HStud 20 (2006) 1, 143–162
DOI: 10.15566/HStud.20.2006.1.12

THE MULTIDISCIPLINARY ANALYSIS OF TALK

BRIAN MACWHINNEY

Carnegie Mellon University, Pittsburgh, PA
USA

For this volume of *Hungarian Studies* dedicated to multidisciplinary contributions of Hungarians around the world, I have chosen to describe my work on a unique multidisciplinary effort called TalkBank. This effort seeks to harness the new information technology to study the great complexities of human talk.

Keywords: multidisciplinarity, statistical analysis, databases of written texts, sample analysis, classroom discourse

Researchers in the brain and behavioral sciences have benefited immensely from the rise of information technology. We now have powerful methods for performing statistical analysis, simulation, modeling, and graphic display. Our computers provide increasingly precise control of eye tracking, brain imaging, and stimulus presentation. We can use computers to analyze large databases of demographic and survey data, as well as huge corpora of written language texts. However, we have not yet succeeded in harnessing web technologies for the study of the most basic of all human social processes – conversational interactions. The TalkBank Project (http://talkbank.org) addresses this problem by providing computational support for the online multimedia collaborative analysis of talk.

To give the reader a sense of what types of materials are currently available in TalkBank, consider this sampler:

1. You can listen to Franciso Ma describe in Nahuatl (with English translations) how he lost his teeth when a bandit shot a gun at his mouth. (Thanks to Jane Hill.)
2. You can view and listen to a discussion of a group of medical students engaged in problem-based learning about diagnosis of a case of an amnesic, dysnomic aphasic. (Thanks to Tim Koschmann.)
3. You can listen to Larry Lessig plead the Napster copyright infringement case at the Supreme Court. (Thanks to Jerry Goldman's SCOTUS Project.)

4. You can watch Bettino Craxi discussing corruption in Italian politics in a TV interview. (Thanks to Johannes Wagner and the MOVIN Project.)
5. You can listen (in Spanish) to Christian Iniguez arguing with a sports talk show announcer about his predictions for the soccer match between Guadalajara and Monterrey. (Thanks to Christian Iniguez.)
6. You can listen to a discussion of card playing and shopping between three Jewish immigrants who resettled in 1938 from Vienna to London, as they code-switch back and forth from German to English. (Thanks to Eva Eppler.)
7. You can watch videos from Alicia between ages 1 and 3 as she interacts with her English-speaking father and other Cantonese speakers in Hong Kong. (Thanks to Virginia Yip and Stephen Matthews).
8. You can watch an elderly gentleman with severe aphasia holding fully effective gesturally-based conversations with his family and friends. (Thanks to Chuck Goodwin).
9. You can watch German schoolchildren from Dresden singing and dancing to Czech songs in their L2 Czech classroom. (Thanks to Angelika Kubanek-German.)
10. You can study transcripts and audio from the CHILDES database from about 1200 children learning 28 different languages.

I cite these ten examples simply to illustrate the enormous diversity in the CHILDES and TalkBank databases. Alongside this qualitative diversity is the quantitative richness of available data, now reaching 300 MB of text and an additional 2 terabytes of sound and video.

Viewing TalkBank along disciplinary lines, we can distinguish 17 fields or research circles that are involved in the study of conversational interactions. For each of these research circles there are corresponding data in TalkBank:

1. child development (mother-child attachment, peer groups)
2. child language development (CHILDES system, test development)
3. language disorders and remediation (aphasia, stuttering, retardation)
4. multilingualism (second language learning, code-switching, acculturation)
5. emergency medicine (ambulance, ER, teaching simulations)
6. legal argumentation (courtroom, Supreme Court, probation hearings)
7. small group dynamics (government meetings, security)
8. psychotherapy (therapist-patient discourse, therapy groups)
9. conversation analysis (rhetorical analysis, sociolinguistics, text and discourse)
10. computational linguistics (parsing, tagging, data-mining, content analysis)
11. speech technology (voice recognition, prosodic analysis)

12. anthropology (field linguistics, film ethnographies, oral tradition)
13. gestural communication (narrative, conversational, crosslinguistic)
14. classroom discourse (science, math, literacy, cultural effects, lectures)
15. tutorial dialog (dyadic, small group, man-machine)
16. human-computer interaction (collaborative dialog, usability)
17. ethology (animal communication and behavior)

The development of a shared database has been a crucial formative step in the maturation of the each of the sciences. In Genetics, projects such as the Human Genome Project (www.ornl.gov/hgmis), GenMapp (www.genmapp.org), or Protein Map (Aisenman – Berman, 2000) are now storing all published genetic sequences in forms that are open to analysis and data-mining through the web. In fact, gene sequences are not accepted for publication until they have been entered in these systems. In Paleontology, museums worldwide preserve fossils whose specific physical structure, radiological dating, and stratificational location are crucial to our reconstruction of the history of life and the earth. Electronic records and scans based on this evidence are now being made available electronically (www.ucmp.berkeley.edu/pdn/) for deeper analysis and data-mining. Internet databases are now fundamental to progress in Astronomy (van Buren, Curtis, Nichols, Brundage, 1995), Physics (Caspar et al., 1998), Economics, Medicine, History, Political Science, Experimental Psychology, Linguistics, and other sciences.

1. TalkBank Data Research Methods

Psychologists often rely on laboratory studies using reaction time methodology, computerized control (Cohen, MacWhinney, Flat, Provost, 1993; MacWhinney, St. James, Schunn, Li, Schneider, 2001), and random assignment to condition to test empirical hypotheses. Although this particular methodology cannot be applied to the study of the real human interactions represented in TalkBank data, there are at least seven equally powerful methodologies that can operate directly on TalkBank data. These include:

1. **Microanalytic studies.** Microanalysis of videos relies on frame-by-frame analysis of linkages of conversation, gesture, proxemics, props, and prosodics. Exemplary applications of this method already in TalkBank include studies of problem-based learning (Koschmann – LeBaron, 2002), Guugu Yimithirr narratives (Haviland, 1993), and professional problem-solving (Goodwin, 1994). These analyses have already served as our first test-bed for published collaborative commentary in journals with CD-ROMs.

2. **Microgenetic studies.** Research programs that study cognitive development (Siegler – Crowley, 1991) use careful video analysis to track subtle changes in learner's strategies across days and weeks as a result of various types of teaching.
3. **Sampled comparisons.** Within the new infrastructure, situations can be sampled across groups and conditions and compared in terms of analytic codings. For example, we could compare the gestures of deaf children of hearing parents with those of normal children in terms of a coding system delineating reference to the here and now (Morford – Goldin-Meadow, 1997). The coding system itself would then be the focus of collaborative commentary.
4. **Error analysis.** We can, for example, distinguish cases of failed perspective-taking in phone call survey data (Schober, Conrad, Fricker, in press) both by sampling across conditions and by microanalysis within conditions.
5. **Longitudinal studies.** In areas where controlled experimentation is not possible, longitudinal analysis is often equally powerful. For example, we can trace the dynamic emergence of the Supreme Court's position in Roe v. Wade and its implementation in 25 years of subsequent decisions. Or we can trace the process of mathematical development across the 12 years of Carolyn Maher's video study of a cohort of learners in Baltimore.
6. **Large sample analysis.** We can assess the effects of oral arguments in the Court's decision-making, by examining outcomes and processes in a large number of oral arguments. For example, attitudinalists (Segal – Spaeth, 1993) view oral arguments as largely irrelevant to the Court's decisions making. Informationists (Johnson, in press) espouse a contrasting view that emphasizes the extent to which justices acquire new information during oral arguments. To adjudicate this issue, we can provide representatives of these competing views with complete access to transcripts and audio of the Court's oral arguments along with simple methods for annotating and coding the transcripts from their respective theoretical positions.
7. **Dynamic modeling.** We can track individual differences in referential compression in dyadic interactions with computerized systems. Because the behaviors being produced in these system are frequent and repetitive, they are amenable to modeling (Anderson – Lebiere, 1998).

Each of these seven methods fits in naturally with TalkBank and our vision of a new community of collaborative commentary. By combining several of these methods, we can begin to understand how processes that operate across very different time scales can become entrained by interactions that occur in observable interactions (MacWhinney, in press-c).

2. Infrastructural Development

The development of the TalkBank database and programs has been facilitated by a series of relatively recent developments in computer and network hardware and software, as well as computerized recording technology. These advances include:

1. **Computational speed.** Current desktop computers run at speeds that make it easy to examine hundreds of megabytes of transcript data in a few minutes. These speed advances make it possible to search quickly for a wide variety of lexical and syntactic patterns in aphasic speech. Today's computers can also be used to compress video material fast enough to allow users to run compression jobs overnight that earlier would have taken weeks. In this project, video compression will be conducted at CMU, where we are already relying heavily on these advances.
2. **Network bandwidth.** Most users now have access to broadband Internet connections that are capable of playing high-quality compressed video without distortion or dropped frames.
3. **Streaming video.** Compressed video can now be configured to permit random access or hinted streaming. Hinting adds time marks to delineate the beginning and ends of small segments of the video. These marks then allow a user to directly access, for example, a clip that begins at minute 18 of the 42 minutes video without having to wait to download the first 18 minutes.
4. **Disk storage.** It is now possible to purchase a terabyte of disk storage for $600. Three or four years ago, this would have cost $10,000. Because these prices have fallen so rapidly, it is easy for the TalkBank Project to store and backup large amounts of high-quality video for distribution through streaming servers.
5. **XML.** Currently, most documents on the web are in HTML. However, over the next few years, materials on the web will move from HTML to the more powerful XML framework. Modern computer software such as Java and C# provides solid support for documents encoded in XML. Moreover XML has strong linkages to the new Unicode standard.
6. **Unicode.** The Unicode character-encoding standard provides a single consistent standard for encoding all of the world's languages, as well as the major classes of symbol systems. Because this encoding is consistent across computer platforms, it allows us to create a single consistent database for aphasia across languages and platforms.
7. **Grid Computing.** Working with Bennett Bertenthal's new NSF Social Informatics Data Grid project, TalkBank can configure data analysis and data sets across a wide array of machines on the Internet that can be accessed by the tools of grid computing. Grid Computing is one component in a set of ongoing ad-

vances in Cyberinfrastructure that can be tracked by consulting the proceedings of the recent NSF Workshop on Cyberinfrastructure for the Social Sciences at http://vis.sdsc.edu/sbe/Sessions.

8. **Recording Technology.** The widespread availability of 3CCD digital video recorders using relatively inexpensive mini-DV cassettes makes the construction of video databases economically feasible. The recent advent of high-quality portable digital recorders also facilitates the collection of excellent digitized audio.
9. **Compression Software.** Particularly on the Macintosh platform, there are excellent systems for quick software video compression. QuickTime 7 and MPEG-7 offer still further advances.

3. TalkBank Tools

TalkBank and the CHILDES Project that predated it have produced six major pieces of software. They are: the CHAT editor, CLAN analysis, the XML convertor, phonological analysis in *Phon*, the TalkBank Server, and the TalkBank Browser.

The CHAT Editor. The CHAT editor is a full text editor written in C++ that runs on Macintosh and Windows platforms. It provides users with four methods for linking transcripts to audio or video media.

1. The *Waveform* method allows users to drag over a segment of the waveform display corresponding to an utterance and then transcribe that utterance in the text window.
2. The *Sound Walker* method simulates the old foot pedal method of transcription that continually replays the current sound and then advances a specified amount.
3. The *Transcriber* method allows the user to play media and hit a space bar whenever an utterance ends. This creates a transcription full of bulleted segments to be transcribed. Then the transcriber can go back and insert transcriptions for the bulleted segments.
4. The *Post Hoc* method allows the transcriber to first create an unlinked typed transcript and then to step through that transcript utterance by utterance linking the transcript to the audio or video media.

Each of these methods allows the user to adjust the borders of the sound segment, replay segments, and relink. The final product resulting from all of these is a transcript with markers that can be used to replay each utterance directly. Using the CHECK program, it is possible at each moment to check the extent to which the transcript is making legal use of CHAT codes. The editor also provides the standard facilities for character search and replacement, automatic line numbering, and hiding or showing coding tiers. The CHAT editor also provides facilities for transcribing in Conversation Analysis (CA) format. CA researchers have found that facilities in CLAN for linkage to audio and video from the transcript improves not only their ability to study details of the interaction, but also their ability to present analyses in lectures and over the web. Between 2000 and 2003, CLAN provided a separate mode of analyses called CA mode that allowed CA transcribers to use standard CA characters in a special font.

CLAN analysis. Once files have been transcribed in CHAT, users can run a wide variety of CLAN analysis programs. The CHAT editor and the CLAN programs are the creations of Leonid Spektor, who has worked on the CHILDES and TalkBank Projects for 21 years. There are 28 CLAN programs, each with a wide variety of subfunctions and options. String-search programs can compute frequency counts, key-word and line profiles, mean length of utterance, mean length of turn, type-token ratios, maximum word length counts, maximum utterance length histograms, VOCD, and so on. CLAN has a subprogram called MOR that applies part-of-speech taggers for English, Spanish, German, French, Italian, Japanese, Cantonese, and Mandarin. The results of these taggers are then disambiguated using the POST statistical disambiguator (Parisse – Le Normand, 2000). These morphological codes can then be used to automatically compute indices such as DSS, IP-Syn (Sagae, MacWhinney, and Lavie, 2004), and a simple version of LARSP.

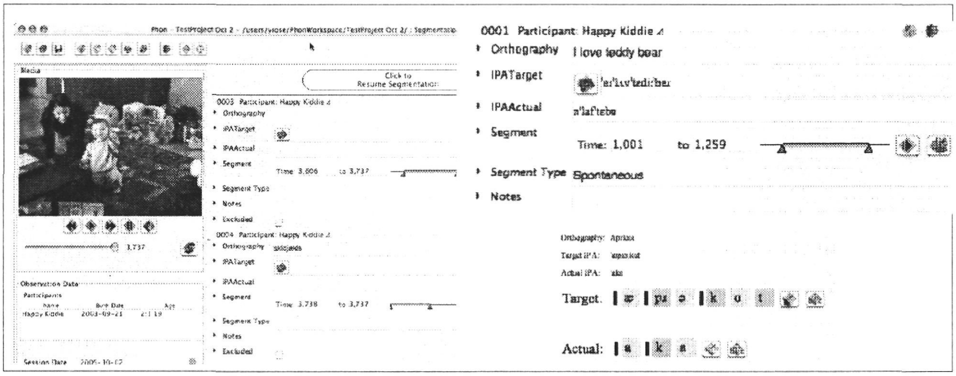
The XML Convertor. Franklin Chen has constructed Java-based tools that convert CHAT files to XML. These files can then be reformatted back to CHAT and the initial and final versions compared to guarantee the accuracy of the roundtrip. Only when the roundtrip runs without differences can we accept the data into TalkBank. The process of converting the database to XML was completed in 2004, after nearly three years of work. An important outcome of this conversion has been the full systematization of the coding system and an increase in consistency in the database. In addition, we were able to convert a wide range of discrepant font and character encoding systems to a consistent Unicode format. This was particularly important for Asian languages that use non-Roman characters, but it was also useful for special Roman characters with diacritics in languages such as French, German, and Spanish.

TalkBank has made extensive use of the XML format as a method for translating between alternative transcription systems, including SALT, HamNoSys, Elan, MediaTagger, SBCSAE, HIAT, ISL, LDC, TRS, and so on. The actual

XML used in this work is too verbose to be useful to users. However, several programmers outside the TalkBank project have used the TalkBank XML Schema (http://talkbank.org/talkbank.xsd) as a simple, well documented, platform for writing their own Java programs to process TalkBank data. In this way, the TalkBank project has maximized the availability of the TalkBank and CHILDES databases on both the file level and the program development level.

Phonological Analysis in *Phon*. The study of phonological processes in aphasia is important for both theoretical and practical reasons. On a theoretical level, studies of phonological processes in aphasia can illuminate theories such as markedness , optimality theory (Kager, Pater, Zonneveld, 2004), and dynamic system approaches (Lindblom, 2000). On a practical level, improvements in the control of articulatory processes can facilitate smoother communication. Earlier versions of the CLAN software for child language analysis provided only marginal support for phonological analysis. The only available program for general phonological analysis was the LIPP program developed by Kim Oller in the 1980s. Unfortunately, that program had not been updated in nearly 20 years and was unable to fit in well with modern systems of phonological analysis. Recently, the construction of the *Phon* program (Rose et al., 2005) by Yvan Rose at Memorial University Newfoundland in collaboration with the TalkBank Project has begun to fill this major gap. *Phon* works directly with CHAT files and allows users to segment and analyze children's productions on the level of the syllable and prosodic unit. Once high-level segmentation has been done, an automatic algorithm conducts syllabification. By making reference to a dictionary of standard and variant pronunciations, the model provides an automatic model-replica (Ferguson, Peizer, Weeks, 1973) alignment of the child's production to the adult target. This can then be used as the basis for analyses of phonological processes (Stampe, 1973), syllabic structures (Vihman, DePaolis, Davis, 1998), and constraint application (Goad – Rose, 2003, 2004). The results of the automatic syllabification and model-replica alignment can be checked and modified by the researcher in a variety of ways. Phon includes facilities for conducting a variety of prepackaged and custom analyses on large data sets. Finally, segments analyzed in Phon can easily be sent to Praat for further detailed phonetic analysis. Here are screenshots from the *Phon* Media Alignment window (left), Transcription window (top right), and Automatic Segmentation window (bottom right).

The TalkBank Server. Recently, TalkBank has configured a software/hardware package at that allows institutions outside of CMU to deploy their own full TalkBank sites. The code for this server can be located at http://www.talkbank.org/tbviewer/local/. This system is particularly useful for projects with tight privacy restrictions or specific local requirements, since it allows them full access to



TalkBank tools without having to contribute their data to an international database. In November 2005, we installed a full TalkBank server configuration at the Medical School of the University of Southern Denmark.

The TalkBank Browser. The most significant recent development in the TalkBank system has been the construction of the TalkBank Browser. Users can download and install this browser using the Java WebStart facility and the Java that is now built in to Windows, Mac, and Linux. Using the TalkBank Browser, users can directly access TalkBank transcripts and play them back interactively over the web. The program is written in Java, with components running in C#. The standard Internet Explorer (IE) browser is embedded within Java on Windows, using an API from Sun. On Macintosh, the embedded browser is Firefox. On both platforms, C# is used to control QuickTime streaming playback. A simpler form of playback can be achieved through the WebData facility built into CLAN. For instructions on the use of WebData and the TalkBank Browser, reviewers can consult http://talkbank.org/aphasia.

The TalkBank Browser is now being elaborated to permit collaborative commentary (MacWhinney, in press-a; MacWhinney et al., 2004). This process allows users to view a segment of an interaction with an aphasic and insert comments or blogs in the dialog. These comments can then be stored on the TalkBank server and subjected to further peer commentary.

4. Research Circles

Our discussion of the software advances underlying TalkBank has temporarily deflected focus from consideration of the intellectual core of TalkBank. This core is represented by the notion that, although human communication is a unified fact, it is analyzed through markedly separate techniques in at least 17 disciplinary re-

search circles. By creating a single, shared database on human communication, we can begin to encourage communication across these disciplines. At the same time, much of the initial dialog that has occurred in the TalkBank framework has been disciplinary. This is because researchers tend to identify with particular research communities that understand their goals and terminology. We can think of these groups as research circles. Although TalkBank research circles are now beginning to use common tools and frameworks, they still continue to focus on very different types of communicative interactions and different subject populations. Because of this, it is important for TalkBank to realize that true interdisciplinary work will only emerge from the increase in communication between parallel research circles that have each reached a high level of technical and theoretical sophistication. In this section, I review progress in the development of seven such circles in the TalkBank framework: classroom discourse, medical education, aphasia, CA, second language learning, legal discourse, child development.

Classroom Discourse

Much current research in science education relies on the distillation of longitudinal video, still-image, and observational data to create rich models of learning-in-context, with specific attention to interactions among tasks, discourse, and systems of representation in classroom settings (Greeno, 1998; Sfard – McClain, 2002). Video work has impacted the study of teacher activities (J. Frederiksen, Sipusic, Sherin E., 1998), international comparative studies of videos of mathematics classrooms (Stigler, Gallimore, Hiebert, 2000), learning of demanding topics in high school physics (Roth – Roychoudhury, 1993), engineering education (Linde, Roschelle, Stevens, 1994), informal learning in science museums (Crowley, Callanan, Tenenbaum, Allen, 2001), interacting with machines (Nardi, 1996), and the role of gestural communication in teaching and learning (Roth, 2001). The pervasive impact of video studies was in evidence at the 2002 American Educational Research Association meetings, which included 44 scientific panels and symposia using video for learning research,teaching, and teacher education.

Video is also used in teacher training programs (Derry, in press; Pea, 1999) and materials illustrating proposed nationwide educational standards (Daro, Hampton, Reznick, 2004). The field also enjoys a great range of high-quality tools for the analysis of video interactions. Systems such as NVivo (www.qrsinternational.com), DIVER, TransAna (www.transana.org), ATLAS.ti (www.atlasti.com), Elan (www.mpi.nl/tools/elan.html), MacShapa (Sanderson – Fisher, 1994), CLAN (childes.psy.cmu.edu)), VideoNoter/C-Video (Roschelle, Pea, Trigg, 1990), Ethnograph (www.qualisresearch.com), Anvil (www.dfki.de/~kipp)), Orion (Baecker, Fono, Wolf, 2006; Goldman-Segall – Reicken, 1989), ePresence (Baecker, Fono, Wolf 2006), Informedia (Wactlar, Christel, Gong, Hauptmann, 1999), and VideoPaper (Beardsley, Cogan-Drew, Olivero, 2006) are allowing researchers to produce large quantities of well-analyzed video interactions.

Despite the high quality of video analysis methodology, the large quantity of data being produced, and the centrality of video to the scientific study of learning and instruction, there has not yet been a community-wide acceptance of the importance of a shared database of instructional interactions. There has been extensive discussion of the formation of collaboratories for the study of instructional interactions (Baecker, Fono, Wolf, 2006; Edelson, Pea, Gomez, 1996). However, without a general method for sharing data across projects, collaboratories are limited to datasets collected from single projects (Abowd, Harvel, Brotherton, 2000). However, many of the most interesting questions in learning and instruction involve comparison between alternative teaching frameworks and situations. This type of diversity in the database can best be achieved by having data from many different laboratories and groups channeled into a uniform, but distributed database.

To address this need, the TalkBank Project has begun an effort to construct a shared database for the Learning Sciences. TalkBank (http://talkbank.org) is an international collaborative effort that has been building a web-accessible database for spoken language interactions. All of the video and audio media in TalkBank are fully transcribed and each transcribed utterance is linked directly to the corresponding segment of the media. The media and transcripts can be downloaded from the web. Users can also open a browser window, scroll through transcripts, play back the corresponding audio or video, and insert commentary regarding their analyses. The current TalkBank database has large collections of data in the areas of child language (CHILDES), aphasia (AphasiaBank), second language learning (SLABank), bilingualism (LIDES), formal meetings, and spontaneous conversational interactions (CABank and MOVIN).

A shared database for the Learning Sciences will have some interesting features unique to this area. It will be important to develop a taxonomy of educationally-relevant activities, events, and interaction types that can serve as metadata for coding and retrieval. It will also be important to supplement video records with additional ethnographic materials such as diaries, notebooks, drawings, and class records. However, the most powerful feature of a shared database in the Learning Sciences will certainly be its availability to collaborative commentary. The idea of scientific collaboratories has been developed and discussed elsewhere in this volume. With the context of collaboratories, projects such as Orion (Goldman, 2006), DIVER (Pea, in press), and WebCast (Baecker, Fono, Wolf, 2006) have shown how a group of educational researchers can work together to analyze interactions and evaluate competing interpretations. However, for the process of collaborative commentary to work as a general model for the learning sciences, it must be linked to a commitment to the process of data-sharing. What is unique about the TalkBank Project is not its emphasis on collaborative commentary, but rather its emphasis on data sharing. However, the greatest value for scientific progress arises when data sharing is joined with collaborative commentary.

Medical Education

Video analysis has also played a major role in the study of case-based instruction (Lampert – Loewenberg-Ball, 1998; Lesh – Lehrer, 2000) in medical education (Koschmann, 1999). Tim Koschmann at SIU has created a database of analyses of "standardized patients" (SPs) by medical students, residents, and medical faculty. The standardized patient cases are widely used as one method for evaluating medical competency. For example, one such case involve a 35-year-old woman complaining of headaches and fatigue. An initial working hypothesis diagnosis is often for migraine headaches with iron-deficiency anemia. However, blood tests later pointed to a diagnosis of *polycythemia vera*. Another case initially seems to involve psychiatric symptoms, but these are later seen to arise from a primary organic lesion. Koschmann has proposed the establishment of a collaborative commentary circle of researchers interested in analyzing these SP video protocols, along with the clinical notes, from a set of complementary frameworks. These frameworks include problem-solving theory (Koschmann – LeBaron, 2002), cognitive discourse analysis (C. Frederiksen, 1999), latent semantic analysis (LSA, Art Graesser), and memory-based reasoning (Seifert – Patalano, 2001).

We will extend this group to include additional viewpoints, with a particular emphasis on medical educators.

AphasiaBank

In May 2005, we organized a three-day meeting of 20 researchers at CMU with the purpose of providing recommendations regarding the construction of AphasiaBank. The group decided to focus initially on aphasia as a core with secondary attention to related disorders. The group also decided that we should formulate a standard protocol for further data collection. This proposal fleshes out these two basic proposals. In the context of preparation for this meeting, each participant contributed an audio or video record, linked to a CHAT transcript. The preparation of the CHAT transcript and the linkage was done in collaboration with workers at CMU. The resultant database can be accessed through the instructions found at http://talkbank.org/aphasia. The current contents of the database are:

1. Four audio recordings of interactions in patients' homes from Beth Armstrong.
2. The crosslinguistic aphasia project data on English, Italian, Chinese, and German from the 1990s, directed by Elizabeth Bates.
3. Chuck Goodwin's video recording of a series of aphasic communications that rely heavily on gestural and deictic communication.
4. Transcripts without audio from 46 patients contributed by Audrey Holland in the 1990s.
5. Video samples of three aphasics contributed by Audrey Holland.
6. Cookie theft descriptions from patients with dementia contributed by Dan Kempler.
7. Filmed interactions in three situations with a single patient contributed by Nina Simmons-Mackie.
8. Filmed interactions with three patients contributed by Lise Menn.
9. Seven group discussions involving aphasic patients contributed by Mary Oelschlager.
10. Picture descriptions, story retells, and personal narratives linked to audio contributed by Gloria Olness.
11. Video interactions in English and Afrikaans contributed by Claire Penn.
12. Picture descriptions and story retellings linked to audio contributed by Barbara Shadden.
13. Cinderella story retellings contributed by Cindy Thompson.

14. Video of problem solving in participants with TBI contributed by Leanne Togher.
15. Video of discussions of participants with TBI contributed by Lyn Turkstra.
16. Classic video of a patient with jargon aphasia contributed by Hanna Ulatowska.

Conversation Analysis

Conversation Analysis (CA) is a methodological and intellectual tradition stimulated by the ethnographic work of Garfinkel (1967) and systematized by Sacks, Schegloff, and Jefferson (1974) and others. Recently, workers in this field and the related field of text and discourse have begun to publish fragments of their transcripts over the Internet. However, this effort has not yet benefited from the alignment, networking, and database technology to be used in TalkBank. The CHILDES Project has begun the process of integrating with this community. Working with Johannes Wagner (http://www.conversation-analysis.net), Brian MacWhinney has developed support for CA transcription within CHILDES. Wagner plans to use this tool as the basis for a growing database of CA interactions studied by researchers in Northern Europe.

Researchers studying gestures have developed sophisticated schemes for coding the relations between language and gesture. For example, David McNeill and his students have shown how gesture and language can provide non-overlapping views of thought and learning processes. A number of laboratories have large databases of video recording of gestures and the introduction of data sharing could lead to major advances in this field. There are also several major groups studying the acquisition of signed languages. One group uses the CHAT-based Berkeley System of Transcription. Other researchers use either the SignStream system developed by Carol Neidle or the Media Tagger system developed by Sotaru Kita. Other groups use adaptations of CHAT and SALT. Because each of these groups is heavily committed to its own current approach, it may be difficult to find a common method for data sharing. However, by relying on XML as an interlingua, it should be possible to store data from all of these formats in a way that will permit movement back and forth between systems. However, the details of this will need to be worked out in a meeting with the various groups involved.

Second Language Learning and Bilingualism

Annotated video plays two important roles in the field of second language learning. On the one hand, naturalistic studies of second language learners can help us understand the learning process. The second use of video in second language learning is for the support of instructional technology. By watching authentic interactions between native speakers, learners can develop skills on the lexical, phonological, grammatical, and interactional levels simultaneously. TalkBank has created a process of data sharing that will address both of these problems. The database now has major corpora from learners of French, Czech, German, English, Japanese, and Spanish. In addition to these new corpora from older second language learners, there are several extensive new video studies of bilingual development in young children. Finally, there are six corpora documenting dual language interaction and code-switching in adult bilinguals.

Legal Discourse

The SCOTUS (Supreme Court of the United States) project, directed by Jerry Goldman at Northwestern University, is currently engaged in digitizing all of the oral arguments at the Supreme Court from 1955 to the present. The CMU component of this work focuses on scanning and reformatting the transcripts into the CHAT format and linking the transcripts to the audio on the sentence level. From this point, Mark Liberman and John Bell at the Linguistic Data Consortium will provide further word-level alignment of the audio. Once this new database is fully constructed and mounted on the TalkBank servers, it will be an outstanding resource for legal scholars and an excellent target for collaborative commentary. Specifically, we plan to first concentrate our efforts on making available a complete set of cases in these four areas: copyright, privacy and reproductive rights, religious expression, and freedom of speech. Working with the University of Michigan Press, we will organize collaborative commentary circles in each of these areas with the goal of compiling volumes on the legal handling of these issues as reflected in Supreme Court oral arguments. Contributors will include legal scholars, historians, conversation analysts, and the attorneys who argued the specific cases. Kevin Ashley at the University of Pittsburgh and Vincent Aleven at CMU will also contribute detailed cognitive analyses of legal argumentation, focusing on the role of hypothetical and conditional reasoning.

Child Development

Finally, we also plan to organize a Collaborative Commentary Circle in the subarea of Child Development that focuses on interactional components of socialization practices. Within this area, we will include both parent-child interactions during the early years and peer group interactions during later years. Many of the members of this circle have already had wide experience with the use of transcripts from the CHILDES database. However, none of them have yet become familiar with our new facilities for web-based browsing of the video database and none have yet been able to think about the application of Collaborative Commentary to these issues. However, this field is a natural for this application. We already have large quantities of parent-child video in the CHILDES database. We have access to peer group data from both school and playground. Catherine Snow and Shoshana Blum-Kulka have organized several meetings and sessions involving workers in the peer-group area. Researchers such as Lois Bloom, Michael Lamb, Grazyna Kochanska, and Inge Bretherton have had extensive experience with video analysis of parent-child interactions. Our task at this point is to organize these two subgroups to produce detailed collaborative commentaries.

5. Collaborative Commentary

The crucial claim underlying the TalkBank Project is that human communication is a unified fact and that this unification will eventually force these disparate fields to engage in the multidisciplinary study of communication. In order to move forcefully in that direction, we have recently been exploring the development of a new mode of scientific investigation called collaborative commentary. We can define collaborative commentary as the involvement of a research community in the interpretive annotation of electronic records. The goal of this process is the evaluation of competing theoretical claims. The process requires commentators to link their comments and related evidentiary materials to specific segments of either transcripts or electronic media.

In order to illustrate how collaborative commentary might work, consider an example based on my own interests in the process of word learning. Experimental studies of children's word learning have become increasingly sophisticated in recent years, providing evidence for causal cue induction (Ahn – Luhmann, in press), analogic mapping (Gentner, in press)), syntactic frame induction (Katz, Baker, Mcnamara, 1974), social referencing (Baldwin, 1993) and attentional shifting (Merriman, 1999; Smith, 1999). However, there has been virtually no attempt to track the application of these proposed processes to natural word learning between parents and children (MacWhinney, in press-b). To explore the neo-Vygotskyan claim (Nelson, 1998) that word meanings are shaped through communicative interactions, I have browsed through online media at the CHILDES (childes.psy.cmu.edu) site, locating several instances of videos of mother-child book reading in the Julie, Maria, and Rollins corpora. In these interactions, mothers help children turn the pages and name the animals or objects in the pictures. In some cases, children call the pictures by the wrong name. Often mothers use these errors as opportunities to provide corrective positive feedback. For example, if the child calls a bear a "doggie," the mother will respond, "no, that's a bear, not a doggie."

Building a system to insert comments or "blogs" on these word learning phenomena only makes sense if it will be quickly picked up by a coherent academic community that is deeply committed to the analysis of learning and development in real-life contexts. In the area of early word learning, such a community does not yet exist. However, there are several research fields where these communities do exist. For these areas, practitioners are already waiting for the development of tools for producing collaborative commentary. The academic groups that are most ripe for the introduction of this tool include: aphasia rehabilitation, medical education, and legal argumentation. There are now active research groups engaging in collaborative commentary in each of these areas. For example, in the group examining the oral arguments of the Supreme Court, commentary is currently focusing both on evidence of argument failure in the Napster case and on errors in transcription that reflect serious gaps in the public record of the Court.

The final goal of TalkBank is the development of an active process of collaborative commentary in each of the 17 research circles studying human talk. This process has just now begun, but it is already an exciting beginning.

References

Abowd, G. D., Harvel, L. D., and Brotherton, J. A. (2000). "Building a digital library of captured educational experiences." *Kyoto International Conference on Digital Libraries*, 395–402.

Ahn, W., and Luhmann, C. (in press). "Demystifying theory-based categorization." In L. Gershkoff-Stowe and D. Rakison (eds.), *Building Object Categories in Developmental Time*. Mahwah, NJ: Lawrence Erlbaum Associates.

Aisenman, R. A., and Berman, R. (2000). "Rethinking lexical analysis." In M. Aparici, N. Argerich, J. Perera, E. Rosado, and L. Tolchinsky (eds.), *Developing Literacy Across Genres, Modalities, and Language*, Vol. 3. Barcelona: University of Barcelona, 187–196.

Anderson, J., and Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Erlbaum.

Baecker, R. M., Fono, D., and Wolf, P. (2006). "Towards a video collaboratory." In R. Goldman, R. Pea, B. Barron, and S. Derry (eds.), *Video Research in the Learning Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Baldwin, D. A. (1993). "Early referential understanding: Infants' ability to recognize referential acts for what they are." *Developmental Psychology* 29: 832–843.

Beardsley, L., Cogan-Drew, D., and Olivero, F. (2006). "VideoPaper: Bridging research and practice for pre-service and experienced teachers." In R. Goldman, R. Pea, B. Barron, and S. Derry (eds.), *Video Research in the Learning Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Caspar, T. A., Meyer, W. M., Moiler, J. M., Henline, P., Keither, K., McHarg, B., et al. (1998). "Collaboratory operation in magnetic fusion research." *ACM Interactions* 5: 56–64.

Cohen, J., MacWhinney, B., Flatt, M., and Provost, J. (1993). "PsyScope: An interactive graphical system for designing and controlling experiments in the Psychology laboratory using Macintosh computers." *Behavior Research Methods, Instrumentation, and Computers* 25: 257–271.

Crowley, K., Callanan, M. A., Tenenbaum, H. R., & Allen, E. (2001). Parents explain more often to boys than to girls during shared scientific thinking. *Psychological Science, 12*, 258-261.

Daro, P., Hampton, S., and Reznick, L. (2004). *Speaking and Listening for Preschool through 3rd grade: Book Discussion and Narratives*. Pittsburgh: LRDC and the National Center for Education and the Economy.

Derry, S. (in press). "STEP as a case of theory-based web course design." In A. O'Donnell, and C. Hmelo (Eds.), *Collaboration, Reasoning, and Technology*. Mahwah, NJ: Erlbaum.

Eckman, F. R. (1991). "The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners." *Studies in Second Language Acquisition* 13: 23–41.

Edelson, D., Pea, R., and Gomez, L. (1996). "Constructivism in the collaboratory." In B. G. Wilson (ed.), *Constructivist Learning Environments: Case Studies in Instructional Design*. Englewood Cliffs, NJ: Educational Technology Publications, 151–164.

Ferguson, C., Peizer, D., and Weeks, T. (1973). "Model and replica phonological grammar of a child's first words." *Lingua* 31: 35–65.

Frederiksen, C. (1999). "Learning to reason through discourse in a problem-based learning group." *Discourse Processes* 27: 135–160.

Frederiksen, J., Sipusic, M., Sherin, M., and E., W. (1998). "Video portfolio assessment: Creating a framework for viewing the functions of teaching." *Educational Assessment* 5(4): 225–297.

Gentner, D. (in press). "The development of relational category knowledge." In L. Gershkoff-Stowe, and D. Rakison (eds.), *Building Object Categories in Developmental Time*. Mahwah, NJ: Lawrence Erlbaum Associates.

Goad, H., and Rose, Y. (2003). "Segmental-prosodic interaction in phonological development: A comparative investigation." *Canadian Journal of Linguistics* 48: 139–152.

Goad, H., and Rose, Y. (2004). "Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic." In R. Kager, J. Pater, and W. Zonneveld (eds.), *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, 109–157.

Goldman, R. (2006). "Orion, an online digital video data analysis tool: Changing our perspectives as an interpretative community." In R. Goldman, R. Pea, B. Barron, and S. Derry (eds.).

Goldman-Segall, R., and Reicken, T. (1989). "Thick description: A tool for designing ethnographic interactive videodisks." *SIGCHI Bulletin* 21: 118–122.

Goodwin, C. (1994). "Professional vision." *American Anthropologist* 96: 606–633.

Greeno, J. (1998). "The situativity of knowing, learning, and research." *American Psychologist* 53: 5–26.

Haviland, J. B. (1993). "Anchoring, iconicity, and orientation in Guugu Yimithirr pointing gestures." *Journal of Linguistic Anthropology* 3: 3–45.

Johnson, T. (in press). *Oral Arguments and Decision Making on the United State Supreme Court*. New York: Cambridge University Press.

Kager, R., Pater, J., and Zonneveld, W. (2004). *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.

Katz, N., Baker, E., and Macnamara, J. (1974). "What's in a name? A study of how children learn common and proper names." *Child Development* 45: 469–473.

Koschmann, T. (1999). "Special Issue: Meaning making." *Discourse Processes* 27(2): 98–167.

Koschmann, T., and LeBaron, C. (2002). "Learning articulation as interactional achievement: Studying the conversation of gesture." *Cognition and Instruction* 20: 249–282.

Lampert, M., and Loewenberg-Ball, D. (1998). *Teaching, Multimedia and Mathematics: Investigations of Real Practice*. New York: Teachers' College Press.

Lesh, R., and Lehrer, R. (2000). "Iterative refinement cycles for videotape analyses of conceptual change." In D. Lesh, and A. Kelley (eds.), *Handbook of Research Design in Mathematics and Science Education*. Mahwah, NJ: Lawrence Erlbaum, 647–664.

Lindblom, B. (2000). "Developmental origins of adult phonology: The interplay between phonetic emergents and the evolutionary adaptations of sound patterns." *Phonetica* 57: 297–314.

Linde, C., Roschelle, J., and Stevens, R. (1994). *Innovative Assessment for Innovative Engineering Education*. Palo Alto: Institute for Research on Learning.

MacWhinney, B. (in press-a). "Opening up video databases to collaborative commentary." In R. Goldman, R. Pea, B. Barron, and S. Derry (eds.), *Video Research in the Learning Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (in press-b). "Studying category learning in natural contexts." In L. Gershkoff-Stowe, and D. Rakison (eds.), *Building Object Categories in Developmental Time*. Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (in press-c). "The emergence of linguistic form in time." *Connection Science*.

MacWhinney, B., Martell, C., Schmidt, T., Wagner, J., Wittenburgh, P., Brugman, H., et al. (2004). "Collaborative commentary: Opening up spoken language databases." In *LREC 2004*. Lisbon: LREC, 11–15.

MacWhinney, B., St. James, J. D., Schunn, C., Li, P., and Schneider, W. (2001). "STEP – A System for Teaching Experimental Psychology using E-Prime." *Behavior Research Methods, Instruments, and Computers* 33(2): 287–296.

Merriman, W. (1999). "Competition, attention, and young children's lexical processing." In B. MacWhinney (ed.), *The Emergence of Language*. Mahwah, NJ: Lawrence Erlbaum, 331–358.

Morford, J. P., and Goldin-Meadow, S. (1997). "From here and now to there and then: The development of displaced reference in Homesign and English." *Child Development* 68: 420–435.

Nardi, B. (ed.) (1996). *Context and Consciousness: Activity Theory and Human-Computer Interaction*. Cambridge, MA: MIT Press.

Nelson, K. (1998). *Language in Cognitive Development: The Emergence of the Mediated Mind*. New York: Cambridge University Press.

Parisse, C., and Le Normand, M. T. (2000). "Automatic disambiguation of the morphosyntax in spoken language corpora." *Behavior Research Methods, Instruments, and Computers* 32: 468–481.

Pea, R. (1999). "New media communication forums for improving education research and practice." In E. C. Lagemann, and L. S. Shulman (eds.), *Issues in Education Research: Problems and Possibilities*. San Francisco, CA: Jossey Bass, 336–370.

Pea, R. (in press). "Video-as-data and digital video manipulation techniques for transforming learning sciences research, education and other cultural practices." In J. Weiss, J. Nolan, and P. Trionas (eds.), *International Handbook of Virtual Learning Environments*. Amsterdam: Kluwer.

Roschelle, J., Pea, R., and Trigg, R. (1990). *VideoNoter: A Tool for Exploratory Video Analysis* (Technical). Palo Alto, CA: Institute for Research on Learning.

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., et al. (2005). "Introducing Phon: A software solution for the study of phonological acquisition." In *30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Roth, W.-M. (2001). "Gestures: Their role in teaching and learning." *Review of Educational Research* 71(3): 365–392.

Roth, W.-M., and Roychoudhury, A. (1993). "The concept map as a tool for the collaborative construction of knowledge: A microanalysis of high school physics students." *Journal of Research in Science Teaching* 30: 503–534.

Sagae, K., MacWhinney, B., and Lavie, A. (2004). "Adding syntactic annotations to transcripts of parent-child dialogs." In *LREC 2004*. Lisbon: LREC, 1815–1818.

Sanderson, P. M., and Fisher, C. (1994). "Exploratory sequential data analysis: Foundations." *Human-Computer Interaction* 9: 251–317.

Schober, M., Conrad, F., and Fricker, S. (in press). "Misunderstanding standardized language in research interviews." *Applied Cognitive Psychology* 17.

Segal, J., and Spaeth, H. (1993). *The Supreme Court and the Attitudinal Model*. New York: Cambridge University Press.

Seifert, C. M., and Patalano, A. L. (2001). "Opportunism in memory: Preparing for chance encounters." *Current Directions in Psychological Science* 10, 198–201.

Sfard, A., and McClain, K. (2002). "Special Issue: Analyzing tools: Perspective on the role of designed artifacts in mathematics learning." *Journal of the Learning Sciences* 11: 153–388.

Siegler, R., and Crowley, K. (1991). "The microgenetic method: A direct means for studying cognitive development." *American Psychologist* 46: 606–620.

Smith, L. (1999). "Children's noun learning: How general processes make specialized learning mechanisms." In B. MacWhinney (ed.), *The Emergence of Language*. Mahwah, NJ: Lawrence Erlbaum Associates, 277–304.

Stampe, D. (1973). *A Dissertation on Natural Phonology*. Chicago: University of Chicago Press.

Stigler, J., Gallimore, R., and Hiebert, J. (2000). "Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies." *Educational Psychologist* 35(2): 87–100.

van Buren, D., Curtis, P., Nichols, D. A., and Brundage, M. (1995). "The AstroVR collaboratory: An on-line multi-user environment for reseasrch in astrophysics." In H. E. Payne, and J. J. E. Hayes (eds.), *Astronomical Data Analysis Software and Systems IV, ASP Conference Series*: Astronomical Society of the Pacific.

Vihman, M. M., DePaolis, R. A., and Davis, B. L. (1998). "Is there a 'trochaic bias' in early word learning? Evidence from infant production in English and French." *Child Development* 69: 935–949.

Wactlar, H. D., Christel, M. G., Gong, Y., and Hauptmann, A. G. (1999). "Lessons learned from building a terabyte digital video library." *IEEE Computer, February 1999*, 66–73.