Lexical Sign Language Resources: Synergies between Empirical Work and Automatic Language Generation

Thomas.Hanke@sign-lang.uni-hamburg.de

Universität Hamburg

Introduction

Sign languages are the preferred communications medium for most Deaf people around the world. In Europe, sign language users form one-out-of-thousand linguistic (and cultural) minorities within the hearing societies. Sign language uses a number of visually distinctively recognisable articulators (hands, facial expression, mouth, body) in parallel and fully exploits spatial and temporal relations to establish grammatical features.

As sign languages have no written form, language resources for sign language often use "phonetic" notations, such as HamNoSys (Prillwitz et al., 1989 and Schmaling/Hanke, 2001). However, the current state-of-the-art for sign language notation is far away from being a full compensation for an orthography (Miller, 2001), which in general is the main access key to language data for written language as well as annotated speech. We therefore consider it essential for sign language corpus annotation to explicitly link tokens to lexical entities. It is obvious that sophisticated tool support is needed for that to become practical.

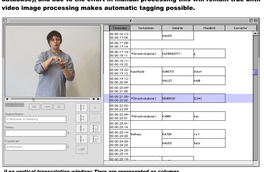
A sign that is glossed HAMBURG18 in our database together with its HamNoSys and SiGML representation

Sign Language Corpus Transcription

Out of the many corpus transcription tools available today, some address the special needs of the sign language transcribers community, esp. tight integration of high-resolution video and support for non-standard fonts (for sign language phonetic notation). The distinctive feature of iLex, the tool we have developed for this purpose (Hanke, 2002 and Hanke et al., 2001), is that signing cannot only be tagged with text, e.g. glosses, but primarily with database references to lexemes. (Glosses are spoken-language labels for signs that match the semantics of a sign as closely as possible. The phonetic form of the sign is not reflected at all. Glossing is widely used not only in academic contexts, but also for note-taking by deaf people. However, thee danger of using another language for describing utterances in one language has to be always kept in mind, cf. Pizzuto/Pietrandrea, 2001.)

This means that tagging is mainly the task of token-type matching. The relational model allows the user to search candidate types by meaning, form description (including support for fuzzy search), or grammatical class. When browsing through the candidate types, the user can immediately access video clips of prototype tokens for the types. At the same time, it is possible to verify tokens assigned to a type to be verified not only by inspecting descriptive features, but also by viewing the source video data.

Size considerations do not render this approach infeasible neither for the time being nor for the near future: Current sign language corpora do not exceed the magnitude of hundreds of thousands of tokens (sizes easy to handle for any database), and due to the effort in manual processing this will remain true until video image processing makes automatic tagging possible.



iLex vertical transcription window: Tiers are represented as columns

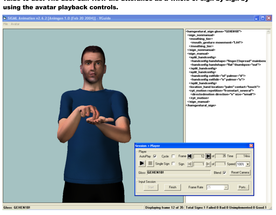
eSIGN Editing Environment

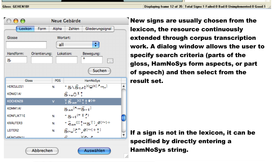
In the eSIGN project , our goal is to lower the barriers for Deaf people from participating in tomorrow's information society by creating efficient means to provide information in sign language. As the information in sign language is provided by an on-screen avatar driven by SiGML, an XML encoding for the HamNoSys sign notation (Elliott et al., 2000 and Kennaway, 2002), bandwidth requirements are minimal. A key point here is the speed in which a translator can adapt the signed text to contents changes in the source material.

The editing environment (cf. Hanke et al., 2003) created for this purpose gives the user an economic approach to create signed sequences without restrictions on word order, grammatical constructions, or the choice of lexical items.

In the editor, each utterance of the signed text is handled separately. In order to create or modify an utterance, the user opens a window to specify the sequence of signs.

Signs are represented by glosses as well as a couple of form aspects. Some columns are governed by the lexical entries. I.e. the lexical entry determines whether a field may be filled by the user (otherwise it is grey) and which default value to use. The user can view the utterance as a whole or sign-by-sign by using the avatar playback controls.



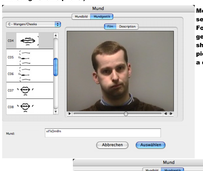


New signs are usually chosen from the lexicon, the resource continuously extended through corpus transcription work. A dialog window allows the user to specify search criteria (parts of the gloss, HamNoSys form aspects, or part of speech) and then select from the result set.

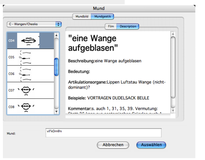
If a sign is not in the lexicon, it can be specified by directly entering a HamNoSys string.

For numbers, fingerspelling as well as discourse structure signals, special editors are available.

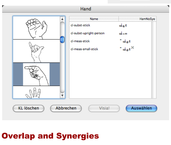
Lexical entries usually have a mouthing, either a mouth picture or a mouth gesture. As there is no strict one-to-one mapping, the user has the option to specify an alternate form. For mouth pictures, a spoken language text can be entered and then be converted to the SAMPA (Gibbon et al., 1997) pronunciation encoding used for viseme description. (The pronunciation database builds on the Bonn machine-readable pronunciation dictionary, Portele et al. 1995. The SAMPA codes are used for reading convenience, they are so to say labels of the visemes, with a viseme having as many labels as it represents visually undistinguishable phons.)



Mouth gestures are selected from a list. For each mouth gesture, the code is shown along with a picture, a movie, and a description.



Other editors allow the user to specify the kind of inflection that is to be applied to the sign. Values chosen here modify the HamNoSys string for the sign finally to be sent to the avatar.



Overlap and Synergies

Certainly the greatest time saver in creating signed contents with the eSIGN editor is the integration of the lexicon. Even if its main use as the repository for empirical work shows and requires some compromises on the side of the eSIGN contents creators, it saves the time consuming ad-hoc notation of signs in more than 90% of the cases. For the transcribers, on the other side, the animation path provides a convenient feedback loop which now forms an essential part of the strategy for quality assurance in the transcription database: The transcriber can immediately send any notation to the avatar and verify that it matches the observed token or the intended type citation form.

To a certain extent, an iLex transcription can be converted into an eSIGN document and vice versa: Both manual and nonmanual form can be converted back and forth. Differences in other tiers are by intention, as the normative use of higher-level concepts for efficient specification of sign modification is not acceptable for the transcription trying to minimise inherent interpretation of the data in the first step. (The transcript in the second figure was created from the eSIGN document in figure 3 and then time-aligned to the video.)

Outlook

The eSIGN approach to describe signed utterances as sequences of signs works quite well in the project's current domain, informative texts. In casual signing, however, co-articulation is a usual phenomenon. While two signs in parallel pose no problem at all for an interlinear transcription tool such as iLex, the eSIGN editor needs to become more flexible without loosing too much of its efficiency.

The eSIGN user interface still relies on written language, the tool is therefore currently targeted at bilingual users, such as sign language interpreters and bilingual Deaf people. Further work is needed to make the tool more suitable for monolingual (sign language) users.

As animation quality improves, the integration of iLex and eSIGN technologies also offers a possibility of practical relevance for any sign language researcher: It allows the faithful reproduction of data where the original data cannot be made available as it is virtually impossible to anonymise sign language video. (Facial expressions and mouth movements, for example, are integral parts of at least European sign languages.)

References

Elliott, R., J. Glauert, R. Kennaway and I. Marshall, 2000. The development of language processing support for the ViSiCAST project. In *The fourth international ACM conference on Assistive technologies ASSETS*. New York: ACM. 101-108.

Gibbon, D., et al. (eds). 1997. *Handbook of Standards and Resources for Spoken Language Systems*. Berlin: de Gruyter.

Hanke, T., 2002. iLex - A tool for sign language lexicography and corpus analysis. In: M. González Rodríguez, Manuel and C. Paz Suarez Araujo (eds.): Proceedings of the third International Conference on Language Resources and Evaluation. Las Palmas de Gran Canaria, Spain; Vol. III. Paris: ELRA. 923-926.

Hanke, T., R. Konrad and A. Schwarz, 2001. GlossLexer – A multimedia lexical database for sign language dictionary compilation. *Sign Language and Linguistics* 4(1/2): 161–179.

Hanke, T., H. Popescu and C. Schmaling, 2003. eSIGN: HPSG-assisted sign language composition. Paper presented at the Gesture Workshop 2003, Genova.

Kennaway, R. 2002. Synthetic animation of deaf signing gestures. In: I. Wachsmuth and T. Sowa (eds.), *Gesture and sign language in human-computer interaction. International Gesture Workshop, GW 2001, London, UK, April 18-20, 2001 Proceedings*. Berlin: Springer. 146-157.

Miller, C., 2001. Some reflections on the need for a common sign notation. *Sign Language and Linguistics* 4(1/2): 11-28.

Pizzuto, E., and P. Pietrandrea, 2001. The notation of signed texts: Open questions and indications for further research. *Sign language and linguistics* 4 (1/2):29-45.

Portele, T., J. Krämer, and D. Stock, 1995. Symbolverarbeitung im Sprachsynthesesystem Hadifix. In: *Proc. 6. Konferenz Elektronische Sprachsignalverarbeitung 1995, Wolfenbüttel*. 97-104.

Prillwitz, S. et al., 1989. *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Hamburg: Signum.

Schmaling, C. and T. Hanke, 2001. HamNoSys 4.0. In: T. Hanke (ed.), *Interface definitions*. ViSiCAST Deliverable D5-1. [This chapter is available at http:// www.sign-lang.uni-hamburg.de/projekte/ HamNoSys/HNS4.0/englisch/HNS4.pdf]

Acknowledgment

The work described here was partially supported by the European Commission in the eSIGN project under the eContent programme. eSIGN is a collaboration of seven partners from three countries: Televirtual (UK), UEA (UK), RNID (UK), Norfolk City Council (UK), Vistaal (NL), EDS (DE), UHH (DE). More information on eSIGN can be found at

http://www.sign-lang.uni-hamburg.de/eSIGN/