Electromyographic studies of facial muscles during speech

Öhman, S. and Leanderson, R. and Persson, A.

journal: STL-QPSR
volume: 6
number: 3
year: 1965
pages: 001-011

http://www.speech.kth.se/qpsr
I. SPEECH ANALYSIS AND SPEECH PRODUCTION

A. ELECTROMYOGRAPHIC STUDIES OF FACIAL MUSCLES DURING SPEECH

S. Öhman, R. Leanderson*, and A. Persson**

Reasons for these studies

Electromyographic (EMG) studies of speech muscle activity are of interest from at least two different points of view.

1) Human utterances consist of sequences of temporally and spatially tightly integrated gestures which apparently are brought about by combining a fixed set of invariant neural motor commands according to strict phonetic rules. This circumstance gives the physiologist interested in complex voluntary muscle activity certain possibilities of obtaining reproducible experimental conditions which are not always available in work with most other kinds of such activity. 2) The sophisticated methods of EMG present an opportunity for the student of speech to approach the problem of describing the articulatory gestures at the level of the neural instructions and thereby to understand the transformations that these gestures undergo owing to mechanical fusion of the responses to temporally and spatially partly overlapping motor commands.

The facial muscles which are easily accessible for EMG work are not the least interesting in these two respects. In fact, as the possibility of decoding spoken messages by so-called lip-reading shows, the facial region transmits a surprising amount of phonetic information in normal conversation.

The facial muscles constitute a speech articulation system in miniature, as it were, which is quite on a par in complexity with that of the tongue. In particular, we notice that separate kinds of gestures are used for vowels and consonants. Thus, the rounding/spreading dimension of motion in pairs of vowels like Swedish y/i as in byta/bita or d/r as in mätamät should be compared with the closing/opening dimension of motion in consonants such as p, b, m of e.g. Swedish tappa, snaäba, kamma. Note further the inward/outward dimension of motion which appears in the difference between the consonants f/p or v/b in e.g. skaffa/tappa and vit/bit. Among the vowel gestures there is also a distinction

*Phoniatric Clinic, Karolinska Sjukhuset
**Central Neurophysiological Laboratory, Karolinska Sjukhuset
between inrounding and outrounding best demonstrated perhaps by the pair of vowels /w/ and /y/ as in stupa/stypa which differ very little in respect of tongue configuration and almost only in respect of in-rounding/outrounding. These gesture components are known to occur variously combined in practically every language of the world.

The fact that separate kinds of gestures are used for different classes of sounds is the physiological basis for the universal phenomenon of coarticulation. This may be exemplified by utterance pairs like bo/bi ([bu:]/[bi:]) where the rounding or spreading gesture of the final vowel is anticipated during (=coarticulated with) the closing gesture of the initial b. This fact is of prime phonetic importance. It implies that the neural "program" which controls speech must contain rules that sometimes allow coarticulation but sometimes also prevent it. The latter case occurs with gestures the phonemic interpretations of which are mutually incompatible. Thus, in an utterance like Swedish uppföra the inward motion of the lower lip associated with the /f/ must not be anticipated during the preceding /p/. If this rule is violated the utterance comes out as ufföra which is incorrect.

In one word, the motor integration of facial muscle activity presents a great number of interesting problems both from the physiological and the phonetic points of view.

Aims

Besides the physiological aspects to labial activity in speech the phonetician is interested in an acoustic aspect, namely, the determination of the precise shape of the mouth orifice for arbitrary articulatory configurations. Knowledge of this shape is essential for the correct prediction of the sound pressure wave that represents the utterance (1). Ultimately, a quantitative model of labial action would allow the derivation of this shape from parameter specifications relating to the dynamical dimensions of the myo-mechanical lip system along with a specification of the position of the jaw. High-speed motion picture studies of the orifice shape are presently being carried out by Lindblom (3). The following report summarizes the results of our first attempts to get a grasp of the physiological problem.
Experiment

The first question is where to put the electrodes. Palpation and visual studies of our own articulations in front of the mirror suggested that (4) the n. orbicularis oris is involved in the rounding gestures of a, o, y, e, u, etc., and that the inrounding/outrounding distinction ([da:] / [dy:]), [do:] / [dy:], etc.) might be due to a differential activity of the marginal and labial portions of this muscle. The p, b, m, closures appeared to involve the muscles of the chin (depressor labii inf., mentalis, and possibly depressor anguli oris) while the spreading gesture of vowels like i, e, a, etc., might be due to contractions in levator anguli oris, risorius, buccinator, depressor anguli oris, or some combination of these. Finally, the e/v type gestures seemed to involve the chin muscles as well as, perhaps, buccinator (to stretch the lower lip and pull it in towards the teeth) as well as levator labii superioris to pull the upper lip upwards somewhat. On this basis we selected the thirteen electrode positions (ep's) shown in Fig. I-A-1. One of the investigators (R.L.) served as subject.

The electrical muscle activity was recorded with a thin concentric needle electrode (DISA 13E59, diameter .30 mm) and the recorded motor unit potentials were amplified by an ac-amplifier (Grass P9) and fed into the vertical beam of an oscilloscope (Tektronix 502) which was photographed by moving film with a speed of 25 cm/sec. The sound pressure wave as picked up by a Sennheiser M2 421/2 microphone placed in front of the subject was displayed below the EMG record on the oscilloscope.

The subject was seated comfortably in a resting chair with a written list of "words" (see below) in front of him. The electrode was inserted at one of the preslected ep's, whereupon the subject read through the list of words. The electrode was then adjusted in the in/out direction and the list was read again. This process was repeated until an electrode depth was found that gave a maximal activity response for those speech sounds that were assumed to employ muscles in the neighborhood of the ep being tested. At this point a film was taken of the oscilloscope output during a complete reading of the list. The electrode was then moved to the next up and so on for all thirteen ep's. The raw records are shown in Figs. I-A-3:1 - I-A-3:19 and will be discussed later.
Fig. I-A-1. Numbers indicate the thirteen electrode positions used in the EMG session. In the figure the subject is articulating the Swedish vowel [u].
It should be pointed out here that the back-and-forth adjustments of the electrode are in general not painful. A weak pain is felt only when the skin is punctured and for this reason only surface anesthesia was used (Xylocain paste).

More important was the finding that if the subject hears his motor unit activity through a loudspeaker monitor, he unconsciously tends to adjust his muscle contractions in such a way as to maximize the loudspeaker noise. If the auditory feed-back is suddenly removed the motor-unit activity previously observed may decrease considerably or even disappear.* Accordingly, the present experiments were conducted without the above-mentioned auditory feed-back for the subject.

The list of "words" read by the subject at each positioning of the electrode is given below. It comprises all long Swedish vowels as well as the short vowels [a] and [e] (as in **katt** and **hund**). The latter vowels were prolonged in the reading of the list. The labial consonants [p], [b], [m], [f], and [v] in intervocalic position [−ː−] are also included. Each vowel is preceded by the (non-labial) consonant [d]. Key words are given within parentheses in orthographic spelling for the vowels.

1. **diː** (hit) 6. **dyː** (hy) 14. **ɛːpː**
2. **ðeː** (het) 7. **ðɛː** (hut) 15. **ɛːbː**
3. **ðɛː** (häl) 8. **ðuː** (hot) 16. **ɛːmː**
4. **ðɛː** (här) 9. **dɔː** (nɔt) 17. **ɛːfː**
5. **daː** (hatt) 10. **dɔː** (båt) 18. **ɛːvː**
11. **dɛː** (hör) 19. **ɛːpfrː**
12. **ðeː** (hutt)
13. **daː** (hat)

*This fact may be of relevance for speech training. The loudspeaker representation of the EMG signal could perhaps complement the defective proprioceptive feed-back of the patient. The EMG signal could also be fed back optically or tactually to deaf and/or blind patients.
Before discussing the TMG data it is appropriate here to summarize briefly the results of two dissections which were made after the experiment and which aimed at establishing tentatively what muscles were in all likelihood hit by the electrode at the 13 ep's described earlier.

Fig. I-A-2:1 shows one of the cadavers. The numbered nails have been fastened at the ep's of the EMG session as estimated from photographs of the live subject. Fig. I-A-2:2 shows the same cadaver as Fig. I-A-2:1 after the skin was prepared free from the underlying tissues. The nails remain in their original positions. The skin was cut along a line going from the nose below the eye over to the ear down to the jaw angle and then forwards along the edge of the mandible. The skin was then folded up towards the mouth and non-muscular tissue was removed. It was possible to distinguish the course and extension of the various muscles rather precisely. Even where they lie closely together their boundaries are quite distinct and it seems to us that with some practice it should be possible to make EMG recordings from any one of them individually.

Further dissections are presently being made and will be reported in a separate publication. The table below shows our tentative identification of the ep's with the various facial muscles as based on the dissections.

| ep 1 | M. levator labii superioris |
| ep 2 | M. caninus |
| ep 3 | Mm. zygomatici |
| ep 4 | M. buccinator |
| ep 5 | M. triangularis = M. depressor anguli oris |
| ep 7 | M. depressor labii inferioris |
| ep 8 | M. mentalis |
| ep 10 | pars marginalis inf. |
| ep 11 | pars labialis inf. |
| ep 12 | pars labialis sup. |
| ep 13 | pars marginalis sup. |
Fig. I-A-2:1. Photograph taken at the beginning of dissection. Nails have been fastened at places corresponding to the electrode positions used with the live subject as shown in Fig. I-A-1.
Fig. I-A-2:Z. Skin and other non-muscular tissue has been removed from preparation shown in Fig. I-A-2:1. The positions of the nails help identifying the muscles hit by the electrode. Further explanations in text.
A qualitative analysis of the spatial arrangement of the above muscle complexes suggests that they may function in antagonistic pairs as follows:

**M. levator anguli oris versus M. depressor labii inf.**

As partly visible in Fig. I-A-2:2 the levator anguli joins the inferior part of the m. orbicularis oris at the mouth corner and continues into the lower lip. It probably lifts the lower lip upwards. Depressor labii inf., on the other hand, seems to pull the lower lip downwards as the name also implies.

**Depressor anguli oris versus levator labii superioris**

M. triangularis = depressor anguli curves up into the upper lip as shown in many text-books of anatomy and could oppose an upward pull of the m. levator lab. sup.

**Orbicularis oris versus buccinator**

When orbicularis contracts as a whole a lip posture results which was fashionable among female movie stars of the 'twenties'. This gesture may be opposed by buccinator contraction which spreads the lips towards the sides of the face.

**M. depressor labii inf. versus M. mentalis**

M. mentalis pulls the skin of the lower portion of the chin upwards and m. depressor labii inf. pulls the lateral edges of the lower lip downwards. Since the latter muscle is located in a more superficial layer than the former one would expect that a simultaneous contraction in both should result in an outward motion of the lower lip away from the teeth of the mandible (protrusion). The two muscles would thus be antagonistic in the cranio-caudal dimension. The anteriorly directed component of the gesture could be antagonized by stretching the lower lip in the ventro-lateral plane by means of that component of m. buccinator which inserts into the lower lip at the mouth corner.

The above statements which are highly tentative reflect the authors' conception that one and the same group of muscles could participate in several antagonistic pairs. M. risorius seems to play an unimportant role. In one of the cadavers it was very small and in the other one it was missing.
Discussion of EMG data

In view of the limited size of our data corpus the following analysis can naturally not be understood as constituting a definitive description of facial muscle function in speech. What is intended is rather the extraction of some plausible hypotheses which may serve as the basis for further experiments.

Figs. I-A-3:1 through I-A-3:19 show the raw data arranged as follows. Each figure displays the oscilloscope output for each of the 13 ops when a given sound was produced by the subject. Figs. I-A-3:1 through I-A-3:5 show the so-called spread vowels [i, e, ø, ʊ, a] as in hit, hot, lůt, hār, and hatt, respectively. Figs. I-A-3:6 through I-A-3:12 show the rounded vowels [y, u, u̯, o, ø, e] as in myt, hut, nôt, bát, hör, and skutt, respectively. Fig. I-A-3:13 shows the vowel [a] as in hat. The bi-labial occlusives [p], [b], and [m] are displayed in Figs. I-A-3:14 through I-A-3:16, the labiodental fricatives [f] and [v] in Figs. I-A-3:17 through I-A-3:18, and the combination [pf] in Fig. I-A-3:19. All the consonants were spoken in the vowel environment [e-ø].

The time reference used in aligning the 13 records for a given utterance was the onset of voicing for the vowels and the onset of voicing of the final vowel for the [e]+consonant+[e] combinations. The sound pressure function is seen in the lower trace of each record.

Summary of spread vowel data

The common trait of these records is the relatively large output at ep 7 presumably associated with m. depressor anguli oris. This muscle is rather wide and its lateral portion could probably pull the corners of the mouth in a downward and lateral direction as required for spread vowels. The downward component might be necessary in vowels with a high jaw position ([i, e]) in order for the lips not to close off the mouth orifice.

A slight activity is observable at ep’s 4, 5, and 6 apparently associated with the mm. buccinator and depressor anguli. This activity seems to decrease with decreasing jaw height (less in [a] than in [i]).

Summary of rounded vowel data

The most apparent features here are the rather strong activity patterns in ep’s 13, 11, 12, and 10 (in order of decreasing
Fig. I-A-3:3
Fig. I-A-3:4
Fig. I-A.3:5
Fig. I-A-3:8
TIME REFERENCE

[Ø]

Fig. 1-A-3:9
Fig. I-A-3:10
Fig. 1-A-3:11
TIME REFERENCE

Fig. I-A-3:12
TIME REFERENCE

Fig. I-A.3:13
Fig. I-A-3:14
Fig. I-A-3:15
Fig. I-A-3:18

TIME REFERENCE

[ɛ:v ɛ:]
impressiveness). Those would all be associated with m. orbicularis oris. Some activity is also visible in ep’s 5, 6, 7, and 9 corresponding to mm. buccinator, depressor anguli, and mentalis, respectively. The combination of 6, 7, and 9 might produce protrusion of the lower lip as argued earlier where now depressor anguli would have the role that we previously also assumed for depressor labii inf. The buccinator contraction suggested by the activity at ep 5 might check the orbicular contraction.

The vowel [a]

At first this vowel seems to belong in the spread class in view of the activity at ep 7 and the lack of activity elsewhere. However, the ep 7 pattern decreases rather abruptly in strength as soon as the voicing of the vowel starts. The vowel [a] may thus either be of the spread type or characterized by no lip-activity at all.

It should be noted that motor unit activity is universally absent at ep’s 1, 2, 3, and 8 in the vowels.

One difficulty with the interpretation of the vowel data is the fact that the contexts in which the vowels occurred were not kept constant enough. Thus, it is at present sometimes difficult to determine whether a given activity pattern should be understood as a constituent of the vowel command or as mere background tonus. In future experiments a linguistic frame such as "säga ad-da igen" or something like that should probably be used for vowels. One could then look for contrasts with the EMG signals of the preceding and following gestures of the frame.

Summary of consonant data

Contrasts with preceding and following patterns can be studied in the consonants of the present body of data. In [p], for instance, (Fig. I-A-3:14) there is a sudden burst of activity at all ep’s except 7 (before and during the period of acoustic silence). Clearly, this activity must start before the finish of the initial vowel since the labial closure is closely correlated with the acoustic silence.
The activity at op's 1, 2, 3, and 8 is particularly notable among the consonants since no activity was recorded here during any of the vowels. According to our earlier assumptions we would associate these patterns with contractions in the antagonistic pair m. levator anguli oris and m. depressor labii inf. Presumably this pair causes a graded upward closing gesture of the lower lip as required for those phonemes.

Most impressive, perhaps, is the activity at the "orbicular" ep's 10, 11, 12, and 13 which apparently is balanced by clear patterns associated with buccinator and depressor anguli (ep's 5 and 6).

Very interesting is the fact that the ep 7 activity which previously was observed to characterize the spread vowels including [ε] which constitutes the context of all the consonants, that this activity disappears during the initial portions of the consonants. When most patterns occurring at ep's other than 7 have finished the "spreading" pattern at ep 7 returns (slightly earlier than the finish of the acoustic segment which signals the consonant). This suggests that the consonantal stricture is released by the following vowel. It is of course well-known that consonants at utterance final position (no phoneme following) tend to be relaxed in many languages.

In this connection it should be pointed out that the patterns at ep's other than 7 do not all seem to finish at the same time as the ep 7 pattern restarts. The patterns at ep's 1, 2, 3, 4, 5, 8, 9, 10, and 13 seem to end earlier and those at 6, 11, and 12 later than that of ep 7. Possibly this means that the release of the consonant occurs in several stages: first, the upward pull on the lower lip and the contraction in the marginal portions of orbicularis is relaxed whereupon the following vowel is turned on; secondly, the contraction in the labial portion of orbicularis as well as that of depressor anguli is also relaxed. An arrangement like this might insure a graded motion from the consonant into the following vowel.

Certain differences seem to prevail between the various consonants. A comparison of [b] and [m] with [p] as well as of [v] with [z] shows an overall reduction in "activity magnitude" in the former members of each pair (i.e. in [b,m,v]). The difference is most clearly marked at ep 2. If the latter ep is
associated with the levator anguli and hence with the upward motion of the lower lip this distinction should correlate with the traditional lax/tense concept. Among the bilabial stops the intraoral pressure is highest in [p] which accordingly should require a firmer bilabial stricture in order to prevent the air from leaking out. In [m] the intraoral pressure is close to the atmospheric pressure since the velum is lowered.

The difference between the bilabial and the labiodental consonants is illustrated by Fig. A-I-3:19 which represents the utterance [ɛ:pʰɛː]. The contrast is most prominent at ep's 8 and 12 (depressor labii inf. and orbicularis sup., respectively). At ep 8 the contrast appears as an abrupt decrease of activity at the p-f junction. At ep 12 it shows as the sudden appearance of a new type of motor unit spike shape which seems to superimpose itself on top of the pattern that started with the onset of the [p] and which extends throughout the consonantal interval (compare ep 12 for A-I-3:14 and A-I-3:17).

The meaning of this could be that depressor labii inf. (ep 8) antagonized by mentalis (ep 9) is instrumental in bringing about the inward/outward motion of the lower lip which characterizes the f/p contrast.

The ep 12 pattern indicates that the upper lip is raised during [f]. This gesture is subjectively apparent if one pronounces the "utterance" [pfpfpf...].

Conclusion

The above suggestions, although highly tentative, increase our optimism as to the possibilities of mapping out the activity of the facial muscles in speech in a meaningful way. Better controlled experiments and more extensive dissections need to be done, however. Also, much more data from each sound as spoken by more than one subject must be collected, and the activity of the different parts of each muscle must be examined in more detail.

The most interesting proposals that emerge from the present pilot study are, perhaps, that both vowel and consonant gestures may be graded motions brought about by antagonizing
pairs of muscle groups, that each gesture may be associated with a temporally delimitable EMG pattern, and that these patterns may indeed be interpreted in terms of contractions in anatomically well-defined muscular structures. Even when a given part of one and the same muscle is active in different ways in different gestures there may be a chance of inferring this from the EMG traces as the case of ep 12 Fig. I-A-3:19 shows.

References:


