

## **New developments in the annotation of gesture and sign language data: Advances in the SignStream™ project**

SignStream<sup>®</sup> is database program for transcription distributed on a non-profit basis to students, educators, and researchers. Version 2 (for Macintosh computers running System 8.1 or later) is now available; version 3 (a Java reimplementation) is currently under development.

The program provides researchers with a way to produce a fine-grained, multi-level transcription of video-based data. Users can transcribe a variety of co-occurring items (e.g. gestures, movements of the head and eyes). The SignStream display makes visible the temporal relations among the coded items. Users can browse through transcriptions and play back selected subparts of the transcribed video. The built-in search facility allows the user to search through the database to find varying combinations of events.

The demonstration will cover the following: current functionality; program development and anticipated features of SignStream version 3; availability of American Sign Language (ASL) data collected in the National Center for Sign Language and Gesture Resources (NCSLGR) and transcribed using SignStream; and current computational research that exploits the SignStream-based corpora.

### **Current Functionality**

While the program was developed to aid in the linguistic analysis of signed languages (specifically ASL), the capabilities and flexibility of the program make it ideal for transcription and analysis of gesture data on video.

A SignStream database consists of a selection of (user-defined) utterances: each utterance has a transcription area (called the Gloss window) as well as associated video and/or audio files. See Figure 1. The program allows a user to annotate video-based data with an intuitive interface and many user-definable options. The program currently includes the following functionality:

- Ability to associate a single transcription with multiple (up to 4) synchronized video files and a sound file
- User-editable fields and values and customizable display
- Built-in Search capability allowing for Boolean searches (AND, OR, NOT) as well as searches for temporal relations (BEFORE, AFTER, WITH)
- Wave form display of audio files
- Export of transcription information to a text file
- Script capability, allowing the playback of utterances in a user-defined order.

### **Program Development**

SignStream 3 is currently under development. This version of the program, implemented in Java, will include the ability to import and export information in XML format and will incorporate tools for detailed transcription of information about hand shape, location, orientation, and movement. While SignStream 2 allows only text annotations, the plans for SignStream 3 include the addition of graphical data types (e.g., graphics showing hand position and shape). It will also include additional functionalities, e.g., facilitating work with durations and other types of numerical data. SignStream 3 will make use of an efficient and intuitive graphical interface and templates to facilitate rapid data entry of fine-grained information.

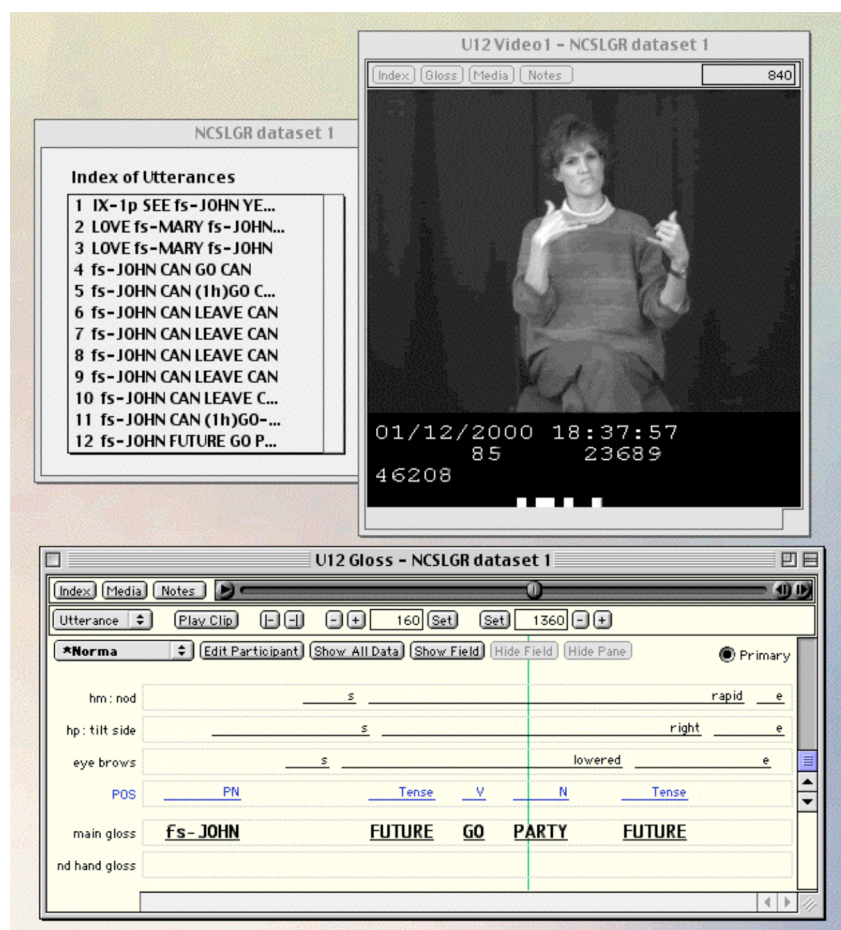


Figure 1: SignStream 2 illustration

### Availability of data

Historically, linguists have presented ASL data solely in the form of written representations, often with no access to the original video data. Having transcriptions separated from the source video makes evaluation of linguistic claims about the data very difficult. However, we make available video data (in a variety of formats) from native ASL signers along with the corresponding SignStream transcriptions and information about our coding schemes. This includes a substantial body of data collected in the NCSLGR (a state-of-the-art digital video data collection facility): multiple, synchronized video files showing the signing from different angles (2 stereoscopic front views, a side view, and a close-up of the face, along with calibration information for computer science researchers).

### Current Computational work

Collaborative work with researchers in the computational area (including Stan Sclaroff's Image and Video computing group at Boston University and Dimitris Metaxas's Computational Biomedicine Imaging and Modeling group at Rutgers) has focused on using data transcribed in SignStream to test computer vision algorithms developed by computer scientists. We are hoping that computer vision recognition algorithms will be able to aid in the (semi-)automation of the coding process for

researchers working in linguistics, and ultimately to progress in the realm of sign language (and gesture) recognition and generation.