

2001; Levinson 2003*b*; Lucy 1992*a*) has interpreted this intuition in a very literal manner (with the book being cognitive psychology), and in my experience, students' fascination often turns into bemusement or incredulity in the face of the peculiar things psychologists ask their experimental participants to do. I think that the approach outlined by Sidnell and Enfield, with its promise of a comparative science of human social action, has the potential to point future students interested in the idea of linguistic relativity in a direction that will allow them to realistically explore the consequences of language diversity.

There are many points of interest in this paper, and I want to focus on the following. I want to ask, What will be the main research question of the program of work sketched by Sidnell and Enfield? As researchers devote their energies to studying the linguistic relativity of social action, this line of work will likely become associated with one guiding question. For example, most neo-Whorfian articles present their big question as follows: Does language influence (affect, restructure) thought? What will the new question be for researchers of grammar-related diversity in social action?

Sidnell and Enfield formulate their question in different ways. They ask, "Does the fact that the words were spoken in English and not in another language have any bearing on how that action [completing a *howareyou* sequence] is accomplished?" They also suggest, "[T]he language you speak makes a difference for the social actions you can perform." These two ways of presenting the matter can be closely related, but they also have their distinct affordances. As an exercise in exploring the linguistic relativity of social action, let us consider the trajectories of thinking (and doing research) that can be projected from these two formulations.

The first question—Does grammar have a bearing on how an action is accomplished?—can suggest that social actions exist independently of the specific verbal and nonverbal acting that brings them about and that local practices of speaking might "affect" their realization. On this view, saying that social action is a "locus" of linguistic relativity can be taken to mean that social action is one of the "dependent variables" on which language exerts its "influence." Along these lines, the research program that Sidnell and Enfield outline could join the neo-Whorfian grooves of thinking, so that the question will become Does language influence social action? I have argued elsewhere that this way of asking the question (Does language influence thought?) gets our thinking about language, culture, and mind on the wrong track. It codifies the abstraction and decontextualization of "language," "thought," and "action" from situated real-time speaking, thinking, and acting, and it is accompanied by an entire glean of red herrings, such as the classic supposed problem of the incommensurability of concepts.

The second question—Does grammar have a bearing on what social actions speakers can perform?—can suggest that talk in a particular grammatical format enters into the real-time constitution of social actions. On this view, saying that

social action is a "locus" of linguistic relativity can be taken to mean that it is in situated (social) activity that talk in a particular grammatical format provides one of the material resources for the moment-by-moment accomplishment of cognition and action (Clark 2006; Goodwin 2000).

It is my impression that Sidnell and Enfield have sympathies for both of these ways in which research on the linguistic relativity of social action might play out. Maybe it is possible to keep both of these ways of thinking together. The line of work discussed by Sidnell and Enfield can open up a radically different way of thinking about the relationship of linguistic to cultural and cognitive diversity, if we manage to resist the temptation of thinking of social action as merely a new dependent variable and treat it instead as the primary site for the situated development of cultural practice and cognitive skills. On this view, situated social action is not so much a third as the primary locus of linguistic relativity.

---

## Reply

### Collateral Effects, Agency, and Systems of Language Use

We sincerely thank the commentators for carefully engaging with our work. We are encouraged by their support of this line of research, and we are especially grateful for their critical input toward clarifying our arguments.

We begin with "collateral effects." While in the article we focus on linguistic structures, both Duranti and Woolard wonder about the breadth of scope of the idea. We expect to see these effects in a broad range of domains and not only the linguistic or "verbal." Collateral effects are caused by dependencies among the multiple features of any structure that one may select for some function. Choosing a structure on the basis of some subset of its features does not mean you are exempt from "choosing" other features at the same time. When you choose what car to buy, your choice may be guided by specifications such as carrying capacity and fuel consumption, but whether or not you care what color it is, you cannot buy a car that is not some color. Suppose that while you truly do not care about the color, the only car available is canary yellow. A collateral effect of this purchase is that the color, something you did not select but rather "settled for," will attract comments in a way that dark blue would not have. Coming closer to language, Duranti asks whether we see collateral effects in "nonverbal" behavior such as gestures. Yes, we do. For instance, if you want to use your hand to depict the path of motion of something you saw, you cannot avoid showing the motion as occurring in a certain direction (expressible as a cardinal direction such as north-northwest), even when this is irrelevant to your expressive purpose (see

Enfield 2009a:17–18). Could such collateral effects of the manual-visuospatial medium have consequences for social action? As Duranti suggests, this is a matter for future research.

On Woolard's query regarding the distinction between our "third locus" of linguistic relativity and "Silversteinian indexicality," we refer the reader to the main article text, and especially footnote 4, where we address this directly. Woolard suggests that "agreeing while asserting authority sounds like the kind of social action accomplishable with classic T/V alternation." Let us clarify why these are different. By "action" here we mean "speech acts" such as requesting, inviting, offering, complaining, excusing, agreeing, and disagreeing. These can be done rudely or politely, with familiarity or distance, with either T or V pronoun forms. A T/V alternation cannot accomplish an action in and of itself, although of course some specific utterance containing a T or a V can accomplish the action of agreement (or K-plus agreement). Unlike the indexical meanings associated with T and V forms, the practices we discuss operate independently of the enduring social relations of the parties and are in principle usable by anyone, precisely because the K-plus versus K-minus distinction is always calibrated relative to some particular thing talked about.

In defining collateral effects, we presuppose that people have goals and that they select from among means to achieve those goals. This strikes us as uncontroversial, although Leavitt questions it. His worry cannot be that people do not have goals. Think of the millions of big, middle-sized, and tiny goals that fill up your life. You need milk and there is none in the fridge. Your goal: get some milk. To achieve it, you could drive to the store, or you could call your friend who is coming over and say, "Can you pick up some milk on the way?" If Leavitt means that the matter is not always so simple, then we agree. For one thing, goals can change as we go along. For another, a single piece of behavior can fulfill multiple goals. For yet another, culture does not just give us ways of meeting our goals; it also specifies the kinds of goals we should, and sometimes must, try to meet. Choosing specific solutions can introduce new, subordinate goals, and this is another example of collateral effects. Leavitt also seems to suggest that the determining of goals, and their selected solutions, can be distributed across individuals and through time in interaction. We agree that the matter of goals is nuanced in these ways, but this does not change our point. People seldom act without reasons or without purpose, and the means they select to meet their goals may introduce secondary, collateral effects.

If linguistic relativity effects are real, then different human groups have different realms of possibility. Since our domain of interest is action, this raises the issue of human "agency," as Duranti says. Our view is that collateral effects can be an agency-reducing force, taking "agency" roughly to mean the degree to which we can determine what we do and how we do it (Duranti 1990; N. J. Enfield, unpublished manuscript; Kockelman 2007). Once you have chosen a strategy for certain

reasons, then your free will is in a sense now used up, and you accept the collateral effects. PC or Mac? Rent or buy? Chinese or Italian? Whatever the reason for your decision, a higher-level choice will determine many other choices for you. Woolard also points to agency in suggesting that strategies for social actions "might be better located in speakers than in languages." But we are not forced to choose between people and practices. Yes, there are people's motives in specific circumstances, but then people have limited conventions to draw on and hence not unlimited agency.

A question raised by both Heritage and Leavitt is whether we expect the K+2A action to be a human universal. Our reason for thinking that it probably is universal is that we expect all human social groups to possess the basic ingredients for this action: an economy of information, normative organization of rights and duties, and individuals' motives to maintain textured sets of social relations through practices of (dis)affiliation. (For the ethnographic grounding of these assumptions, see Sidnell 2005, chap. 2.) Cultures differ widely, but there is a basic common infrastructure for social life that is characteristic of our species (Enfield and Levinson 2006a). We can expect that in all cultures, individuals will be motivated to form and maintain enduring personal relationships of different types and that there will be information-related practices such as agreeing with an evaluation someone else just made ("They're good kids," "She can't be trusted," "This is delicious"). Further, we can expect that in all cultures people are motivated to "police" rights and duties (Henrich and Boyd 2001), in line with local norms, including those associated with epistemic territory (Stivers, Mondada, and Steensig 2011). The nature and subtlety of such policing will vary, but it will always occur, wherever norms are contravened (perhaps most visibly in practices of socialization).

Managing epistemic rights and duties is surely done differently in different cultures, and as Silverstein suggests, the ethnographic context is important. But we do not share his pessimism about the possibility of understanding the actions exemplified in our data. While Silverstein feels unsure about "what is going on in the way of actual social action" in the examples, there is no reason to think that the participants in these interactions are any surer than he is. There is always a rich ethnographic background, but what is its role in online interpretation of social action? People's understandings of what is happening in the fast-moving enchronic context of social life are never consummate and are seldom more than adequate. Even in the coded meanings of words, people's understandings can differ without issue (Barr and Keysar 2005; Enfield 2012b). Interaction is made possible by a bounded form of rationality based on fast and frugal heuristics (Gigerenzer, Hertwig, and Pachur 2011). There is no time for elaborate ratiocination, especially as one needs to avoid the implications associated with delayed response in interaction (Stivers et al. 2009). So while ethnography is indispensable to a full vision of social life, it cannot be that an encyclopedic

knowledge of cultural context is invoked in an exhaustive online interpretation of every bit of conduct in interaction.

A second issue concerning ethnography is the need to distinguish between ideologies regarding territories of knowledge, which are sure to vary greatly across cultures, and actual, nonreflective practice regarding territories of knowledge. In discussing the proposed universality of the action type we are examining, Leavitt offers a possible counterexample. He has been told that in traditional “northern Athabaskan-speaking band societies, . . . the appropriate response to an assertion made by someone with less authority to make it than you have is silence.” If silence is indeed the proper way to react to a K-minus first assertion in those societies, then it should be recognizable as such. It would be readily describable in the terms we suggested: while we showed, for example, that a Creole solution to the K+2A problem is to treat the prior turn as if it had been a question, this “silence” solution would presumably treat the prior as if it had not been uttered at all. We do not know whether this is what is happening in the Athabaskan case, and indeed Leavitt implies that such sequences would not be found anyway, telling us that, as someone tells him, they “would have represented an unacceptable violation of personal autonomy.” For now, we can only exercise the standard cautions in assessing this claim. Is it a statement about what members of these societies say about their linguistic practices, or is it about what they actually do? At best, a known cultural ideology might lead to predictions about what happens in interaction. To find out, we would need access to a corpus of recorded interaction in these societies, and we would begin by examining the corpus for expressions of evaluation, especially the most mundane (e.g., “John’s new goats are pretty unruly”), and seeing how these evaluations are expressed and taken up in relation to different epistemic gradients (e.g., where “John” is related to the addressee, not the speaker).

Zinken is unhappy with our separation of action from language, but such a separation is necessary for the simple reason that social action is possible without language. Along lines argued by Lucy (1992*b* and elsewhere) for linguistic relativity in relation to thought, it can be methodologically useful to keep language and action apart so as to avoid tautology or not to simply describe the same thing in two ways. Lee (1996 and elsewhere) argued against Lucy’s methodological separation of language and thought (see also Hill and Mannheim 1992:382–385), proposing instead that everything is “langua-thought.” Maybe this is what Zinken is going for as well: everything is “langua-action.” But while language and action are always connected in practice, so are fuel and vehicles, guitars and music, or tools and carpentry. It does not mean that you cannot distinguish the two conceptually in order to study the phenomena. Social action is an important locus for linguistic relativity, but it need not be the primary locus of linguistic relativity (although perhaps there are arguments for its primacy from specific perspectives, e.g., ontogenesis). The phenomena we observe in social interaction would not take

place, and could not be understood, without all of these: social action, language, culture, cognition, and more. Our concern in writing the paper was to address a neglect of social action, not to replace this with an equally problematic neglect of some other equally necessary component of the story.

Finally, a recurring theme of the commentaries concerns the systems in which sets of semiotic practices are embedded. Duranti, Heritage, Michael, and Woolard all point out that any language will supply not one but many possible forms to choose from in formulating a social action and that individuals’ agency should therefore be greater than we have implied. While we agree that any language will provide a range of options, central to our point is that the choices in one language system are limited (see note 18). The scope of our article was to describe the most prominent or common or idiomatic ways of doing K-plus agreement in three language communities. For the Creole, for example, there are other ways (e.g., *oh*-prefacing), but these are less common and less central to the vernacular (see note 19). Michael says that a full understanding of what we are claiming “requires examination of the complete set of communicative resources that exhibit practical closure with respect to the K+2A communicative function.” We agree, and indeed, the situation is worse: we would need to see *all* things that can happen after *any* first-position assessment.

This raises the issue of what a “grammar for interaction” would look like, incorporating not only grammatical structures with their interactional uses but also interactional patterns of preference and structural conformity and actions defined not in terms of morphosyntactic structures but in terms of their distribution in conversational sequence (e.g., following a K-minus initial assessment). Michael notes that conversation analysis “does not usually emphasize ‘paradigmatic’ relationships among various interactional strategies.” It is true that this terminology has not been used, but in fact, the logic of selecting from among members of paradigm sets is well established in that literature. An example is the account of person reference in terms of default versus marked forms (Enfield 2012*a*; Schegloff 1996*b*; Stivers, Enfield, and Levinson 2007:9). Another is the analysis of the continuer *uh huh*, which is in structural and functional opposition to two other things that could happen in the same slot: initiating repair and taking a speaking turn (Schegloff 1982; Sidnell 2009*a*).

Critically, the set of choices in any language is a subset of the choices available in all the world’s languages together. Whorf urged us to study languages so as to broaden our worldviews. While nobody is “free to describe nature with absolute impartiality,” he argued, the person who would come closest “would be a linguist familiar with very many widely different linguistic systems” (Whorf 1956 [1940]:214). This is why we study human diversity: to know more about the ways we can think and about the kinds of social systems we can live in, and as we have argued here, to know more about the human potential for action.

—N. J. Enfield and Jack Sidnell