Finding identity: Theory and data

MARY BUCHOLTZ and KIRA HALL

Abstract

This commentary responds to the papers in the special issue ‘Accomplishing identity in bilingual interaction’ and particularly to the use of Bucholtz and Hall’s (2004a, 2004b, 2005) framework for the linguistic analysis of identities in interaction. The commentary focuses on the relationship between theory and empirical work, with attention to the role of ethnographic context in the analysis of both microlevel interaction and macrolevel socio-political and sociohistorical processes, the place of language ideologies in the interactional construction of bilingual identities, and the necessity to ground theoretical claims in rigorous empirical analysis.

Introduction

It is all too rare for researchers to have the opportunity to revisit their past work and to see and respond directly to the ways that other scholars have taken up their ideas, and so we are very grateful to the editors of this special issue for their invitation to us to enter into dialog with the contributors about the relationship between their work and our own. We are still enjoying the novelty of having developed a framework that others find useful enough to apply to their own data, and it is interesting, often exciting, and sometimes surprising to see the varied ways that people are engaging with the ideas we have put forward regarding the relationship between language and identity. In stimulating us to reflect on what we had in mind when we developed the series of papers we have published on this topic (Bucholtz and Hall 2004a, 2004b, 2005), the foregoing articles have led us to think more deeply about a number of issues, but perhaps most fundamentally about theorizing and its empirical underpinnings, a relationship that lies at the heart of all our work on language and identity. In our commentary, then, we focus not only on the common themes that emerge from these papers but more generally on...
the issues they raise both individually and collectively regarding how to
find identity in interactional data and the role of theory in this process.

We are mindful of the many perils and challenges of conducting re-
search on identity, from treating it as an unanalyzed explanation for
social phenomena rather than an object of study in its own right, to
flattening the complexity of identity through ahistorical, decontextu-
alized, or overly deterministic analyses. We are therefore heartened to see
the authors of these articles taking seriously the many levels on which
identity operates: the interactional, the ethnographic, the historical, the
political.

We begin our discussion by highlighting the importance of simulta-
neously grounding the study of identity in the empirical evidence of in-
teraction and contextualizing the study of interaction through ethnogra-
phy. Such close attention to the workings of everyday life both in the
moment and over time is necessary in order to discover the specificity
and social meaning of local identity practices and categories as well as
the ways in which these are built up through the course of repeated
social actions and interactions. Conversely, this foundation in local prac-
tice allows researchers to understand how higher-level sociohistorical
and sociopolitical processes come to impinge on and be deployed by
social actors in the process of identity work. We demonstrate this point
by examining the central role of language ideologies of various kinds in
constructing identity in bilingual contexts, which emerges as a common
theme in all the articles. Finally, we reflect on the relationship between
theory and data in our own work and in this special issue. We argue for
the necessity of advancing theory within sociocultural linguistics while
keeping theories responsible to data, which must be the starting point
for any investigation of language in social life.

Identity in interaction

The contributors to this special issue share with us a belief that the social
world is constituted at the most basic level in and through interaction,
and their articles demonstrate a commitment to giving interaction the
serious analytic attention it requires in order to make empirically
grounded claims about precisely how this happens. Not surprisingly
given this emphasis, a number of the authors are inspired by conversa-
tion analysis and ethnomethodology, which provide rigorous tools for
analyzing the sequential structure of talk and the organization of mem-
bership categories. Moreover, as several of the articles demonstrate and
as we have argued in our own work, conversation analysis and other
forms of discourse analysis are often made more productive when com-
bined with ethnographic methodologies that allow researchers to make
sense of locally specific elements of interaction as well as to get at the sociocultural context in which interaction unfolds and to which it contributes. Not only does ethnography support and extend the conversation-analytic commitment to understanding interaction from the point of view of those who participate in it, but it also ensures that researchers view talk not as a chunk of text removed from any broader context but as a dynamic interactional process embedded in and inseparable from the social and cultural world from which it emerges.

With respect to bilingualism, the sociolinguistic phenomenon that is the focus of this special issue, a commitment to interaction promotes an interpretation of code choice that is firmly rooted in the endogenous concerns of participants rather than in the theoretical goals of the analyst. Thus several of the articles offer support for the recent reconceptualization of codeswitching as only those instances of language alternation that are demonstrably salient within the interactional context — a functional rather than structural definition (e.g., Gafaranga and Torras 2002). As for the sociocultural phenomenon of interest in this special issue, identity, we are also pleased to see that the contributors do not limit themselves to investigating the well-worn categories of social identity that are deeply familiar to sociolinguists — race and ethnicity, gender, age, class, and so on — important as these continue to be. Instead, they often move to a finer level of analytic granularity to consider such fleeting yet consequential subject positions as bully and bullied, advice giver and advice rejecter, interpreter and recipient of interpretation, positions that in turn may (or may not) sediment into something connected to those higher-order categories.

The question then arises as to whether categories at these different levels are all the same kinds of identities, and if not, what sort of analytic work is necessary to move from one level to the other. Our own answer to this question is that the notion of identity, fundamental as it is to sociolinguistic analysis, may not, on its own, be sufficient to the task of explaining the social dynamics of interaction, particularly in its traditional conceptualization as a set of enduring social categories such as gender, race, and class. A finer degree of analytic delicacy is required, and this is provided by the concept of stance. Stance is fundamental to identity construction; as Ochs (1992) has shown, direct indexicalities that link linguistic structures to interactional stances in turn form the basis for indirect indexicalities between stances and the social groups that (are thought to) typically take them. Hence moving too quickly from linguistic structures to identity categories without considering how interactional stances mediate between the two short-circuits the analysis and diminishes much of the utility of considering identity in the first place.

As researchers, we need to start with what speakers are accomplishing
interactionally and then build upward to the identities that thereby emerge. At the same time, in order to ensure that our analyses are cognizant of the rich intertextual layers that resonate between these different levels, we need to ground our interactional analyses both in the ethnographic specificities that endow interactions with social meaning and in the broader social, cultural, and political contexts in which social actors are imbricated. In short, neither identity categories nor interactional analyses alone are enough to account for how social positioning is accomplished through language; the two levels of analysis are most effective when they work in unison, and in conjunction with a focus on the larger social, cultural, and political contexts in which identity work is carried out.

**Finding identities and ideologies in bilingual interaction**

Identity work is a highly politicized process in which social actors claim, contest, and negotiate power and authority. When identities are forged in relation to language, they become bound up with language ideologies, which, as historically rooted and publicly articulated statements of cultural belief about language and its users, mediate between the interactional moment and broader sociopolitical structures. By the same token, ideology mediates between the interactional stances taken by speakers and the indexical relationships to identity that are thereby produced. The contributors illustrate through their data and analyses the varied and intricate ways in which ideology becomes interactionally alive for participants in bilingual talk.

Ideology plays multiple roles in Tim Greer’s examination of bilingual teenagers’ translations between Japanese and English at an international school in Japan, which skillfully uses close interactional analysis to demonstrate that such translations are designed to include all participants. Greer’s argument thus convincingly debunks the widespread and highly paranoid ideology held by many monolinguals that bilinguals use another language in order to exclude them. But we would be wrong to assume that inclusive talk of this kind lacks a power dimension. To get at this dimension requires the tools of ethnography, for while the great strength of conversation analysis is its ability to illuminate the taken-for-granted world of the participants, this commonsense understanding often requires considerable knowledge of local ideologies and practices. A speaker’s work in positioning herself as linguistically competent and another as a linguistic novice does not end at the interactional level; rather, it enters a chain of ideologization involving locally specific beliefs about who is and is not able to speak a language fluently. In Greer’s data, this ideology plays out in the dynamics of who is targeted as a recipi-
ent of translations (as it happens, the researcher himself is often positioned in this way, on the basis of ideologies of language and nationality rather than his actual fluency). In other words, although translation serves to include speakers in the interaction regardless of language ability, from another perspective it serves to exclude or at least marginalize these same speakers by singling them out as linguistically limited: adequation, the positioning of self and other as sufficiently similar, is often not far from distinction, or social differentiation. Thus accommodation through translation, while ensuring the inclusion of all participants, can also be understood as partly hierarchical, while lack of accommodation may symbolically signal inclusion by not foregrounding linguistic difference.

Another perspective on how translation ideologically creates interactional inclusion through social distinction is found in Lisa Del Torto’s investigation of intergenerational language brokering among Italian immigrant families in the United States and Canada. Del Torto rightly points to the importance of temporality in identity production, noting that identities do not emerge solely within a single interaction but rather are forged through the accretion of interactional stances and positionings over time. She shows that in immigrant families, generational identities may also be linguistic identities, with members of the second generation inhabiting the role of language broker between the first and third generations. Strikingly, this role, which has mostly been studied among children in families that have recently immigrated, may endure throughout the adult years, and it may be so integrated into the speaker’s habitus that any indication of interactional trouble between the first and third generations will trigger the second-generational speaker to step into the brokering role. As Del Torto argues, this process is ideological as well as interactional, for it is based on the second generation’s beliefs about its own ‘in-between’ status as bilingual and bicultural relative to the other two generations. The study offers a valuable reminder that while the linguistic distinctiveness of immigrant generations may be an ideological construct as much as a sociolinguistic fact, it is one that allows members of the second generation (and perhaps especially women) to keep families unified across cultural and linguistic spaces.

The role of the language broker appears in very different form in Holly Cashman’s sobering study of children’s teasing practices in a second-grade bilingual classroom in the American Southwest, where it shapes both the form and the content of the interaction. This dynamic is seen most dramatically in the configuration of participation frameworks — as when the English monolingual teacher, who inexplicably has been charged with the education of a child who speaks no English, assigns two of his bilingual classmates to serve as linguistic mediators, and in so doing enables them to torment him with impunity. It is therefore
not surprising that the bilingual children conduct their attacks on the new student entirely in Spanish — not only does their code choice ensure that the target of their insults will understand them but it likewise ensures that the teacher will not. Further, the children’s grasp of exactly what is at stake in this battle rises to the interactional surface when one hurls the metalinguistic insult ‘You don’t even speak English.’ It is, indeed, precisely the new student’s Spanish monolingualism that makes him such a vulnerable target. In addition, his Spanish too comes in for negative evaluation, especially the use of güey, a highly ideologized indexical feature of Mexican Spanish (cf. Bucholtz 2007). Through such bullying practices it is paradoxically the targeted child who is positioned as a ‘troublemaker’ by both peers and teacher alike. Cashman’s appeal to the local ethnographic context of the community of practice as well as the broader institutions and ideologies of language and nationality that are implicated here starkly reveals how longstanding histories of social inequality are reproduced through the cumulative force of social interaction (cf. Goodwin 2002).

The ways that power and inequality manifest themselves in bilingual interaction are also central to the article by Ashley Williams, whose finely observed analysis of Cantonese speakers’ orientation to Mandarin in San Francisco at a time when they have become outnumbered by Mandarin speakers indicates that language ideologies are rarely monotonic. In her data, speakers both valorize Mandarin as necessary and easy to learn and condemn it as different and its speakers as annoying. These language ideologies lead to remarkably complex stances of affiliation and disaffiliation with Mandarin. On the one hand, ‘necessary’ and ‘easy’ do not inevitably signify entirely positive evaluations. Within the United States, for example, there is an ideology that Spanish is both easy to speak and increasingly necessary, as attested by Mock Spanish on the one hand (Barrett 2006; Hill 1993), and on the other hand by the existence of phrasebooks in Southern California to help well-to-do Anglos talk to their household help (Schwartz 2006). In this situation, an ideology of simplicity and necessity coexists harmoniously with more explicitly negative and xenophobic attitudes. Likewise, ‘different’ does not necessarily imply disaffiliation and may in some situations even be a way of expressing admiration via exoticization, as with Americans’ evaluation of French as more cultured than English (Kinginger 2004; Lippi-Green 1997) or of British English as superior to American English (Jones 2001). Thus evaluations of Mandarin are tied to the interactive goals speakers may have in expressing such views, and hence to stance taking. As John W. Du Bois (2007) has discussed, a stance is not a dyadic relationship between speaker and stance object (in this case, Mandarin) but a triadic relationship through which a speaker simultaneously evalu-
ates an object, positions herself, and aligns or disaligns with the addressee. The polyvalent stances the speakers take up are connected to their alignments to other speakers and hence are less self-contradictory than they may initially appear. In moving from interaction to broader social processes, Williams’s use of the idea of ‘brought-along’ identities is particularly helpful for thinking about the temporal and spatial displacement of identity in immigrant communities, which helps shed light on the ideological dynamics at work in this interaction.

Related to Williams’s concern with the ideologies that attach to specific languages, Helena Bani-Shoraka’s study of bilingual women’s alternations between Persian and Azerbaijani in Tehran points to the ideological dimension of interaction, and how this is best uncovered by starting sociolinguistic analysis not with identities but with interactional moves, and particularly with the stances that speakers take in order to accomplish their goals. As her data show, the meanings of Azerbaijani, Persian, or Stylized Persian cannot be entirely given in advance but must emerge within the discourse itself, informed both by the ethnographic meanings of being an ethnically Azerbaijani bilingual in Iran and by each speaker’s history of observing and using these codes. Such semiotically complex linguistic phenomena do not bring the same indexical bundles into every interaction; indeed, their value for speakers is their ability to index different ideological configurations in different moments, thereby requiring that the speaker’s projected meaning be arrived at through negotiation among participants. Here ethnography can be put to use by showing empirically what speakers know, and how they know, about the ideologies underlying their bilingual practices.

The ideological ground on which bilingual interaction occurs is well illustrated by Katherine Chen’s examination of ideologies about code-switching in Hong Kong, which focuses on code choice not at the interactional level but rather at the indexical level, as a semiotic marker of social identity. Through a nuanced analysis of the distinctive language alternation patterns of two different groups of Hong Kong young adults, Chen shows that codeswitching is closely tied to ideologies regarding English and how it is most appropriately integrated into Hong Kong speech: those returnees who violate the local norm and use English more extensively are viewed by locals as pretentious, an indication of the relative valorization of English vis-à-vis Cantonese. Chen’s findings regarding the indexicality of language choice and its connection to social privilege and hierarchy resonate with Hall’s current work on English as an index of sexual modernity in India, where it is used to separate lesbian elites from less highly educated transgender ‘boys’, who are bilingual but prefer Hindi (Bucholtz and Hall under submission). It is certainly striking that, contrary to the usual expectation regarding linguistic accommo-
dation, in Chen’s data it is the more privileged group — returnees, who have had access to international schools and to overseas education — that must accommodate the less privileged group, the locals. This is not a strategy of condescension (Bourdieu 1991) but a strategy of self-protection in order to avoid peer sanctioning. It is likely, then, that this indexical process shapes code alternation within interaction, whether through metalinguistic commentary and explicit sanctioning of returnees’ code-switching practices by locals or through less overt forms of orientation to the markedness of extensive English use in the local context.

The theme of language ideology that runs through these articles gives them added depth by reminding readers that identity does not operate in isolation from other social dynamics. This focus on ideology also helps move the analysis of identity work from the interactional level, where it must begin, to the wider social world where it gains its meaning and force.

**Interaction and sociocultural linguistics**

When we articulated our understanding of the relationship between interaction and identity in our 2005 article, we characterized our approach as *sociocultural linguistics* (see also Bucholtz and Hall under submission). Part of our goal in advocating a richly sociocultural approach to linguistic phenomena is to promote an understanding of interactional moves as the building blocks of social life. To continue this architectural metaphor, although we are certainly interested in what these blocks are made of and how they are put together, the ultimate goal must always be to get to the edifice that they construct. In other words, interaction is a valuable starting point for the analysis of sociolinguistic practice, but it cannot be the end point. Interaction gives us the empirical grounding we need to move from the moment-to-moment unfolding of social activities within language to the larger social, cultural, political, and historical world that we seek to explicate as scholars of culture and society. So whereas in our 2005 paper we urged sociocultural linguists to think small, in the sense that we need to attend to the micro details of interactional and linguistic structure that serve as the semiotic resources for identity work, we also want to emphasize the importance of thinking big, in the sense that as analysts of the sociocultural realm of language use, we need to link individual linguistic forms to interactional moves and then to broader identities, social structures, cultural processes, and ideologies. This is only possible if we have a deep cultural understanding of local meanings and practices, which usually means that we need a significant ethnographic component in our work. Crucially, we need to give equal weight to each of these contexts, from the interactional to the
ethnographic to the larger sociocultural realm beyond the face-to-face, real-time, everyday activities that constitute our recorded data. In so doing, we need to resist the urge to approach our data from the top down, setting out to find interactional manifestations of social structures or power asymmetries or identity categories, or for that matter, theoretical concepts such as adequation, distinction, authentication, denaturalization, authorization, and illegitimation. The danger of such an approach is that we will force our data to fit into the structures, categories, and concepts available to us rather than discovering the ways in which they do not fit and thus developing new structures, categories, and concepts and refining or revising the existing ones.

What, then, is the value of ethnography, of identity, and of theory in sociocultural linguistics? To answer the first part of this question, we briefly review the sorts of functions that the papers in this issue propose regarding particular instances of codeswitching. For Cashman, code choice serves to separate and differentiate the self from a negatively evaluated other: a new kid, a bossy kid, a bully. Thus the kinds of people that individual kids are seen to be within the classroom is crucial information for the analyst, and this sort of information can only be gained through ethnography. Chen demonstrates another dimension of codeswitching: the indexical meaning of a particular pattern of language alternation. Her attention to the ethnographic categories of local and returnee allow her to discern the powerful social meanings of structural regularities, which might have been overlooked in an approach that failed to recognize and investigate the important differences between these two groups. Greer and Del Torto show in different ways that codeswitching to provide translation can be a way of repartitioning the participation framework to ensure the inclusion of (presumed) linguistic novices, and hence paradoxically to overcome linguistic difference by foregrounding linguistic difference. Ethnography complements conversation analysis by highlighting how locally established practices between Japanese and non-Japanese or between immigrant generations shape any given interaction between these categories. Conversely, Williams observes that a speaker’s use of the same code can give rise to different sorts of stances, even from one utterance to the next. Here ethnography can help us understand the complexity of speakers’ language ideologies, based on the need to reposition their ‘brought-along’ identities as they negotiate, over the course of everyday interactions, a new ethnic configuration in which they are now a linguistic minority within a linguistic minority. Finally, Bani-Shoraka reminds us that not only is codeswitching multifunctional, it may also at times be nonfunctional, with little interactional or social significance for the participants (although it may be extremely significant to the analyst interested in establishing patterns
of language use). She notes the danger of exoticizing codeswitching and bilingualism, cautioning against the analytic proclivity to find meaning in every change of code. Ethnography can demonstrate why such analyses are off base by moving beyond single interactions to show that the code choices of speakers at a particular moment are situated in relation to their habitual practice.

In sum, ethnography provides the temporal depth we need to be sure that what we as analysts see in the data is how speakers as cultural members understand it. It allows us to understand the categories to which speakers orient. It allows us to know if what someone does in a particular interaction is what they habitually do, or a radical departure, and it allows us to gauge such affective subtleties as whether people are joking or serious, angry or engaged, friendly or mocking. It also allows us to know who people are, who their friends and enemies are, what they care about, what they do every day, what they believe and think and feel. The sort of interaction we analyze as sociocultural linguists is often a single instantiation of broader social patterns, frozen in time; it is at best a slice of life, not life itself. Ethnography helps us to flesh out the picture.

An ethnographic approach also requires more of us as researchers than do other methodologies: it demands that we put ourselves into the picture and examine our own role in the research process. Whether the ethnographer initiates and directs the interaction (Williams) or simply contributes to it as one participant among others (Del Torto); whether she is a member of the community under study, as with Chen and Bani-Shoraka; whether, like Greer, the researcher unexpectedly becomes ‘the data’ (cf. Dumas 2005) or even becomes unintentionally complicit in larger structures of power and inequality, as Cashman courageously acknowledges — in every case, a complete and ethical analysis will consider how we ourselves are inextricable from the research we do. Ethnography, then, helps keep researchers honest as well as humble.

Turning to the second part of our question, what is identity good for, both for sociocultural linguists and for the people we study? That is, with what goal do we undertake the investigation of identity construction through language use? Undoubtedly every researcher will answer this question somewhat differently, but here are some suggested possibilities: (1) to describe an identity that has been unrecognized or misrecognized by researchers or cultural members, as we ourselves have done with nerds in the United States (e.g., Bucholtz 1999) and hijras in India (e.g., Hall and O’Donovan 1996), respectively, in our own work; (2) to demonstrate the importance of a particular interactional resource for identity work that has previously been overlooked, as in Greer’s discussion of the importance of eye gaze in determining recipiency and hence
the interactional goal of translations into Japanese or English; (3) to add
greater nuance to the conceptualization of identity and its construction,
for example, by detailing the instability of identity categories from mo-
ment to moment, as Williams’s work suggests, or the interactional speci-
ficity of these categories, as Bani-Shoraka’s and Cashman’s studies show
in different ways; or (4) to show how identities are tied up with larger
sociopolitical processes, institutions, histories, and ideologies that are
consequential beyond the interaction itself, as in both Del Torto’s and
Chen’s situated investigations of widely shared language ideologies.
Given this range of possibilities, and others as well, it is always worth
stepping back to reflect on why we are interested in examining identity
and how it helps us better understand our data.

Our final question, and one that all researchers of language, culture,
and society must often ask ourselves, is what, if anything, is gained by
using theory in empirical analysis. While the two of us have often been
inspired by theory ourselves, both within linguistics and in other fields,
we also want to sound a note of caution: our framework, like all concep-
tual models, is only as useful as the insights it yields. It was developed
as a way to incorporate under a single rubric a wide range of empirically
documented linguistic patterns that were often discussed separately; our
goal was less to set an agenda for future research than to summarize
and bring together the findings of diverse lines of research in order to
provide a foundation for investigations yet to come. As gratifying as it
is to see our work cited and applied to new kinds of data, we hope that
the concepts we have developed will be used primarily as tools or vehi-
cles that facilitate new sociocultural linguistic work, and will not be
viewed as the ultimate goal of such work. Just as we would not feel that
we had learned anything by identifying all the nouns and verbs in our
data, or all the first pair parts and second pair parts, likewise, as socio-
cultural linguists, we can learn very little about language, culture, and
society, by merely identifying identities or even the tactics of intersub-
jectivity that produce them. Yet the analytic process of doing so is a
crucial step in understanding how people use language to perform social
actions. Thus just as interaction is an important starting point for socio-
cultural linguistic analysis, theory is best appealed to later in the analytic
process, for it is our data, not our theories, that are our best guide to
how language is used to create local social worlds.

Conclusion
The articles in this issue clearly demonstrate that identity is an indispen-
sible concept for making sense of language as a fundamentally sociocul-
tural phenomenon, and that bilingual language use is an especially re-
revealing and complex situation in which to investigate identity. Bringing together the interactional, ethnographic, political, and historical contexts in which bilingual speakers negotiate their own and others’ identities, these authors offer a variety of ways to find identities in data, and to tie these findings to larger theoretical questions of ideology, power, and history. As sociocultural linguists continue to grapple with the question of identity and when, how, and why it matters, the analytic toolkit will be further expanded, providing new empirical bases for theory and new directions for the study of language as a resource for the construction of identity.

_University of California, Santa Barbara_  
_University of Colorado_

**References**


Bucholtz, Mary and Kira Hall (under submission). All of the above: From paradigms to coalitions in sociocultural linguistics. *Journal of Sociolinguistics*.


