Recognition of Non-Manual Expressions in Brazilian Sign Language

Emely Pujólli da Silva and Paula Dornhofer Paro Costa
School of Electrical and Computer Engineering, University of Campinas, Campinas, Brazil

***Abstract*— Individuals with some degree of hearing impairment typically face difficulties in communicating with hearing individuals and during the acquisition of reading and writing skills. Sign language (SL) is a language structured in gestures that, as any other human language, present variations around the world and that is widely adopted by the deaf. Automatic Sign Language Recognition (ASLR) technology aims to translate sign language gestures in to written or spoken sentences of a target language with the goal of improving the communication between deaf and hearing individuals. Recently, ASLR research community started to appreciate the importance of non-manual signs, since they are related to grammatical and affective meaning of a sentence. The present research project aims to develop an automatic recognition system of non-manual signs expressed by the face. Our object of study is the Brazilian Sign Language, also called LIBRAS. This paper presents our first steps to approach the problem.**

I. Introduction

Sign language is the linguistic system used by the group of people with hearing impairments to communicate with members of the group or others [1]. Frequently, the access of deaf people or hard of hearing individuals to school, public services and medical assistance, is possible only through a sign language interpreter.

In this context, the Automatic Sign Language Recognition (ASLR) technology emerges as a liberating technology to help deaf people in situations that require privacy or a sign language interpreter is not available. In addition, ASLR can be applied as an assistive writing tool for young learners that were born deaf and typically face great difficulties during the acquisition of reading and writing skills [2].

One of the challenges involved in the development of ASLR technology is that sign languages, in the same way that modern spoken languages, emerged spontaneously, evolved naturally and reflected the worldwide sociocultural differences, giving origin to a wide range of variations such as the British Sign Language (BSL), the American Sign Language (ASL), the Japanese Sign Language (JSL), the Brazilian Sign Language (LIBRAS) and many others.

However, analogously to spoken languages, in which is possible to combine phonemes to form words, it is possible identify a set of parameters that are combined simultaneously to create signs [3]. Such parameters can be divided into manual signals and non-manual signals. Manual signals (MS) correspond to parameters such as the hand-shape of the sign, the location of the sign execution in front body region and the path and speed of the hands while the sign is made. On the other hand, non-manual signals (NMS) correspond to head and shoulders movements, facial features such as eye brown motion and lip-mouth movements.

During the last decade, many efforts have been made to explore the automatic recognition of MS gestures [4], [5]. However, the non-manual signals have great impact in sign appearance and they are an essential component of this type of communication. For this reason, recent approaches have been also exploring the NMS recognition [6], [7] and the mechanisms to combine them with MS recognition [8], [9]

The objective of the present work is to develop a NMS automatic recognition methodology capable of contributing to MS recognition rates and improving the recognition of the affective states communicated by the interlocutor. Our object of study is the Brazilian Sign Language (LIBRAS). The underlying concepts adopted in the work are common to other sign languages. The present project is part of the TAS[1] initiative, a multidisciplinary research group composed of deaf individuals, sign interpreters, linguists, engineers and computer scientists, that aims to advance the development of assistive technologies for the deaf.

II. Stage of the Research

This doctoral research project is organized into the following phases:

1) Study of the state-of-art of ASLR techniques;
2) Survey of facial expressions and their role in LIBRAS;
3) Proposal of a taxonomy of facial expressions in LIBRAS;
4) Definition of a LIBRAS corpus;
5) Proposal of a model for the recognition of non-manual signals in LIBRAS;
6) Evaluation of the model through objective computational experiments and with deaf users.

The project is entering its fourth phase, going on towards the definition of a corpus in Libras, composed of video clips that contains samples of the facial expressions identified during the second and third phases of the project.

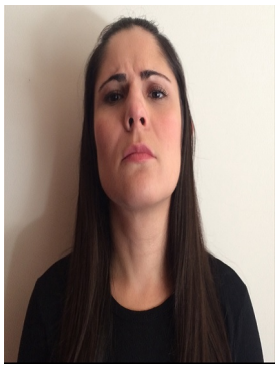
In the following sections we present our partial results regarding the study of the characteristic facial expressions in LIBRAS and we propose a novel taxonomy of such expressions.

III. Facial Expressions in LIBRAS

In Brazil, LIBRAS has the status of second official language since 2002 and its grammar has been subject of intensive study during the last decade.

The research for this paper was financially supported by the National Council for the Improvement of Higher Education (CAPES).

[1]Brazilian Portuguese acronym for *Tecnologias Assistivas para Surdos* (Assistive Technologies for the Deaf). For additional references: http://www.tas.fee.unicamp.br

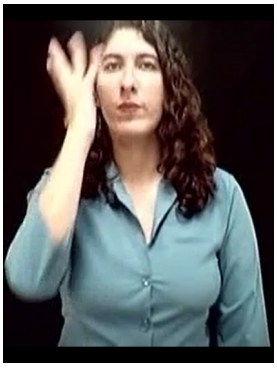


(a) Facial expression: Brief and upward movement of the head, frown and crooked mouth down

(b) Facial expression: Projected lips and frown.

(c) Sign for “increase”.

(d) Sign for “to fill”.



(e) Sign for “beautiful”.

(f) Sign for “very beautiful”.

Fig. 1. In the performance of the signs in LIBRAS we can analyze the variation of the facial expressions by the images (a) and (b). In images (c) and (d) the signs “increase” and “to fill” are performed with the same manual gesture. Their distinction is only based on the cheeks action and of the mouth blowing. Also in images (e) and (f) the interpreters performed the sign “beautiful” and “very beautiful”. The intensity of the sign is displayed by the face expression which passed from neutro to raised eyebrows and crooked mouth up.

Our approach consisted of conducting a systematic review of references dedicated to the description of LIBRAS language and, in particular, we dedicated special attention to those references that describes the production of non-manual signals. The definition of the relevant references for this study was done through numerous interactions with deaf individuals, sign interpreters and linguistic researchers that are part of the TAS initiative.

Facial expressions may change widely in LIBRAS (see Fig. 1). A unique facial expression may be combined to many different manual signs and the meaning of the same manual sign can vary significantly depending on the facial expression that accompanies it. Our studies unveiled that there is still limited documentation regarding the list of facial expressions that carry relevant meaning in LIBRAS and how they combine with manual signals.

In LIBRAS, facial expressions that convey an idea of feeling and emotion, are called Affective Facial Expressions (AFE). Affective facial expressions can start before a specific sign and end after the sentence has been completed. In other words, AFEs modulate the whole sentence, modifying the full meaning of a sequence of signs. AFEs are adopted, for example, when the signer communicates ideas sarcastically or when he/she is describing a sad event. A visual characteristic of AFEs is that they employ an integrated set of facial muscles.

Grammatical facial expressions in Libras are expressions that typically occur at specific points of a sentence or they are associated to a specific sign execution [10], [11]. Observing the different properties of grammatical facial expressions we can categorize them into Grammatical Facial Expression for Sentence (GES), Grammatical Facial Expressions of Intensity (GEI) and Grammatical Facial Expressions of Distinction (GED).

Grammatical facial expression for sentence defines the type of sentence that is being signed. In LIBRAS, there are GES markers that are expressed by the face and head movements:

- **WH-question** (WH): generally used for questions with What, who, when, why, where and how;
- **Yes/No question** (YN): used when the question being asked has a yes or no answer;
- **Doubt question** (DQ): it is used to emphasize the information that will be supplied;
- **Topic** (T): when one of the sentence’s constituents is displaced to the beginning of the sentence;
- **Negation** (N): used in negative sentences;
- **Assertion** (A): used when making assertions;

- **Conditional clause** (CC): used in subordinate sentence to indicate a prerequisite to the main sentence;
- **Focus** (F): used to highlight new information into the speech pattern;
- **Relative clause** (RC): used to provide more information about something.

In addition, Grammatical Facial Expressions of Intensity differentiate the meaning of the sign assuming the role of quantifier. For example, the same sign associated to the word "beautiful", can have its meaning attenuated to "cute" or "very beautiful", depending on the signer's facial expression (see Fig. 1(e) and (f)).

Finally, without its characteristic Grammatical Facial Expression of Distinction a sign is incomplete and cannot be distinguished from other signs with the same manual signal. In other words, GFDs helps to define the meaning of a sign.

IV. Proposed Taxonomy for Facial Expressions in LIBRAS

As a first contribution of our work, we conducted a careful survey in the literature to identify the elementary characteristics of facial expressions and head movements that are associated to affective and grammatical facial expressions in Libras. Table I summarizes and standardizes our findings considering the following references: Quadros and Karnopp [12], [13], Araújo [14], Freitas et al. [10] and the recent findings of TAS project [15], [16].

Table I highlights that different authors refers to different sets of relevant expressions and there are some expressions that are not mentioned by one or more authors. For example, the semi-open mouth (blowing) expression was considered relevant by the linguists of the TAS project, but it is not explicitly mentioned by other references. Also, different authors use different ways to describe the relevant NMS in Libras. Table I presents a standardized nomenclature that takes into account the movement performed by the face and the head.

Alternatively, it is possible to group the facial and head NMS regarding their meaning in the sentence. In other words, we construct the relation between the facial expression and its semantic function. In this case, it is clear that the result is a composition of the non-manual signals described in the Table I. As a result, we built a more comprehensive list of non-manual signals and classified the NMS into AFS and GFS. The results are shown in Table II.

In sign gesture classification there are two main approaches, which is either employ a single classification stage, or represent the gesture as consisting of simultaneous components which are individually classified and then integrated together for sign-level classification [17]. With the classification described in Table II, we proceed to the second approach, where we will be based on intersections and associations. Thus, according to the NMS picked up by a system, we can indicate which list of expressions that signal belongs to, making integration with manual signal recognition methods easier. Table II was constructed from the analysis of works [10], [13], [15], [18]. Additional combinations can be added to it as our study advances.

TABLE I

Comparative summary of facial expressions among the related works in linguistic literature.

Facial Expressions	Authors			
Face				
Upper part of the face				
Frown	Q	D	T	A
Raised eyebrows	Q	D	T	A
Left / Right eyebrow raised			T	
Wide open eyes	Q	D	T	A
Slightly closed eyes	Q		T	A
Closed eyes		D		
Left / Right eye closed			T	
Look at the speaker		D		
Direct the eyes	Q		T	
Lower part of the face				
Inflated cheeks	Q	D	T	
Contracted cheeks	Q	D	T	
Contracted cheeks and projected lips	Q	D		
Contracted lips			T	A
Projected lips			T	A
Only left / right cheek inflated	Q		T	
Run the tongue against the lower part of the cheek	Q		T	
Smile with apparent teeth			T	
Crooked mouth up		D	T	A
Crooked mouth up laterally			T	
Crooked mouth down		D	T	A
Crooked mouth down laterally			T	
Contraction of the upper lip	Q		T	
semi-open mouth (blowing)			T	
Sibilant tongue[2]		D	T	
Open mouth		D	T	A
Clenched teeth			T	
Nose wrinkle	Q		T	
Swinging lingua-alveolar[3]			T	
Tongue in lisp position			T	
Mouth movement		D		
Head				
Balance back and forth (yes)	Q	D	T	A
Quick nod			T	
Balancing sideways (no)	Q	D	T	A
Brief and upward movement of the head		D	T	
Forward lean	Q		T	A
Tilt to the side	Q	D	T	
Tilt back	Q		T	A

Q-(Quadros and Karnopp, 2009) [12], D-(Araújo, 2015) [14], T-(TAS project, 2016) [13], [15] A-(Freitas, Peres and de Moraes, 2014) [10].
2-In other words, open mouth and tongue in movement.
3-This term refers to the region within the mouth in the inside margin of the upper central incisors.

V. CONCLUSIONS AND FUTURE WORK

In a sign language, non manuals cues are used as modifiers for specific signs and sentences. The grammatical and affective function of these cues and means of associating them with facial expressions are still being investigated for Brazilian Sign Language. The summarization and classification of facial expressions according to syntactic functions are the main results of this work so far.

The next step in our research is built a corpus, with video images from interpreters performing examples of some features discussed in this paper. In addition, we will perform a tracking of the facial points during the execution of each facial expression, so that in the future we can use this data

TABLE II

TYPES OF FACIAL EXPRESSIONS ACCORDINGLY WITH THE NON-MANUAL SIGNS CLASSIFICATIONS.

AFE	Left / Right eyebrow raised; Raised eyebrows and wide open eyes; Raised eyebrows, wide open eyes and open mouth; Slightly closed eyes and crooked mouth up; Smile with apparent teeth; Smile with apparent teeth and open mouth; Lowered eyebrows and crooked mouth down; Frown and contraction of the upper lip; Crooked mouth up laterally.		
GFE	GES	WH	Brief and upward movement of the head and frown.
		YN	Brief and upward movement of the head and raised eyebrows.
		DQ	Frown, slightly closed eyes and contracted lips.
		T	Brief upward and forward movement of the head, raised eyebrows, open mouth, projected lips; Quick nod, brief upward movement and wide open eyes; Quick nod, brief upward movement, raised eyebrows and wide open eyes; Quick nod, brief upward movement, raised eyebrows, open mouth and projected lips.
		N	Crooked mouth down; Quick nod, frown and crooked mouth down; Head balancing sideways.
		A	Balance back and forth of the head.
		CC	Brief and upward movement of the head and raised eyebrows.
		F	Brief upward and forward movement of the head, raised eyebrows, open mouth, projected lips; Quick nod, brief upward movement and wide open eyes; Quick nod, brief upward movement, raised eyebrows and wide open eyes; Quick nod, brief upward movement, raised eyebrows, open mouth and projected lips.
		RC	Raised eyebrows.
	GEI		Frown < Frown and Slightly closed eyes; Inflated cheeks and semi-open mouth < inflated cheeks, semi-open mouth and frown; Contracted cheeks and frown < contracted cheeks; Contracted lips and frown < contracted lips; Projected lips and frown < projected lips; Open mouth and frown < open mouth; Crooked mouth up< Smile with apparent teeth; Quick nod < balance back and forth of the head;
	GED		Left eye closed; Inflated cheeks; only right cheek inflated;

Semantic Functions: AFS- Affective Facial Expression; GFS- Grammatical Facial Expression; WH- WH-question; YN-Yes/No questions, DQ- Doubt question, T- Topic, N- Negation, A- Assertion, CC- Conditional Clause, F- Focus, RC- Relative Clause.

in the construction of a model for automatic recognition of facial expressions in Libras.

REFERENCES

[1] George Caridakis, Stylianos Asteriadis, and Kostas Karpouzis. Non-manual cues in automatic sign language recognition. *Personal and ubiquitous computing*, 18(1):37–46, 2014.

[2] Richard Kennaway, John RW Glauert, and Inge Zwitserlood. Providing signed content on the internet by synthesized animation. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 14(3):15, 2007.

[3] Scott K Liddell and Robert E Johnson. American sign language: The phonological base. *Sign language studies*, 64(1):195–277, 1989.

[4] Helen Cooper, Brian Holt, and Richard Bowden. *Sign Language Recognition*, page 539562. Springer London, London, 2011.

[5] Thad Starner, Joshua Weaver, and Alex Pentland. Real-time american sign language recognition using desk and wearable computer based video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1371–1375, 1998.

[6] Epameinondas Antonakos, Anastasios Roussos, and Stefanos Zafeiriou. A survey on mouth modeling and analysis for sign language recognition. In *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, volume 1, pages 1–7. IEEE, 2015.

[7] Malladihalli S Bhuvan, Vinay Rao, Siddharth Jain, TS Ashwin, Ram MR Guddetti, and S Pramod Kulgod. Detection and analysis model for grammatical facial expressions in sign language. In *Region 10 Symposium (TENSYMP), 2016 IEEE*, pages 155–160. IEEE, 2016.

[8] Hee-Deok Yang and Seong-Whan Lee. Robust sign language recognition by combining manual and non-manual features based on conditional random field and support vector machine. *Pattern Recognition Letters*, 34(16):2051–2056, 2013.

[9] Luis Quesada, Gabriela Marín, and Luis A Guerrero. Sign language recognition model combining non-manual markers and handshapes. In *Ubiquitous Computing and Ambient Intelligence: 10th International Conference, UCAmI 2016, San Bartolomé de Tirajana, Gran Canaria, Spain, November 29–December 2, 2016, Proceedings, Part I 10*, pages 400–405. Springer, 2016.

[10] Fernando A Freitas, Sarajane M Peres, Clodoaldo A de Moraes Lima, and Felipe V Barbosa. Grammatical facial expressions recognition with machine learning. In *FLAIRS Conference*, 2014.

[11] Fernando de Almeida Freitas. Reconhecimento automático de expressões faciais gramaticais na língua brasileira de sinais. Master's thesis, Universidade de São Paulo (USP), 2011.

[12] Ronice M de Quadros and Lodenir B Karnopp. *Língua de sinais brasileira: estudos lingüísticos*. Artmed Editora, 2009.

[13] Nelson Pimenta and Ronice M de Quadros. *Curso de LIBRAS 1*. 2006.

[14] Adriana DS de Araujo. As expressões e as marcas não-manuais na língua de sinais brasileira. Master's thesis, Universidade de Brasília (UnB), 2013.

[15] José M De Martino, Ivani R Silva, Carmen Z Bolognini, Paula D Paro Costa, Kate MO Kumada, et al. Signing avatars: making education more inclusive. *Universal Access in the Information Society*, pages 1–16, 2016.

[16] José M De Martino, Paula D Paro Costa, A Benetti, Luciana A Rosa, Kate MO Kumada, and IR Silva. Building a brazilian portuguese-brazilian sign language parallel corpus using motion capture data. In *Proceedings of Workshop of Corpora and Tools for Processing Corpora*, pages 56–63, 2016.

[17] Sylvie CW Ong and Surendra Ranganath. Automatic sign language analysis: A survey and the future beyond lexical meaning. *IEEE transactions on pattern analysis and machine intelligence*, 27(6):873–891, 2005.

[18] Fernando C. Capovilla, Walkiria D. Raphael, and Aline C. Mauricio. *Novo dicionário enciclopédico ilustrado trilíngue da Língua de Sinais Brasileira (Novo Deit-Libras)*. Edusp, 2008.