August 1993

Rule-Structured Facial Animation System

Catherine Pelachaud
University of Pennsylvania

Marie-Luce Viaud
INRIA

Hussein Yahia
INRIA

Postprint version. Presented at *Fuzzy Logic In Artificial Intelligence: IJCAI '93 Workshop*, August 1993, 6 pages.

Rule-Structured Facial Animation System

Abstract
Our overall goal is to produce as automatic as possible facial expressions with wrinkles from spoken input. We focus on two aspects of this problem: integration of the expressive wrinkles and generation of synchronized speech-animation. Our facial model integrates facial muscles deformations and bulges. We have produced a high level programming language to automatically drive 3D animation of facial expressions from speech. Our system embodies rule-governed translation from speech and utterance meaning to facial expressions. We are concerned primarily with expressions conveying information correlated with the intonation of the voice, some of which are also correlated with affect or emotion. We apply our automatic animation model to a new facial animation system which integrates effects of the facial motion as expressive wrinkles and muscles dependencies. We obtain then with this subtle criteria of modeling and motion an animation much more expressive and natural.

Comments
Postprint version. Presented at *Fuzzy Logic In Artificial Intelligence: IJCAI '93 Workshop*, August 1993, 6 pages.

Rule-Structured Facial Animation System

Catherine Pelachaud[*]
University of Pennsylvania
pelachau@graphics.cis.upenn.edu

Marie-Luce Viaud[†]
INRIA
luce@bora.inria.fr

Hussein Yahia[‡]
INRIA
hussein@bora.inria.fr

Abstract

Our overall goal is to produce as automatic as possible facial expressions with wrinkles from spoken input. We focus on two aspects of this problem: integration of the expressive wrinkles and generation of synchronized speech-animation. Our facial model integrates facial muscles deformations and bulges. We have produced a high level programming language to automatically drive 3D animation of facial expressions from speech. Our system embodies rule-governed translation from speech and utterance meaning to facial expressions. We are concerned primarily with expressions conveying information correlated with the intonation of the voice, some of which are also correlated with affect or emotion. We apply our automatic animation model to a new facial animation system which integrates effects of the facial motion as expressive wrinkles and muscles dependencies. We obtain then with this subtle criteria of modeling and motion an animation much more expressive and natural.

1 Introduction

Facial animation is part of 3D character animation concerned with the modeling, motion control and rendering of human faces. As the main communication channel, human face is a very sensitive object which is moreover never still. Facial and audio channels complement and support each other: visual channel can clarify and make unambiguous what is being said. While talking, people not only use their lips but their eyebrows may raise and create expressive wrinkles which will amplify the content of the speech; they may turn the head to underline a specific word; their eyes may move or blink, etc. It is the accumulation of such synchronized details which give expression and life to synthetic animation.

Earlier systems [Parke, 1989] used a set of conformation and expression parameters to perform the animation. Simulation of muscles of the face and of the propagation of their movement [Platt, 1985], [Waters, 1987] produce more accurate facial actions. Modelling the external and interface structures of the skin leads to a computer simulation of some plastic surgery operations [Pieper, 1992]. Multi-layer approach [Kalra *et al.*, 1991] or addition of some speech parameters [Hill *et al.*, 1988], [Lewis and Parke, 1987], [Nahas, 1988] were included in animation systems to produce automatic lip synchronization. Their approach was to express the animation by a set of rules where every element (of the sound and of the face) can be modified interactively through the use of parameters. To gain higher-level control over facial animation, techniques of extracting contours over designated areas on a real actor's face recorded on video were developed [Terzopoulos and Waters, 1991].

Facial expressions are often very subtle. The observation of details such as expressive wrinkles, and conversational signals in the speech, is an important feature of understanding human face. Thus, in order to improve facial animation systems, perceiving this facial language and its interaction with intonation is one of the most important steps. The consideration and synchronization of visual and audio channels enhance animation system.

We present in this paper a system in which secondary effects of the movement are modelled according to the primary deformation of the muscles and their effects on the skin. We first characterize the various channels of facial expressions and the intonational notation we are using. We explain the assumption of our system and its methodology. We then present our facial model. We define in our context wrinkles and bulges. We describe how we have included then in our animation system.

1.1 Facial Notation

We are using the **FACS** notation (Facial Action Coding System) created by P. Ekman and W. Friesen [Ekman and Friesen, 1978]. It is based on anatomical studies. Since facial movement is due to muscle action, a system was derived to describe visible facial expressions (emotional or conversational) at the muscle level. An action unit **AU**, basic element of this system, describes the action produced by one or a group of related muscles.

This notation is relevant in our case since it allows us to work at the **AUs** level and not at the regions or muscles level; it implies then the independence of the facial model we are using.

[*]CIS department, Philadelphia, PA 19104 U.S.A
[†]Domaine de Voluceau, BP 105, 78153 Le Chesnay, France
[‡]Domaine de Voluceau, BP 105, 78153 Le Chesnay, France

2 Analysis of Facial Action and Vocal Parameters

Our main goal is to look at the coordinated motion during conversation. Our method is based on finding the link between the spoken intonation, the transmitted information in the given context, and the facial movements.

To counterbalance the difficulties involved in manipulating manually the action of each muscle, our system offers to the user a higher level of animation by lip synchronization and by an automatic computation of the facial expressions related to the patterns of the voice. Not all facial expressions are used to characterize emotions. Some can highlight a word, underline a pause; some are tied to the intonation of the voice. Our first step was to elaborate a repertory of such facial actions called *channels*.

2.1 Facial Channel

Facial actions can be clusterized depending on their functionality [Ekman, 1979].

phonemic channel : lip shapes and coarticulation.

intonational channel :

CONVERSATIONAL SIGNALS: when occurring on accented items or on emphatic segments, they clarify and support what is being said. A stressed element is often accompanied not by a particular movement but by an accumulation of rapid movements (such as more pronounced mouth position, blinks or rapid head movements).

PUNCTUATORS: they can reduce the ambiguity of speech, grouping or separating sequence of words into discrete unit phrases. A boundary point (such as a comma) will be underlined by slow movement and a final pause will coincide with stillness [Hadar *et al.*, 1983]. Eyeblinks can occur also during pauses [Argyle and Cook, 1976].

REGULATORS: they help the interaction between speaker-listener, they control the flow of speech. Head and eye movements are coordinated to synchronize speech and organize communication. Speaker-Turn system groups them into four sets [Duncan, 1974].

MANIPULATORS: they correspond to the biological needs of the face (like blinking to keep the eyes wet).

emotional channel : Emotions are mainly displayed on the face and through the voice. Six emotions (anger, disgust, fear, happiness, sadness and surprise) were found to have universal facial expressions [Ekman, 1979] corresponding to prototypes. We have chosen to study all of them [Pelachaud *et al.*, 1991a].

We have excluded all movements not related directly to the pattern of the voice such as intentional, deliberate actions used to communicate.

2.1.1 Separability of the Channels

We differentiate facial actions by their functions (conversational signals versus punctuators) and not by the facial features it involved in the actions (eyebrows versus nose wrinkling). Therefore channels are functionally independent with each other even though they might involve the same facial regions (i.e. an eyebrow movement can accompany an accented item or be part of an emotive signal). This independence allows us to express any facial action of a channel as a set of **AUs** which has the consequence to make the computation independent of the facial model.

2.2 Vocal Parameters

A parallel between the syntax of sentences and suprasegmental features is defined. Suprasegmental features are added by the speaker to the text and their meanings are directly related to the strength of the conveyed affect as a linear function of the degree of arousal. The listener detects the speaker's emotion from prosodic features. Those are the pitch (while frequency is a physical property of a sound, pitch is a subjective one), loudness (the perceived intensity of a sound), pitch contour (the global envelope of the pitch), tempo (rate of speech) and pause.

2.2.1 Specification of the Intonation

To define the syntactic structure of intonation, we are using an extension of Janet Pierrehumbert's notation [Pierrehumbert and Hischberg, 1987]. Under this definition, intonation consists of a linear sequence of accents made from two tones (**H** and **L** for high and low tones respectively). Utterances are decomposed into *intonational* and *intermediate* phrases. Different intonational "tunes" composed of these elements are used to convey various discourse-related distinctions of "focus": old/new information, contrast and propositional attitude.

We can represent the decomposition of an utterance into intonational (or intermediate) phrases by brackets. The choice of intonational bracketing is obtained by the context the sentence is uttered and on the meaning of the utterance (what the speaker wants to focus on, what s/he considers as new information versus old) [Steedman, 1991].

The input utterance "Julia prefers popcorn" has the following bracketing and set of accents in the context defined by the question:

```
Question: Well, what about JUlia?
          What does SHE prefer?
Answer: (JUlia prefers) (POPcorn).
Accent: (L+H*      LH%) (H*  LL%).
```

(Nota Bene: * and % represent respectively pitch and boundary accents).

3 Relation Voice/Facial Movement

We present here the principles and parametrized schemes we are using in our system.

3.1 Synchrony

An important property linking intonation and facial expression is the existence of synchrony between them. That is changes occurring in speech and facial movements should appear at the same time. Synchrony occurs at all levels of speech (a blink is more adapted to phoneme level, while eyebrows action to the word level).

This is the basic principle which regulates the computation of our facial animation.

3.2 Modularity

The computation over the different channels is performed by a set of rules. These rules are derived from physiological, psychological, and linguistic considerations [Pelachaud, 1991b]. This scheme allows the user to modify, add, or delete a rule for one channel without modifying the other part of the system. Considering different channels for facial movement offers the possibility for researchers to experiment with their individual significance. Switching off one of the component channel is a way to analyze the meaning and information it conveys. For the same reasons, the system will allow modular refinement of each channel.

3.2.1 Parameters of Facial Action

Facial expressions follow the flow of speech but they are also tied to the speaker's individuality. Speakers differ in their way to punctuate speech and in the number of displayed action. Therefore two independent variables define a facial action, its *type* (set of **AUs**) and its *time of occurrence* (in the spoken utterance).

In order to vary the manifestation of speaker's individuality (submissive, persuasive, or ...) and attitude (what s/he wants to convey, such as, for instance, irony or politeness) the user can specify particular set of type and timing data for the facial actions.

4 Rule-Based Program

In these sections we describe the system of rules we are using to compute the facial expressions involved in the animation.

4.1 Different Levels of Rules

Following our principle of synchrony, each rule involves a particular level. All the rules defining the lip shapes are at the phoneme level. Blinks (appearing as conversational signals or punctuators) involve the same level. The various rules computing the existence and timing of blinks determine on which phonemes the eyes start to close, on how many phonemes the eyes remain closed, and finally on which phoneme the eyes start to open again.

On the other hand, rules related to conversational signal work at the word level (or syllable level depending on the speech-rate). They compute the beginning and end of an action not by scanning the utterance phoneme by phoneme but word by word (or syllable by syllable).

Regulators involving mainly head shifts act on parts of the utterance; for group of related utterances, head movements show identical pattern.

The rules defining the types of facial action to express an emotion is at the utterance level; i.e. it alters the entire animation.

Emotion does not modify the shape of the contour of an utterance, i.e., it does not affect either the type or the placement of the accents (which are defined by the context of an utterance, what is new/old information to the speaker). This property allows us to compute every facial action corresponding to the given intonational pattern.

4.2 Rules of Occurrence of Action

The rules corresponding to the emotion are evaluated on the top of the already found rules. They determine the effective existence of a facial movement (a sad person shows less movement than a frightened person). Some types of facial action are emotion dependent and are therefore computed at this level. For instance, a disgusted person has tendency to punctuate silences, with nose wrinkling, while a happy person smiles.

Choosing interactively an emotion not belonging to the basic set (of the six defined emotions) is not available at the present state of this study yet. But it can be obtained by adding to the main set of rules, the rules defining the occurrence of facial actions and the list of **AUs** for the chosen emotion; that is to specify for example if blinks should occur on pauses and accents.

4.3 Coarticulation Rules

There is not a unique equivalence between a phoneme and a particular lip shape due to the phenomenon of coarticulation. Coarticulation occurs due to the overlap of phonemic segments during their production. The boundary between segments are blurred. No complete set of rules exists. We have implemented *look-ahead model* [Kent and Minifie, 1977] which considers articulatory adjustment on a sequence of consonants followed or preceded by a vowel. To counterbalance some of the unsolved problems due to the incompleteness of this model, *geometric and temporal constraints* were integrated in the set of rules. They are applied on every phonemic element belonging to context-dependent clusters [Jeffers and Barley, 1971]. Some rules were defined to look at how consecutive actions are performed; they modify the intensity of an action accordingly (geometric constraints). Relaxation and contraction time of a muscle (temporal constraint) consider if the current speech posture has time to relax before the next speech posture (or, respectively, to contract after the previous one) otherwise the next speech posture is influenced by the relaxation of the current one (some action of the current speech posture remains over the next one) [Pelachaud, 1991b].

4.4 Attenuation Rules

The final animation is obtained by adding the list of facial action (set of **AUS**) occurring for each channel (**AUs** as defined in **FACS** work in an additive manner). Channels might involve the same facial regions. In such a case, attenuation rules are applied (it is done mainly for lips, head and eyes gestures). These rules attenuate conflicting movement rather than eliminate one action. For instance, the head may turn slowly during a sequence of words, while nodding (changing ups-and-downs direction at every phoneme). The intensity of these movements are modified (they decrease) but not their existence (both movements are still occurring).

5 Execution Step

The input of our program is explained here as well as the various steps of computation to produce the final set of facial expressions for the animation.

Figure 1: "Julia prefers popcorn" slow speech-rate.

Figure 2: Algorithm of Main Program

5.1 Assumption for the Input

The input of our program is an utterance with its intonational pattern specified. For the moment, we are using recorded natural speech to guide our animation. After recording a sentence, we extract from its spectrogram the timing of each phoneme and pause. We plan to use analysis-and-resynthesis methods to automate the determination of paralanguage parameters and phoneme timing.

At the beginning of the file, the user should specify the emotion. The intensity of emotion (a number between 0, for minimum intensity, and 1, for maximum intensity) should also be indicated. Furthermore, we assume that the input contains the decomposition of the utterance into its bracketed elements, the placement and type of accents. The sentence is written as a list of strings corresponding to the phonetic representation of the utterance and whose notation is compatible with Dectalk's ascii-keyboard notation [Dectalk, 1985].

5.2 Computation for each Channel

The program runs through the set of rules for each channel computing the value of the two parameters of actions (see Fig. 2). The program retrieves all valid rules corresponding to the intonational pattern defined in the input file. The chosen rules define the type and timing of occurrence of facial actions for each channel.

After reading in the input file, the computation of lip shapes is done (see Fig. 1). We apply speechreading technique [Jeffers and Barley, 1971] to group phonemes depending on their lip shapes. This clustering varies with the speech-rate. Coarticulation rules are then performed (see section 4.3).

To each **emotion** corresponds a facial expression which serves as a base with whom the other facial actions are combined. Each emotion is defined by three parameters: onset, offset (manner of appearance and disappearance of an action), and apex (time an action is stabilized).

Conversational signals are then computed. The beginning and end of an action are computed depending on the speech-rate. To follow the principle of synchrony, the program finds the closest set of phonemes whose duration from the starting point and ending point of the action are equal to the onset, respectively offset, values.

When an action occurs on a pause (**punctuators**) its onset and offset coincide with respectively the beginning and ending of the pause. The type of movement varies with the emotion (for comma, frown will be chosen in the case of anger) or with the type of pause (period will be marked by a frown, while a question mark with a raise of the eyebrows, especially when the question is not stated linguistically) [Pelachaud, 1991b].

When a blink occurs as conversational signal or punctuator (called voluntary blink) the program follows again the principle of synchrony to find its starting point. But to compute the duration of a blink, the program looks for the phonemes which have the closest timing to the average speed of a blink (it lasts about 1/4 of a sec., with 1/8 sec. of closure time, 1/24 sec. of closed eyes, and 1/12 sec. of aperture time).

Manipulators, in the present stage of this study, consider only blinking (called periodic blink). Blinks should occur periodically in order to keep the eyes wet. The program adds any necessary ones if too much time is elapsed between two consecutive voluntary blinks. We find the phonemic segment whose time from the last blink is the closest to the period of occurrence of blinks. The remaining parts of the blink are synchronized with the following segments.

Regulators are mainly specified by head and eye actions. Looking toward the listener is part of a Speaker-Turn-Signal and is emitted when the speaker wants to keep the floor. On the other hand, the speaker looks away from the listener when s/he starts speaking (i.e. Speaker-State-Signal).

Head motions are distinguished by their amplitude and frequency. Rapid movements of the head occur on accents while pauses coincide with slowing down of head motion. After the computation of all existing head movements, if no motion is specified for some phonemic elements, the head is forced to go back to its starting position.

For a first approach, we decided to implement a simpler simulation of eye movements. We supposed that when an action occurs (mutual gaze, breaking eye contact, and so on) head and eyes follow the same behavior. This means we interpret eye position as head position. Otherwise, when no action is occurring, the eyes will scan the listener's face in a random way [Argyle and Cook, 1976].

6 Our Facial Model

The goal of the facial animation model we are presented in this paper is the generation of a more expressive and natural animation of the human face. The face is a common sensitive object whose movements are composed of primary actions, such as the direct muscle deformations, and secondary actions, such as the expressive wrinkles and the subtle interaction between muscles. The secondary effects of the muscle actions take part in a non negligible way in the perception of a face.

Primary deformations correspond to more visible actions: they are the bulges, the creases, and displacements of parts of the face that are involved during the contraction of a muscle. The opening of the mouth under the action of the masseter, the frowning of the eyebrows under pyramidal muscles, the inflating of the cheek when laughing under the action of the masseter are easily noticeable examples.

Secondary actions we are considering here and which derive from muscle contractions are of two types. The more visible ones are the expressive wrinkles. These are folds created by the skin when muscles compress it. Indeed the capacity of compression of the skin is not unlimited and varies with the age, creating rolls. The contraction of the frontal which appears during surprise for example creates horizontal folds all along the forehead. We notice also as secondary actions, the secondary movements linked to the fact that muscles are not completely independent with each other. Physiologically some muscles have interleaved fibers, the action of one of the muscles can then partially act on the others. For example it is not easy to contract the upper lip and nasal wing elevator without contracting to some extent the nose wrinkler. We remark also the dependency of some symmetrical muscles: it is very difficult to frown from just one eyebrow. Some of the muscles contraction are symmetric, even if a face is never symmetric.

6.1 Generation of Primary Muscle Deformations

The muscle deformations are controlled by a dynamical physical system which computes the result of muscles action in real time.

The database of the face is a spring network (2D or 3D).

The muscles are modelled by forces of different types : the shape of attraction or repulsion field created by a force on points of the spring network belonging to its zone of influence depends on its type. Six basic types are right now implemented and can be gathered to model more complex deformations. If we write P_i^{rep} the rest position of a point P_i of the network, the forces on P_i are:

- The forces produced by the neighbor points of the spring network f_r^i
- The feedback spring force of the point to its initial position f_{rap}^i
- The external forces f_{ext} corresponding to active muscles in P_i
- The viscosity force of the environment $f_v^i = h\vec{v_i}$

where k_{ij} is the value of the stiffness of spring P_iP_j, k_i the stiffness of the feedback spring of P_i, h the viscosity coefficient of the environment, $\vec{v_i}$ the velocity vector of the point P_i, and f^{type} the computational function of the force depending on its type. the dynamical system is then written as:

$$m_i\vec{a_i} = f_r^i + f_{rap}^i + f_{ext} + f_v^i$$

where m_i is the mass of the point P_i and $\vec{a_i}$ its acceleration. This system is solved in real time by temporal relaxation.

6.2 Generation of the Expressive Wrinkles

To the spring network representation is associated a spline representation (the knots of the spring network correspond to control points of the spline surface). [1] The expression wrinkles are creases perpendicular to the action line of the muscular force which is applied on the skin. They are geometrically modelled by creating tangency discontinuities between two isolines of the spline surface.

The coefficient of skin compression is computed in function of the intrinsic age of the database, and determines the length of the curvilinear abscissa of the bulges. From the value of this curvilinear abscissa, we compute the new values of the tangents by tri-linear interpolation in a precomputed table of curvilinear abscissas based on the three parameters:

- The deviation of the tangent from its initial position around the axis of the wrinkle.
- the amplitude of the tangents.
- the distance between two control points of the two isolines that described the wrinkle.

As the distance between control points of the isolines defining a wrinkle varies with the action of muscular forces, the shape of a wrinkle varies with the muscles deformation.

6.3 Rules of Muscles Combinations

The rules of combination are given by a 2D table for which both inputs are the muscles list. The value of the links between muscles are given by a real number varying

[1] We have defined in an earlier publication how a database of the face can be rebuilt such that the isolines of the spline surface associated to the face are blended into the wrinkles lines of the face [Viaud and Yahia, 1992]. We will assume the spline surface has these characteristics.

between 0 (no link) and 1 (fully linked). The user has the possibility to act in two ways on the face:

- the geometric modeller ACTION3D: each muscle of the face is associated with a dial whose rotation activates the muscle (standard input used: dial box of the Iris Silicon Graphics workstations).
- the animation modeller ACTION3D: the user defines amplitude graphs in time for each facial muscle. The animation is then precomputed and a line test can visualize the result.

6.4 Facial Animation of the Speech

The **AUs** refer to positions. The combination rule of the **AUs** is the addition, i.e. it generates an average position between the positions defined by each of the **AUs**. The muscles are independent physiological entities acting on the face. Their combination is done by adding the amplitude. We define a table of correspondence between **AUs** and muscles. This conversion is made at the end of the program generating the list of movements linked to the speech. The output keyframe are then transformed into graphs for each muscle and a line test of animation is precomputed to visualize in real time the final animation.

7 Conclusion

To improve facial animation techniques, our study embodies subtle facial actions as expressive elements and coordinated motions. We have presented some tools allowing the generation and animation of 3D wrinkled faces, and a method of characterizing any facial movements by separating them into phonemic, intonational and emotional channels. The coordination of these various facial motions with the intonation is done by rules. This method allows us to define various speaker individualities by specifying particular sets of type and timing parameters for the facial actions and their effects on the skin (main wrinkles can be specified according to the personality of the speaker).

8 Acknowledgements

We would like to thank particularly Dr. Norman I. Badler and Dr. Mark Steedman without whom this research will not have been done. We are also grateful to Steve Platt for his very useful comments. We are thankful to Prof. Alberto Paoluzzi and his students for their kindness. We would like to thank also A. Gagalowicz for the interest and the enthusiasm he gives to this study.

References

[Argyle and Cook, 1976] M. Argyle and M. Cook. *Gaze and Mutual gaze.* Cambridge University Press, 1976.

[Dectalk, 1985] Digital Equipment Corporation. *Dectalk and DTC03 Text-To-Speech System Owner's*, 1985.

[Duncan, 1974] S. Duncan. *Nonverbal Communication*, Some signals and rules for taking speaking turns in conversations. Oxford University Press, 1974.

[Ekman and Friesen, 1978] P. Ekman and W. Friesen. *Facial Action Coding System.* Consulting Psychologists Press, Inc., 1978.

[Ekman, 1979] P. Ekman. *Human ethology: claims and limits of a new discipline: contributions to the Colloquium*, About brows: emotional and conversational signals, pages 169–248. Cambridge University Press, Cambridge, England; New-York, 1979.

[Hadar *et al.*, 1983] U. Hadar, T.J. Steiner, E.C. Grant, and F. Clifford Rose. Kinematics of head movements accompanying speech during conversation. *Human Movement Science*, 2:35–46, 1983.

[Hill *et al.*, 1988] D.R. Hill, A. Pearce, and B. Wyvill. Animating speech: an automated approach using speech synthesized by rules. *The Visual Computer*, 3:277–289, 1988.

[Jeffers and Barley, 1971] J. Jeffers and M. Barley. *Speechreading (lipreading).* C.C. Thomas, 1971.

[Kalra *et al.*, 1991] P. Kalra, A. Mangili, N. Magnenat-Thalmann, and D. Thalmann. Smile: A multilayered facial animation system. In T.L. Kunii, editor, *Modeling in Computer Graphics.* Springer-Verlag, 1991.

[Kent and Minifie, 1977] R.D. Kent and F.D. Minifie. Coarticulation in recent speech production models. *Journal of Phonetics*, 5:115–135, 1977.

[Lewis and Parke, 1987] J.P. Lewis and F.I. Parke. Automated lip-synch and speech synthesis for character animation. *CHI + GI*, pages 143–147, 1987.

[Nahas, 1988] M. Nahas, H. Huitric, and M. Saintourens. Animation of b-spline figure. *The Visual Computer*, 3:272–276, 1988.

[Parke, 1989] F.I. Parke. Parameterized models for facial animation - revisited. *Siggraph'89 Course Notes, State in the Art in Facial Animation*, 1989.

[Pelachaud *et al.*, 1991a] C. Pelachaud, N.I. Badler, and M. Steedman. Linguistic issues in facial animation. In N. Magnenat-Thalmann and D. Thalmann, editors, *Computer Animation '91*, pages 15–30. Springer-Verlag, 1991.

[Pelachaud, 1991b] C. Pelachaud. *Communication and Coarticulation in Facial Animation.* PhD thesis, Computer and Information Science Department, University of Pennsylvania, Philadelphia, Pennsylvania, 1991.

[Pierrehumbert and Hischberg, 1987] J. Pierrehumbert and J. Hischberg. The meaning of intonational contours in the interpretation of discourse. Technical memorandum, AT&T Bell Laboratories, 1987.

[Pieper, 1992] Steven D. Pieper *CAPS: Computer-Aided Plastic Surgery.* PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1992.

[Platt, 1985] S.M. Platt. *A Structural Model of the Human Face.* PhD thesis, Computer and Information Science Department, University of Pennsylvania, Philadelphia, Pennsylvania, 1985.

[Steedman, 1991] M. Steedman. Structure and intonation. *Language*, 67:260–296, 1991.

[Terzopoulos and Waters, 1991] D. Terzopoulos and K. Waters. Techniques for realistic facial modelling and animation. In N. Magnenat-Thalmann and D. Thalmann, editors, *Computer Animation '91*, pages 45–58. Springer-Verlag, 1991.

[Viaud and Yahia, 1992] M.L.. Viaud and H. Yahia. Facial Animation with Wrinkles. *Eurographics'92*, September 1992.

[Waters, 1987] K. Waters. A muscle model for animating three-dimensional facial expression. *Computer Graphics*, 21(4):17–24, July 1987.