



*A Database Tool for Research
on Visual-Gestural Language*

**Carol Neidle
Boston University**

*Report No.10
American Sign Language
Linguistic Research Project
<http://www.bu.edu/asllrp/>*

August 2000

SignStream™: A Database Tool for Research on Visual-Gestural Language

Carol Neidle, Boston University

1 Introduction

SignStream is a multimedia database tool designed to facilitate research on signed languages and gesture. A SignStream database consists of a collection of utterances, where each utterance is associated with one or more synchronized media clips (selected by the user) and a detailed, fine-grained, multi-level transcription. SignStream provides a single computing environment for manipulating media files and for transcribing and analyzing utterances. The program allows simultaneous access to digital

SignStream software, ©1997-2000 Trustees of Dartmouth College & Trustees of Boston University & Rutgers the State University of New Jersey, has been developed by Carol Neidle, Dawn MacLaughlin, Benjamin Bahan, Otmar Foelsche, Judy Kegl, and David Greenfield. The programming has been carried out at Dartmouth College by David Greenfield in the Department of Humanities Resources (Otmar Foelsche, Director). This work has been supported in part by grants #SBR-9410562, #IIS-9528985, #IIS-9912573, and #EIA-9809340 from the National Science Foundation to Boston University.

This material is adapted from documentation prepared in collaboration with Dawn MacLaughlin (including MacLaughlin, Neidle & Greenfield, 2000). Thanks to Debra Aarons and Erica Hruby for comments on an earlier draft of this document.

video and audio files along with representations of the data in linguistically useful formats (see **Figure 1**).

SignStream enables a user to annotate a segment of video by entering information into distinct fields. Manual signs and non-manual behaviors of various kinds (including, e.g., head nods, head shakes, eye gaze, raised or lowered eyebrows) may be represented in distinct fields. Within each field (e.g., *eyebrow position*), a value may be assigned (e.g., *raised*) that corresponds to a portion of the video; the start and end frame of the relevant behavior are identified by the user (and stored by the program). SignStream dynamically aligns information visually on screen so that relative position along the horizontal axis corresponds to relative position in time, thereby providing a comprehensive visual representation of temporal relations among linguistic events.

In addition, the program allows for the coding of multiple conversational participants. Search capabilities provide instant access to coded video data, enabling linguistic researchers to test hypotheses efficiently. A script facility makes it possible to select subsets of utterances to view in a given order.

SignStream runs on MacOS systems and makes use of QuickTime video technology. However, information from a SignStream database can be exported into a system-independent text format. SignStream is distributed on a non-profit basis to students, educators, and researchers. Further information about the program is available at <http://www.bu.edu/asllrp/SignStream>. See also the SignStream User's Guide (MacLaughlin, Neidle & Greenfield 2000) and Neidle *et al.* 1997.

2 Data entry and manipulation of annotated items

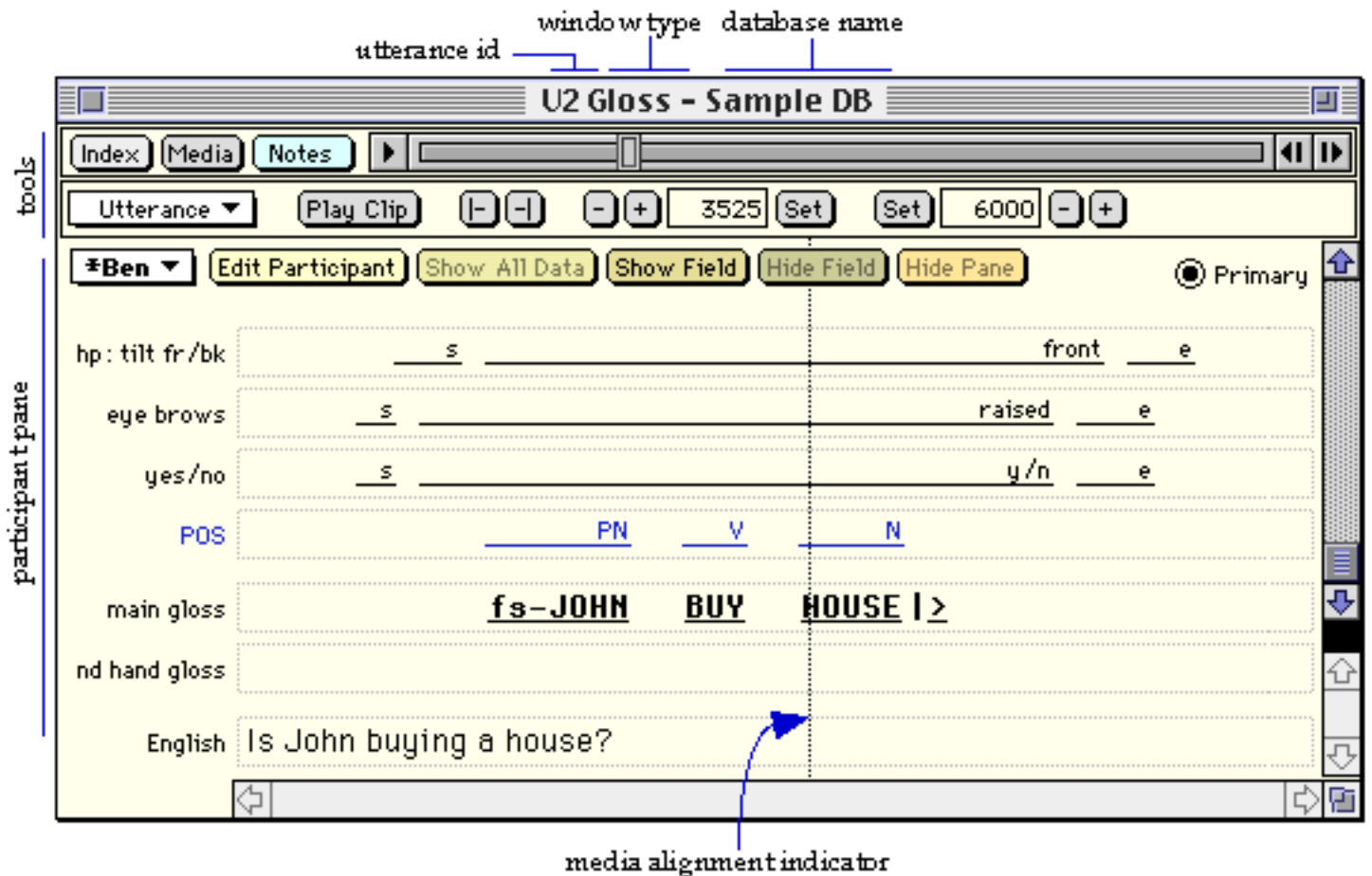
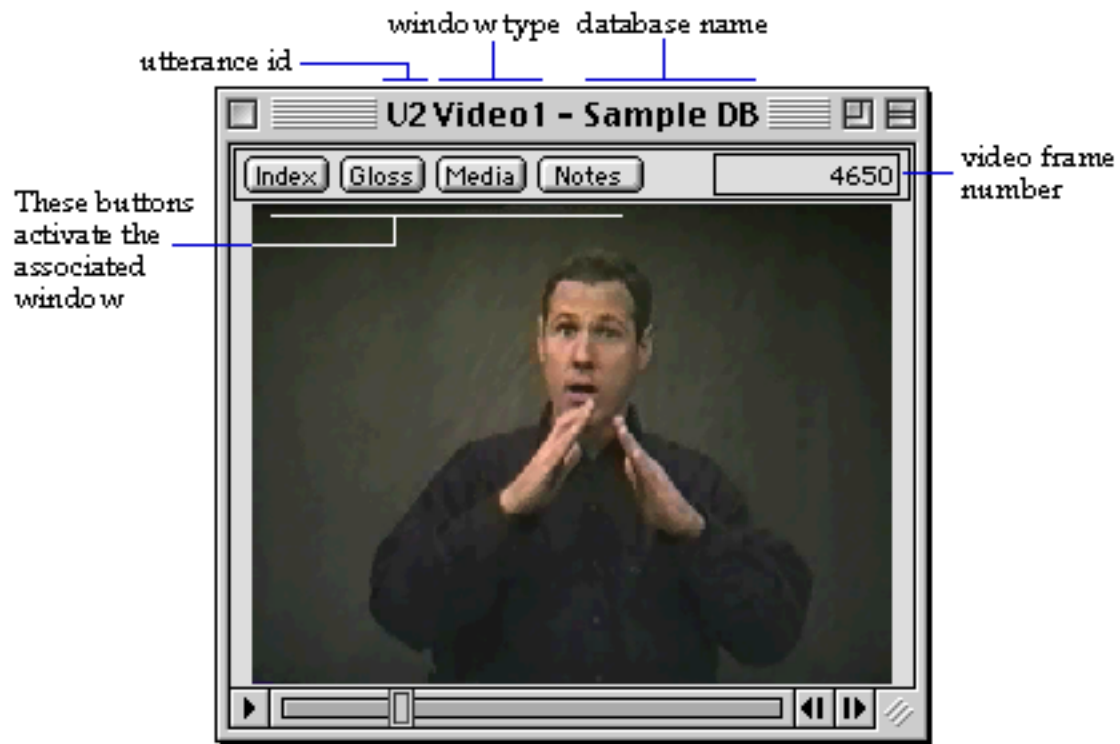
The user controls the video display and enters linguistic information in a variety of fields, linking this information to specific frame sequences of the

video by identifying the start and end point for each annotation. The user may enter data in one of several intuitive ways, depending on the type of information being coded. For example, non-manual items may be entered by drawing lines (where the video advances automatically as the user moves the mouse forward or backward to draw the line), selecting from menus, and typing text.

The program provides an extensive set of fields and field values. The user may also create new fields and values and edit those that have been defined. This flexibility allows users to code different types of data.

Items can be manipulated through the use of a variety of SignStream tools. The user can, for example, advance the video to the start or end point of any object, or the user can play the media clip associated with any object. As information in the transcription changes, the program automatically adjusts the data display accordingly, so that simultaneous events are vertically aligned. The result is a fine-grained, multi-level, temporally calibrated transcription, as illustrated in **Figure 1**. The media alignment indicator marks the current frame. It is possible to have windows from several utterances (potentially from different SignStream databases) open on screen at once, to facilitate comparisons.

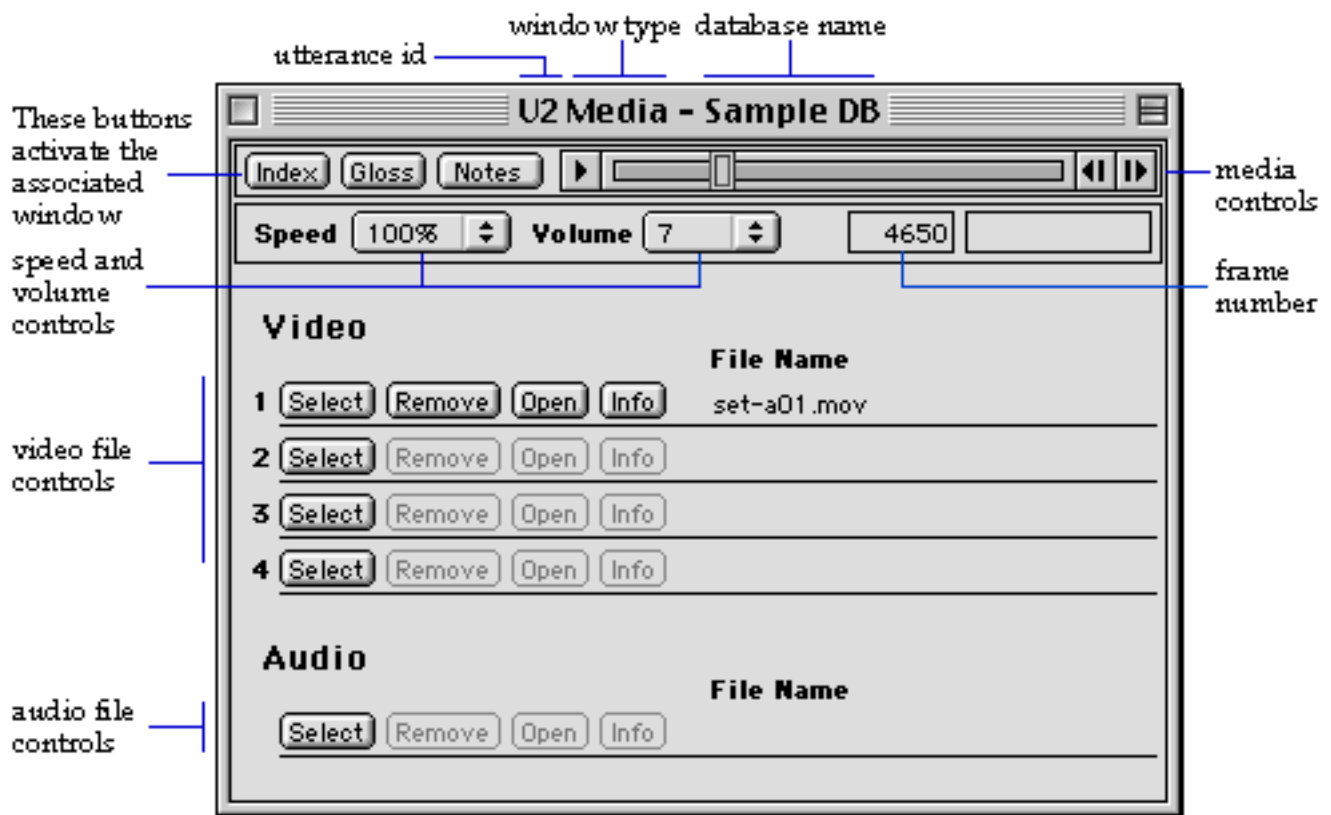
Figure 1. SignStream video and gloss windows



3 Integration of clips from movie files and sound files

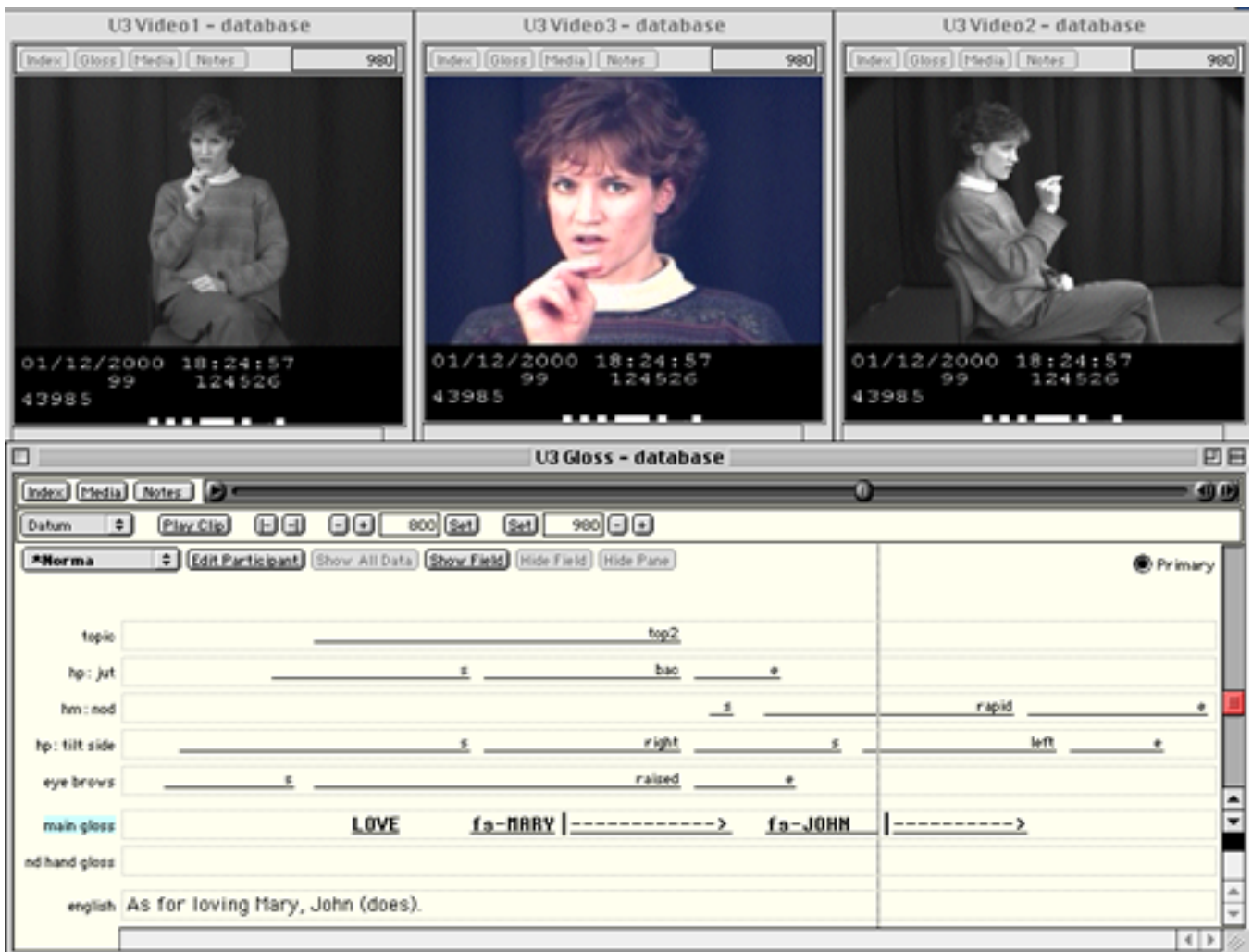
A media controller window allows for selection and control of media files (QuickTime video files and AIFF sound files), as shown in **Figure 2**. To select a movie file or sound file, the user simply clicks on the "Select" button in the media controller window and then locates the desired file (using the standard Macintosh interface for opening a file).

Figure 2. Media controller window



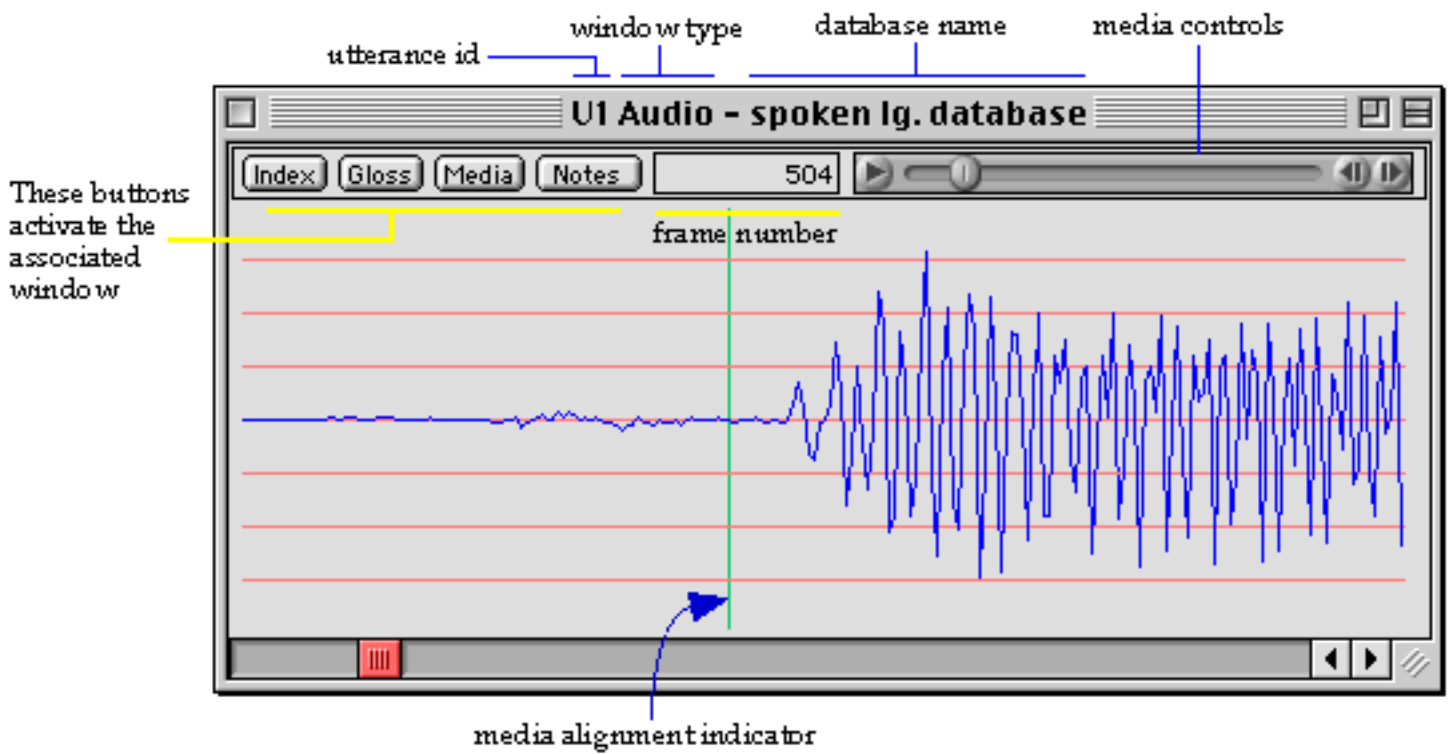
The media clips associated with a given utterance may include up to four synchronized QuickTime video files as well as one synchronized AIFF sound file. In **Figure 2**, a single movie file has been associated with the utterance in question. **Figure 3**, however, shows an utterance associated with three synchronized movie files, which present the signing from different angles; these data were collected in the National Center for Sign Language and Gesture Resources (see **section 10** for more information about that).

Figure 3. Display of multiple video windows



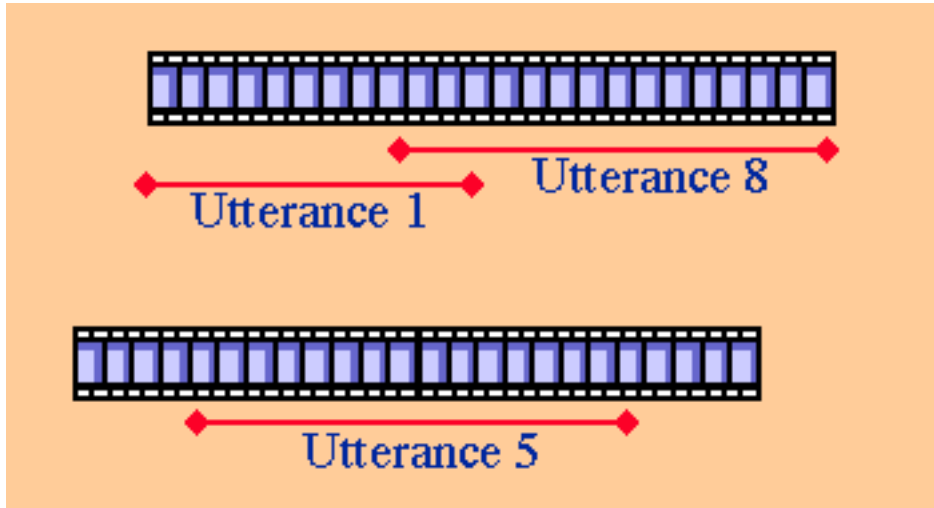
For sound files, a window displaying the sound wave is available, as shown in **Figure 4**. The media alignment indicator indicates the position of the current frame within the display. The user can play the media at varying speeds and make use of normal QuickTime controls.

Figure 4. The audio window



Once the user has selected the media clip(s) for an utterance, the database file will contain a link to those movie and sound files, which will be automatically accessed each time that utterance is opened. Different utterances within a single database need not be associated with the same movie files. Different movie files may be associated with different utterances, although it is also fine for different utterances to be linked to (potentially overlapping segments of) the same movie file, as illustrated schematically in **Figure 5**.

Figure 5. Association of movie clips with utterances



One consequence of this is that, although media files may be quite large, database files themselves are small in size. In any case, once the media files have been selected for a set of utterances, the links are transparent to the user; each utterance opens up with the appropriate media files displayed.

Although the video and transcription are displayed in separate windows, the windows are linked in various ways to facilitate working with the associated pieces of an utterance. For example, the video may be manipulated via controls in the gloss window. In addition, a vertical hairline indicator in the gloss window indicates the position within the transcription that corresponds to the current frame; this indicator is draggable, allowing the user to reposition the media (video and audio).

4 Display of multiple participant panes

As shown in **Figure 6**, it is possible to display separate panes for different conversational participants. All coded information will be vertically aligned in the display.

**Figure 6. Multiple participant panes
(video clip courtesy of DawnSignPress, used with permission)**



UI Gloss - DSP 3signers DB

Video Notes

Datum Play Clip (-) (-) (-) (+) Set Set (-) (+)

B (center) Edit Participant Show Field Hide Field Hide Pane Primary

hm: nod

main gloss **a: IX-2p-dual: c WANT DRINK |> IX-2p: c |----->**

nd hand gloss **"WELL" |-----**

tied xlate **Do either of you want a drink? How about you?**

A (left) Edit Participant Show Field Hide Field Hide Pane Primary

hm: nod **rapid**

main gloss

nd hand gloss

tied xlate

A Participants editor is available, enabling information about each participant to be stored. It is also possible to restrict searches (discussed below) to a particular participant.

5 Notes

A notes window is available and may be opened by clicking on the "notes" button. Each utterance is associated with a separate set of notes, and this notes field is searchable.

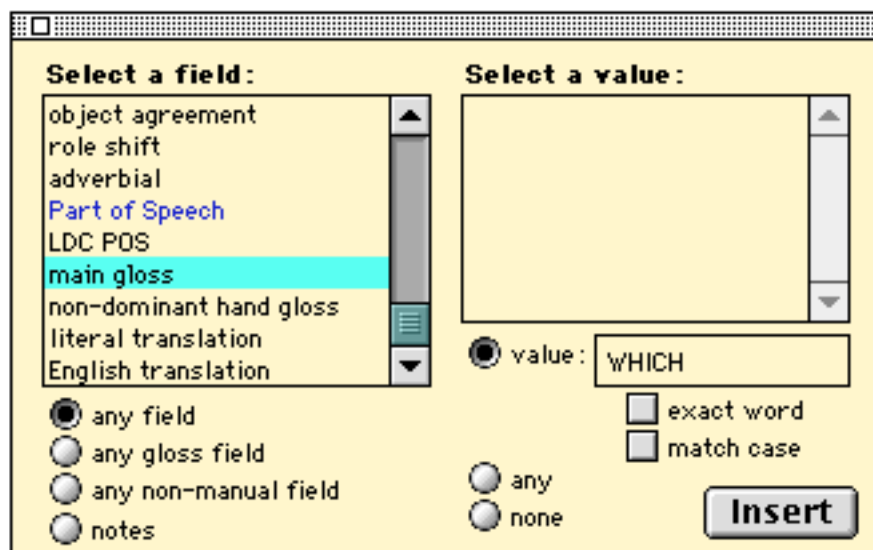
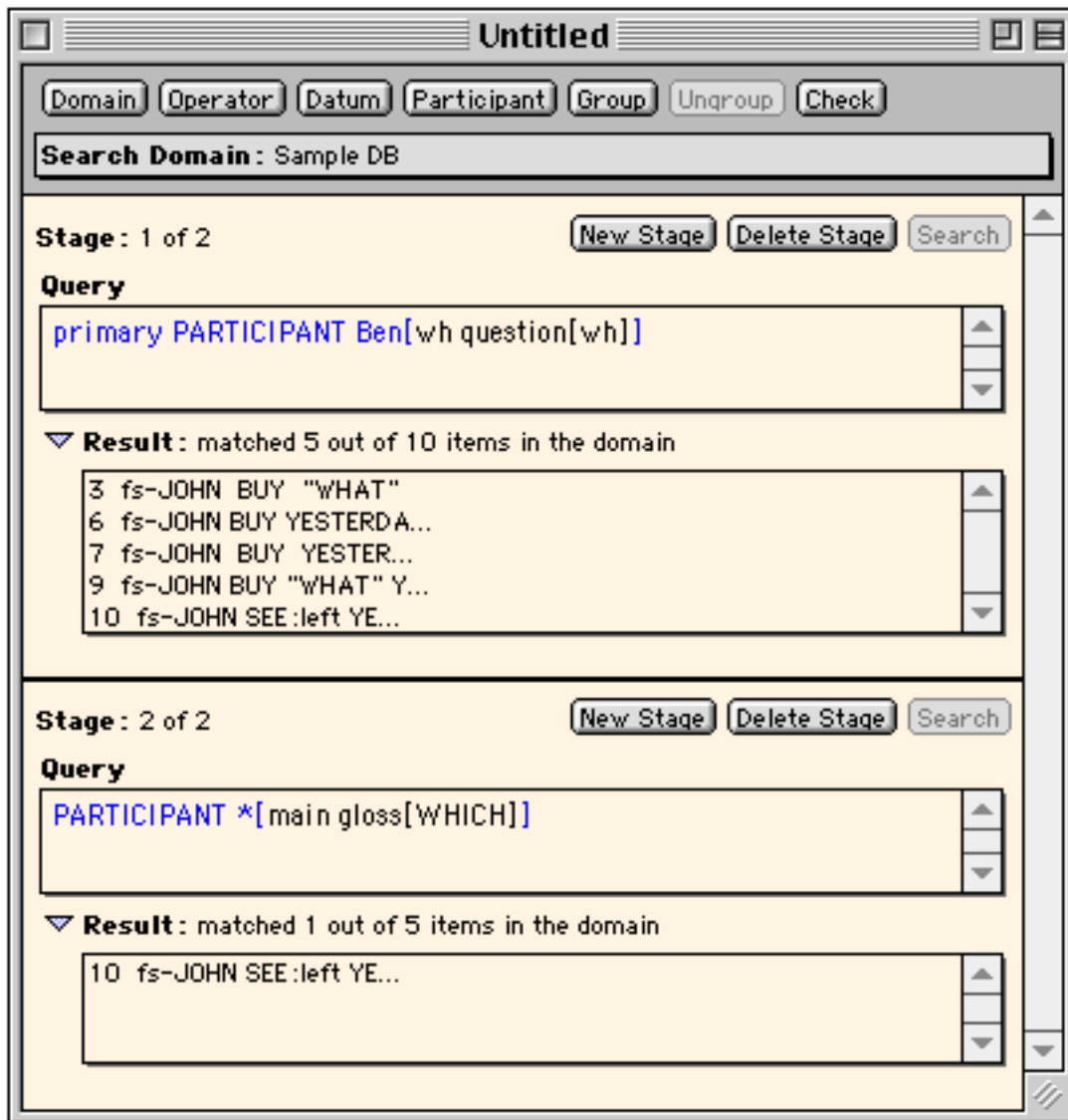
6 Searching

SignStream includes an integrated search facility, which allows the user to search over a collection of utterances in a database. A SignStream user can formulate complex queries, combining datum specifications using both standard Boolean operators (AND, OR, NOT) and specialized temporal operators (such as WITH, BEFORE, AFTER). For example, one can search for a pointing sign occurring simultaneously with eye gaze.

Searches can be conducted in multiple stages, allowing the user to successively refine the search. This is illustrated in **Figure 7**: the first stage of the search shown finds utterances with non-manual wh-question marking, and then the second stage restricts the search to those which contain the gloss WHICH.

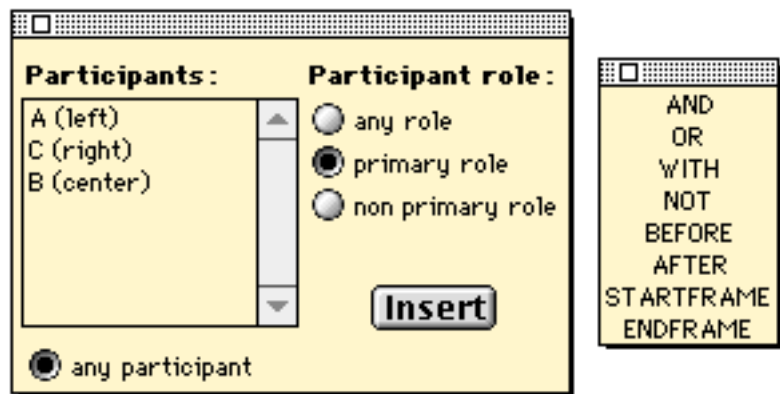
Various palettes are provided to facilitate the formulation of search queries. The Datum palette, allowing for selection of specific fields and values, is shown in the bottom of **Figure 7**. When a particular field is selected in this palette, all values for the field defined within the given database are displayed in the window to allow the user to make a selection of the value to be used in the query.

Figure 7. Multi-stage search (top) - facilitated by use of the datum palette (bottom)



Additional palettes are available to restrict the search to a given participant or to combine datum specifications with the operators just described (AND, OR, WITH, etc.), as shown in **Figure 8**.

Figure 8. Participant and operator palettes

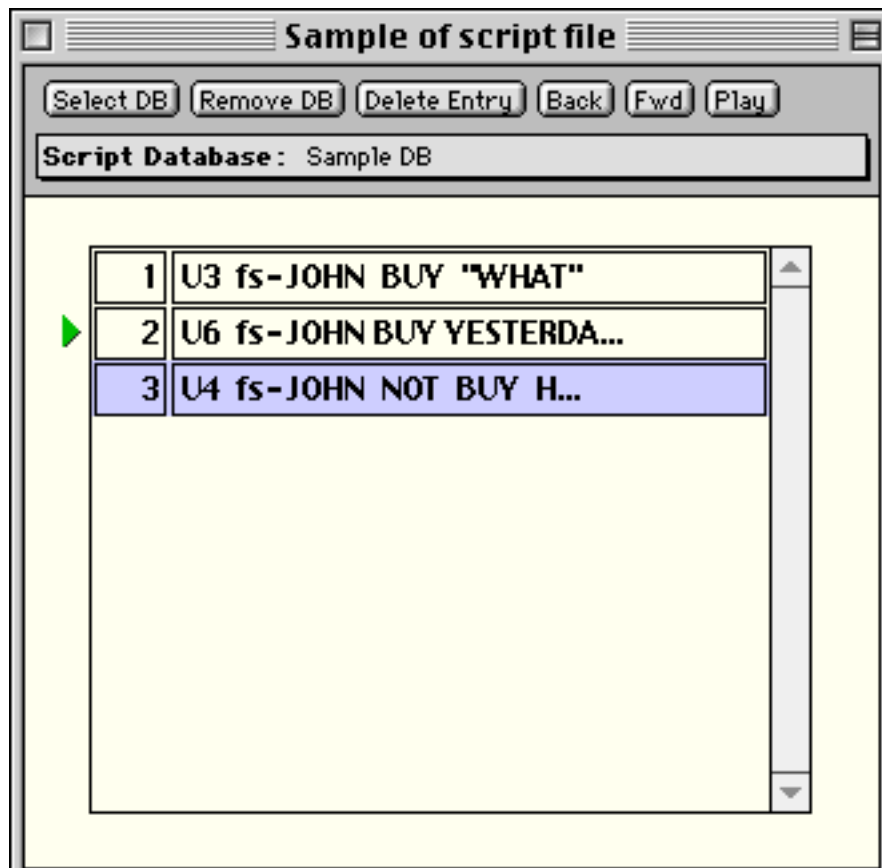


Conducting the search results in a list of utterances. The user may click on an utterance to open it. Searches may be saved and reopened at a later time. It is also possible to search through another database file without retyping the query. The user can simply select a different search domain and re-execute the search.

7 Script capability

New in version 2.0 is the capability to select (and save) a subset of utterances to be played in a given order. Utterances may be dragged into a script window, illustrated in **Figure 9**, where they may be reordered (if desired) and played. This capability is useful for selecting a subset of utterances for research or teaching purposes.

Figure 9. Script window



8 Advantages of SignStream

SignStream has several advantages over other existing computational tools geared toward the coding of video data. See <http://www.bu.edu/asllrp/SignStream/comp.html> for a comparison of SignStream with MediaTagger and SyncWriter, two Macintosh applications, and CHILDES. SignStream provides unique capabilities important for the coding of visual language data, including those listed here:

- SignStream provides an intuitive interface for viewing, manipulating and coding data.
- SignStream represents both start and end points of coded items.
- SignStream offers a large set of pre-defined fields and values; however, these are editable and new fields and values may be created by the user.
- SignStream provides a visual representation of temporal relations among items.
- SignStream allows direct access to video and audio segments corresponding to utterances or to specific items that have been coded within utterances.
- SignStream utterances may contain separate panes for distinct participants in a conversation; all information is time-aligned across participant panes.
- Distinct utterances (potentially from different SignStream databases) can be viewed on-screen at the same time.
- SignStream enables the user to incorporate multiple movies into a single database and allows potentially overlapping segments of video to be coded in distinct utterances.
- SignStream includes an integrated search capability.
- SignStream provides a script facility, making it possible to save and play subsets of utterances in a given order.

9 Export of SignStream data in text format

It is possible to export SignStream data as text, delineated by tabs and carriage returns. Plans for future development include implementation of import/export to XML format.

10 Availability of data

In conjunction with the SignStream project, an Internet data repository has been established to facilitate the distribution and sharing of SignStream-encoded data. It is hoped that users of SignStream will contribute data to this repository. As part of the American Sign Language Linguistic Research Project, we have been using SignStream as a tool for our linguistic research on American Sign Language (see, e.g., Neidle *et al.* 2000) and are distributing several large SignStream databases involving elicited data, dialogs, and stories in ASL. These are available on CD-ROM. Further information is available at

<http://www.bu.edu/asllrp/SignStream>.

Some of the data for which SignStream databases are available have been collected in our new digital data collection facility established as part of the National Center for Sign Language and Gesture Resources, a collaborative project with the University of Pennsylvania (supported by a grant from the National Science Foundation: #EIA-9809340). There are two facilities, one located at Boston University and another at the University of Pennsylvania, each containing four synchronized digital video cameras enabling capture of multiple views of the signing. Part of the goal of this project is to make public high quality data from ASL in a variety of video formats along with SignStream annotations. Additional SignStream databases associated with data collected in this center will be made public as they become available. See

<http://www.bu.edu/asllrp/csigr/> for further information.

11 Current status and directions for future development

SignStream is still under development. We welcome comments and suggestions; e-mail may be sent to carol@bu.edu.

Version 2.0 of SignStream, to be released in 2000, is being distributed to students, researchers, and educators on a non-profit basis. Updated versions will be distributed as they become available. Any database files created with this and subsequent versions of SignStream will be upward-compatible; future versions of the program will be capable of reading database files created by an earlier version. There is a listserv for SignStream users and others who may be interested; see our Web site for more information.

Future areas of development include:

- fields designed specifically for representing fine-grained phonological information
- a graphical interface for efficient entry of phonological information
- capability for importing and graphing of numerical data
- additional tools to facilitate the coding of audio
- XML import/export
- capability for merging of databases
- integration with other computational tools for research on sign language (such as dictionaries)
- printing capabilities
- extended search capabilities
- cross-platform access to SignStream databases

In addition, through collaborative work with computer scientists, we hope eventually to be able to (semi-)automate aspects of the transcription.

References

- MacLaughlin, D., C. Neidle & D. Greenfield. (2000). *SignStream™ User's Guide, Version 2.0*. American Sign Language Linguistic Research Project, Report Number 9, Boston University, Boston, MA, December 2000:
<http://www.bu.edu/asllrp/reports.html#RPT9>.
- Neidle, C., J. Kegl, D. MacLaughlin, B. Bahan & R.G. Lee. (2000). *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: The MIT Press.
- Neidle, C., D. MacLaughlin, B. Bahan, R.G. Lee & J. Kegl. (1997). *The SignStream™ Project*. American Sign Language Linguistic Research Project, Report Number 5, Boston University, Boston, MA, July 1997:
<http://www.bu.edu/asllrp/reports.html#RPT5>.