Coming to Grips With Variation in Sociocultural Interpretations: Methodological Considerations

Marina Terkourafi[1] iD

Abstract

Empirically capturing sociocultural interpretations—situated interpretations of linguistic expressions shared among members of a group—can be difficult for two reasons: First, the interpretations themselves cannot be directly observed and, second, the contexts that enable these interpretations cannot be defined independently of them. Yet, the reality of such interpretations attested in piece after piece of empirical research calls for an explanation. This article outlines a bottom-up methodology that seeks to extract context-sensitive definitions of, on one hand, sociocultural interpretations and, on the other hand, the context variables that covary with them, from the data itself. Uptake-based definitions of sociocultural interpretations are empirically verifiable and include speaker, context, and addressee contributions to the bringing about of a certain sociocultural interpretation. Dynamic definitions of macro-social variables (gender, age, class, ethnicity, region, etc.) can emerge by gradually abstracting over the minimal contexts that are found to enable particular sociocultural interpretations. The article outlines with examples how this methodology can be applied to spoken conversational data, as well as some of its limitations.

Introduction: Offering a Cup of Coffee as an Interactional Minefield

Let me start with a personal anecdote. Over a decade of living in the States as a (female) home-owner, I often welcomed to our home (male) workmen who came to fix or maintain various home appliances. As a Mediterranean hostess, I always offered these guests a cup of coffee or some water and a slice of homemade cake. Not once was this offer accepted. When I casually mentioned this during a wedding banquet we were attending as something I found surprising, another guest at the table, who was local and happened to be a psychologist, pointed out that they probably took these offers as a come-on. Shocked at how my hospitality had been misconstrued (and at my own naïveté), I immediately stopped this practice. Since moving to the Netherlands 2 years

[1]Leiden University, The Netherlands

Corresponding Author:
Marina Terkourafi, Leiden University, Van Wijkplaats 4, 2311 BX Leiden, The Netherlands.
Email: m.terkourafi@hum.leidenuniv.nl

ago, I have again welcomed several of these workmen to our home. One of the first things new-comers to the Netherlands learn is that guests are always greeted with a cup of coffee. Sure enough, this time my offers of coffee are accepted without exception—even if a workman's job is done within minutes. What is more, acceptance is immediate and minimal (usually a simple "yes, please!" before I have even finished asking the question), to the point that I feel the absence of such an offer would have been perceived as a serious omission. I have even come to know how these workmen typically take their coffee: black.

Readers from other parts of the world may be surprised to find out about these two diametri-cally opposed understandings of the "same" event (a come-on vs. a routine offer)—reflected in the format of the corresponding exchanges (refusal vs. minimal acceptance)—by two cultures, the American and the Dutch, that are often considered to be "sister cultures." Conversely, readers from other parts of these two countries may not even recognize themselves in these experiences: I can only speak for the two cities in which we have lived. It is precisely because I expect both kinds of reactions to be forthcoming that this example nicely illustrates the fact that sociocultural knowledge is both (contextually) situated and (culturally) local. Not only does such knowledge not travel well, it is also tied, within the local, to specific situations and the roles found in them. And it takes a native to tell—if they have an eye for observation and can articulate it free from self-presentational concerns—how a situation has been interpreted and (less likely) why.

While this can make navigating cultures in daily life a virtual interactional minefield, things are no simpler for the analyst studying these behaviors. In a nutshell, the problem is this: On the one hand, sociocultural interpretations of linguistic expressions are not a matter of linguistic form or of the speaker's intention alone; rather, they are discursively constructed between speaker and addressee and are "owned" by neither—nor can they be directly observed (except through their eventual consequences on subsequent behavior). On the other hand, the parameters that covary with these interpretations (the situations and the roles found in them) are also discur-sively—including sociohistorically—constructed: The roles of work(wo?)man and host(ess) and their rights and obligations are likely to be understood differently at different places and times, yet only real-time observation of how these roles are brought into existence, including through language, can grant us access to these variable understandings. Exactly because they are enabled at the nexus of (language/social) agency with (language/social) structure, both sociocultural interpretations and the situations in which they come about do not exist wholly independent of, or prior to, the specific interactions during which they emerge. How can the relationship between them be charted when both sides are equally inscrutable and mutually constitutive of each other? In what follows, I try to sketch a scientifically accountable approach to analyzing interaction in different cultures that does not lose sight of the cultural variability inherent in every step of this endeavor.

A Brief Sketch of the Proposed Approach

In the previous sentence, I referred to "interaction *in* different cultures." This is significant. Any discussion of communication *across* cultures starts from the assumption that different cultures are at play and that the phenomena talked about arise (only, or more acutely) when interactants do not share the same cultural background. Yet, oftentimes, little time is spent on finding out exactly what each interactant's cultural background entails. Rather, cultures are often treated as internally homogeneous and identified with national cultures, which are in turn identified with national languages. These are all questionable assumptions.

Research on linguistic variation over the past 50 years has shifted from external/objective assignment of speakers to predetermined speech communities based on macro-social variables such as location, ethnicity, and social class (e.g., Labov, 1972; Trudgill, 1974), to practice-based models of sociolinguistic identity, where the degree of membership of a language user in a

Community of Practice depends on the degree to which he or she partakes of its goals, repertoire, and practices (Lave & Wenger, 1991). Such models allow more space for the speaker's agency and subjective self-determination. It follows that the ensuing linguistic Communities of Practice are brought into existence through the observed practices of their members and can only be identified through observation of these linguistic practices in conjunction with nonlinguistic patterns of behavior (Eckert, 2000).

In line with the latter tradition, the view taken in this article is that "cultures" are best understood as Communities of Practice. Cultures do not preexist interaction but rather are actively performed during interaction and only identifiable in its course. Moreover, one and the same language user can belong to more than one culture and perform different cultural identities at different times in response to different "environmental" conditions. The approach developed in this article, therefore, places itself one step before any investigation of cross-cultural communication proper. To analyze and understand what goes on during cross-cultural communication, we first need to pin down the (different) respective starting points of the interactants. We need empirical investigations of intracultural communication that can reveal the linguistic practices of the communities they are deemed to belong to, where "practice" refers to goal-oriented linguistic behavior endowed with certain (social) meaning in specifiable contextual conditions. Two notions will emerge as central to this empirical endeavor: the notion of "uptake" from linguistic pragmatics (Austin, 1962/1975; H. H. Clark, 1996), which can help delimit socioculturally significant behaviors, and that of a "minimal context" (Terkourafi, 2005, 2009), which can serve to refer to the surrounding situational context as a conceptual prime or gestalt (Lakoff, 1977). To repeat, the approach described in this article is *not* directly applicable to the analysis of cross-cultural communication. It is, however, a necessary prerequisite for it.

In brief, the methodology described in this article consists of the following steps. It all begins with audio/video recording and transcribing large amounts of spontaneously produced conversations in a variety of settings: at home, at work, in shops, at public services, during volunteer activities, or at the playground. Without excluding anyone, the idea is that the majority of people will be brought to these places by an interest in the activities that take place there and will therefore constitute, during and through that engagement, a Community of Practice characterized by its own (linguistic) patterns of interaction, which the investigation aims to uncover. Having determined the linguistic phenomenon of interest (e.g., a type of speech act), the researcher identifies all its occurrences in the transcribed data by looking for evidence of understanding in the next speaker's turn that such and such type of speech act has occurred. They then look back to the previous speaker's turn that generated this understanding in context and note aspects of its linguistic form (lexical items, sentence type, and other grammatical features). In addition, the researcher annotates the previous speaker's turn with extralinguistic information about the speaker and the addressee (age, class, gender, ethnicity, role, and other categories salient in the interaction) as well as the setting of the interaction (physical surroundings and time of day) and the sequential placement of the utterance in the flow of the conversation (first occurrence or subsequent repetition of the speech act). The next step is to quantify over the linguistic form of utterances and their extralinguistic contexts of use in parallel. What we are trying to identify is eventual preferences for particular forms realizing the same speech act (e.g., linguistic constructions of the form "I want X" or "you will do Y" realizing requests) in relation to specific combinations of the extralinguistic features noted during the annotation of the data.

It is not expected that such preferences will be found over the data set as a whole. Rather, one linguistic realization is expected to prevail in one combination of extralinguistic features and another in another. It is through repeated attempts to aggregate and group the data that the researcher tries to identify those combinations of extralinguistic features that go hand in hand with specific linguistic realizations of the phenomenon investigated (and vice versa) and that should therefore be considered "minimal contexts" stored in memory alongside the linguistic

expressions most frequently generating specific sociocultural interpretations therein. Only some of the combinations of extralinguistic variables recorded are expected to constitute culturally recognizable scenes or "minimal contexts" in this strict sense. These will involve frequently repeated, context-bound activities, whose frequency justifies having a "ready-made" linguistic solution for dealing with them. The end-goal of this exercise is to uncover qualitative differences in how different groups deal with frequently repeated activities. For it is only by gaining a sound understanding of insiders' shared knowledge that we can begin to account for what happens when insiders become cultural outsiders as they encounter each other in cross-cultural communication.

This is still a tall order and to avoid appearing to promise more than I can deliver, two limitations of this approach should be pointed out from the outset. First, while undoubtedly a desideratum, etic generalization may not always follow from such an emically grounded approach. Whether generalization can follow—whether, that is, the contextual parameters identified as relevant for one data set can be used to analyze another—is an empirical matter and depends on the data itself. In other words, while not excluding this possibility, this approach does not aspire to yield a universal inventory of variables that can be considered equivalent across cultures and directly transplanted between them. This relates back to the well-known problem of isomorphism, which is all too often assumed but rarely empirically established (cf. Fischer, 2011). Considering that cultures are not designedly commensurable, this limitation, inconvenient as it may be, is not necessarily a bad thing.

Second, only some types of data are amenable to this type of analysis. Specifically, data must include evidence of both interactional partners' behavior (i.e., speaker and hearer, for verbally realized behavior) as well as ethnographic information about the context itself (who is talking to whom, what is the occasion, place, time of day, etc.). This requirement follows from another observation by Fischer (2011), namely that "situations are not that easily decoded and meaning needs to be learned and abstracted in rich contextual settings" (p. 11). In other words, what we need are "thick descriptions" of communicative events (Geertz, 1973). Large amounts of transcribed face-to-face conversational data collected in naturalistic settings are necessary for this type of analysis. These can be found in spoken corpora,[1] while multimodal corpora are also becoming increasingly available (e.g., the Nottingham Multimodal Corpus; Knight, Adolphs, Tennent, & Carter, 2008). And while written texts, especially those written for a larger audience in which reader responses may be lacking or displaced, are less suitable for this type of analysis, data sets from online interactions (e.g., an article, YouTube video, or blog entry followed by reader comments) can also meet this desideratum, when the information available to the analyst is the same as that which is available to the participants themselves (i.e., information about the age and gender of language users may not be available to the analyst but then this is also the case for the participants themselves).

Variation in Sociocultural Interpretations: What Are We Talking About?

In my example of offering a cup of coffee to workmen in the American Midwest versus the Randstand,[2] what varied was not the language used nor its primary interpretation. In both cases, the same question ("Would you like some coffee?")[3] was interpreted as a speech act of offering a cup of coffee. It was the perceived appropriateness of this offer that varied between the two groups, creating the affective undertones described and eliciting two different sets of reactions—rejection versus immediate acceptance. In other words, what differed was the subtext—technically, the implicated (pragmatic) meanings of uses of this question on different occasions.

Such unspoken interpretations are typically thought to be fleeting, situated, and, first and foremost, subjective. To make matters more complicated, as they are unspoken, they are not open

to observation or, indeed, quantification. When Grice (1975/1989) proposed his notion of conversational implicature to capture precisely such type of context-dependent meanings, he attributed them to the speaker's intention vis-à-vis a particular addressee. As intentions cannot be directly observed and may not even be accurately reportable or accessible through introspection—as Fischer (2011) notes, "Do we always know why we did things? Probably not, but we are exceptionally good at rationalizing about plausible reasons afterwards" (p. 15)—this would seem to place conversational implicatures beyond the possibility of generalization altogether.

At the same time, however, Grice (1975/1989) did allow for implicated meanings that arise "normally . . . (in the absence of special circumstances)" (p. 37). He called such meanings *generalized conversational implicatures* and attributed their existence to "the use of a certain form of words" (Grice, 1975/1989, p. 37) rather than the speaker's intention. Developing this notion further, Levinson proposed that generalized conversational implicatures belong to a third level of utterance-*type* meanings, intermediate between semantics (what is encoded) and pragmatics (what is inferred), which he defined as "a level of systematic inference based not on direct computations about speaker-intentions, but rather on general expectations about how language is normally used" (Levinson, 1995, p. 93). What is important about this level of meaning is that it does not depend on specific information about who the speaker and the hearer are and what is the context of utterance, but rather on the type of language used. It is at this level that pragmatic meanings can be expected to vary systematically along macro-social dimensions (gender, class, ethnicity, region, etc.) and to show the kind of "orderly heterogeneity" (Weinreich, Labov, & Herzog, 1968, p. 100) that can bestow social significance upon linguistic forms.

The notion of generalized conversational implicatures is especially apt for accounting for variation in sociocultural interpretations for two reasons. First, it raises the possibility that regularity of form may be accompanied by regularity of meaning. Even if such meanings remain, in principle, defeasible (they can be denied or canceled), they do not require access to inscrutable speaker's intentions but arise in virtue of the speaker's choice of words or "putting it that way" (Grice, 1975/1989, p. 39). This gives us a first handle on meanings that cannot otherwise be directly observed. Our second handle comes from their presumptive nature: generalized conversational implicatures occur "normally" (in the absence of special circumstances). That is, rather than waiting to be enabled by one-off contexts, they are assumed *all else being equal*. This second feature amounts to a quantitative claim that can be empirically verified: A generalized conversational implicature should, in principle, be present when a certain form of words is encountered unless there are reasons to think that it is not. In other words, their occurrence should be more frequent than their nonoccurrence, and this last one should be somehow (interactionally) marked.

The frequent realization of a certain meaning through a certain form of words can be further linked to appropriateness—although the direction of this influence remains a separate matter. In the coffee-offering example above, it is the frequency of these offers in a Dutch context that is arguably constitutive of their appropriateness (for participants at the micro level) and vice versa in an American one: in a setting where such offers are infrequent, there can be more uncertainty about how to evaluate them. The claim made here, then, is twofold: both speech acts (offers, requests, complaints, apologies, etc.) and the linguistic forms[4] by which they are realized are monitored for frequency. Positive evaluations emerge when the speech act selected is frequent in the context at hand and additionally so is the linguistic means (the words) by which it is realized; this is the most unmarked case, what is assumed to happen in native-to-native communication. Uncertainty is introduced when either the speech act selected is *not* frequent in the context at hand (which ipso facto means there can be no frequent means for performing it in that context) or the speech act itself *is* frequent but the linguistic means selected to perform it is not. In the former case, we can expect evaluation to be negatively affected; in the latter, it is interpretation itself that can be affected.

The opening example in this essay is a case of the former: Offers to workmen on home visits being infrequent in a U.S. Midwestern context, my words were correctly interpreted as an offer (drawing on the fact that these same words can perform offers in other contexts) but negatively evaluated in the current context of use. It could have been otherwise: The first time that I encountered the phrase "How is life treating you?" my lack of familiarity with this expression led me to misinterpret it as a genuine question, although it occurred in a context where greetings are frequent and it was meant as no more than that. This is then a case of the latter: a situation in which the speech act itself *was* frequent in the context at hand but the linguistic means selected to perform it was not frequent (to the recipient's—that is, my—experience), with the result that interpretation itself was affected. In sum, as Lavandera (1982) insightfully pointed out,

> different social groups or different social situations find certain communicative modes (ways of speaking, communicative intentions) more appropriate than others. The marked preference for a certain communicative mode explains the higher frequency of these forms corresponding to the signifiers behind this particular mode. (p. 94, my translation)[5]

Culture as Enactment: An Uptake-Based Account

While the frequent presence of a certain form of words may increase the likelihood of a certain (sociocultural) interpretation and its concomitant positive evaluation, claims of an interpretation must be empirically verifiable. When it comes to linguistic behavior, this can be achieved with recourse to the hearer's uptake. The notion of the hearer's uptake was proposed by Austin (1962/1975) to refer to the understanding of the speaker's utterance as a certain kind of act with a certain kind of (referential) meaning by the addressee.[6] As Austin (1962/1975) famously pointed out,

> Unless a certain effect is achieved, the illocutionary act will not have been happily, successfully performed . . . I cannot be said to have warned an audience unless it hears what I say and takes what I say in a certain sense . . . the performance of an illocutionary act involves the securing of uptake. (pp. 116-117)

By highlighting the securing of uptake as an integral part of the performance of a speech act, Austin effectively made the hearer's uptake constitutive of the speech act itself. That does not mean, however, that uptake can only be secured through recognition of the speaker's intention, as theorists such as Strawson and Searle subsequently proposed. Austin saw speech acts as irreducibly social acts and the centrality of the hearer's uptake to his account reflects that. Unpacking his views in this regard, Sbisà (2009) convincingly argues that shared norms and conventions are crucial to the securing of uptake. On this view, the successful performance of a speech act can be a matter of social coordination relying on shared norms and can proceed automatically, so long as expectations are met. Given this link to conventionally expected outcomes, there is reason to expect that those constituting a Community of Practice will interpret speech acts in a similar way; indeed, similarity of uptake can provide behavioral evidence of their shared group membership.

The notion of uptake thus emerges as central to an understanding of culture as enactment. This is an important step forward for cross-cultural research that has, for a long time, favored national cultures as an easily tractable way of dealing with the variable of culture (Fischer, 2011). This choice is becoming increasingly inappropriate in a world where transnational movement (of oneself or of others) and new technological affordances—notably, the possibility of building and maintaining ties with a large and diverse group of people over social media—make daily life multicultural and multilingual, even for those who are not themselves geographically mobile. As national cultures are called into question by these developments and redefined on new grounds

(Pew Research Center, 2017), variables other than country of origin, such as generational cohort, political affiliation, or professional expertise, are gaining momentum in generating shared understandings and like-mindedness among people. The importance of these other factors explains at least in part why it is difficult to extrapolate from group norms to individual behaviors as "any results found [by aggregating data at the nation-level] cannot be applied to individuals living within these nations" (Fischer, 2011, p. 5). An enactment view of culture provides an answer to this problem. In the practice-based understanding of culture advocated here, group-belonging is not presumed based on external attributes (e.g., nationality) but rather built from the bottom up, through specific behaviors *and* their having been interpreted in particular ways. This paves the way for a recurrent type of uptake to be used as evidence that a certain sociocultural interpretation of a prior turn is present, *without* building this interpretation into the encoded meaning (the semantics) of this prior turn.

All this makes uptake a powerful tool for methodologically implementing culture as enactment. Yet uptake, consisting in the hearer's understanding of the speaker's utterance, is a mental act and cannot be directly observed. Are we back to square one? Not exactly. Although uptake itself is impossible to observe directly, further perlocutionary effects of the speaker's utterance *are* open to observation. As some of these can occur only if the speaker's utterance has been understood in a certain way, they can be used to identify indirectly the kind of uptake that has occurred. Take Austin's example above: What the hearer who has been warned decides to do with that warning—whether they heed it, ignore it, or end up being scared because of it—makes no difference to this understanding per se. The securing of uptake is distinct from the occurrence of these perlocutionary effects. Yet, they all presuppose first understanding the speaker's utterance as a warning and that is why they provide useful pointers that this understanding has occurred.

In the field of Conversation Analysis, this process is called "validation through 'next turn'" and it is considered fundamental to ensuring the validity of the researcher's analytic claims (Peräkylä, 2004, p. 291). In short, what validation through next turn provides is evidence that the understandings claimed by the analyst are the participants' own. As H. H. Clark (1996) notes, "second parts of adjacency pairs serve both functions—uptake and evidence of understanding—and [that is] why they are expected to be adjacent" (p. 200). Whether one puts it down to rationality and Gricean cooperation (Grice, 1975/1989), or conditional relevance (Schegloff, 1968), what happens after a speaker's utterance is relevant to understanding what came before.

Identifying Uptake Displays in Conversational Data

An uptake-based definition of utterance-type meanings holds that what is responded to as a certain kind of act (say, a request) counts as that kind of act. In other words, in a process that uses displays of hearer uptake as heuristic devices for identifying recurring sociocultural interpretations, we start with the uptake. Only after having determined that a certain kind of understanding (for instance, as a request) has taken place, do we then look *back* at the prior speaker's turn to find the particular combination of linguistic form and extralinguistic context that enabled this understanding.

Note that this is different from a standard corpus-based (or, possibly these days, Big Data) approach, which takes the expression used by the speaker (rather than its interpretation by the listener) as the starting point for analysis. In the approach advocated here, it is the listener's interpretation, as evidenced in his or her verbal or nonverbal reaction, which provides the starting point, based on which the expression that preceded it is identified as a realization of the phenomenon investigated (e.g., requests and how they are linguistically cast in different contexts). This means that only positive understandings will be identified. Utterances where the ensuing discourse fails to provide evidence (verbal, nonverbal, or implicit) of positive understanding will not be identified for analysis during this process. By repeating this process over large data sets,

our goal is to identify form/context combinations that regularly enable a particular interpretation on different occasions. This methodology thus acknowledges the potential multifunctionality of all utterances (Pichler, 2010, p. 597)—that is, it does not assume that all occurrences of "Would you like to X?"[7] are offers but only those that were responded to as offers in our corpus—while also affording us some insight into the regularities of in situ interpretation by a specific community or group.

Uptake or validation through next turn can take many forms. One's conversational contribution can be assented to explicitly (verbally and nonverbally) or implicitly, checked, challenged, rejected, misinterpreted, or plain ignored. The exchanges below provide some examples.[8]

(1)   Explicit assent – nonverbal
(21:12; A is a volunteer collecting money for disabled children during an outdoor campaign organized by a major radio station; B is the campaign treasurer)
    A: Can I give you? Can I give you ((the money)) because I'm full?
→ B: ((receives money)).

(2)   Explicit assent – verbal, immediate
(21:37; C and D are volunteers collecting money for disabled children during an outdoor campaign organized by a major radio station)
    C: We need one ((more person)).
→ D: Yes. She's coming now.

(3)   Explicit assent – verbal, delayed
(24:3; E and F are young Constantino's grandparents, the scene is taking place in their living room, which E is entering as she speaks)
    E: Have you seen Constantino's Santa Claus, grandpa?
    ((no uptake; E is now in the room))
    E: Grandpa have you seen Santa- Constantino's Santa Claus?
→ F: Oh my goodness!

(4)   Implicit assent
(21:1; G is a volunteer collecting money for disabled children during an outdoor campaign organized by a major radio station; B is the campaign treasurer)
    B: I want twenty-two pounds.
→ G: Do you want it in cash? (.) I can cut you a check.

(5)   Checking
(22:6; H is a customer at a pharmacy, I is the service provider)
    H: I want a (librax)
→ I: You want a librax?

(6)   Challenging
(21:35; J and K are volunteers collecting money for disabled children during an outdoor campaign organized by a major radio station)
    J: Listen, you'll lea- your bag you'll leave it here
→ K: And where'll I put ((the money))?
    J: You'll keep it in your hand.

(7)   Rejecting
(21:10; M is a volunteer collecting money for disabled children during an outdoor campaign organized by a major radio station; B is the campaign treasurer)
    B: Rebecca, do you want me to go do it? ((referring to bank deposit))
→ M: No let me go the first time.

As these examples show, (1) to (6) were understood as requests, (7) as an offer. To receive the money from A in (1), B must first understand A's question "Can I give you ((the money))" as a request to take the money—as that is not what A's words (a question about ability) literally mean. Through the nonverbal act of receiving the money, then, B displays his understanding of A's utterance as a request at the same time as assenting to that request. Not all requests can be nonverbally assented to, however. In (2), D's reply to C's assertion "We need one ((more person))," which simply states the need for something to happen, indicates that D has understood C's assertion as a request and proceeds to outline how it will be met. Note the two-part structure of D's reply: It is the opening "Yes." in D's reply—which is structurally unnecessary, as C's utterance is not a yes/no question—that explicitly signposts her understanding of C's assertion as a request. In (3), on the contrary, E's question, "Have you seen Constantino's Santa Claus, grandpa?" is initially met with silence by the intended addressee (grandpa), and it is not until it is repeated a second time that grandpa finally reacts with "oh my goodness"—indicating both that he has interpreted E's question as a request to look and that he is assenting to it by looking and taking delight in what he sees (his toddler grandson dressed up as Santa). Example (4) is especially interesting, in that this time assent is only implicitly provided by moving on to the next point in the conversation. G's question-cum-suggestion "Do you want it in cash? (.) I can cut you a check." suggests different ways of providing B with £22, indicating that he has interpreted B's statement "I want twenty-two pounds" as a request. This type of exchange is frequent in planning common activities and highlights the fact that uptake displays need not be verbalized: not only can such displays be nonverbally performed, as in (1), they can even be skipped in an application of Zipfian economy (Zipf, 1949), whereby "minimal forms [in this case, absence of explicit uptake] warrant maximal [in this case, preferred] interpretations" (Levinson, 2000, p. 35). In such cases, the absence of explicit uptake (whether verbal or nonverbal) can indicate that an act has been performed in the most run-of-the-mill, expected way. If so, implicit uptake displays, where the action simply moves on to the next step in the shared activity, should be especially common when an interpretation is shared among a group, that is, with precisely the type of sociocultural interpretations we are trying to identify here. Based on this reasoning, if we find that the expression "I want X" (from B's preceding turn) is interpreted as a request in informal situations frequently enough in our corpus, that would support identifying this expression as a form for performing requests appropriately (i.e., evaluated positively) in an informal situation in Cypriot Greek (as indeed it was found to be).

In Examples (5) and (6), the speaker's utterance is not assented to; in (5), the addressee asks for clarification by echoing H's statement ("You want a librax?"), while in (6) the addressee identifies a potential problem with the speaker's request ("And where'll I put ((the money))?"). Nevertheless, it is precisely by identifying obstacles to compliance that both of these addressee responses indicate that the preceding statements were understood as requests to begin with. As the understanding of the speaker's utterance as a request is what we are interested in, rather than whether the request is granted, again, this understanding is enough to identify the speaker's utterance as a request in this (informal) context. And if it turns out that request understandings of the expressions used ("I want X" in [5] and "you'll do Y" in [6]) are frequent in informal situations in our corpus, this would suggest that these are appropriate (positively evaluated) realizations of requests in this context. Finally, in (7), the addressee again does not comply with the speaker's utterance. This time, however, it is the second part of her utterance "let me go the first time" that construes B's yes/no question, "do you want me to go do it?" as an offer by providing a reason for turning it down. Examples (1) to (7), then, all provide evidence that a speaker's conversational contribution has been understood in a certain way (as requests in the cases of [1]-[6] and as an offer in the case of [7]). Moreover, both explicit and implicit displays of uptake are relevant.

But what about misinterpreting and ignoring? What, if anything, can we tell about the hearer's understanding of the speaker's utterance (his or her uptake), if the next turn provides no evidence

of how the speaker's utterance has been understood or, on the contrary, indicates that it has been misunderstood? Several things might be going on here—all having specific methodological consequences for the type of data required and how to deal with them. To begin with, violations of adjacency can indicate a need for further information (Example [6]) or momentary shifts of attention to attend to something else in the environment (Example [3]). If not immediately displayed, evidence of uptake may be provided later on in the conversation. Although not strictly "next turn," such displays can still furnish evidence that a certain utterance has been understood in a certain way. Nonetheless, delayed uptake displays should be treated with caution for two reasons. First, if it is no longer possible to relate this understanding to a particular prior turn and the "form of words" used, they can be of little use in identifying sociocultural interpretations, understood here as canonical (group-level) interpretations of particular expressions (see the section "A Brief Sketch of the Proposed Approach"). Second, it is part of the definition of utterance-type meaning that it is identified automatically (presumed all else being equal). If an interpretation appears to be a matter of negotiation over several turns, then we are probably not dealing with utterance-type meaning at all, but rather with more particularized, one-off types of meaning.

Rather than being delayed, evidence of uptake may be completely missing. We may, in such cases, be dealing with purposeful breaches of Relation,[9] whereby failure to engage with the speaker's utterance indicates that something else is going on (possibly at the interpersonal level). Such breaches of Relation can, for instance, be face-saving (when a topic shift indicates, without saying so, that the previous speaker's turn was somehow inappropriate) or face-threatening (as when someone's contribution, although not inappropriate, is blatantly ignored); the presence of third parties can be decisive in such cases. Given such instances provide no evidence for how (or whether) the speaker's utterance has been understood, they are of little use in identifying sociocultural (shared) interpretations of linguistic expressions.

Finally, the speaker's utterance may be misunderstood. If no evidence of a misunderstanding becomes apparent in the rest of the conversation (and here, we are unavoidably limited by the realities of data collection, which is bounded in time and place), then there are no grounds to claim that misunderstanding has occurred[10] as this would imply prioritizing the analyst's interpretation over that of the participants' themselves.[11] If, on the contrary, evidence of a misunderstanding does become apparent in subsequent behavior (verbal or nonverbal) by the original speaker or by someone else, subsequent discourse may (but need not, as self-presentational concerns can again intervene) clarify the original speaker's intention. Nevertheless, this means that more work was needed by both participants to get the point across. Communication, in such cases, is likely to rely on explicit reasoning about the speaker's intentions rather than shared linguistic norms. Similar instances can be expected to occur when interlocutors do not share a set of linguistic norms or some other type of renegotiation of rights and obligations is taking place. While interesting in their own right, such cases can have little to tell us about the norms that interlocutors do share. This brings to the fore an important point: An uptake-based identification of utterance-type meaning is only relevant when (there is reason to think that) the speaker and the hearer share some linguistic norms. As we are tracking both parties' behavior over large amounts of data and using this process to tap into their shared norms, if they share no norms and misinterpretation is the result of a difference in linguistic conventions (as in the case of L2 learners, or of lexical error), then an uptake-based process is unlikely to lead to the identification of shared norms because none are present.

This result is desirable for a couple of reasons. First, it safeguards the validity of the analysis by providing evidence (which may well be implicit, as in Example [4]) that a certain understanding has taken place. It thus serves to ground the analyst's claims and avoids mentalizing interpretations and attributions of intentions that cannot be independently checked. Second, it serves to focus attention on interpretations that are likely to be shared among a group. Although the analysis of cases of misunderstanding or lack of understanding can enrich our interpretive accounts or

serve to expand them, understanding in such cases is likely to be actively achieved based on the speaker's intention and not with reference to shared norms. Moreover, among users who share similar norms, misinterpretation should be less frequent and hence not regular across a large data set. If in fact misinterpretation appears to be frequent in such a data set (as evidenced by uptake displays signaling misunderstanding), we may be dealing with a shift in norms (language change).

Having determined, via tracking the uptake, all the utterances in our data set that achieved certain illocutionary goals (were understood as, say, requests), the next step is to identify the linguistic expressions that were used to realize these requests. We saw a variety of such expressions in Excerpts (1) to (6): "Can I VP?" "We need NP," "Have you VP-ed?" "I want NP," and "You will VP."[12] How often and, crucially, in which contexts is each expression used to realize requests? Note that it is unlikely that one expression will prevail across all contexts; if that is the case, we are likely dealing with a "convention of the language" (Morgan, 1978), such as can often develop grammatical reflexes of their use (cf. English *gonna* as a marker of futurity rather than literal motion). More likely, each of these request-realizing expressions will be preferred in one subset of situations (informal interaction at home, service encounters in small shops, service encounters in big supermarkets, working-group meetings, live interviews, etc.), what were referred to in the section "A Brief Sketch of the Proposed Approach" as "minimal contexts" (see also the next section). It is this frequency-relative-to-a-context that makes these expressions "conventions of usage" (again, in Morgan's 1978 terms) yet not interchangeable without affecting their appropriateness (i.e., positive evaluation) relative to the context in which they actually occurred.

The identification of linguistic expressions based on their lexicogrammar as described above means that what we are focusing on during this process are conventions of form. It is, however, possible that this approach can also be applied to identify conventions of content. A convention of content would be a preference for a certain sequential format (e.g., greeting-greeting or invitation-denial-repeat invitation-acceptance), which may not involve the same lexicogrammar each time but nevertheless instantiates a stable sequence of conversational steps. While this extension has not been empirically implemented yet, there is nothing preventing the discovery of such sequences being frequent-relative-to-a-context in the manner described above, so long as via uptake it is established that the utterances involved are, indeed, understood in a consistent way (as greetings, or as invitations, denials of invitations, etc.), despite the fact that the actual words realizing them each time are different.

This type of extension would bring the proposed approach in close dialogue with the Conversation Analytic (CA) methodology, with which it already shares a lot, as it builds on the notion of the uptake as an observable behavioral outcome of sociocultural interpretations that are otherwise not open to scrutiny. However, CA usually stops at identifying the uptake on a per case basis, without being interested in generalizations about the lexicogrammar of the linguistic expressions (what we called "conventions of form") that are used to achieve these uptakes. On the contrary, the approach outlined here uses uptake as a heuristic means of finding out what interpretation has taken place, while its ultimate goal is to find out what linguistic expressions regularly achieve a certain type of uptake in different contexts. It is in this awareness that different linguistic expressions may achieve the same uptake in different contexts, and in its quest to discover these context-bound regularities in linguistic expressions, which are connected to perceptions of the appropriateness of each expression relative to its actual context of occurrence, that the proposed approach sets itself a different set of goals from those of CA—indeed, goals that make it an appropriate tool for cross-cultural research.

In closing, it should be noted that the uptake-based account outlined here imposes some important requirements both on suitable data and types of analysis. Specifically, it requires a *large* body of *conversational* data, with *video-recorded* data being preferable. This ensures that both the speaker and the addressee contributions can be readily observed, and that nonverbal

features, including the phrasal stress, facial gestures, and body orientation of the speaker available to the hearer and playing an important (if only confirmatory, in the case of generalized conversational implicatures) role in their interpretation of the speaker's utterance will be equally available to the analyst. Furthermore, as uptake displays can take several forms, correct identification of uptake by the analyst(s) is paramount. This can be ensured through an *emic understanding of the cultural settings*, or via gauging *interannotator agreement* (or, preferably, both). Finally, a large amount of such data is crucial, if we are to identify trends over many encounters and pairs of users, to substantiate the quantificational claims relating to appropriateness that are central to this kind of approach.

Situating Culture: Minimal Contexts as Conceptual Primes

In the previous section, I highlighted generalized conversational implicatures as a promising methodological tool for capturing sociocultural interpretations, that is, interpretations of linguistic expressions that are shared among members of a group. A defining feature of generalized conversational implicatures, as I noted there, is that they rely "on general expectations about how language is normally used" (Levinson, 1995, p. 93). This insight is especially relevant for sociocultural interpretations, which similarly rely on expectations shared among a group, and therefore one we should try to keep. However, precisely in the case of sociocultural interpretations, what is "normal" can vary among groups. Can the notion of generalized conversational implicatures still be useful for analyzing these interpretations, and how? I suggest that it can, provided that what constitutes normal circumstances for a group can be defined. This is where the notion of "minimal context" (Terkourafi, 2005, 2009) comes in.

Minimal contexts encapsulate language users' expectations about who is talking to whom, when, and where. This means abstracting away from the actual one-off encounters in which language is used to schematic representations of contexts that covary with the linguistic expressions used therein. Such schematic representations combine pertinent information about participants (their gender, age, class, ethnicity, etc.), the relationship between them, the setting of the exchange, and the stage in the conversation where a particular utterance occurs. Ethnographic interviews suggest that, compared with more abstract dimensions such as power and distance, these dimensions of context are intuitively graspable and correspond to participants' own internalized categories. Moreover, the literature on language acquisition shows that by age 4 children are paying attention to the categories of gender, age, and profession (the latter being one determinant of social class), talking differently to mothers versus fathers, adults or peers versus younger addressees, and nurses versus doctors (reported in E. Clark, 2004). Concepts of ethnicity and race are also cognitively acquired starting around the same age (Quintana, 1998). The fact that these parameters are acquired early and recur in participants' explanations of what is going on makes this an emically oriented approach that satisfies the ethnomethodological injunction to take "as a starting point of departure for the analysis of context the perspective of the participant(s) whose behaviour is being analyzed" (Goodwin & Duranti, 1992, p. 4).

This does not, however, mean that these contextual parameters are directly perceived or that they are the same for all. In actual fact, macro-social variables such as gender, age, race, and class are but analytical abstractions that do not exist independently of each other: In the physical world, there are no such things as walking and talking genders, ages, races, or social classes. All that exists are flesh-and-blood speakers and listeners who embody all of these simultaneously. In other words, macro-social variables can never be perceptually apprehended directly or, in isolation, separated from each other. This observation suggests a possible explanation for why macro-social variables can be understood differently by different groups, as well as a way of capturing this methodologically.

An important part of this endeavor is its holistic character: Linguistic expressions are associated not with particular characteristics of speakers or settings directly but rather with minimal contexts as a whole. Much like Wittgenstein's (1953/1976) well-known duck–rabbit illusion, coinstantiation within a whole constrains understanding of the parts. Barsalou (1999) outlines the cognitive underpinnings of this process, whereby symbolic representations emerge out of frequently repeated patterns of perceptual input. By virtue of this process, macro-social variables coinstantiated in the perceptual input (understandings of "female," "young," "white," or "working class") can vary across minimal contexts and among language users based on their experience—with similarity of experience, as might be expected among members of a Community of Practice, accounting for similarities in understanding. This is especially useful if we are to allow macro-social factors (gender, ethnicity, class, etc.) to retain their dynamicity and to avoid essentializing them in our descriptions of the distribution of sociocultural interpretations across contexts. An account that prioritizes minimal contexts is thus in line with sociolinguistic findings that definitions of, for example, gender, race, age, or class are rooted in particular Communities of Practice: Within a minimal context, these macro-social variables are defined relative to—rather than prior to—one another, and "color" one another. Methodologically, similarities in participants' understandings of these variables can be first identified through ethnographic analysis, including metalinguistic commentary, and then further probed experimentally, for instance, by manipulating specific dimensions of the context and observing the impact of these manipulations over interpretations shared within a group.

An account that treats minimal contexts as conceptual primes is also dynamic in a further sense, in which an account that takes macro-social variables to be primary, such as multivariate analysis (Pichler, 2010), cannot be. Precisely because they come out of ethnographic analysis that takes the participants' perspectives into consideration, the macro-social variables coinstantiated in a minimal context are not a closed group. As Peräkylä (2004) notes, "*where* the workings of context will be found in a single piece of research cannot be predicted in advance" (p. 295, emphasis added). The fact that there is no predetermined, fixed set of objective variables to look out for allows for the identification of new, potentially previously unnoticed dimensions of context to emerge as important to the bringing about of a particular sociocultural interpretation. Eckert (2000) described how a Jock (school- and sports-oriented) versus Burnout (antischool and antiauthority) identity is the driving force behind the interpretation of behaviors (including language) of teenagers at a Detroit high school, while Mendoza-Denton (1996) did the same for the categories of Sureñas (Mexican-identified) versus Norteñas (U.S.-identified) among Latinas at a Northern California urban public high school. Jocks, Burnouts, Sureñas, and Norteñas are situated, agentive identities that distinguish among young people who otherwise share the larger demographic categories they belong to in terms of age, gender, ethnicity, and class and would therefore be indistinguishable on the basis of those. Moreover, they are observable (they can be tracked through nonlinguistic behaviors, including ways of dressing up and wearing one's makeup) emic categories, categories which participants themselves acknowledge as useful for self- and other-identification and thus have come up with names for. A researcher going into the field cannot know about the existence and remit of these categories in advance. They will simply note as much information about contexts, including the participants' nonlinguistic behaviors and own categorizations, as they can and it is in trying to systematize this information over repeated interactions that patterns of correlation with language use will emerge. It is here in particular that the treatment of minimal contexts as conceptual primes proves especially powerful.

By the same token, an approach based on the notion of minimal contexts imposes its own methodological constraints (or limitations). As the point is to capture those variables that users themselves notice and which guide their perception of what is normal in a certain situation, the analyst should be equipped with an emic understanding of the community they are working with to be able to identify and include in the minimal context those variables the participants themselves are

paying attention to (those they are orienting to in their exchanges). Uptake displays discussed earlier (see the section "Identifying Uptake Displays in Conversational Data") are of course also relevant to this. This further means that the variables that together make up a minimal context must be identified anew each time. This process may not yield the same number of variables or even the same variables for different data sets (cf. Pichler, 2010): Results of analyses in different communities may not be directly comparable and different variables may be obtained from different data sets. Although this means some loss in generalization, this is counterbalanced by a gain in faithfulness to the effects of culture. Cultures should, after all, be described in their own terms, at least initially. That said, when the results of analyses do turn up similar sets of variables from different data sets, the resulting comparisons can be a lot more powerful, as they will have emerged bottom-up, from the conceptualizations of the participants themselves.

Concluding Remarks

Capturing the ways in which particular, situated interpretations of linguistic expressions are shared among members of a group—what I have here called "sociocultural interpretations"—can be difficult for two reasons: First, the interpretations themselves cannot be directly observed, and, second, the contexts that enable these interpretations cannot be defined independently of them. Yet, the reality of such interpretations attested in piece after piece of empirical research calls for an explanation.

In this article, I have outlined a bottom-up methodology that seeks to extract context-sensitive definitions of, on the one hand, sociocultural interpretations and, on the other hand, the context variables that covary with them, from the data itself. Uptake-based definitions of sociocultural interpretations are empirically verifiable and include speaker, context, and addressee contributions to the bringing about of a certain sociocultural interpretation. Dynamic definitions of macro-social variables (gender, age, class, ethnicity, region, etc.) can emerge by gradually abstracting over the minimal contexts that are found to enable particular sociocultural interpretations. By allowing macro-social variables to be defined relative to each other rather than in isolation, minimal contexts represent an improvement over previous decontextualized definitions of such variables.

The approach outlined here differs from a range of other approaches often used to construct understandings of culture from linguistic data. In the preceding sections, I briefly touched on three of these: corpus approaches, exit interviews, and CA methodologies. Corpus approaches take as their starting point for analysis the expression used by the speaker rather than its interpretation by the listener. In a standard corpus approach, the researcher searches for certain expressions (words, lemmas, or longer collocations) in the corpus and it is this predetermined set of expressions that delimits the scope of the investigation. On the approach advocated here, on the contrary, it is the listener's interpretation (the uptake), as evidenced in his or her verbal or non-verbal reaction, that serves as the starting point for identifying the expression that preceded it as a realization of the phenomenon investigated. This allows for the identification of novel expressions, expressions that the researcher may not have envisaged as realizations of a certain phenomenon, in a bottom-up fashion.

The approach advocated here also differs from a methodology incorporating exit or post hoc interviews as a way of cross-checking participants' interpretations. Identifying more abstract phenomena such as speech acts based on linguistic data alone is certainly complex and the researcher must be alert to different types of uptake, as shown in the section "Identifying Uptake Displays in Conversational Data." Moreover, relying on the listener's uptake can only help identify positive understandings: Instances where a speaker may have intended to carry out an act but there is no evidence in the listener's uptake that they interpreted the act as intended will not be

identified through such a process. Exit interviews have been argued to provide precisely such a window into speaker's intentions (He, 2012) as they can help clarify what the speaker was trying to achieve when the available behavioral evidence falls short of making that clear. However, I would argue that prioritizing the speaker's intention is inappropriate when trying to identify sociocultural interpretations, that is, interpretations that are shared between the speaker and her addressee. This is so for two reasons: First, as explained at the start of the section "Variation in Sociocultural Interpretations: What Are We Talking About?", sociocultural interpretations are not interpretations that rely on the speaker's intention but rather on practices that are shared among members of a Community of Practice. The speaker's intention must be discharged through an expression recognizable also to the addressee if such interpretations are to be identified; information about the speaker's intention is one-sided and not adequate for that. Second, it is actually not unusual that participants have different understandings of what is going on; this only becomes problematic if subsequent interaction furnishes evidence of misunderstanding. Lacking such evidence, whether the speaker had a certain intention or not has no impact on the subsequent unfolding of the interaction and no real-world consequences. It is therefore not of interest to an analysis of how sociocultural interpretations of the type that are taken for granted within a group are reached. Post hoc interviews also raise other issues, relating to their being separate events with their own dynamics and varying depending on the identity of the eventual interviewer(s) (compare Mills, 2003 and Note 11). In sum, while they can be a useful tool for triangulating data gathered from a small number of participants, they are less so if the goal is to identify patterns of understanding shared among a group.

Finally, the approach outlined here also differs from CA methodologies. The main difference between CA and the current approach is that CA is interested in the structure of actions at the micro level but not so much in generalizations about the lexicogrammar of the linguistic expressions (what I earlier called "conventions of form") used to bring about these actions at the meso level. On the contrary, the current approach ultimately aims to find out what linguistic expressions regularly achieve a certain type of uptake in different contexts. Awareness that different linguistic expressions may achieve the same uptake in different contexts and an emphasis on uncovering which linguistic expressions do that in which contexts set the current approach apart from CA—conceptually more than methodologically.

To return to my opening example of offering a cup of coffee to workmen in the Midwest versus the Randstand, the approach proposed here would entail tracking several of these interactions with different work(wo)men and different host(esse)s in each place (the working assumption being that a majority of them are familiar with the respective service cultures of those two places—an assumption which, to the extent that patterns *are* discovered, is subject to empirical confirmation by the investigation itself); recording how they unfold linguistically, the types of speech acts that take place and the linguistic expressions used to perform these speech acts; and generalizing over the findings to determine what are the prevailing ways in which such exchanges are (linguistically) handled in each place. Only after such an investigation has been independently undertaken in each community, will it be possible to compare the (linguistic) format of these exchanges to gain insight into whether and how appropriate the offer of a cup of coffee in those circumstances is, and, in the end, what caused the cross-cultural miscommunication I experienced.

Acknowledgments

Declaration of Conflicting Interests

Funding

ORCID iD

Marina Terkourafi (iD) https://orcid.org/0000-0002-4400-9336

Notes

1. For a searchable catalog of transcribed corpora, including many spoken ones, see https://catalog.ldc.upenn.edu/search. Another list of corpora currently available can be found at https://www1.essex.ac.uk/linguistics/external/clmt/w3c/corpus_ling/content/corpora/list/index2.html

2. This is the area of the central-western Netherlands circumscribed by the four largest Dutch cities (Amsterdam, Rotterdam, The Hague, and Utrecht) and their surrounding areas.

3. Widespread societal bilingualism in the Netherlands (European Commission, 2012) means daily communication is possible in English with people of all ages and from all walks of life.

4. "Would you like some coffee?" "Can I make you some coffee?" "Will you have a coffee?" and "I just made coffee" are some of the linguistic forms that can perform the speech act of offering a cup of coffee in English.

5. "Différents groupes sociaux ou différentes situations sociales, trouvent certains modes de communication (façons de parler, intentions d'énoncés) plus apropriés que d'autres. La préférence marquée pour un certain mode de communication explique la fréquence plus élevée de ces formes comprenant les signifiés à l'origine de ce mode particulier" (Lavandera, 1982, p. 94).

6. Although in what follows I refer to speech acts, the mechanism developed here can be applied to all types of utterance-type meaning listed by Levinson (1995), including generalized conversational implicatures, conversational presequences, and presuppositions, inasmuch as they all involve understanding, that is, some kind of hearer uptake.

7. "Would you like to try some cake?" uttered by a host to a dinner guest and "Would you like to follow me?" uttered by a service provider, for instance, are two examples where this expression may be realizing different speech acts—an offer in the first case, and a request in the second.

8. All examples drawn from Terkourafi (2001; for a description of the corpus see Chapter 2). The original Greek examples are cited in the English translation. The numbers in parentheses refer to (Recording Session: Utterance within the Session). The following transcription conventions are used: (.) = short pause, ? = rising intonation, . = falling intonation, ! = animated speech, (text in parentheses) = not clearly audible words, and ((text in double parentheses)) = material added for clarity or completeness. The arrow ($\rightarrow$) indicates the turn where uptake display occurs.

9. I use the term *Relation* here as a cover-term for a family of concepts related to Grice's (1975/1989) maxim by the same name ("Be Relevant"; p. 27).

10. Lack of evidence of misunderstanding is not the same as understanding, of course, and may be due to a variety of reasons. Issues of power or self-presentation may prevent a speaker from indicating that something has gone wrong. While that is certainly possible, the fact remains that a certain interpretation of an utterance is left to stand on record and, so long as it is not replaced by a different one, there are no other interpretations to which both the speaker and the addressee have mutually assented and which could be taken to be part of their shared understanding. If we take seriously the constitutive role of the uptake and Austin's view that without it there is no speech act, and if you did not take what I said as a warning, then no warning has occurred; conversely, if you did take what I said as a warning (and I didn't contest that), then warning has taken place (even if that is not what I had intended).

11. One exception here is the possibility of post hoc interviews, which may provide the necessary evidence that the speaker's utterance was taken differently than intended, when this cannot be gleaned from the transcript itself. Although this technique has been successfully applied in intercultural communication

research (e.g., He, 2012; Spencer-Oatey & Xing, 2003; Tyler, 1995), I do not discuss it here for several reasons. First, evidence of misunderstanding is often visible in the conversational transcript itself, that is, it is not the case that the interview is needed to reveal that a misunderstanding has taken place but only to help clarify the speaker's original intention. Second, as Mills (2003) has argued, postevent interviews create "another text, another conversation, only this time the interaction is with the analyst" (p. 45); if so, analyses of misunderstandings provided during such interviews may be (additionally) reflective of self-presentational concerns vis-à-vis the analyst rather than uniquely representative of the original event itself. Finally, on a practical level, post hoc interviews can only be undertaken on a limited scale and not with the often hundreds of participants involved in collection of the large-scale conversational data needed for the approach outlined here.

12. VP and NP stand for Verb Phrase and Noun Phrase, respectively. This means that the linguistic expressions referred to here are akin to constructions with some lexical elements fixed (Can I, We need, Have you, etc.) and others being open variables (VP, NP), as constructions are understood in the framework of Construction Grammar (Goldberg, 1995, 2005).

References

Austin, J. L. (1975). *How to do things with words* (2nd ed.). Oxford, UK: Oxford University Press. (Original work published 1962)

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577-660.

Clark, E. (2004). Pragmatics and language acquisition. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 562-557). Oxford, UK: Blackwell.

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Eckert, P. (2000). *Linguistic variation as social practice: The linguistic construction of identity in Belten high*. Malden. MA: Blackwell.

European Commission. (2012). *Special Eurobarometer 386: Europeans and their languages*. Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf

Fischer, R. (2011). About chicken and eggs: Four methods for investigating culture-behaviour links. In F. J. R. van de Vijver, A. Chasiotis, & S. M. Breugelmans (Eds.), *Fundamental questions in cross-cultural psychology* (pp. 190-213). Cambridge, UK: Cambridge University Press.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz (Ed.), *The interpretation of cultures: Selected essays* (pp. 3-30). New York, NY: Basic Books.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: The University of Chicago Press.

Goldberg, A. E. (2005). *Constructions at work: The nature of generalization in language*. Oxford, UK: Oxford University Press.

Goodwin, C. M., & Duranti, A. (1992). Rethinking context: An introduction. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon* (pp. 1-42). Cambridge, UK: Cambridge University Press.

Grice, H. P. (1989). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics, Vol. 3: Speech acts* (pp. 41-58). New York, NY: Academic Press. (Reprinted from *Studies in the way of words*, pp. 22-40, by P. Grice, Ed., 1975, Cambridge, MA: Harvard University Press)

He, Y. (2012). Different generations, different face? A discursive approach to naturally occurring compliment responses in Chinese. *Journal of Politeness Research*, *8*, 29-51.

Knight, D., Adolphs, S., Tennent, P., & Carter, R. (2008, May 28-30). The Nottingham Multi-Modal Corpus: A demonstration. In *Proceedings of the 6th Language Resources and Evaluation Conference (LREC), Palais des Congrés, Marrakech, Morocco* (pp. 1-7). European Language Resources Association. Retrieved from http://orca.cf.ac.uk/78855/

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.

Lakoff, G. (1977). Linguistic gestalts. In W. A. Beach, S. E. Fox, & S. Philosoph (Eds.), *Papers from the Thirteenth Regional Meeting, Chicago Linguistic Society* (pp. 236-287). Chicago, IL: The University of Chicago Press.

Lavandera, B. (1982). Le principe de réinterprétation dans la théorie de la variation [The principle of reinterpretation in the theory of variation]. In N. Dittmar & B. Schlieben-Lange (Eds.), *La Sociolinguistique dans les Pays de Langue Romane* (pp. 87-96). Tübingen, Germany: Gunter Narr.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral Participation*. Cambridge: Cambridge University Press.

Levinson, S. C. (1995). Three levels of meaning. In F. R. Palmer (Ed.), *Grammar and meaning: Essays in honour of Sir John Lyons* (pp. 90-115). Cambridge, UK: Cambridge University Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Mendoza-Denton, N. (1996). "Muy macha": Gender and ideology in gang-girls' discourse about makeup. *Ethnos: Journal of Anthropology*, *6*, 47-63.

Mills, S. (2003). *Gender and politeness*. Cambridge, UK: Cambridge University Press.

Morgan, J. L. (1978). Two types of convention in indirect speech acts. In P. Cole (Ed.), *Syntax and semantics, Vol. 9: Pragmatics* (pp. 261-280). New York, NY: Academic Press.

Peräkylä, A. (2004). Reliability and validity in research based on naturally occurring social interaction. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice* (pp. 283-304). London, England: SAGE.

Pew Research Center. (2017). *What it takes to truly be "one of us": In U.S., Canada, Europe, Australia and Japan, publics say language matters more to national identity than birthplace*. Retrieved from http://assets.pewresearch.org/wp-content/uploads/sites/2/2017/04/14094140/Pew-Research-Center-National-Identity-Report-FINAL-February-1-2017.pdf

Pichler, H. (2010). Methods in discourse variation analysis: Reflections on the way forward. *Journal of Sociolinguistics*, *14*, 581-608.

Quintana, S. M. (1998). Children's developmental understanding of ethnicity and race. *Applied & Preventive Psychology*, *7*, 27-45.

Sbisà, M. (2009). Uptake and conventionality in illocution. *Lodz Papers in Pragmatics*, *5*, 33-52.

Schegloff, E. (1968). Sequencing in conversational openings. *American Anthropologist*, *70*, 1075-1095.

Spencer-Oatey, H., & Xing, J. (2003). Managing rapport in intercultural business interactions: A comparison of two Chinese-British welcome meetings. *Journal of Intercultural Studies*, *24*, 33-46.

Terkourafi, M. (2001). *Politeness in Cypriot Greek: A frame-based approach* (Unpublished doctoral thesis, Department of Linguistics, University of Cambridge). Retrieved from https://www.ideals.illinois.edu/handle/2142/9573

Terkourafi, M. (2005). Pragmatic correlates of frequency of use: The case for a notion of "minimal context." In K. Nikiforidou, S. Marmaridou, & E. Antonopoulou (Eds.), *Reviewing linguistic thought: Converging trends for the 21st century* (pp. 209-233). Berlin, Germany: de Gruyter.

Terkourafi, M. (2009). On de-limiting context. In A. Bergs & G. Diewald (Eds.), *Contexts and constructions* (pp. 17-42). Amsterdam, The Netherlands: John Benjamins.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Tyler, A. (1995). The coconstruction of cross-cultural miscommunication: Conflicts in perception, negotiation, and enactment of participant role and status. *Studies in Second Language Acquisition*, *17*, 129-152.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics* (pp. 95-195). Austin: University of Texas Press.

Wittgenstein, L. (1976). *Philosophical investigations* (G. E. M. Anscombe, Trans.) (3rd ed.). New York, NY: Macmillan. (Original work published 1953)

Zipf, G. (1949). *Human behavior and the principle of least effort*. Cambridge MA: Addison-Wesley.