative methods, such as independent contrasts are needed for studies including multiple taxa (Harvey and Pagel 1991).

There are several solutions to pseudoreplication. In the design stage of an experiment, one solution is to choose the appropriate factor to sample and take replicates. Although this sounds trite, much aggravation is saved with appropriate design prior to data collection. After data have been collected, analyzing them with either a repeated-measures design using the replicates as repeated measures of the random variable, or simply averaging replicates over a level of organization, is a legitimate statistical approach (Kirby 1993; Neter et al. 1996; Doncaster and Davey 2007).
Materials and methods

Most of the data used in this study have been previously analyzed using different techniques and published elsewhere (Thexton et al. 1998; Thexton et al. 2007); however, we also include unpublished data obtained using the same preparations and methods. The details of the experimental methods for the collection of the data and for the initial signal processing are included in Thexton et al. (2007), and are briefly outlined here. All experiments that generated data for the current article were approved by the Harvard University IACUC (23-05).

Animal models and behavior

There were two distinct samples, decerebrate animals (Thexton et al. 2007) and animals with an intact CNS (Thexton et al. 1998). Several reasons underlie the decision to use a decerebrate model. First, we wished to replicate Doty’s and Bosma’s (1956) study on EMG pattern during swallowing. They used a number of unanesthetized preparations with a reduced CNS, of which decerebration is the only current ethically acceptable method. Second, decerebration removed high-level sensory-motor integration. We hoped that this would limit variation that might otherwise arise because of higher level sensory processing and conscious motor control. Finally, decerebration removed most of the rhythmic oral function associated with suckling and intraoral transport, normally leaving only the pharyngeal swallow. For this study, we compared the decerebrate and intact models to see if there was significantly reduced variation in the decerebrate model, or a difference in the patterns of variation.

Videofluoroscopic and EMG data were collected in a series of experiments using both intact and decerebrate infant pigs between 15 and 30 days of age (preweaning). We attached a radio-opaque epiglottal marker to each animal to assist in visual monitoring of the swallow. Electrodes were surgically implanted for chronic recording. Under general anesthesia and aseptic conditions, the supra-hyoid and infra-hyoid muscles were exposed and identified following a standard atlas of pig anatomy (Sack 1982). We implanted duplicate bipolar wire electrodes into selected muscles with the separation between the recording surfaces oriented parallel to the long axis of the muscle fibers, and with the emerging wires sutured to nearby epimysium (Thexton et al. 2007). The muscles selected for study were *m. cricothyroidoideus*, *m. omohyoideus* (superior belly), and *m. thyrohyoideus*. In addition, we examined the sheet-like *m. mylohyoideus*, with duplicate bipolar ‘patch’ electrodes (Loeb and Gans 1986) sewn to the surface of the muscle midway between the median raphe and the mandible. In all cases, the duplicate electrodes inserted into each muscle were laterally displaced from each other, with respect to the orientation of the muscle fibers.

However, surgical considerations could limit placement of electrodes. While some adult muscles have clearly delineated separate bellies and others have histologically defined compartments (Balice-Gordon and Thompson 1988; Bennett and Ho 1988; Lev-Tov et al. 1993; Mu and Sanders 2008), we were attempting to obtain recordings from multiple closely packed hyoid muscles in preweaning miniature pigs. In this situation, to minimize potential damage to fine nerves supplying the individual muscles or their close neighbors, the surgical exposure of individual muscles was kept to a minimum. In an obvious case such as the cricothyroid, where more than one belly was present, duplicate electrodes were inserted primarily into the oblique fibers but, because of the limited size of the muscle and sometimes indistinct delineation of the bellies, there could be no guarantee that the more caudal of the two electrodes did not record from other, differently oriented, fibers.

Intact animals were fed infant pig formula from a standard baby bottle fitted with a special pig nipple (Nasco, Fort Atkinson, WI, USA). The decerebrate individuals were held in a custom-designed body sling with the animal in a nearly normal prone position and the head supported approximately in the same position as the intact animal suckling on an artificial teat. Most of the decerebrates lacked anterior oral function and the ability to suckle. They were fed using a small catheter (inserted into the posterior oral cavity), so that we could deliver milk containing barium from a syringe in 1 ml aliquots directly into the valleculae. In both intact and decerebrate animals, the movements of swallowing were recorded using digital video radiography (Siemens Tridoros 150G3 cineradiographic apparatus with a Sony DCR-VX1000 digital video camera) in the lateral (sagittal) plane. EMG signals from the selected muscles were amplified (×400 to ×10000) using an MA-300 EMG System (Motion Lab Systems Inc, LA, USA) with a band pass of 20 Hz–2kHz and a 60 Hz notch filter. The EMG signals were then recorded on a TEAC RD-145T digital data recorder together with the synchronization signals from the X-ray apparatus; an effective digitization rate of 6kHz was used. The unit of analysis for this study was a single pharyngeal or ‘reflex’ swallow. We used the start of the rapid, caudally directed movement of
the epiglottal tip (as recorded by lateral-view digital video radiography) as the marker for the swallow. A period of multichannel EMG activity was subsequently aligned to this event using video frame count and the synchronization signal.

Individual decerebrate swallows were elicited serially over a period of 1–3 min with a minimum of a few seconds delay between each swallow. A set of serially elicited individual swallows was termed a sequence and consisted of 12–35 swallows. The sequences varied in length of time for technical reasons (e.g., stopping recording to reset catheter position or to refill delivery syringe). We recorded one or more sequences in each animal two or three times a day. In the intact animals suckling on a bottle, continuous activity was recorded for several minutes at a time, so that large numbers of swallows were available for analysis.

**Data processing**

We processed the bandpass filtered and amplified EMG activities in several steps. The initial step was rectification and constant time (10 ms) reset integration (Thexton et al. 1998; Thexton et al. 2007) after which two additional processes were applied to the data. The first statistically defined a noise threshold and rejected the background activity (Thexton 1996), and the next extracted the “average” EMG swallow responses from the data as detailed subsequently. We worked with 400 ms time blocks in the intact animals and 1000 ms time blocks in the decerebrates. This produced a matrix of selected EMG data consisting of ‘n’ swallows and 100 time units or bins (each representing either a 4 ms or a 10 ms reset resampling period for each electrode, depending on the animal preparation). The time difference was required because of the different speeds at which the swallow was performed in the two cases. The timing of the start of epiglottal flip during each swallow was identified from the frame count on the videotape and the corresponding time established in the EMG data stream, using the synchronization signals. We first collected the EMG signals for a period starting 300 ms prior to each epiglottal flip and ending 700 ms after the flip in the decerebrate and starting 100 ms before and 300 ms after the epiglottal flip in the intact animals. These intervals were resampled to generate 100 time periods that formed the basic time unit of this study and represented an entire cycle of swallow behavior. The processed EMG activity recorded by each electrode in each swallow was then normalized to a maximum amplitude of 100 units. This equalized signal amplitudes across electrodes, despite inherent variation in amplitude due to differences in the distance of the electrodes from the nearest active muscle fibers (Buchthal and Schmalbruch 1980; Ertas et al. 1995).

**Data analysis**

The levels of variation in this study were (1) among swallows (cycles) with respect to a single electrode over time, (2) between multiple electrodes in one muscle (tissue level) that were matched pairwise or as repeated measures over time, (3) among sequences of swallows (using identical electrodes) at different times either within a day or between days, (4) among individuals, (5) among named muscles (e.g., *m. mylohyoideus* versus *m. omohyoides*), and (6) between biological groups (intact and decerebrate). Sample sizes are shown in Table 1.

The experimental design was a multifactorial repeated-designs model with nested factors and interactions among factors. The response variable

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Muscle</th>
<th>Condition</th>
<th>N animals</th>
<th>Cycles per animal</th>
<th>Sequences per animal</th>
<th>Total swallows</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mylohyoid</td>
<td>Intact</td>
<td>3</td>
<td>20</td>
<td>–</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>Thyrohyoid</td>
<td>Intact</td>
<td>3</td>
<td>20</td>
<td>–</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>Omohyoid</td>
<td>Intact</td>
<td>2</td>
<td>20</td>
<td>–</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>Cricothyroid</td>
<td>Intact</td>
<td>3</td>
<td>20</td>
<td>–</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Omohyoid</td>
<td>Decerebrate</td>
<td>3</td>
<td>34, 36, 44</td>
<td>2, 3, 4</td>
<td>114</td>
</tr>
<tr>
<td>6</td>
<td>Mylohyoid</td>
<td>Decerebrate</td>
<td>4</td>
<td>34, 36, 44</td>
<td>2, 3, 4</td>
<td>114</td>
</tr>
<tr>
<td>7</td>
<td>Omohyoid–Cricothyroid</td>
<td>Intact</td>
<td>2</td>
<td>20</td>
<td>–</td>
<td>40</td>
</tr>
<tr>
<td>8</td>
<td>Mylohyoid–Thyrohyoid</td>
<td>Intact</td>
<td>3</td>
<td>20</td>
<td>–</td>
<td>60</td>
</tr>
<tr>
<td>9</td>
<td>Omohyoid–Mylohyoid</td>
<td>Decerebrate</td>
<td>3</td>
<td>34, 36, 44</td>
<td>2, 3, 4</td>
<td>114</td>
</tr>
<tr>
<td>10 (type I error)</td>
<td>Omohyoid–Mylohyoid</td>
<td>Decerebrate</td>
<td>3</td>
<td>34, 36, 44</td>
<td>2, 3, 4</td>
<td>114</td>
</tr>
<tr>
<td>11 (type II error)</td>
<td>Omohyoid–Mylohyoid</td>
<td>Decerebrate</td>
<td>3</td>
<td>34, 36, 44</td>
<td>2, 3, 4</td>
<td>114</td>
</tr>
</tbody>
</table>
was EMG level, and the unit of analysis was the EMG activity in one of the 100 time bins. Time bin was treated as a repeated measure with 100 repeats, and electrode-pair (in the same tissue) was also treated as a repeated measure. Thus, a single case in this study was EMG activity in a single cycle or swallow. The between-individual variation measures the variation among cycles. The two repeated factors are both within-individual factors. Sequence, a random factor, was nested within an individual animal, which in turn was nested within named muscle, which was crossed with condition (intact/decerebrate). The null hypothesis was equivalence of EMG within a factor or an interaction. We used the $F$-value to estimate the portion of variation explained by a given factor.

The advantage of this design is that it considers the entire EMG signal, i.e., the activity in each of the 100 time bins. Rather than just on, off, duration, or time of peak activity, this method will distinguish between signals with different burst patterns that may have identical durations. For example, even when scaled to their respective maximum amplitudes, two muscles with the same duration could exhibit very different rates at which the signals rose to that maximum. The disadvantage of this study is that the sample size quickly becomes very large and issues of pseudoreplication arise. For example, an analysis with one sequence each of for three animals with two duplicate electrodes, where each sequence includes 20 cycles of EMG activity, will ultimately lead to 120 cases of on/off time measurements. Using 100 time bins further increases the sample size to 12,000. Even with a repeated-measures design for time and for electrode, the error degrees of freedom are nearly 9000.

A complete design was not possible. Not all animals had multiple electrodes in the same muscles. Therefore, 11 different analyses were conducted (Table 1) on several different datasets. The first six had only ANIMAL as a random factor. Three different sets of individuals were in these analyses, so that pairwise comparison of the muscles in the first six analyses were possible (Table 1). These analyses (7–9, Table 1) include MUSCLE as a fixed factor, and the ANIMAL × MUSCLE interaction. In contrast with the first nine analyses, we performed incorrect analyses on two of these datasets to demonstrate the problems that result either from not performing a repeated-measures analysis (analysis 10, Table 1, producing a type I error) or from not including all levels of analysis (analysis 11, Table 1, producing a type II error).

Interpretation of results concerning different hierarchical levels in complex designs with several repeated factors, with multiple interactions is not easy (Kirby 1993; Neter et al. 1996; Doncaster and Davey 2007). In this design that used 100 time bins for the EMG values, the repeated-measures TIME factor adds complexity to the interpretation. A number of specific interpretations are possible.

1. The factor ANIMAL tests for differences among the individuals in the overall amount of EMG signal during the entire burst.
2. The factor TIME tests for the differences in EMG level among the 100 time bins. Thus, significant TIME means that at least one of the 100 bins is different in EMG value from the others. We would expect TIME always to be highly significant; if it were not, it would mean that the EMG signal was flat and did not change over the 100 time bins.
3. The factor ELECTRODE tests for pairwise differences in the total amount of EMG signal between the duplicate bipolar electrodes in the same muscle. It does not test anything concerning the timing of those signals. The ELECTRODE factor is nested within ANIMAL, and cannot be crossed with it.
4. The interactions with the TIME factor test hypotheses that are biologically important. The interaction TIME × ANIMAL tests if differences among the 100 time bins vary among the individuals. This can also be interpreted as testing whether differences in the timing and shape of the EMG burst existed among the animals. The interaction TIME × ELECTRODE tests whether the timing and shape of the EMG burst varies between the two bipolar electrodes in the same muscle.
5. The factor SEQUENCE, in the decerebrate animals, represents data collected in different recording sessions for the same animal. SEQUENCE is also nested within ANIMAL, but can be crossed with TIME. A significant SEQUENCE factor tests if there are session-to-session differences in the total amount of EMG signal for a given muscle, and SEQUENCE × TIME tests if the shape and timing of the EMG burst differs between sequences.
6. In analyses 7–9, there is also a TIME × MUSCLE interaction, testing if the timing is different between the two muscles (e.g., m. omohyoideus versus m. cricothyroidus).

The interpretation of a significant main effect is different from the interpretation of the
interaction (Fig. 1). A main effect, such as ANIMAL or ELECTRODE measures the average difference between the total integrated area of the burst (Fig. 1A). In this example, two bursts occurred for approximately the same duration, with the same starting times and peak times, and thus there would be little interaction between TIME and the other factor. However, if these bursts were representative of the entire sample, there would be a significant factor effect because one burst would contain significantly more activity than the other. In Fig. 1B, the two bursts are nearly identical in their total amount of activity, but the start and the time of peak activity of the two bursts were offset in time. If the two curves represented samples from two individuals or from two electrodes, there would be a significant ANIMAL × TIME or ELECTRODE × TIME interaction, although no main factor effect.

**Results**

In all samples, there was significant variation at nearly all levels (Tables 2 and 3). In analyses 1, 2, and 4–6, which included only one muscle, the factor TIME had the largest $F$-value (Table 2). This repeated-measure factor, accounting for differences among the 100 time bins, indicated that the signal was not constant across all 100 time bins, which included more time than the actual EMG burst duration. In most cases, the next two highest $F$-values were TIME × ANIMAL or ANIMAL (Fig. 2), indicating that there were significant differences in the total EMG signal between animals as well as differences among animals in the timing of the burst of EMG activity. There was no consistent pattern of $F$-value in the remaining factors and interactions over the different analyses (Table 2). The only non-significant component in any analysis was the ELECTRODE factor in the decerebrate m. omohyoideus analysis. In all cases, the TIME × ELECTRODE interaction was significant, although often much less so than the other factors (Fig. 2).

Some of the biological differences of the $F$-values within each analysis are shown in Fig. 3. In the case of the cricothyroid (analysis 4), the ANIMAL factor had a higher $F$-value than did the ELECTRODE factor (Table 2) because, while the bursts for the two bipolar electrodes were nearly identical, the two animals shown by dotted lines both had visibly more extended bursts of EMG activity than did the animal

**Table 2** $F$-statistics for repeated-measures analysis

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Source of variation</th>
<th>1 Intact mylohyoid</th>
<th>2 Intact thyrohyoid</th>
<th>3 Intact omohyoid</th>
<th>4 Intact cricothyroid</th>
<th>5 Decerebrate omohyoid</th>
<th>6 Decerebrate mylohyoid</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANIMAL</td>
<td></td>
<td>9.4</td>
<td>56.3</td>
<td>4.7</td>
<td>20.5</td>
<td>55.3</td>
<td>3.5</td>
</tr>
<tr>
<td>TIME</td>
<td></td>
<td>54.01</td>
<td>173.8</td>
<td>91.3</td>
<td>109.0</td>
<td>330.3</td>
<td>228.2</td>
</tr>
<tr>
<td>ELECTRODE</td>
<td></td>
<td>2.3</td>
<td>24.0</td>
<td>6.4</td>
<td>1.2</td>
<td>0.8 ns</td>
<td>17.8</td>
</tr>
<tr>
<td>TIME × ANIMAL</td>
<td></td>
<td>47.9</td>
<td>47.2</td>
<td>76.8</td>
<td>12.4</td>
<td>24.9</td>
<td>28.00</td>
</tr>
<tr>
<td>TIME × ELECTRODE</td>
<td></td>
<td>17.8</td>
<td>4.9</td>
<td>9.6</td>
<td>3.9</td>
<td>14.8</td>
<td>27.6</td>
</tr>
<tr>
<td>SEQ (ANIMAL)</td>
<td></td>
<td>8.8</td>
<td>7.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEQ (ANIMAL) × TIME</td>
<td></td>
<td>10.7</td>
<td>14.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
represented by the solid line (Fig. 3A). That is, between duplicate electrodes in criocothyroid, there was little difference in the amount of EMG activity but a much greater difference among the three animals. There was a difference between the interactions of the ANIMAL and ELECTRODE factors with TIME. Within each pair of duplicate electrodes there was little difference in timing, and the F-value was relatively small for TIME × ELECTRODE. However, there were larger differences in the times of peak occurrence among the three animals, and the TIME × ANIMAL interaction was larger than the TIME × ELECTRODE interaction (Table 2). Figure 3B shows that signals detected by two electrodes in the omohyoid were very similar to each other but that, between animals, there were obvious timing differences. The same was true for m. thyrohyoideus (Fig. 3C).

The EMG patterns and the results for m. mylohyoideus were quite different from the other muscles (Fig. 3D). There were two distinct time patterns of m. mylohyoideus activity. In the first animal, both electrodes detected the same bimodal pattern (Fig. 3D1). In each of the other two animals, one electrode detected bimodal activity, but the other detected a single large burst of activity (Fig. 3D2 and 3D3). In these representative cycles, the single burst also appeared to differ significantly in shape between the second and third animals. However, the total amount of activity of the bimodal bursting pattern was roughly equal to the amount in the single modal pattern. In this case, both the ANIMAL and the ELECTRODE factors had relatively small F-values (Table 2). Yet, the difference in timing, both of the start and of the peak activity, was different between electrodes in the same muscle and also among the three animals. For example, the start of the first bimodal burst occurred much later in the first animal. This pattern was reflected in the relatively high F-values for both interactions TIME × ANIMAL and TIME × ELECTRODE, relative to the main effects.

In the three analyses (7–9) with MUSCLE as an additional factor and which compared two muscles for the same animals, there were differences among the analyses (Table 3, Fig. 4). In the comparison between the m. omohyoideus and m. cricothyroides (analysis 7), again with intact animals, TIME and MUSCLE × TIME had the highest F-values. In the mylohyoid–thyrohyoid analysis using intact animals, MUSCLE × TIME and MUSCLE were the most significant components. In the comparison of the

<table>
<thead>
<tr>
<th>Analysis Source of variation</th>
<th>Analysis 7 Omohyoid–cricothyroid</th>
<th>Analysis 8 Mylohyoid–thyrohyoid</th>
<th>Analysis 9 Omohyoid–mylohyoid</th>
<th>Analysis 10 Omohyoid–Mylohyoid</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUSCLE</td>
<td>0.3 ns</td>
<td>146.9</td>
<td>30.7</td>
<td>109.7</td>
</tr>
<tr>
<td>ANIMAL</td>
<td>42.1</td>
<td>38.2</td>
<td>57.0</td>
<td>203.9</td>
</tr>
<tr>
<td>MUSCLE × ANIMAL</td>
<td>9.7</td>
<td>7.2</td>
<td>14.1</td>
<td>50.2</td>
</tr>
<tr>
<td>TIME</td>
<td>54.5</td>
<td>62.4</td>
<td>526.2</td>
<td>453.76</td>
</tr>
<tr>
<td>ELECTRODE</td>
<td>10.7</td>
<td>6.3</td>
<td>2.44 ns</td>
<td>31.6</td>
</tr>
<tr>
<td>TIME × ANIMAL</td>
<td>39.8</td>
<td>24.0</td>
<td>16.24</td>
<td>13.4</td>
</tr>
<tr>
<td>TIME × ELECTRODE</td>
<td>9.1</td>
<td>12.1</td>
<td>45.0</td>
<td>33.3</td>
</tr>
<tr>
<td>TIME × MUSCLE</td>
<td>152.9</td>
<td>120.2</td>
<td>3.5</td>
<td>3.0</td>
</tr>
<tr>
<td>SEQ (ANIMAL)</td>
<td>8.5</td>
<td>30.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEQ (ANIMAL) × TIME</td>
<td>17.3</td>
<td>15.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2 F-values for six muscles from repeated-measures ANOVA. Values are taken from Table 1. Note that the two muscles from the decerebrate model have very high F-values for time, and are not plotted.

Variation in EMG
m. mylohyoideus with the m. omohyoides, analysis 9, ANIMAL and TIME \times ELECTRODE had higher \( F \)-values than did other components. However, in analyses 7 and 8, with the intact animals, the pairs of muscles were clearly different in the total amount of EMG signal, and to a lesser extent in the timing of those muscles' activities. The only nonsignificant component of these analyses was again the ELECTRODE factor in the comparison between m. omohyoides and m. mylohyoides in decerebrates.

In analyses 5, 6, and 9, there were sufficient replicates to test differences among sequences of data recorded at different times. SEQUENCE, which was nested within animal, was also significant, but had the lowest or near lowest \( F \)-value in each analysis. This indicated that the overall amount of EMG activity for each muscle was different in different sequences. The SEQUENCE \times TIME interaction, or the timing of EMG activity in different sequences, was also significant, and in each case, more so than was true of the SEQUENCE factor.

The results from the final two analyses (10 and 11) differ from the results of analysis 9, even though they use the same data. In analysis 10, a nonrepeated measures, multifactor analysis of the differences between Omohyoid and Mylohyoid had near equal or higher \( F \)-values for all factors and interactions (Table 3). In particular, the factor ELECTRODE is now significant. The degrees of freedom associated with the \( F \)-value for this factor are now (1, 4.45 \times 10^4), as compared with (1, 216) in the correct model (analysis 9). The results from the reduced model that lacked the hierarchical factors (analysis 11) also differed from the full model (analysis 9). In this case, TIME was still significant (\( P < 0.001 \)), but the neither MUSCLE (\( F = 0.2, P = 0.62 \)) nor the MUSCLE \times TIME interaction (\( F = 1.4, P = 0.22 \)) were significant.

**Discussion**

**Variation in EMG studies**

Doty and Bosma (1956) were among the first to point out temporal variation in EMG signals recorded by a single electrode during repeated identical actions. Thexton et al. (2007) tested whether this early report of EMG variation in mammalian swallowing was just a function of the stage of
development of EMG technique in 1956 or whether it represented a real biological phenomenon; they found that the variation was real and demonstrated significant differences between signals recorded by duplicate electrodes in the same hyomandibular muscles, as well as those between individuals.

Shaffer and Lauder (1985), and several subsequent papers influenced by Shaffer and Lauder (Sanderson 1988; Wainwright 1989), used nested ANOVA to test for biological differences, usually among species. The results of these studies, across a variety of species, muscles, and functions, documented a consistent pervasiveness of variation in EMG signals across all levels of experimental organization. However, the majority of papers using EMG methods either have not used, or have not documented the use of, any hierarchical analysis to account for other levels of variation. It is likely that type II errors (false acceptance of the null hypothesis) are frequently made in studies without a hierarchical or nested design. As shown in analysis 11, such results can be damaging to biological conclusions. Had analysis 11 been the definitive analysis, there would be no statistically significant difference between the activities in Omohyoid and Mylohyoid, either in timing or in overall amount of activity. If one can account for “nuisance” variation (Doncaster and Davey 2007) in experimental factors, such as in the individual or the trial, then that variation will not accrue to the error term and valid rejection of the null hypothesis is more likely.

Between-electrodes/within-muscle variation

Motor units in compartmentalized muscles can behave differently in different compartments of the same muscle (Herring et al. 1979; Lev-Tov et al. 1993; Gorassini et al. 2000). However, motor units in muscles that are not obviously compartmentalized can also show functional heterogeneity (Murray et al. 1999; van Lunteren and Dick 2001; Tsuruyama et al. 2002; Phanachet et al. 2003). These findings, not surprisingly, suggest that a single intramuscular electrode, recording from a limited sample of the motor units, cannot be relied upon to characterize the activity of an entire muscle.

Heterogeneity of histochemically identified types of myofiber appears to be particularly marked in the case of branchial and cranial muscles (Sciote et al. 2003), where additional hybrid myofiber types have also been identified in both humans and animals (Mu et al. 2004; Ren and Mu 2005). The general finding in mammals that a number of submandibular muscles are predominantly composed of fast-contracting fibers (Cobos et al. 2001), does not reduce the functional heterogeneity in such muscles (van Lunteren and Dick 2001), any more than it does in the medial and lateral gastrocnemius (Gorassini et al. 2000), both of which are also predominantly composed of fast-contracting fibers.

A priori knowledge of the exact distribution of the different histochemically identified types of fiber in a muscle would be greatly advantageous for the precise insertion of intramuscular recording electrodes into the defined regions of the muscle. This approach is, however, impractical unless a stereotaxic level of accuracy is available for the electrode placement (particularly in small mammals) and so is currently not feasible. However, it possible to insert multiple recording electrodes and then to use statistical methods to derive the time-varying levels of activity that best represent the predominant patterns of activity and perhaps also of any well delineated outlying activity groups. Such activity groupings might then be correlated with the presence of different groups of fiber type in the muscle.

Few EMG studies duplicate the insertion of electrodes into a single muscle in an individual animal. When dictated by the hypothesis under consideration, left and right sides may be recorded, e.g., in asymmetrical behaviors, such as in mammalian chewing or in turning during locomotion (Hylander et al. 2000; Crompton et al. 2007; Hedrick and Biewener 2007). That is not the equivalent, however, of a replication of the same muscle. This lack of replication is understandable. Collecting EMG data in any animal is not easy. Ethical limitations on the number of allowable animals and surgeries, physical limitations of muscle size, and limitation on the number of potential recordings possible in any given study put a premium on collecting the biologically interesting data. This imperative usually suggests recording from different muscles rather than duplication of potentially the same information.

The results of the present study, albeit for only a few muscles recorded during swallowing in mammals, suggests that there is biologically significant variation among electrodes in the same muscle. Whereas consistent and moderate temporal differences may be explained by the order of recruitment of the motor units, inconsistent and large temporal differences of peaks of EMG activity are quite a different matter (Fig. 3, D2, and D3). Such differences in timing of the signals recorded by electrodes in the same muscle strongly suggest different tasks/functions of motor units within that muscle.
The signals recorded by duplicate electrodes were, in some muscles, similar to each other, but with significant differences among individual animals (Fig. 3A–C). While the latter could obviously arise if there were systematic errors in experimental timing between different animals, it is also possible that the stratagems of neuromuscular coordination used by different animals were themselves different. This is the case in human swallows, where individuals can exhibit different patterns of muscle activation (Perlman et al. 1999). The excitability of motoneurons supplying the muscles involved are affected by vestibular and cervical afferents (Griffiths et al. 1983) so that small variations in posture among free feeding animals could be a factor in the differences in swallowing EMG activity between animals.

**Measuring EMG activity**

Many EMG studies extract statistics from the EMG bursts of interest by determining the times of onset and termination of the bursts, their durations, and times of peak activity (Shaffer and Lauder 1985; Canu and Falempin 1997). Other studies have employed cross correlation as a method for determining relationships among muscle timings (Wren et al. 2006).

An alternative strategy is to consider the entire burst as a repeated series of time bins, as outlined in the Methods section earlier. If the time bins are considered repeated measures, then the issue of lack of independence among the bins is not statistically relevant. This gives a finer assessment of differences among bursts than does simple time of onset and termination. One critical aspect of such an analysis is that all the cycles must be in register, preferably with an ‘external’ event, in our case the epiglottal flip that accompanies swallows. Although the epiglottal flip in the infant pigs studied is a consequence of EMG activity, it is primarily a consequence of the gross activity of the m. palatopharyngeus (Crompton et al. 2008), not a direct consequence of small samples of EMG activity in other muscles.

**Statistical and experimental design considerations**

One aspect of experimental design that is frequently neglected is the lack of independence among levels of nested factors or interactions among factors. Although Hurlbert (1984) explicitly outlined this problem and Doncaster and Davey (2007) discussed it, few published papers using EMG methodology address the issue. Frequently, large degrees of freedom in an analysis are pointed out as being “potential pseudoreplication,” but the statistical diagnosis of pseudoreplication lies in the independence of the levels within a factor. Thus, EMGs measured on multiple days in the same animal are not independent. The solution to the potential problem of pseudoreplication is a repeated-measures design that explicitly accounts for this lack of independence. Using the wrong design, i.e., a factorial as opposed to a repeated-measures model, is statistically incorrect and will give the wrong biological answer. In analysis 10, relative to analysis 9, many of the F-values are significantly higher, due to the increased degrees of freedom associated with the error term. The conclusions from this model would be that ELECTRODE is a significant factor, when in the correct model, it is not.

Doty and Bosma (1956) documented real variation in their EMG recordings. In 1956, however, the level of statistical and computational development precluded a detailed analysis of their data. Now, even if only one muscle is studied in one species, it is possible to quantify and account for multiple hierarchical levels of the experimental design. Consideration and statistical analysis of multiple levels of analysis of the factors associated with the EMG recordings ensures appropriate tests of the highest organizational level, such as among species or among muscles. Such an analysis also provides insight into biological questions about variation that can be answered at lower levels of analysis, including within muscle behavior.

**Acknowledgments**

We thank Nicolai Konow and Shannon Gerry for organizing the symposium of which this article was part. This work was supported by NIH DE6953 to R.Z.G.

**References**


in two species of macropodines (*Macropus eugenii* and *Macropus rufus*). Comp Biochem Physiol A Mol Integr Physiol [Epub ahead of print].


