

Transcription Issues in Current Linguistic Research

Organizers:

Mary Bucholtz and John W. Du Bois
University of California, Santa Barbara

Linguistic Society of America
80th Annual Meeting
Albuquerque, New Mexico
January 5-8, 2006

Session Proposal

Methodological and theoretical developments in several subfields of linguistics have led many linguists to draw increasingly on the data of spoken language as a primary source for a wide range of scholarly investigations. From the emerging focus on language documentation, field linguistics, and endangered languages, to the growing emphasis on corpus-based analysis of grammar, to phonologists' longstanding interest in prosody, many are finding that naturalistic recordings of spoken language are of vital importance for addressing their current research questions. As linguists take up new kinds of research projects, many involving multimodal digital technologies that have only recently become widely available, they are beginning to confront with greater urgency the issue of how spoken data are to be represented in visual form. For researchers who have not previously drawn extensively on ordinary conversation as a primary source of data, such questions may be new; by the same token, although well-established tools of transcription already exist in subdisciplines that use elicitation techniques in data collection, new issues arise when data sets are extended to include unelicited spoken language and longer stretches of discourse. Yet transcription remains largely undiscussed in many areas of the field. For example, advance publicity for the recent 2005 LSA Summer Institute conference on "Language Documentation: Theory, Practice, and Values" noted that "the quality of documentation available for an endangered language can determine the success" of revitalization and other programs. Although the conference addressed such important documentation issues as digital technologies, archiving, and preservation of data, it did not, apparently, include any sessions or individual papers devoted specifically to problems of transcription. In the field of linguistics, the need is there—though not always recognized—for a thoroughgoing reassessment of the state of the art in the transcription of spoken language for linguistic research. The time is ripe to focus attention on the role of transcription in creating the data representations that our linguistic research depends on. These are the needs that will be addressed by this session on "*Transcription Issues in Current Linguistic Research*."

On one level, transcription seems simple: you write down what you hear. But linguists know that there is more to it, to the extent that linguistic theory is inherently bound up in the process. For example, among the theoretical issues which impact transcription are the representation of contrasts, categories, hierarchies, predictable patterns, rules, and so on. Transcription is not merely a preliminary to analysis but is itself part of the analysis. This is evidenced by the fact that different research communities often use different forms of transcription. Transcription systems differ in the specific sets of analytic choices they offer for representing the spoken language data, and the set of available transcription choices may lead different analysts to see different things in the same recorded data. Nevertheless, it is possible to identify certain broad classes of issues that all transcription systems must come to terms with in one way or another, including representations of words and their temporal sequence, phonetic variation, prosody, and other features. In addition to theoretical concerns, a number of methodological issues arise in all transcription systems, including issues of iconicity and ease of learning of transcription symbols, reliability of data management, and so on. For these reasons, it is useful for this panel to bring together

representatives of diverse transcription practices to compare notes on the various problems and solutions that various communities of transcription users have identified to date.

Participants

This session brings together a wide-ranging group of scholars who have interest and expertise in the transcription of spoken language, to focus on transcription as a fundamental analytic tool for linguistic research. Panel members include representatives of several distinct subfields, and bring diverse theoretical and methodological commitments to their research in linguistics. Panel members include linguists with expertise in grammar, prosody, discourse, corpus linguistics, computational linguistics, field linguistics, endangered languages, archiving, orthography, literacy, psycholinguistics, pragmatics, sociolinguistics, and other areas.

We anticipate that the audience members will be attracted from across the various disciplines represented on the panel and more broadly across the field of linguistics. On a more general level, we anticipate that the panel will appeal to three different types of audiences across the discipline:

- (1) researchers embarking on a kind of research which is new to them, who do not currently have a transcription system suitable for this kind of research, including especially graduate students but also more established scholars venturing into new territory
- (2) researchers who currently use some transcription system, but recognize the need to adapt it to the demands of new methodological or practical considerations
- (3) researchers who wish to develop fundamentally new approaches to transcription in order to explore new theoretical dimensions of language

As an indication of the current level of interest in the topic, with little advance advertising a one-day symposium on transcription held at the University of California, Santa Barbara in Spring 2005 attracted an audience of nearly 100 participants from neighboring institutions. The proposed panel will cover a wider range of topics and involves different personnel and so does not duplicate this earlier event.

Goals

This event's central goal is to provide participants with the analytic tools and theoretical resources they need to develop methods of transcription that serve their own current research needs. The session does not seek to impose a transcription orthodoxy or even to move toward a single standard for transcription. Rather, it seeks to provide a forum for discussing a wide range of intellectual issues that confront anyone who uses transcription as a critical research tool for linguistic research.

Given current research trends, an in-depth examination of the wide range of issues involved in linguistic transcription is a timely and valuable undertaking for the entire field of linguistics. Specifically, the goals of the session are to:

- (1) present a capsule description of a range of transcription systems from several subfields
- (2) survey the most salient current transcription issues for several different transcription systems
- (3) allow audience members to become acquainted with the array of available transcription systems, whether they wish to:
 - adopt an existing system
 - learn from the successes and mistakes of others
 - adapt or create their own transcription system
- (4) foreground issues of the role of transcription design in the building and management of data sets, including spoken language corpora and language archives
- (5) address the transcription concerns of those who create corpora as well as those who use them
- (6) place the issues of transcription design in a broader perspective of ethics, politics, and users' goals

Presentations

The session is divided into three broad topics that are of central importance in current linguistic research: (1) general principles of transcription design; (2) the representation of prosody; and (3) the representation of linguistic variation and diversity. Although these topics are addressed by presenters throughout the session, this organizational scheme promotes focused dialogue between papers and allows for a natural progression from one theme to the next.

Following a brief introduction by the session organizers sketching the main issues in transcription as a tool for linguistic research and outlining the goals of the session, the first set of presentations examines the development of transcription systems for use in the creation of corpora, which are increasingly important resources for linguistic research of all kinds. In the first paper in the session, John W. Du Bois presents the concept of the delicacy hierarchy, a procedure for calibrating the precision with which transcription is carried out for particular analytic purposes. This paper provides an overview of the range of discourse phenomena that can potentially be represented in transcription, and offers some principles for deciding when and how to include such phenomena in a transcript. The second paper, by Jane Edwards, describes the development and transcription of a corpus of spoken discourse for use in two different types of linguistic research: acoustic analysis and discourse analysis. Through the inclusion of tagging and annotation, the transcripts of this corpus are adapted for multiple kinds of analysis, thus allowing for a single corpus to be used for a wide range of linguistic research. Thomas Schmidt's paper provides another example of transcription of a corpus for multiple uses, this time involving research on a variety of issues related to multilingualism. Schmidt describes the EXMARaLDA transcription system, which includes a set of tools for transcription, maintenance, and analysis; the system was developed to enable researchers to use their own project-specific transcription while sharing data with other research teams and thus allows for flexibility and variability in transcription practice.

The second set of papers examines in depth the issues raised in the transcription of prosody, a dimension of spoken language whose representation in written form presents unique challenges to linguists. Stefanie Shattuck-Hufnagel, Nanette Veilleux, and Alejna Brugos offer an overview of the ToBI (Tones and Break Indices) system of transcribing and labeling prosodic constituents. The authors document the influence of ToBI in studying prosody in a number of different languages and dialects as well as surveying the system's broad range of uses, including discourse analysis, phonetic analysis, computational linguistics, and speech/gesture coordination. Daniel Hirst's presentation offers another approach to the transcription of prosody: via acoustic modeling and coding of the original speech signal. Hirst's methodology, which is being tested in a number of languages, involves obtaining a phonetic representation semi-automatically and creating a transcription using a prosodic alphabet and coding tonal segments, duration, and alignment; this procedure has the advantage of not requiring any appeal to prosodic function in the representation of prosodic form.

The final set of papers concentrates on questions of representing linguistic variation and diversity in research transcripts, an issue that inevitably confronts every transcriber. The first paper in this group, by Robert Englebretson, argues on the basis of comparative transcription of Indonesian and American English that transcription systems must be adapted to the specific language being transcribed. The paper highlights historical, cultural, and linguistic issues that must be considered when making transcription choices, and additionally offers recommendations for how to train native-speaking research assistants in transcription and how to incorporate glossing into transcription design. In the next presentation, Alexandra Jaffe focuses on the representational effects of orthographic variation in the transcription of linguistic varieties for which there is no standard orthography. Noting that nonstandard spellings in transcripts may be interpreted either as symbolic of in-group identity or as associated with sociolinguistic stigma, Jaffe discusses the potential of transcription to either challenge or reinforce dominant language ideologies. In the last paper in the session, Mary Bucholtz addresses a set of transcription practices that present dilemmas for linguists because they so often go unnoticed yet have powerful representational effects. Through use of published research transcripts, Bucholtz demonstrates that rarely examined practices such as speaker tag assignment, the representation of turn taking, uses of capitalization and

punctuation, and translation have both analytic and ethical consequences for transcription as a tool for linguistic research. The session concludes with a general discussion among the panelists and the audience.

Session Program

The session will feature 9 speakers, with presentations of 15 minutes followed by 5 minutes for discussion, for a total duration of 3 hours. It will begin with a 10-minute introduction and orientation by the session organizers, and end with a 10-minute period for additional general audience discussion.

Organizers: Mary Bucholtz (UC Santa Barbara) and John W. Du Bois (UC Santa Barbara)

Chair: John W. Du Bois (UC Santa Barbara)

- 0:00 Mary Bucholtz (UC Santa Barbara) and John W. Du Bois (UC Santa Barbara)
“Introduction: Transcription Issues in Current Linguistic Research”
- 0:10 John W. Du Bois (UC Santa Barbara)
“Transcription and the Delicacy Hierarchy: What Is to Be Represented?”
- 0:30 Jane Edwards (UC Berkeley) (not presented)
“Transcription Design: The Multi-Use Corpus in Current Linguistic Research”
- 0:50 Thomas Schmidt (Hamburg)
“Transcription and Tools: Creating and Analyzing Spoken Language Corpora with EXMARaLDA”
- 1:10 Stefanie Shattuck-Hufnagel (MIT), Nanette Veilleux (Simmons), and Alejna Brugos (Boston)
“Transcription of Prosody: The ToBI Framework”
- 1:30 Daniel Hirst (CNRS and Université de Provence)
“Transcription of Prosodic Form: A Cross-Language Approach to Prosodic Function”
- 1:50 Robert Englebretson (Rice)
“Transcription in the Field: Transcribing Discourse across Languages”
- 2:10 Alexandra Jaffe (CSU Long Beach)
“Transcription in Sociolinguistics: Nonstandard Orthography, Variation and Discourse”
- 2:30 Mary Bucholtz (UC Santa Barbara)
“Transcription and Identity: Dilemmas in the Written Representation of Speech and Speakers”
- 2:50 General discussion
- 3:00 Session ends

Transcription and the Delicacy Hierarchy: What Is to Be Represented?

John W. Du Bois (University of California, Santa Barbara)

This paper examines the role of delicacy in the transcription of spoken language use. After defining the concept of delicacy, the paper proposes a delicacy hierarchy based on 10 criteria for assigning transcription features to levels of delicacy. The resulting hierarchy organizes 84 categories of transcription features into 7 hierarchical levels of transcription delicacy. Finally, the paper considers how the delicacy hierarchy can influence the production and interpretation of transcription data in the context of linguistic research.

As researchers we ask that a transcription be informative, accurate, readable, and, nowadays, searchable. In particular, we ask that it contribute a level of informativeness regarding salient aspects of the original recorded occasion of language use which is appropriate to the current purposes of our research. So delicacy enters the picture.

What is delicacy? *Delicacy* (or *granularity*) concerns the degree of fineness of categorial discriminations and the quantity of informational detail specified in a representation of language use. Informally speaking, if a transcription is said to have *higher delicacy* (or *finer granularity*), this means it makes finer discriminations, has more detail, and includes more different types of transcription categories. It is important not to confuse delicacy with accuracy or correctness—even a transcription that is low in delicacy can be highly accurate for the categories it claims to represent.

What are the criteria for delicacy? 10 criteria are proposed which motivate placing a feature in a relatively lower (vs. higher) delicacy level: the feature is more basic, more essential, more widely transcribed, less marked, foundational for later categories, less differentiated, earlier in transcribing, easier to transcribe, easier to learn, and carrying broader significance. These criteria were used to place 84 categories of transcription features into 7 levels of increasing delicacy. For example, a *Level 1* transcription incorporates a representation of words, speaker labels, and intonation units:

Level 1 transcription

JEFF; it's been a long time

JILL; it's been so long

Moving higher in the delicacy hierarchy, a *Level 3* transcription adds representations of simultaneous speech ([]), intonational finality (.), prosodic lengthening (:), and in-breath ((H)), among other features. The following represents the same recorded interaction, as transcribed with higher delicacy:

Level 3 transcription

1 JEFF; It's been a [long time].

2 JILL; [(H)] It's been so: long.

Note that once the marking of Jill's in-breath has been introduced, it becomes relevant to show its temporal relationship to Jeff's words. But if a lower delicacy level is used for transcribing, breathing is not transcribed, and there is nothing to overlap with.

What is a delicacy-granularity hierarchy good for? It serves at least six key functions: (1) to organize the complexity of the transcription process; (2) to organize the sequential introduction of transcription features into transcripts; (3) to organize the sequential learning of transcription concepts; (4) to prioritize transcription categories according to value (e.g. for research); (5) to create a framework for systematic transition between degrees of delicacy; and (6) to modulate the interpretation of transcription data at a given delicacy level.

Transcription Design: The Multi-Use Corpus in Current Linguistic Research

Jane Edwards (UC Berkeley)

Corpus construction is notoriously time consuming and expensive. For this reason, “multi-use” corpora are highly desired. However, transcripts are never theory-neutral. Is it possible to meet the needs of two different communities within a single transcript?

The ICSI Meetings Corpus project was designed to accommodate speech researchers on the one hand and discourse researchers on the other. In addition to this, its transcription conventions needed to be internally consistent so that they could be separated out via XML markup and still yield a sensible representation of the meeting. This paper summarizes the main features of this transcription.

The ICSI Meetings corpus consists of 75 research meetings, of 4 to 9 participants each, in which all participants wore headset microphones. The transcripts are word-level transcripts, time-synchronized to digitized audio recordings. In addition to words, there are several types of tags, annotations and comments.

Transcripts were prepared by means of a computer interface in which each person’s channel is displayed as a horizontal ribbon, which can be subdivided into smaller segments of time, called “time bins.” The times bins on one person’s channel are independent of those on other speakers’ channels. The basic strategy is to view each ribbon as capturing the actions (spoken words, non-word vocalizations, non-vocalized sounds or silence) of a particular meeting participant, as heard via the headset microphone.

The main transcription-related differences between the two target groups of researchers are as follows: Speech researchers are interested mainly in acoustic events, whereas discourse researchers are mainly interested in communicative events. In addition, the two communities differ in the techniques used to analyze the data. Speech researchers train mathematical models to find patterns in large amounts of data. Linguists often search for expressions or read through the transcript looking for patterns in the data.

To facilitate speech research, tags were added to distinguish sounds which were produced by a vocal tract (e.g., breath, laugh, cough) from sounds which were not (e.g., door slam, knock at the door, microphone spike). In addition, conventions were added to help identify word forms which would likely not match the stored lexical representations. These included: foreign language expressions, non-canonical pronunciations caused by speaking while laughing or excessive lengthening of the vowel, etc.

The conventions which were added for linguistic research, had to do with clarifying communicative acts. For example, distinguishing orthographically between words which speech researchers would not distinguish: “wanna” vs. “want to”, and tagging other word forms to enable easy retrieval (e.g., interjections and vocal gestures, such as a vocalized outbreath in the context of frustration or difficulty). Also, intonational phenomena and contextual information were noted in comments fields.

The differences between the two research groups were not incompatible; they simply involved different areas of detail. It is likely that multi-use corpora will increase in the future. It is hoped that the observations mentioned here are helpful in that effort.

Transcription and Tools: Creating and Analyzing Spoken Language Corpora with EXMARaLDA

Thomas Schmidt (Hamburg)

EXMARaLDA (Extensible Markup Language for Discourse Annotation) is a system of data formats and tools developed at the Research Centre on Multilingualism at the University of Hamburg. In this research centre, 14 different projects carry out research on multilingualism from a variety of theoretical perspectives and with quite diverse aims of knowledge. Their different transcription practices, i.e. the symbols, categories, layout principles and computer software used to create representations of spoken language interaction on the computer or on paper, reflect this diversity. Data exchange, and hence data-driven cooperation between projects, thus becomes a non-trivial matter. EXMARaLDA is designed to bridge a part of this gap. It formulates a common data model for different approaches to transcription and provides input, output and analysis tools that operate on this common data model, but still allow for a theory or project specific handling of the data. Thus, researchers are able to use whichever symbols (IPA, Roman or other alphabets etc.), categories (turns, utterances, intonation units, syllables etc.) and layout principles (score, line, column notation etc.) their theories and research aims require, and still draw from a common set of data formats and software tools.

In this talk, I will give an overview over the system's components. More specifically:

- A data model based on the annotation graph approach by Bird/Lieberman (2001), designed to achieve a maximum degree of interoperability between different transcription systems and different software tools.
- Several XML based data formats that implement this data model and that are suitable for long-term archiving and reuse of transcription data.
- The Partitur Editor, a transcription editor optimized for the transcription of multi modal interaction and/or spoken interaction with many participants. This is the core software tool of the EXMARaLDA system. Its user interface is a "Partitur", i.e. a musical score where each tier (or layer) represents one line of action for one speaker. As the following example demonstrates, this makes possible an intuitive representation of sequential and simultaneous events in spoken discourse as well as the adding of analytical information to the transcribed material.

DS [sup]	<i>faster</i>
DS [v]	Okay. D'accord d'accord.
DS [en]	Okay. Agreed, agreed.
DS [nv]	<i>right hand raised</i>
FB [v]	Alors ça dépend ((cough)) un petit peu.
FB [en]	That depends, then, a little bit
FB [pho]	[ɛ̃tipø:]

- The Corpus Manager, a tool for creating, maintaining and searching meta data for corpora of spoken language. This tool is intended to help researchers create and maintain larger collections of transcription data.
- ZECKE, a query tool for locating transcribed phenomena inside a corpus of spoken language. This tool is based on what classical concordance software does for written language corpora, but has additional functionality for dealing with the particularities of spoken language transcriptions.

Transcription of Prosody: The ToBI Framework

Stefanie Shattuck-Hufnagel (MIT), Nanette Veilleux (Simmons), and Alejna Brugos (Boston)

The ToBI system (Tones and Break Indices) is a labelling system for transcribing the intonation and prosodic constituent structure of spoken utterances, based on the intonational theory of Pierrehumbert (1980) and Pierrehumbert and Beckman (1986, 1988) (the autosegmental-metrical approach) and the break index method for labelling constituent structure proposed in Price et al. (1991). It specifies both the phonological elements that are presumed to govern the f₀ contour (e.g. High, Low and complex pitch accents, High and Low phrase accents, and High and Low boundary tones), and 5 levels of constituent boundary. For example, X1_SBC015_0862-94.wav (10.18-11.31 seconds) is labelled

	H*	L+H*		L-L%		tones
It's	not	at	all	like	me.	words
2	0	1-	1	1	4	breaks

ToBI was developed in a series of workshops in the early 1990's, to address the need for a generally-accepted convention for labelling prosody (Ostendorf et al. 1995). Since then a number of speech databases have been ToBI-labelled. These prosodically labelled databases have facilitated a wide variety of studies of prosody and its interaction with other aspects of spoken utterances. These include inquiries into cues to discourse structure (Hirschberg and Nakatani 1996, Nakatani 1997), the prosody of read vs. spontaneous speech (Hirschberg and Grosz 1992), studies of the co-occurrence patterns among intonational elements (Dainora 2002), comparison of boundary cues across languages (Fon 2002), and the effect of prosodic structure on surface phonetic variation (Cole et al. 2003, Dilley et al. 1996, Shattuck-Hufnagel et al. 1994, Veilleux and Shattuck-Hufnagel 1998, Bell et al. 2002, Jurafsky et al. 2002). These labelled databases have also formed the basis for development of algorithms to automatically annotate prosody for automatic speech recognition and text-to-speech projects. In addition, extensive use of these annotation systems for spontaneous speech has identified a number of questions which are the subject of intensive current research in linguistics, speech engineering and cognitive science.

ToBI-labelled corpora have enabled the study of constituent-edge-related phonetic phenomena, such as articulatory strengthening and voice quality changes, which have made sense of many patterns of phonetic variation which do not align with traditional syntactic structure. In addition, these corpora permit the investigation of prosodic meaning, of the frequency and distribution of different types of prosodic elements, and of the alignment of prosodic structure with other events such as speech-accompanying gestures.

Development of the ToBI system for transcribing the prosodic structure of Mainstream American English (MAE-ToBI, Beckman et al., 2005) has served as a model for developing similar resources for other language systems (Jun 2005). There are now a number of ToBI systems or proposed ToBI tag extensions for other dialects of English (e.g. Mayo et al. 1997 for Glasgow English, Fletcher et al. 2001, 2002 for Australian). ToBI systems developed for other languages include J-ToBI for Japanese (Venditti 2005), K-ToBI for Korean (Jun 1993), G-ToBI for German (Grice, 2005), and related autosegmental-metrical systems for Mayali (Bishop and Fletcher 2005), Chickasaw (Gordon 2005) etc.

Transcription of Prosodic Form: A Cross-Language Approach to Prosodic Function

Daniel Hirst (CNRS and Université de Provence)

This paper presents a framework for the multilingual analysis by synthesis of speech prosody which has been developed over a number of years in Aix-en-Provence and is currently being evaluated with a number of Indo-European and non-Indo-European languages.

The analysis is based on the assumption that in all languages a certain number of prosodic functions are expressed by prosodic forms. While there is evidence that both the functions and forms of prosody present quasi-universal characteristics, the mapping between the two levels is clearly highly language (and dialect) specific.

Instead of combining the representation of functions and forms, as in most other prosodic transcription systems, in this approach the two levels are clearly separated.

As a first step, a phonetic representation is obtained semi-automatically from the acoustic signal. Segmental duration is normalised with respect to the statistical distribution of each phoneme. Fundamental frequency is modelled semi-automatically as a sequence of target points using the MOMEL algorithm (Hirst et al. 2000).

A quasi-phonological surface transcription of prosodic form is then obtained without reference to prosodic function, using the prosodic alphabet INTSINT (Hirst 2005), coding the target points by means of a set of discrete tonal segments and the segmental durations with a scalar lengthening factor. These short-term characteristics are accompanied by longer term factors of tempo, pitch key and range. The alignment of the tonal segments with relation to the prosodic structure is finally coded by a system of diacritics which, like the segmental lengthening factor, can be applied to any appropriate level of prosodic structure (phonemes, syllables, rhythm units etc) making it possible to apply and evaluate different models of rhythm and tonal alignment.

The reversibility of the acoustic modelling and coding means that the only question the linguist need ask is whether the resynthesised speech expresses the same prosodic functions as the original signal: it is not necessary at this stage to identify or classify these functions, many of which have proved notoriously elusive to investigators.

Once a sufficient degree of approximation to the original message has been achieved by an appropriate symbolic transcription, it then becomes a task of relating these transcriptions to more abstract representation of prosodic functions, a task which it is hoped will prove amenable to techniques of natural language processing. It is further hoped that the fact that no a priori assumptions about the inventory of prosodic functions are incorporated into the transcription of prosodic forms will make it possible to evaluate the adequacy of different models of prosodic functions.

References

- Hirst, D.J., Di Cristo, A. & Espesser, R. 2000. Levels of representation and levels of analysis for intonation. in M. Horne (ed) *Prosody: Theory and Experiment*. Kluwer Academic Publishers, Dordrecht. 51-87.
- Hirst, D.J. 2005. Form and function in the representation of speech prosody. *Speech Communication* 46 (3-4), 334-347.

Transcription in the Field: Transcribing Discourse across Languages

Robert Englebretson (Rice University)

This talk provides a discussion of transcribing conversational data in a field language of which the researcher is not a native speaker. Based on my own specific experiences in Indonesia, I address both practical issues of transcription methodology, as well as theoretical implications for designing a transcription system. In terms of practical “how-to” issues, I briefly discuss the training of research assistants for recording and transcribing naturally-occurring conversational discourse in the field, as well as database design for interlinear glossing of conversational data. In terms of theoretical implications for transcription design, based on excerpts from my corpus of colloquial Indonesian (described in Englebretson 2003) and excerpts from the Santa Barbara Corpus of Spoken American English (Du Bois et al.) I demonstrate subtle yet specific ways in which the target language itself influences the process and product of discourse transcription. I argue that language researchers need to pay greater attention to the dynamic relationship between transcribing and the specific language being transcribed.

Previous published work on transcription design (most notably papers in Edwards & Lampert, ed. 1993. *Talking Data*) has focused primarily on how theoretical orientation and research methods shape the design of transcription and coding systems. While such work has indeed proven invaluable in prompting language researchers to choose an optimal transcription system appropriate to particular research goals and questions, I argue that, in addition, language researchers must also optimize the transcription system based on the particular language being transcribed. With regard to my own work on colloquial Indonesian conversation, I explore three such aspects.

First, I argue that the very act of transcription is always contextualized within the historical and sociocultural framework of the target language. In colloquial Indonesian, the diglossic contrast of formal standard (written) Indonesian on the one hand, and informal colloquial (spoken) Indonesian on the other, problematizes the act of transcription, and what it is that native-speaking research assistants understand their task to be. The act of producing written language calls for the formal variety, yet the language being written down (transcribed) violates these expectations.

Secondly, I argue that particular transcription practices which are inappropriate for one language may indeed be appropriate for another. For example, many researchers of conversational English argue against the use of re-spelling (or “eye dialect”) as it is often difficult to interpret, difficult to use in an automated concordance, and indexes social stigma. Yet, in the context of colloquial Indonesian, I show that re-spelling is an important analytic tool which native-speaking transcribers employ in the representation of consistent features of the colloquial variety.

Thirdly, I show the necessity of having a discourse transcription system which is flexible enough to accommodate unique aspects of the target language. Specifically, there are unique prosodic features of colloquial Indonesian which a transcription system developed for English does not accurately capture.

In sum, this talk contributes a cross-linguistic perspective to the panel, and seeks to initiate a discussion of the ways in which transcription design and implementation are directly affected by the specific language being transcribed.

Transcription in Sociolinguistics: Nonstandard Orthography, Variation and Discourse
Alexandra Jaffe (California State University, Long Beach)

This presentation focuses on the connection between orthographic variation in transcription and processes of social identification, differentiation and stratification. Following a line of argument opened by Dennis Preston (1982), it examines the political and ideological nature of nonstandard orthographic representations of speech and how it is related to the larger issue of how writing practices reflect and reproduce sociolinguistic hierarchies. The first set of data considered concerns the debates over identity and representation of regional variation that have surrounded the elaboration of Corsican (a minority language) as a written language. These debates underscore the intimate connection between orthography and linguistic identity, and the extent to which orthographic practices authorize and legitimate particular linguistic varieties. In Corsica, these issues led to the elaboration of a "polynomic" orthography in which speakers of different varieties of the language use different spellings for the same words. That is, the speaker who pronounces the word for "ours" as [noʎu] spells it "nosciu"; but it would be spelled as "nostru" by speakers who pronounce it [nostru] in their regional variety of Corsican.

The Corsican data is used as a point of departure for a discussion of the use of nonstandard orthographies in languages like English that do not have polynomic orthographies that institutionalize the representation of difference. This includes a review of 1) categories of nonstandard spellings found both in academic transcripts and popular representations of speech and 2) research on the production and interpretation of nonstandard spellings. The former includes *eye dialect* forms like "cuz" that provide no phonetic information; *allegro forms* like "wanna" that reflect fast or casual pronunciations; *dialect respellings* like "Ah" (for a Southern American English pronunciation of "I") and *prosodic spellings* like "baaaad" to indicate vowel lengthening. Research on 2), above, shows that on the one hand, nonstandard orthographic practices can be powerful tools of self-representation and markers of in-group identity. At the same time, they invariably carry sociolinguistic stigma.

With this strong risk of stigma in mind, the rest of the presentation explores the sociolinguistic information value of nonstandard orthography in transcriptions with respect to variation and its predictability, and with respect to principles of consistency. This discussion shows that the ideological consequences of using nonstandard spellings in the representation of others have to be understood with reference to specific, situated discursive contexts of production and interpretation/consumption. This raises complex questions for linguists relative to their relationship to the speakers that they represent, and with respect to the multiple purposes and audiences for their transcriptions. Nonstandard spellings have the potential to expose and call into question dominant language ideologies and the social hierarchies in which they are embedded. But this subversive power is always held in check by the very power of those ideologies, and respellings of others are always problematic with respect to the legitimacy and authority of those they represent.

Reference

Preston, Dennis R. 1982 'Ritin 'Fowklower Daun 'Rong: Folklorists' failure in phonology. *Journal of American Folklore* 95(377):304-326

Transcription and Identity: Dilemmas in the Written Representation of Speech and Speakers

Mary Bucholtz (University of California, Santa Barbara)

Following Ochs's (1979) groundbreaking article on discourse transcription as a theoretical as well as a methodological undertaking, linguists who work with discourse data have become increasingly aware of the implicit analytic commitments embedded in any given transcription system. As Ochs showed, decisions about how to transcribe the speech of different kinds of speakers—in her example, adults versus children—can lead analysts to overlook important aspects of discourse structure. Although Ochs's focus is on the analytic consequences of transcription, it is clear from her discussion that transcripts are representations not only of speech but also of speakers. That is, in writing speech down in particular ways, transcribers also (often unwittingly) locate those whose speech they transcribe as embodying particular identities or social positions (such as dominant adult and responsive child), that may not be justified by the data. This paper argues that several widely used linguistic transcription practices may unintentionally promote empirically unwarranted representations of speakers and thus raise both analytic and ethical questions for researchers.

Linguists have long been concerned with ethics in creating research transcripts. Sociolinguists in particular have focused on the important ethical issues involved in the nonstandard orthographic representation of nonstandard linguistic varieties (e.g., Jaffe and Walton 2000; Preston 1985). However, much less attention has been given to ethical issues in transcription beyond orthography. Such practices may not be obvious, and it is therefore difficult for researchers to recognize them and to develop explicit principles for ethical representation.

Drawing on published examples of researcher transcripts within discourse analysis and sociolinguistics, the paper examines four such practices: (1) the assignment of speaker tags (e.g., role-based ["teacher"] versus name-based ["Jane"]); (2) the representation of turn taking (e.g., marking a researcher/interviewer's speech with a different font or style); (3) the creation of a naturalized (i.e., writing-like and product-oriented) versus a denaturalized (i.e., speech-like and process-oriented) transcript via capitalization, punctuation, spatial formatting, and other devices; and (4) the representation of other languages and their translation into the researcher's language (e.g., using a very formal or very informal style). In each instance, the potential analytic and ethical consequences of transcriber choices are assessed.

Example 1 illustrates practices (1) and (3); it is a denaturalized transcript that uses lowercase and stanza-like formatting; in addition, the speaker is assigned no speaker tag. The researchers who published this transcript used it to illustrate an "oral" style, but the transcription itself in part creates this effect and positions the nameless speaker as a generic representative of such a style.

- (1) cool dūde, you know catch wōmen, this and that.
 but he, must get his nò:se, wide open, behind
 some ôther girl, and this and that

Through such small but powerful transcription decisions, researchers may unintentionally bias their analysis and in some cases may even represent speakers in ways that are ethically problematic. The paper argues that transcribers should strive not for a neutral transcription system, which is impossible to achieve, but for greater awareness and explicitness about transcription decisions and their consequences.