Three-dimensional motion analysis of facial movement during verbal and nonverbal expressions in healthy subjects

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Clinical Anatomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>Draft</td>
</tr>
<tr>
<td>Wiley - Manuscript type:</td>
<td>Original Communications</td>
</tr>
<tr>
<td>KeyWords:</td>
<td>3D motion analysis, health, speech, facial expression, symmetry</td>
</tr>
</tbody>
</table>
Three-dimensional motion analysis of facial movement during verbal and nonverbal expressions in healthy subjects

Original research manuscript submitted to Clinical Anatomy on August, 15, 2016.

Running head: 3D facial movements

Key-words: face, mimicry, motion analysis, 3D.

Number of abstract words: 205
Number of keywords: 4
Number of text words: 3394
Number of figures: 5
Number of references: 31
Three-dimensional motion analysis of facial movement during verbal and nonverbal expressions in healthy subjects

ABSTRACT

Introduction. Functional impairments of facial expressions alter the quality of life and their quantitative analysis is a key step in the description and grading of facial function and dysfunction. The aim of the current investigation was to quantitatively assess 3D lip movements in a group of young healthy subjects.

Materials and Methods. Lip movements of five repetitions of usual and random sequences of vowels and nonverbal expressions were recorded using an optoelectronic 3D motion analyser. The mean value of facial landmarks maximum displacement was used to compute 3D unilateral mobility and the symmetry indices SI, separately for each sex.

Results. Facial mobility was significantly larger during open-mouth than closed-mouth smiles. The subjects performed the largest facial motion for the articulation of vowel /a/ and the smallest for vowel /i/. Lip purse was done with a SI very close to 95%, while smiles were slightly more asymmetric. Vowel /a/ was articulated with significantly better symmetry than vowels /e/, /i/ and /u/.

Conclusions. The outcomes suggest that the proposed method well characterised normal facial animations, and it could be a useful tool to evaluate patients with facial lesions, and dentofacial deformities. Patient assessment would profit from this quantitative approach, thus reducing the discordance among several clinical examinations.

Key words: 3D motion analysis; health; speech; facial expression; symmetry.
INTRODUCTION

Face expressions characterise human beings. In particular, mouth and lips are dynamically involved in both verbal and nonverbal expressions by the synergistic or cooperative action of many different facial muscles and their relation to underlying bones (Mishima et al., 2014; Al-Hiyali et al., 2015). Smile is one of the most frequent facial expressions, and it is used to transmit positive emotional state, as well as to serve social functions such as greeting. Like many other facial expressions, smile can be produced either deliberately by voluntary movement of the zygomaticus major muscles or spontaneously (Frank et al., 1993; Schmidt et al., 2006).

Anatomical literature suggests that lip purse expression requires symmetrical movements of bilateral orbicularis oris muscles around the mouth, which act as four substantially independent quadrants. Furthermore, the vermilion thickens in lip protrusion, and its dental surface often detaches from the teeth. A certain three-dimensional (3D) stress distribution within the entire labial tissue must be considered to account for the causal factors of lip protrusion. This expression is generally associated with anger emotion (Fukunaga et al., 2009; Al-Hiyali et al., 2015).

Speech production requires bilateral and symmetrical oral muscle movements as well as lip protrusion. The difference between lip protrusion (nonverbal) and speech production is that the former is a simple movement without sound, words, images or complex tongue movements; the two activities involve the activation of different brain areas (Fukunaga et al., 2009; Kent, 2015). According to preceding investigations, lip movements during verbal expressions can be influenced by the presence of various environmental stimuli (Gentilucci and Bernardis, 2007). Additionally, a previous knowledge of the sequence of words may change the activity of facial muscles (Scheme et al., 2007). To the best of our knowledge, no studies seem to have analysed if facial
muscle activity may also result in cinematic differences comparable to those induced by external stimuli.

Spontaneous and instructed facial movements are usually used to assess healthy orofacial functions and dysfunctions in several medical and dental fields (Giovanoli et al., 2003; Ferrario and Sforza, 2007; Popat et al., 2011, 2012; Kent, 2015), during diagnosis, treatment, and follow-up (Proff et al., 2006; Mishima et al., 2014; Al-Hiyali et al., 2015). Quantitative methods have been proposed to overcome some limitations of clinical evaluations, like reduced inter-examiner repeatability and problems in data sharing among different caretakers or research centres (Proff et al., 2006).

Nowadays, 3D motion analysers allow a non-invasive quantitative assessment of soft-tissue facial movements, without interfering with the subject’s performance. Among the others, optoelectronic motion analysers and laser scans offer a valuable support for the extraction of accurate numeric values (Weeden et al., 2001; Giovanoli et al., 2003; Ferrario and Sforza, 2007; Popat et al., 2008, 2011, 2012; Sforza et al., 2010b, 2012; Sidequersky et al., 2014; Lee et al., 2015). Patients’ facial range of motion can be compared to reference values obtained from healthy individuals according to age, gender and ethnicity (Sforza et al., 2010b; Popat et al., 2011, 2012); longitudinal assessments can also be performed during treatment and rehabilitation (Proff et al., 2006; Sforza et al., 2012, 2015; Hontanilla and Marre, 2014; Al-Hiyali et al., 2015). Facial movement symmetry is another important parameter, which can be measured both at rest and during mimicry. Considering the well-known fluctuating asymmetry existing in all subjects, the definition of clinical values for “asymmetry” should be preceded by normative reference data that can be used for patient assessment (Sforza et al., 2010b; Popat et al., 2011, 2012).
Most of the previous investigations focused on standardised nonverbal expressions, like smile and lip purse, but to obtain a more complete evaluation of the orofacial mimic functionality, beside nonverbal facial movements, also verbal expressions should be investigated (Gentilucci and Bernardis, 2007; Okudaira et al., 2008; Mishima et al., 2011; Popat et al., 2011, 2012; Schötz et al., 2013; Kent, 2015).

In the current investigation we quantitatively assessed 3D facial movements in a group of young healthy subjects performing a set of verbal and nonverbal expressions. The aim was to provide a reference database of standardised indices characterising healthy facial kinematics in terms of range of motion and symmetry, and for future use, in comparison with data of diagnosis and treatment of patients with facial lesions, and before and after Orthognathic surgery. We also searched for sex differences and verified if the previous knowledge of the vowel sequence induced a different kinematic output.
MATERIAL AND METHODS

The project was approved by the local ethical committee in accordance with the standards of the 1964 Declaration of Helsinki, and all subjects gave written informed consent to participate.

Subjects

Twenty healthy young adults (10 men and 10 women) aged 20 to 41 years, natural speakers of Italian language, participated in the study. They were recruited among students and staff attending the Medical School. All subjects had a clinically normal facial function, no previous facial trauma, paralysis or surgery, no known neurological diseases, presence of full natural permanent dentition (28 teeth at least), and no current orthodontic treatment. Also, to be recruited, the subjects had to manifest no temporomandibular joint disorders (TMD) according to the Research Diagnostic Criteria for TMD (RDC/TMD, Dworkin and LeResche, 1992), and no orofacial myofunctional disorder (OMD) according to the protocol of orofacial myofunctional evaluation with scores (OMES, de Felicio et al., 2012). They were examined while sitting on a dental chair in a room with appropriate lighting, by the same examiner (specialist in TMD/orofacial pain and myofunctional therapy).

Optoelectronic motion analyser recordings

Facial movements in verbal and nonverbal expressions were recorded at 60Hz using an optoelectronic 3D motion analyser (SMART-E system, BTS S.p.a, Garbagnate Milanese, Italy) (Sforza et al., 2012, 2015). Nine cameras were deployed around a stool to create a 60 (width) x 60 (height) x 60 (depth) cm$^3$ working volume. Metric
calibration and correction of optical and electronic distortions were performed before each acquisition session according to manufacturer’s guidelines. The subjects sat on a stool inside the working volume and were asked to perform series of facial movements. Before acquisitions, a set of facial landmarks were identified by 2 mm round reflective markers stuck on the skin: n, nasion; ft, right and left frontotemporale; ng, right and left naso-genian; cph, right and left crista philtri; ch, right and left cheilion; li, right and left lower lip midpoints (Fig. 1). The position of the markers was carefully controlled to avoid any interference with lip movements (Sforza et al., 2010b, 2012).

Three-dimensional landmark movements were obtained during the execution of four standardised nonverbal expressions and two sequences of verbal expressions (Fig. 1):

- open-mouth smile (with unsealed lips),
- closed-mouth smile (with sealed lips),
- spontaneous smile,
- lip purse (protrusion of the lips),
- usual sequence of five vowels (/a/, /e/, /i/, /o/, /u/),
- randomised sequence of the same five vowels.

Each expression was repeated five times, starting from a relaxed facial posture, without inter-session modifications of markers’ position. All animations were explained and shown to the subjects, who familiarised with the protocol before data acquisition; spontaneous smiles were induced by funny videos shown on a monitor positioned in front of the subject. The articulated vowels are the five cardinal vowels of the Italian language, and both their usual (conventional) and random sequences were steered by the monitor.
Data analysis

For each subject, facial movements were resolved in the head local reference system, defined by nasion and frontotemporale landmarks (Sforza et al., 2012). Subsequently, for each of the 8 facial markers (naso-genian, crista philtri, cheilion, lower lip of the right and left sides), the 3D movements during both verbal and nonverbal expressions were computed, and the magnitude of the 3D vector of maximum displacement from rest was calculated. Then, the unilateral facial mobility was estimated as the sum of the maximum displacement of the four unilateral facial markers; besides, the inter-side difference was quantified by an index of symmetry (SI), calculated as the ratio between the smaller unilateral 3D displacement and the larger of the two. SI indices ranged between 0% (the worst symmetry condition) and 100% (perfect symmetry).

Method error

Within- and between-session repeatability of facial mimicry was previously assessed in healthy subjects, using the same optoelectronic 3D motion analyser. Within session, the technical error of the measurement for single landmarks (random error) ranged from 0.5 to 3.38 mm, showing good reproducibility. Between sessions, all facial movements had standard deviations lower than 1 mm (Ferrario and Sforza, 2007).

Statistical analysis

For all subjects, the five repetitions of verbal (usual and random sequences of vowels production) and nonverbal expressions (open-mouth smile, closed-mouth smile, spontaneous smile and lip purse) were averaged, and the mean value of each landmark’s
maximum displacement was used for the subsequent computation of the 3D unilateral mobility and the SI.

Inter-subject descriptive statistics (mean and SD) were obtained for the mean 3D unilateral mobility (each subject’s right and left side mobility were averaged) and the SI, separately for each sex. The normal distribution of data was checked with the Kolmogorov-Smirnov test, and no significant deviations from normality were found.

A 2-way ANCOVA with repeated measures was applied to test the influence of gender (between-subject factor) and expression (within-subject factor) on the 3D mean unilateral mobility during the nonverbal expressions; the inter-cheilion distance at rest was used as covariate to control for sexual dimorphism in lip size. A 2-way ANOVA with the same independent variables was used to compare the SIs of the nonverbal expressions. A 3-way ANCOVA with repeated measures was applied to test the influence of gender (between-subject factor), expression and sequence type (within-subject factors) on the 3D mean unilateral mobility during the verbal expressions; the inter-cheilion distance at rest was used as covariate for the gender factor. A 3-way ANOVA with the same independent variables was used to compare the SIs of the verbal expressions. In case, post-hoc Student’s t-paired tests were performed between the expressions. The significance level was set at 5% for all analyses of variance (p<0.05), with Bonferroni’s corrections for multiple testing (nonverbal expressions, p<0.008; verbal expressions, p<0.005).
RESULTS

On average, 3D facial mobility resulted significantly larger during open-mouth smiles than during closed-mouth smiles. For each nonverbal expression, women had smaller mean landmarks’ displacement than men, but the difference was not significant, considering inter-cheilion distance as covariate (Fig. 2). Comparing the five verbal expressions, the subjects performed the largest facial motion for the articulation of vowel /a/ and the smallest for vowel /i/ with respect to all the others (Fig. 3). A statistically significant interaction was found between vowel and sequence type: in the random sequence, facial motion for /i/ was not significantly smaller than for /u/ (t-paired test, p=0.131). No sex difference was observed.

On average, lip purse was done with a symmetry index (SI) very close to 95%, while smiles were slightly more asymmetric. The lowest SI was recorded for the open smile in men (approximately 90%). The difference among the four nonverbal expressions was significant (ANOVA, p=0.031), but none of the post-hoc comparisons was statistically significant (Fig. 4). No sex difference was found.

For vowels’ production, vowel /a/ was articulated with significantly better symmetry than vowels /e/, /i/ and /u/ (on average, 95% for /a/, and around 87-92% for the other vowels, Fig. 5). Additionally, a significant interaction was observed between expression and sequence type: the difference between /a/ and /u/ was significant only in the usual sequence (t-paired test, p=0.005), while /a/ had significantly better symmetry than /e/ and /i/ only in the random sequence (t-paired tests, p<0.005). No gender effect was found.
DISCUSSION

In the present study, instructed and spontaneous facial movements of young healthy men and women were recorded by a 3D motion analyser, and the range of motion and symmetry of selected labial landmarks were quantified. When the effect of facial dimension is removed, the range of facial motion in both verbal and nonverbal expressions of healthy subjects had no sexual dimorphism; also the degree of symmetry of the movements was not gender-dependent. The sequence of vowels production had some effect on the kinematics of labial movements.

A special care was used to select the participants to the investigation. In particular, all subjects were screened for temporomandibular joint and orofacial muscle disorders (Dworkin and LeResche, 1992; de Felicio et al., 2012), two conditions that can significantly alter facial muscle activity. Temporomandibular joint disorders are the most common cause of facial pain, provoking pain and dysfunction in the jaw joint, associated muscles and supporting tissues (Dworkin and LeResche, 1992), and affecting about 5-12% of the population (NIH, 2016). Orofacial Myofunctional Disorders are conditions with altered functions and incorrect behaviours involving the facial muscles, in particular the tongue and lips (de Felicio et al., 2012); they have an estimated prevalence of 38% in the general population, raising to 81% of children with speech/articulation problems (International Association of Orofacial Myology, 2016). The use of highly selected participants is necessary to provide reference values for the evaluation of pathologic conditions.

Among the various indices proposed for the assessment of facial mobility, two standardised parameters were used: the 3D unilateral mobility and the symmetry index (SI). The marker set chosen for their computation was the same previously used in
longitudinal assessments of facial paralysis (Sforza et al., 2015), due to its simplicity and sensitivity to the movement of the facial lower third.

Indeed, landmark displacement can depend from facial dimensions: persons with bigger faces are expected to perform larger movements. Typically, male faces are larger than female faces; in young adults, lip area is approximately 6% bigger in men than in women (Sforza et al., 2010a). Some previous studies eliminated the effect of differences in facial dimensions by standardising the magnitudes of all vectors of maximum displacement by some estimate of facial size (Sforza et al., 2010b). In the current investigation, the inter-cheilion distance at rest was used to control the sexual dimorphism in mouth size, and no sex-related difference was found either in range of motion or symmetry. This finding, observed for both smiles, lip purse and vowels articulation, confirmed what have been previously reported about other nonverbal facial expressions (Giovanoli et al., 2003; Sforza et al., 2010b). Some researchers observed larger displacements in men than in women during lip purse expression (Weeden et al., 2001), while other investigators reported larger lower lip movements on the left side in men and on the right side in women during verbal expressions, but the differences were not practically significant (Popat et al., 2011).

Spontaneous and posed smiles have different neural pathways: the emotional contraction of the mimetic muscles originates from subcortical brain areas and supply excitatory stimuli to the facial nerve nucleus in the brainstem via extrapyramidal motor tracks, often involving a concomitant contraction of the orbicularis oculi muscles. Posed smiles, instead, are controlled by pyramidal tract firing from the motor cortex (Frank et al., 1993; Schmidt et al., 2006; Mishima et al., 2014). Both kinds of smiles should be assessed in patients: posed smiles offer information about voluntary, maximum contraction, while spontaneous smiles are critical for a positive social interaction (Frank
et al., 1993; Schmidt et al., 2006). However, in the current group of healthy subjects their kinematic output was almost the same in terms of range of motion of the facial lower third and inter-side symmetry. In contrast, Schmidt et al. (2006) described a greater amplitude of movement in deliberate than in spontaneous smiles, but with similar lip corner asymmetry. Differences were also observed in patients with facial palsy rehabilitated with a gracilis free flap, where the spontaneous smiles had smaller and more asymmetrical displacements than posed maximum smiles (Sforza et al., 2015). Indeed, Schmidt et al. (2006) did not specifically assess “enjoyment” smiles, but included in their analysis smiles recorded during social interaction, while both Sforza et al. (2015) and the current study elicited spontaneous smiles via a funny stimulus. Unfortunately, in the current study we did not assess movement speed, a characteristics that seem to differentiate the two animations (Frank et al., 1993), and that will be the topic of future investigations.

Among posed smiles, open-mouth smile required more landmarks’ displacement than closed-mouth smile, on average approximately 1.48 (women) and 1.26 (men) times larger. The difference is likely to be caused by the vertical component of displacement characterising \(cph\) and \(li\) when the lips unseal during open-mouth smile. Similar differences were reported in our previous study (Sforza et al., 2010b): the free smile was performed with larger displacements of almost all labial landmarks. Differences were also observed in six facial expressions in health volunteers, where the males had larger displacements [ranging from 14.31 mm (fear) to 41.15 mm (anger)] than females (Lee et al., 2015).

Among verbal articulations, in this study we used the Italian alphabet vowels (/a/, /e/, /i/, /o/, /u/), which are distinct from phonetic vowels, where accents on the vowels /e/ and /o/ should also be taken into account. According to the classification
provided by Kent (2015), our sequences can be categorized as paraspeech. No differences between males and females were found, as previously reported in Swedish speakers (Schötz et al., 2013). The vowel /a/ had the largest range of landmarks’ displacement and the better degree of symmetry than vowels /e/, /i/ and /u/; actually its speech is accomplished by both labial detaching and mandible descent. It is difficult to compare the present results with literature findings, as there are differences in both spoken sequences and analysis methods. In general, in agreement with the current findings, symmetrical labial movements were reported by Popat et al. (2012), while no-sex related differences were found by Schötz et al. (2013). In contrast, Popat et al. (2012) found that women had a more protrusive articulation of words than men.

In the current study, the production of vowel /a/ was made with the largest facial motion, a finding in good accord with previous data collected in Japanese speakers (Okudaira et al., 2008). In contrast, the vowels with the smallest movements were /i/ for Italian speakers, and /u/ for Japanese ones (Okudaira et al., 2008).

Gentilucci and Bernardis (2007) measured kinematic parameters of maximal lip movements during the production of phonemes containing the vowel /a/ by using a 3D optoelectronic system. They found that lip movements were influenced by the experimental conditions, with different opening amplitudes in presence of either reading, visual or acoustic stimuli. In their experiment, no automatic sequences of letters were given, as they wanted a conscious recognition of the actual stimulus. In contrast, in the current study we wanted to assess possible differences between usual and unusual sequences, that should involve different degrees of automatic performance. While Scheme et al. (2007) found that a known sequence of words may modify natural speech inducing anticipatory articulation, with different electromyographic activity in the facial muscles, we observed some significant interactions between vowel and
sequence type (usual vs. random). In particular, facial motion for /i/ resulted smaller than for /u/ only in the usual sequence, while no difference was found for the random sequence. Considering that the labial articulatory gesture for /i/ and /u/ is different (vowel /i/ necessitates a lip corner pull or stretch while vowel /u/ is obtained by lip purse), and that for /o/ is somehow midway between the other two vowels, it seems that in the usual vowel performance subjects do not perform a full labial movement, but reduce it anticipating the subsequent vowel. Subject testing should therefore include also random vowel sequences to ensure assessment of the entire movement.

One limitation of this study is that the degree of symmetry of facial movement was assessed at the instant of maximum global landmarks’ displacement, lacking to quantify the degree of synchrony. Indeed, the choice of this reference instant could affect the actual symmetry value, even if the current data sampling (60 Hz) has already been reported to be sufficiently sensitive (Popat et al., 2013). Future research should evaluate also this aspect, assessing the actual trajectories of the various verbal and not verbal animations. Furthermore, we assessed only single vowels, a test that can be considered as paraspeech (Kent, 2015). More complex and natural speech production should be tested as well.

The outcomes suggest that the proposed method well characterised normal facial animations, and it could be a useful tool to quantitatively evaluate patients with history of alterations in the facial hard- and soft-tissue structures, reducing the discordance among several clinical examinations.
REFERENCES


Hontanilla B, Marre D. Differences between sexes in dissociation and spontaneity of smile in facial paralysis reanimation with the masseteric nerve. 2014. Head Neck 36:1176-1180.


FIGURE CAPTIONS

Figure 1. Landmarks’ position at rest and simulation of their displacement during the nonverbal and verbal (vowels) analyzed expressions: n, nasion; ft, right and left frontotemporale; ng, right and left naso-genian; cph, right and left crista philtri; ch, right and left cheilion; li, right and left lower lip midpoints.

Figure 2. Three-dimensional mean unilateral mobility (right and left sides averaged) in nonverbal expressions (mean±1SD). Two-way ANCOVA (factors: gender and expression; covariate: inter-ch distance), p=0.008 for expression. Open-mouth smile (∗) > closed-mouth smile (∗), post hoc t-paired test, p=0.000.

Figure 3. Three-dimensional mean unilateral mobility (right and left sides averaged) in verbal expressions (mean±1SD). Three-way ANCOVA (factors: gender, expression and sequence type; covariate: inter-ch distance), p=0.000 for expression, p=0.017 for expression × sequence type. The black arrows indicate the largest (up pointing arrow, /a/ > /e/, /i/, /o/, /u/, post hoc t-paired tests, p<0.005) and the smallest (down pointing arrow, /i/ < /a/, /e/, /o/, /u/, post hoc t-paired tests, p<0.005) vowel’s facial movement.

Figure 4. SI of nonverbal expressions (mean±1SD). Two-way ANOVA (factors: gender and expression), p=0.031 for expression. Post hoc t-paired tests, p>0.008.

Figure 5. SI of verbal expressions (mean±1SD). Three-way ANOVA (factors: gender, expression and sequence type), p=0.015 for expression, p=0.032 for expression × sequence type. /a/ (*) > /e/, /i/, /u/ (*), post hoc t-paired tests, p<0.005.
Landmarks’ position at rest and simulation of their displacement during the nonverbal and verbal (vowels) analyzed expressions: n, nasion; ft, right and left frontotemporale; ng, right and left naso-genian; cph, right and left crista philtri; ch, right and left cheilion; li, right and left lower lip midpoints

figure 1
90x53mm (300 x 300 DPI)
Three-dimensional mean unilateral mobility (right and left sides averaged) in nonverbal expressions (mean±1SD). Two-way ANCOVA (factors: gender and expression; covariate: inter-ch distance), p=0.008 for expression. Open-mouth smile (*) > closed-mouth smile (*), post hoc t-paired test, p<0.001.

figure 2
80x43mm (300 x 300 DPI)
Three-dimensional mean unilateral mobility (right and left sides averaged) in verbal expressions (mean+1SD). Three-way ANCOVA (factors: gender, expression and sequence type; covariate: inter-ch distance), p<0.001 for expression, p=0.017 for expression x sequence type. The black arrows indicate the largest (up pointing arrow, /a/ > /e/, /i/, /o/, /u/, post hoc t-paired tests, p<0.005) and the smallest (down pointing arrow, /i/ < /a/, /e/, /o/, /u/, post hoc t-paired tests, p<0.005) vowel’s facial movement.

Figure 3
94x50mm (300 x 300 DPI)
SI of nonverbal expressions (mean+1SD). Two-way ANOVA (factors: gender and expression), $p=0.031$ for expression.
Post hoc t-paired tests, $p>0.008$.

figure 4
80x43mm (300 x 300 DPI)
SI of verbal expressions (mean+1SD). Three-way ANOVA (factors: gender, expression and sequence type), p=0.015 for expression, p=0.032 for expression x sequence type. /a/ (*) > /e/, /i/, /u/ (*), post hoc t-paired tests, p<0.005.

figure 5
92x48mm (300 x 300 DPI)