TOWARDS CONVERGENCE OF METHODS FOR SPEECH AND SIGN SEGMENTATION

André Xavier [1], Oksana Tkachman [1], Bryan Gick [12]
[1]Department of Linguistics, University of British Columbia, Vancouver, BC, Canada
[2]Haskins Laboratories, New Haven, CT, USA

1   Présentation (Français)

Comme la parole, déterminer la limite d'un signe dans un flot de signes n'est pas une tâche facile. Pour certains chercheurs, le signe commence lorsque la main atteint l'emplacement où le signe est produit et termine lorsque la main quitte ce lieu [1-3]. Par conséquent, ces chercheurs laissent les mouvements transitoires hors des limites du signe. D'autres chercheurs pensent que les transitions doivent être partiellement ou entièrement considérées comme partie du signe, parce que (1) certaines caractéristiques articulatoires du signe sont visibles même avant ou encore après le signe est produit, et (2) les percepteurs sont capables de deviner les signes avec seulement l'information transmis dans les transitions [4-8]. L'étude présente utilise les données de video du langue des signes brésilienne (Libras; Xavier 2014) pour évaluer de façon critique les critères traditionnellement utilisés pour délimiter les éléments lexicaux dans le flot de signes. Utilisant la théorie de commande de moteur, nous proposons un traitement unifié de la production de langue, quelle que soit sa modalité.

2   Presentation (English)

Similar to speech, determining sign boundaries in a signing stream is not an easy task. For some researchers, the sign starts when the hand reaches the position where the sign is produced and ends when the hand leaves that position [1-3]. Thus, these researchers leave transitional movements out of the sign limits. Other researchers think that transitions should be partially or entirely regarded as part of the sign, because (1) some articulatory features of the sign are visible even before or still after the sign is produced, and (2) perceivers are able to guess signs solely drawing on information conveyed during transitions [4-8]. The present study uses video data of Brazilian Sign Language (Libras; Xavier 2014) to critically evaluate the criteria traditionally used to delimit lexical items in the sign stream. Using motor control theory, we propose a unified treatment of language production, regardless of its modality.

3   The study

3.1   Background

According to the traditional view, the location of the hand in sign language delimits the boundaries of a sign: the sign begins when the hand reaches the position where the sign is initially or entirely produced, and ends when the hand leaves that position [1-3]. This is also how signs are usually depicted in dictionaries, where only features crucial for sign recognition are depicted [7-8]. As a result, the *transitional movements*, that is, movements produced between signs to place the hand at a certain location or to move it to the rest position [10], are left out as irrelevant which results in multiple gaps in the signing stream.

The alternative view claims that the signs last longer, from the moment when some articulatory features can be recognized and until these features disappear in the transitional movements, and signers can guess the sign based solely on these features of the transitional movements [6-8]. Jantunen (2010b) also shows that mouthing can also indicate sign boundaries [11]. Gesture scholars [4] go even further and analyze the gesture/sign as production phases of preparation, stroke, and retraction.

3.2   Method and analysis

We used video data of Brazilian Sign Language (Libras) [9] to critically evaluate the traditional criteria of sign segmentation. The data have been analyzed and annotated using ELAN software [12].

We follow the alternative view [4, 6-8] in considering transitional movements as parts of the signs, but we further refine this claim by taking into account speech motor control theory, also successfully applied to spoken language production analysis [13]. In this approach, articulatory movement coincides with the end of an articulatory task, after which the articulator loses the control over the movement. Thus, when the hand has clearly accomplished its task, the sign ends. Therefore, the sign consists of the preparation and stroke phases only. We keep this distinction based on the qualitative differences between transitions and signs demonstrated in [6]. Retractions are annotated but not considered part of the sign.

This approach makes delimiting of sign boundaries easier compared to Jantunen's proposal [6-8], because coarticulation effects make it hard to distinguish between articulatory features of one sign and of the following sign, which makes determining a boundary between them problematic. In our transcriptions, the signing stream is broken into a sequence of preparation and stroke phases with no gaps between signs. We place the sign boundary at the frame where the hand starts to exhibit at least one of the features of the sign, which we take as the beginning of the sign production. Determining the beginning of the sign production is easier when the sign is at the onset of an utterance than when it follows some other sign. In the latter case, we place the sign boundary on the frame that

immediately follows the last frame of the stroke, similar to the beginning of retraction phases.

3.3 Preliminary results and discussion

The preliminary results show that the adopted approach is effective in determining sign boundaries as well as enables us to observe hand behaviour left out in the traditional approach to sign segmentation. Here we give examples of coarticulatory effects in one- and two-handed signs.

In two-handed signs, the dominant hand of a sign in the utterance onset position starts its preparation before the nondominand hand; however, both hands start the stroke synchronously. On the other hand, if the two-handed sign follows another two-handed sign, the hands act synchronically both in preparation and stroke phases. If a one-handed sign follows the two-handed sign, however, the nondominant hand of the two-handed sign moves back to the rest position before the dominant hand has finished its stroke. This effect is not observed if the two-handed sign is utterance-final; both hands produce the stroke and retract synchronously.

We also found some counter-evidence to Jantunen's claim [8] that mouthing coincides with sign boundaries: in some cases mouthing preceded or lags behind the preparation phase of the hand(s), or lasts across more than one sign.

4 Conclusion

Comparing natural human languages of different modalities is important in order to determine which aspects of those languages are modality-specific, and which are characteristic of all human languages. Likewise, the strength of a theory can be evaluated by applying it not only to data for which the theory was proposed, but also to additional data that were not originally meant to be accounted for by the theory. In this study we have successfully applied speech motor control theory to sign language data and the preliminary results show that methods used by speech researchers to delimit units in the speech stream are a good fit for delimiting units in the sign stream as well. By analyzing signs as a combination of preparation and stroke phases we not only were successful in identifying sign boundaries without leaving any gaps in the annotation, but also were able to capture coarticulation effects left out in the traditional approach to sign segmentation. This study points to the possibility of a unified account of articulatory movements of human language, regardless of its modality.

References

[1] O. Crasborn, and I. Zwitserlood. *Annotation of the Video Data in the* "Corpus NGT." Department of Linguistics and Centre for Language Studies, Radboud University Nijmegen, the Netherlands. 2008. http://hdl .handle.net/1839/00-0000-0000-000A-3F63-4

[2] T. Johnston. Guidelines for Annotation of the Video Data in the Auslan Corpus. Department of Linguistics, Macquarie University, Sydney,Australia. 2009. http://media.auslan.org.au/media/upload/attachments/Annot ation Guidelines_Auslan_CorpusT5.pdf (accessed January 15, 2010).

[3] R. E. Johnson, and S. K. Liddell. Toward a Phonetic Representation of Signs: Sequentiality and Contrast. *Sign Language Studies* 11: 241–74, 2011.

[4] S. Kita, I. Van Gijn, and H. Van der Hulst. Movement phases in signs and co-speech gestures, and their transcription by human coders. In *Gesture and sign language in human-computer interaction*, pp. 23-35. Springer Berlin Heidelberg, 1998.

[5] J. Bressem, and S. H. Ladewig. Rethinking gesture phases: Articulatory features of gestural movement?. *Semiotica*, 184: 53-91, 2011.

[6] T. Jantunen. On the role of transitions in signed language (or: What's wrong with the sign?). Paper presented at the 10th conference on the Theoretical Issues in Sign Language Research (TISLR10), Purdue University, West Lafayette, Indiana (USA), September 30–October 2. 2010.

[7] T. Jantunen. Signs and transitions: Do they differ phonetically and does it matter? *Sign Language Studies* 13(2): 211–237, 2013.

[8] T. Jantunen. How long is the sign? *Linguistics* 53(1): 93-124, 2015.

[9] Xavier, André Nogueira. "Uma ou duas? Eis a questão! Um estudo do parâmetro número de mãos na produção de sinais da língua brasileira de sinais (libras)". 158 f. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2014.

[10] S. K. Liddell, and R. E. Johnson. American sign language: The phonological base. *Sign language studies*, *64*(1): 195-277, 1989.

[11] T. Jantunen. A comparison of two linguistic sign identification methods. In Philippe Dreuw, Eleni Efthimiou, Thomas Hanke, Trevor Johnston, Gregorio Martínez Ruiz & Adam Schembri (eds.), *Proceedings of the 4th workshop on the representation and processing of sign languages: Corpora and sign language technologies* (organized as a part of the 7th Language resources and evaluation conference LREC at Valletta, Malta, May 22–23, 2010), 129–132. Paris: ELRA. 2010b

[12] Technical Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, ELAN – Linguistic Annotator (Version 4.6.2) [Computer Program]. First retrieved November 1, 2013, from http://www.lat-mpi.eu/tools/elan/

[13] C. P. Browman, and L. Goldstein. Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-180, 1992.