Research Methodology in Second-Language Acquisition

Edited by
Elaine E. Tarone
University of Minnesota

Susan M. Gass
Michigan State University

Andrew D. Cohen
University of Minnesota

LEA
LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1994 Hillsdale, New Jersey Hove, UK
How well do applied linguists and second-language acquisition (SLA) researchers understand theory construction in our field? As attested by a number of books, edited collections, journal articles, and a recent conference devoted to theory construction, SLA researchers are becoming increasingly sophisticated in such matters. We have not only reached the stage where we are actively exploring what philosophy of science has to say about theory construction; we are also defining the goals of SLA research ever more precisely.

These developments are welcome signs that SLA studies are coming of age. However, this sense of growing maturity is being achieved at some cost to the field. As even a casual reading of the SLA literature will attest, the overwhelming majority of SLA studies are of the logico-deductive, experimental variety. Thus, whether they recognize this or not, most SLA researchers subscribe to a nomothetic epistemology that has, in my opinion, prematurely achieved the status of a dominant orthodoxy.

Although the results of nomothetic science are in many instances impressive, we should not be seduced into accepting the validity of this epistemological position without question. I therefore first present a critique of the current nomothetic orthodoxy in SLA research from an ethnomethodological perspective and then demonstrate how Conversation Analysis may be used to motivate the theoretical position that conversation is the sociocultural context of second-language learning.
A CRITIQUE OF THE CURRENT ORTHODOXY IN SLA RESEARCH

Let me begin with a summary of some technical terms. Ochsner (1979) differentiates between nomothetic and hermeneutic scientific traditions by noting that the former is concerned with explaining and predicting how natural phenomena work, whereas the latter focuses on understanding and interpreting how these phenomena are organized. More specifically, nomothetic science (the Greek prefix *nomos* means "lawful") assumes the existence of a single, discoverable reality that causally obeys the laws of nature. In contrast, hermeneutics (meaning "the art of interpretation") assumes that multiple realities exist and that human events in particular can be interpreted only according to their outcomes. In terms of research methodology, an experimental, quantitative methodology is associated with the nomothetic tradition, and a naturalistic, qualitative methodology is associated with the hermeneutic tradition.

As I noted earlier, the dominant paradigm in SLA research today is that of nomothetic science. The value of this tradition has been forcefully, even relentlessly, articulated by Mike Long and various colleagues and collaborators (see Crookes, 1991, 1992; Larsen-Freeman & Long, 1991; Long, 1985, 1990, in press, among others). Specifically addressing issues of theory construction, Larsen-Freeman and Long (1991) differentiated between two types of theories in nomothetic science. These are causal-process theories and sets-of-laws theories. They noted that sets-of-laws theories consist of collections of (often) unrelated generalizations or laws regarding SLA that are based on observed and tested relationships between variables. The problem with this type of research, according to the authors, is that it does not provide an explanation for the processes being studied; it merely establishes that they are in need of explanation. On the other hand, causal-process theories consist of definitions and operationalizations of abstract theoretical constructs, existence statements, and deterministic and/or probabilistic statements, which are mathematically tested to explain how and why SLA occurs. Larsen-Freeman and Long therefore argued that causal-process theories are the most desirable type of theory from a theory construction perspective because they lead to predictions and generalizations of great explanatory power.

A few writers, principally Schumann (1983, in press), supported by Candlin (1983), Guiora (1983), and van Lier (1988), have valiantly attempted to stem the nomothetic tide. But the debate has unfortunately been somewhat lopsided. For the most part, these have been voices crying in the applied linguistic wilderness.

Various statements by Crookes, Larsen-Freeman and Long give a flavor of the confidence with which the nomothetic cause is being advanced in SLA studies. Thus, Crookes (1991) stated: "It may be that SL production research is at the point where the field is ready to move from primarily descriptive research to greater use of experimental investigations of a more obviously hypothesis-testing, theory-developing nature" (p. 125). Similarly, although acknowledging the utility and necessity of descriptive studies as means of establishing baseline data, Larsen-Freeman and Long (1991) labeled these studies as limited and maintained that only a causal-process approach to theory construction can explain SLA.

I disagree with these statements for two reasons. First, although nomothetic science does not preclude descriptive studies, these statements have the undesirable (if unintentional) effect of severely undercutting the viability of hermeneutics in SLA studies by pinning a negative (and incorrect) label ("limited") on qualitative research.² And second, these statements gloss over the fact that a nomothetic epistemology is itself open to serious criticism, a fact that its proponents rarely mention, much less address in any serious fashion.³

Before I develop this latter point, I should make it clear that I am not interested in reopening the argument whether qualitative and quantitative approaches to SLA research are mutually exclusive. There is, in fact, considerable agreement, among SLA researchers on both sides of the methodological divide, that qualitative and quantitative studies are in reality complementary ways of creating new knowledge (see Allwright & Bailey, 1991; Chaudron, 1988; Cohen & Olshtain, this volume, chapter 8; Ellis, 1984; Grotjahn, 1991; Larsen-Freeman & Long, 1991; Long, 1983; Selinker & Douglas, this volume, chapter 6; Shohamy, this volume, chapter 7; van Lier, 1988; and Watson-Gegeo, 1988). I wholeheartedly endorse this assessment. What I wish to do in this chapter is reexamine the relative value that we currently pursue.

²In this regard, Long (in press) called for a culling of existing theories of SLA so that the field can progress to a period of paradigmatic stability, in which a so-called Dominant Theory will rule the SLA roost, and demanded that relativists explain why such a culling should not take place (see also Beretta, 1991). As mentioned earlier, we need only look at the preponderance of experimental research published in SLA journals to realize that a culling of SLA theories probably has already taken place. More specifically, the dominant nomothetic paradigm has in fact preempted the development of hermeneutic theories of SLA before their potential worth could be evaluated. This is a loss to the field, not an indication of theoretical maturity, because researchers (particularly doctoral students, on whom the field's capability for self-renewal rests) are thereby subtly discouraged from committing themselves to nonexperimental research.

³Thus, Crookes (1992) mentioned in passing that there are objections to experimental research but does not say what they are. And Larsen-Freeman and Long (1991) referred to hermeneutics only three times (in notes) in their entire book, never considering the possibility that hermeneutics can offer an explanatory theory of SLA.

---

¹As Ochsner (1979) correctly noted, most of the social sciences (including SLA) are actually preparadigmatic and are likely to remain so for quite some time. For this reason, following Giddens (1988), I prefer to speak of a tradition rather than a paradigm.
ascribe to understanding versus explaining SLA phenomena and question whether explanation can be achieved only through a causal-process approach to SLA theorizing.

What are the weaknesses of a causal-process approach to SLA? I address this question from the perspective of ethnomethodology, a radical sociology in the hermeneutic tradition proposed by the sociologist Harold Garfinkel, who defined ethnomethodology as "the investigation of the rational properties of indexical expressions and other practical actions as contingent ongoing accomplishments of organized artful practices of everyday life" (Garfinkel, 1967, p. 11).

The ethnomethodological critique of nomothetic science focuses crucially on experimentalists' rejection of ordinary language and day-to-day experience as valid ways of knowing and organizing the world, both these ways of accessing knowledge are viewed by experimentalists as too value-laden and subjective to be useful tools for science, which must above all else be value-free and objective. In order to go beyond the perceived inadequacies of the lay talk of ordinary social actors and the "shallow" explanations of social phenomena that a dependence on ordinary language generates, experimentalists have developed a highly technical language to describe their subject matter and to provide "deeper" explanations that are cast in terms of underlying causes and effects. Thus, they attempt to define and operationalize abstract theoretical constructs and develop logico-deductive hypotheses about a given phenomenon (such as SLA), which they then set out to falsify using the pure, value-free language of mathematics. But in their efforts to develop this objective, value-free language, experimentalists are confronted with the contradiction that the technical language and causative explanations of social science cannot be anything but parasitic upon the ordinary language of social actors and their lay explanations of everyday experience (Giddens, 1988).

More specifically, I wish to develop the following theses: The operationalization of constructs and the technical definitions that experimentalists develop are contaminated by the notion of reasonable agreement concerning the defining characteristics of a given phenomenon (Lynch, 1991). Furthermore, experimentalists depend on lay talk as a component of any technical definition (Coulter, 1991) and/or depend on highly metaphorical language to develop their arguments. In addition, the degree to which a discipline has been mathematized (i.e., the extent to which it routinely uses inferential statistics to test predicted interactions between variables) is not necessarily an indication of maturity. Not only can the use of some statistical procedures result in unnecessarily opaque discourse whose actual conclusions may in the end be comparatively meager, but the supposed objectivity of such discourse also frequently camouflages the many practical decisions, judgments, and subjective interpretations that inform this attempt to objectivize the language of science (see Benson & Hughes, 1991; Sharrock & Button, 1991). Therefore, there can be no absolute guarantee that an abstract, mathematical explanation of a phenomenon is necessarily superior to a hermeneutic explanation that is constructed in terms of lay participants' real-time understanding of the same phenomenon (Sharrock & Button, 1991).

Let me now illustrate what I mean by critiquing some undeniably good examples of recent experimental SLA research. This research has been selected to show that, notwithstanding its careful research design and attention to detail, it is nonetheless vulnerable to the criticisms outlined previously.

In line with the position developed by Long (1983), Long and Porter (1985) and Porter (1986) used the qualitative work on conversational repair in Schegloff, Jefferson, and Sacks (1977) to generate their own operational definition of repair as a prelude to quantification and experimentation. Briefly, Long and Porter defined repair as a composite variable consisting of six subcategories including confirmation checks, clarification requests, comprehension checks, verifications of meaning, definition requests, and expressions of lexical uncertainty. They used this new definition of repair to test hypotheses about the relative importance of different types of input and interaction that native and nonnative speakers provide to each other in task-centered talk. These hypotheses were inspired by Long's (1981) hypothesis that negotiated comprehensible input is the necessary and sufficient cause of language learning.

The first problem that Long and Porter came up against was that their operationalization of the construct of repair was based on the notion of reasonable agreement. That is, they accepted that Schegloff et al.'s treatment of repair was an accurate and trustworthy account of how participants actually

---

4Of course, I do not mean to imply that experimental research on classroom discourse is not based on transcripts of actual classroom interaction. But these transcripts are merely the raw material for the subsequent coding and statistical analyses, which become the primary data for the experimental researcher.

5In his critique of orthodox social science, Lynch (1991) pointed out that for Kuhn (1961), "the only possible criterion for assessing the agreement between the accepted numbers and observed measures is the mere fact that they appear, together with the theory from which they are derived, in a professionally accepted text" (p. 78). For example, physicists can decide, given current measurement techniques, that a discrepancy between the predicted and actual orbit of a planet is acceptably accurate or that this discrepancy is due to the effect of gravitational forces from an as yet undiscovered planet, only on the basis of what other respected researchers have said on the subject. For the purposes of this chapter, I maintain that the notion of reasonable agreement underlies all attempts to infer, interpret, or argue a particular point of view.

6See the second part of this chapter (in particular, footnote 23) for a discussion of the methodology that is used to identify participants' understanding of the conversations they construct. I wish to emphasize that the methodology of Conversation Analysis does not use the kind of triangulated self-report procedures that are a familiar part of ethnographic research.
achieve repair. Note that I am not denying the high quality of Schegloff et al.'s work here (which would be foolish, because it constitutes a foundation for the analysis presented in the second part of this chapter). What I am pointing out is that, in terms of their own epistemology, the supposedly objective research enterprise described by Long and Porter in their various papers is founded on the quicksand of opinion, however well informed this opinion may be. The ostensibly scientific definition of the composite variable of repair that they propose cannot escape being contaminated by the so-called imprecisions of everyday language, because our understanding of notions like comprehension checks and clarification requests (as shown by the italicized turns in the following extracts) must necessarily depend ultimately on examples of ordinary talk extracted from transcripts of the participants' original conversations:

Comprehension check

L: To sin- uh . . . to sink
N: Do you know what that is?
L: To go uh-
N: To go under . . .

Definition request

L: . . . what is the meaning of research?
N: Um, study? You study a problem and find an answer.

(Porter, 1986, p. 207, emphasis original)

With respect to the question of camouflage in this particular research, notice that the statistical analyses reported later to test the hypotheses of the study depend crucially on such categories as comprehension checks and definition requests being objectively distinct. But what Porter's analysis glosses over is, from a participant's intersubjective perspective, these categories have little or no psychological validity. They are in fact superficial artifices of the analyst that distort the pragmatic intent of the participants. More specifically, there is no real-world justification for treating these two categories as discrete illocutionary acts. The interlocutors in both extracts orient to (a) a need to resolve some trouble in their conversation, and (b) a resolution by means of some definitional work on the lexical items that are problematic. Arguably, therefore, the single category of defining, in which the meaning that the participants themselves attach to what they are doing is intersubjectively constructed over several turns, seems better motivated by the data than the two categories proposed by Porter. The analyst's subjective and unwarranted decision to treat these categories as discrete entities therefore inevitably taints the subsequent statistical analysis.

I wish to emphasize that these kinds of problems are not peculiar to the particular research I have just reviewed. For example, with respect to the issue of how important it is to mathematize a discipline if it is to have anything worthwhile to say, the work of Bayley (this volume, chapter 9) also illustrates some of the problems faced by researchers who work in the nomothetic tradition rather well.

Bayley explained the interlanguage variation exhibited by Mandarin Chinese speakers of English as a second language with respect to their use of the phonetic form of the past tense and their use of grammatical aspect by using the method of maximum likelihood estimation, a form of multivariate statistical analysis. This is expressed by the following mathematical formula:

$$
\frac{P}{1 - P} = \frac{P_0}{1 - P_0} \times \frac{P_1}{1 - P_1} \times \ldots \times \frac{P_n}{1 - P_n}
$$

The exact meaning (or, rather, the lack of meaningful to perhaps the majority of his readers) of this supposedly objective but hardly very transparent equation required Bayley to devote about a quarter of his chapter to explaining what all this actually means. Yet, in order to explain this highly abstract communication, Bayley had to explain himself through language!

The persistent reader was rewarded with the conclusion that "variation in interlanguage tense marking is indeed systematically conditioned by a range of linguistic, social, and developmental factors" (which are then outlined in greater detail). I leave it to readers to decide whether their investment of time and energy into an understanding of the statistical procedure was worth the conclusions that Bayley reached.

Finally, the question of metaphor in "objective" research is highly problematic for the nomothetic tradition (and indeed for researchers who work in the more logical-deductive tradition that informs much formal linguistic research). Consider the following citation from Bley-Vroman and Chaudron (this volume, chapter 13), in which the authors set out their ideas on how language is stored in the brain:

In keeping with nearly all present thinking on language processing, we assume that, in the native speaker, the language processor automatically and obligatorily produces representations of the input and does not itself require the use of short-term memory. To borrow the evocative allusion of Fodor (1983), parsing is not "sickled o'er with the pale cast of thought" (p. 64). In the conception of current generative linguistics, the language processor is "encapsulated" in a language module (see Garfield, 1987, for papers in this tradition). Parsing thus cannot affect imitation accuracy directly by "filling up" short-term memory. It is important to be clear about this matter. The claim is not that the parser has no memory store: No doubt it has at least some sort of "look-ahead buffer" (Bergwick & Weinberg, 1984). Rather, whatever the store used by the parser, it is
not the same short-term store that is used by a subject in remembering what was said in order to repeat it. (quotation marks original)

What is remarkable about this extract is that it draws on no less than three different metaphors in one paragraph to clarify the authors' position. These metaphors are identified explicitly by the use of quotation marks. First, there is the Shakespearean metaphor, which is used throughout the chapter to provide one explanation of how short-term memory may work. Second, there is the formalist metaphor of generative linguistics, which suggests that the language processor (whatever that is, neurobiologically) is "encapsulated" in something called a language module (again, whatever this might be, from a neurobiological perspective). And third, there are the artificial intelligence/computer science metaphors of "filling up" memory banks and putative "look-ahead buffers."

The reader may wonder which of these three metaphors is the most important in these authors' discussion. Furthermore, it is instructive to consider that if this chapter had been written in the 17th century, we might have presented with a metaphor involving language as a well-regulated mechanism that could be explained by analogy with the chronometer (the dominant metaphor of the day, used by Newton and others to explain the universe). Thus, a fundamental epistemological question for researchers working in this kind of tradition is: Does the use of such metaphors serve to clarify the phenomenon of SLA or does it serve merely to obfuscate it?

Experimental researchers interested in developing causative theories of SLA may reply that these kinds of issues are merely technical problems that should not sidetrack us from the real business of SLA research, which is to provide causative explanations of SLA. Furthermore, so the argument runs, many of these objections can be met by developing better coding procedures.

The development of better coding procedures is indeed always a possibility. However, the elaboration per se of better analytical categories does not circumvent the basic problem of "objective" technical definitions being dependent on "subjective" lay talk and/or the inescapable use of metaphor. But even if we accept the technocratic argument that these deeper epistemological matters can be set aside temporarily in the interests of scientific progress (an argument that, of course, I do not accept), it is by no means an irrefutable, objective fact that only nomothetic science can explain SLA phenomena.

As we have seen, the categories that constitute Porter's composite variable of repair by themselves tell us next to nothing about the fundamental properties of second-language learning. For example, we do not know whether, by repairing their speech, the learners in the two extracts actually understood, much less learned, what the words to sink or research mean. Thus, the use of quantified data cannot tell us whether a particular conversational mechanism (such as turn-taking or repair) is actually available as a resource for second-language learning; it merely tells us how often this mechanism occurs in a conversation and whether this frequency of occurrence is statistically significant. Thus, SLA researchers are left in the peculiar position of positing that negotiated comprehensible input is an important variable in SLA, without really knowing on the basis of empirical evidence what successful input (i.e., input that results in demonstrable learning) actually looks like in context. Ironically, therefore, the kind of predictive explanation (Watson-Gegeo, 1988) that researchers working in the nomothetic tradition seek to construct seems to be quite limited in its ability to capture the essence of the phenomena I have just been discussing.

In contrast, the originality of the ethnomethodological respecification of social science (and therefore of SLA studies) offered by Giddens (1988) and others is that this respecification (a) problematizes the methodological assumptions in traditional social science about the process of theory construction as intersubjectively achieved phenomena that are worthy of analysis in their own right; and (b) embraces ordinary language as the indispensable medium for analyzing participants' constructions of their everyday experience, whether this analysis is done by participants or by analysts. Thus, as Benson and Hughes (1991) put it, ethnomethodology is an attempt to make the world investigable in the participants' own terms. Consequently, from an ethnomethodological perspective, the scientific rigor of a study is not evaluated in terms of the sophistication of the statistical techniques that might be employed, because ethnomethodologists rarely use such techniques. Rather, a study is rigorous to the extent that it explains the intersubjective achievement in which participants understand the locally recognizable and locally adequate turns-at-talk that they engage in to organize their world.

At the risk of being accused of setting up a straw man and/or of being redundant, I predict that many experimentalists will raise the familiar objections: (a) such a program of research is useful only to describe baseline SLA phenomena but, cannot explain SLA for the reason given in (b); (b) ethnomethodologists do not attempt to make a priori predictions that can be falsified experimentally; and (c) ethnomethodologists do not work with large groups of subjects, which is the only methodology that can explain SLA phenomena in such a way that the results can be generalized from the subjects tested to a broader population of learners.

It should be obvious by now that I believe that objections of this kind entirely miss the point of the arguments developed in this chapter. Perhaps an example from anatomy will demonstrate the fallacy of such arguments. It would be a very strange version of anatomy indeed that insisted that the function of the human heart as a pump could be reliably explained only by dissecting a large number of hearts to check whether the pumping hypothesis was true of a statistically significant sample of hearts, so that this fact could
be generalized to the entire population of human beings. Dissecting a single cadaver is sufficient to demonstrate (i.e., explain) that the heart does indeed function as a pump. Similarly, researchers interested in the "anatomy" of conversation (i.e., its structure) can make valid generalizations from single cases to the broader population because, as Benson and Hughes (1991) argued, "the point of working with 'actual occurrences,' single instances, single events, is to see them as the products of 'machinery' that constituted members' cultural competence enabling them to do what they do, produce the activities and scenes of everyday life...the explanation, say, of some segment of talk in terms of the 'mechanism' by which that talk was produced there and then, is an explanation of some part of culture" (p. 130, emphasis original). Thus, although the explanation is of a different type (what Watson-Gegge, 1988, p. 576, calls "interpretive explanation"), it is no less powerful and generalizable in its own terms than is predictive explanation.

AN SLA PHENOMENON RESPECIFIED:
A CONVERSATION-ANALYTIC EXPLICATION
OF SPEAKED DEFINITIONS

I now demonstrate how Conversation Analysis (CA), a manifestation of ethnomethodology that has already been successfully used in SLA research (see van Lier, 1988, and others), can be used to provide an interpretive explanation of the mechanisms that enable language learners to use spoken definitions as a resource for language learning. Four basic assumptions govern CA work (Heritage, 1988): (a) conversation has structure; (b) conversation is its own autonomous context—that is, the meaning of a particular utterance is shaped by what immediately precedes it and also by what immediately follows it; (c) there is no a priori justification for believing that any detail of conversation, however minute, is disorderly, accidental, or irrelevant; and (d) the study of conversation requires naturally occurring data.

Because the rules of evidence used by conversation analysts are not as well understood as those used by experimental researchers, let me briefly review what counts as evidence in CA and the kinds of claims made by conversation analysts (in this regard, see Jacobs, 1986, 1987, whose work forms the basis of the following summary). The methodology of CA is qualitative and subject to the usual evaluation criteria for qualitative research. Beyond this, however, the methodology of CA attempts to explicate the lay knowledge of conversationalists by "unpacking" examples that demonstrate the participants' orientations to the conversations they construct in real time. Such examples provide the primary evidence for the asserted existence of particular conversational mechanisms identified by analysts; a case is convincing to the extent that it is directly motivated by the conversational data presented for analysis. Thus, CA makes no appeal to ethnographic knowledge to make an argument. Furthermore, conversation analysts do not develop arguments about the structure of conversation on the basis of quantitative analyses of frequency data, because such analyses cannot tell us anything about the underlying structure of conversation per se. Instead, conversation analysts seek to demonstrate that conversation could not be conversation if such universal interactional resources for constructing meaning as turn-taking, repair, or preference rules did not exist.

In order to demonstrate the existence of such phenomena, conversation analysts use prototypical examples that give discursive form to the phenomena being analyzed. But such examples are not by themselves sufficient to make a convincing argument. Analysts must be able to corroborate their claims by pointing to a convergence of different types of textual evidence or by showing that a single structure identified by the analyst plays a role in different types of cases. Note that the use of convergent evidence, like the use of related data, is a particularly important resource in countering the charge that an analysis is merely an artifact of the examples collected and chosen for presentation to readers. Thus, for example, a preceding preinvitation turn shows that reading the following turn as an invitation is contextually warranted. Finally, analyses must be subject to critical falsification. That is, analysts must demonstrate that potential counterexamples and different accounts for the same data set have been anticipated and that other researchers can replicate findings with different transcripts. Let us now move on to an analysis of classroom data that demonstrates how CA might be used to motivate an ethnomethodological respecification of SLA studies.

THE DATABASE

The complete database for this project consists of 14 lower-intermediate to upper-intermediate English as a Second Language (ESL) classes at a large research university located in the Midwest of the United States. These classes were video- and audiotaped during spring semester, 1990. Each class lasted...
50 minutes. Currently, the conversations of 3 teachers and 33 learners (11 in Class 1, 10 in Class 2, and 12 in Class 3) interacting in ordinary classrooms\(^1\) have been fully transcribed, using transcription conventions that are based on those developed by Jefferson (1978) and van Lier (1988) (see Appendix). The data analyzed in this chapter come from Class 2 only.

**DATA COLLECTION PROCEDURES**

Two video cameras were used to film the participants, who were visually identified by a number pinned to their clothing (i.e., L1, L2, L3, etc.); the video signals were fed into an electronic switcher operated by an assistant. Camera 1 (the main camera) recorded learners interacting in groups or whole class activities, and Camera 2 filmed teacher-fronted activities and/or presentations by students using the blackboard or overhead projector. These video recordings were used primarily to check visually who was speaking to whom when this information could not be determined from the audio data (see Fig. 5.1 for a graphic representation that shows the layout of the class, the position of the two cameras, and the composition of the groups in Class 2).

The audio recordings are the primary sources of data. Each participant was issued a numbered Walkman-sized stereo cassette recorder and a lapel microphone. The number on each recorder (and on each cassette) corresponded to the number pinned to a participant's clothing. This set-up allowed the analyst to identify participants visually on the videotape and aurally on their audiotapes. The portability of this equipment also enabled teachers and students to move around the class without having to worry about problems like tripping over cables. The availability of multiple audio recordings meant that such technically significant information as the difference between pauses and inhaled or exhaled breaths, the specific number of laughter tokens, the precise onset and resolution of overlaps, and the content of muttered commentaries could be distinguished and therefore transcribed with a high degree of confidence. Transcripts for each group and teacher in every class were produced on the basis of these multiple recordings; in the case of Class 2, this yielded a total of six transcripts for the lesson (four parallel transcripts for the four groups, one teacher's transcript, and one transcript consisting of collections of definitions excerpted from the class interaction).

\(^1\)That is, the teachers could teach what they wanted, how they wanted. This contrasts with the more experimental nature of much classroom research, in which the kinds of activities teachers use are prespecified and variables such as the composition of groups (in terms of characteristics like gender and proficiency level) are carefully controlled by the researcher.

**FIG. 5.1.** Seating plan.

C1 = camera 1 (main camera)  
C2 = camera 2 (secondary camera)  
b/b = blackboard
THE TASKS

The tasks learners had to complete in Class 2 involved an open-ended four-way exchange of information. Students first read and discussed, in four small groups, one of four thematically-related magazine articles on the greenhouse effect. A representative or representatives from each group then presented the information contained in each reading with an overhead projector in an oral, whole-class activity. The end product was some written work, which was done in a later class that was not recorded. Approximately 30 minutes (which included 5 to 10 minutes of silent reading, depending on the group) were allocated to small group discussion and about 20 minutes were given over to three oral presentations. The first seven excerpts cited here were produced collaboratively by Group 3 during the group work phase, and Excerpt 8 was produced by L10 (an erstwhile member of Group 3) as part of the oral report to the whole class.

THE DATA

Excerpt 1: Group Work Phase

1 L10: <hh> hhhh what is th- what is the (+) coral (+) what’s ((whisper)) (+) I don’t know (h)
2 (+)
3 L11: just- look at it (+) as a (+) an m- material that’s all
4 (+)
5 L9?: uhm don’t worry about it
6 (+)

Excerpt 2: Group Work Phase

1 L10: excuse me what is c-o-r-a-l ((L10 spells out the word))
2 (+)
3 T: can I: (+) open //((h))/ <h> (+ +) get an idea (+) see where’s that
4 L10: //((h))/
5 L10: <h> I don’t know whether the-
6 (+)
7 T: corals (+) does anyone know? (+) where you find corals?
8 L9: corals (+) u- underwater //you mean? under the-
9 T: //uh huh,/
10 (+)

Excerpt 3: Group Work Phase

1 T: ok when you’re explaining that to the: uh to the class or (+)
2 whoever is <h> probably //you’ll// have to
3 L9: //((unintelligible whisper))/
4 T: say (+) what corals (+) are //= because//
5 L10: //what is// corals
6 T: yeah give a definition //uh of corals//
7 L10: //oh okay what is corals maybe uh//
8 (+)
9 L10: or say-
10 T: or give an example say //for// instance in //Austra://lia (+)
11 L9: //((unintelligible whisper))/
12 L10: //oh yeah,/
13 T: around Australia you get lots of barrier reefs
14 L9: (barrier reefs)
15 T: //o// otherwise people might not know (+) might not be
16 L10: //oh//
17 T: familiar with that word/
18 L10: //oh ok, ok coral reefs/
19 (l)

Excerpt 4: Group Work Phase

1 L10: what is uh corals I don’t know
2 L11: did you read (in) about corals
Excerpt 5: Group Work Phase

3 L9: no
4 (+ +)

Excerpt 6 (simplified): Group Work Phase

36 (+)
37 L9: what
38* L10: c/orals/
39* L11: //corals/
40 L9: corals oh okay
41 L10: yeah

Excerpt 7 (simplified): Group Work Phase

4 T: <h> yeah what would be another word for a habitat then
5 (+) it's like (1) //it's bli-/
6 L11: //I ha/ve no idea
7 L9: //situation/
8 T: //s like the whole// (+) situation (+) //home//
9 L10: //home//

Excerpt 8: Whole Class Oral Discussion Phase

1 L10: <h> in my section I think the main point is the: <hh> raising of
the sea (+) sea level <hh> we are accompanied uh global
3 tempera increasing <hh> (+) so these (+) put the co[l][a]l (+) at
[ll]isk. <h> I think the co[l][a]l is the kind of fossil (+) <h>
fossil at the: botto of the sea. <hh> the: co[l][a]l reef
you are one of the imp- very important, <hh> habitats (+)
7 for fish that support th[a]:m more than (+) <h> one (+) more
than one third of topic-topical species. \(<hh>\) the **habitats**

is the: **home** (+) is the **home** for animal living \(<hh>\) so this is

very important (+) very important \(<hhh>\) so: now, the

conversationalist uh want to find some [ylay: (+) to save

these uh (+) to: (+) solve this problem.

---

### A Definition of Spoken Definitions

For the purposes of the analysis of the larger corpus from which these data were taken, the defining characteristics of spoken definitions\(^\text{12}\) were determined not on the basis of a priori categories but on the basis of an analysis of the interaction in all three fully transcribed classes. Thus, spoken definitions are not defined here as linguistic products (see Sacks, 1985, for an analysis of the linguistic structure of definitions) or as logical forms (see Flowerdew, 1991, 1992, for an analysis of definitions inspired by speech act theory). Rather, they are defined as any turn(s)-at-talk that are hearable by participants as explanations of lexical items or phrases whose meaning is actually or potentially unclear.\(^\text{13}\) More specifically, participants achieve definitions by simultaneously orienting to the resources of turn-taking and repair available to them as conversationalists and using a range of vocabulary elaboration strategies (Chaudron, 1982) to resolve the problems they are confronted with. These strategies (which may be used singly or in combination) include the use of iconic, nonverbal means of defining, such as pointing, acting, drawing, and showing pictures. They also may be explicit verbal strategies such as simplification, synonymy, antonymy, classification, approximation, exemplification, comparison, and translation. Finally, these explicit strategies are complemented by such implicit verbal elaboration strategies as apposition, parallelism, and paraphrase.\(^\text{14}\)

\(^\text{12}\)Field notes indicate that although definitions are not unknown in ordinary conversations between fully competent native speakers, they are infrequent. For example, an analysis of three transcripts of native-speaker ordinary conversation, “Two Girls,” “SN-4,” and “Auto discussion,” showed that no definitions occurred in these conversations. But the present data indicate that they can be frequent in SL classrooms. Definitions may therefore be understood as a typically pedagogical form of talk.

\(^\text{13}\)Note that for a stretch of talk to count as a definition, analysts cannot rely on their subjective impressions regarding participants’ possible psychological motivations for doing a definition. They must be able to show, on the basis of empirical data from the transcript, that the participants are demonstrating to each other (and thus to the analyst) that a term is actually or potentially problematic and that it therefore needs to be explained.

\(^\text{14}\)All of these elaboration strategies are attested in the three classes that have been fully transcribed. In the eight related examples given in the previous section of this chapter, we find a more restricted subset of these strategies, which includes the use of categorization in Excerpt 1 (lines 3); exemplification in Excerpt 2 (lines 7–22) and Excerpt 3 (lines 10–18); synonymy in Excerpt 6 (line 2); comparison in Excerpt 7 (lines 7–9); a combination of simplification (lines 11–17),

---

### 5. Ethnomethodological Respecification of SLA Studies

The data were first examined to establish the prototypical conversational structure of definitions. Prototypical definitions are achieved as sequences that consist of a question-and-answer adjacency pair (Sacks & Schegloff, 1973), followed by an evaluation turn in which participants indicate whether they have understood the definition. In the question turn(s) of the adjacency pair, participants predominantly use **wh**-questions such as **what does X mean**, **what's the meaning of X**, **what is X**, (or interlingual variations thereof) to initiate definitions; however, a number of other forms (such as **yes/no** questions or plain **X with rising intonation**) are also found. In the answering turn(s), participants define problematic terms using the kinds of elaboration strategies identified by Chaudron (1982). And in the third and final commenting turn(s), participants typically use change-of-state tokens like **oh (ok)**, which assert understanding and which may close the sequence (Heritage, 1984).

In order to guard against inadvertent analyst bias, the data were also checked (a) to see if participants achieved definitions in ways that did not conform to this prototypical pattern, and (b) to confirm that participants did indeed orient to sequences that superficially conformed to the prototypical pattern as bona fide definitions. As the teacher’s talk in Excerpt 3 (lines 1–10) demonstrates, Excerpt 8 includes talk (at lines 4–9) that is intended by L10 to be heard as a definition, even though it is not done as a sequence and does not display the three-part structure I have just alluded to.\(^\text{15}\) And with respect to the issue of talk that superficially resembles prototypical definitions, untapped classroom observations and field notes from a fourth class that is not fully transcribed at the present time indicate that sequences that begin with a **What is X** turn followed by the prototypical answering and commenting turns need not necessarily count as definitions of unknown terms. More specifically, a learner in this fourth class (which was discussing euthanasia) rhetorically asked, “What is death?” His interlocutor initially oriented to this question as a request to define this word and began to provide a definition in his answering turn. But it immediately became clear, when the first student interrupted this definition-in-progress with his own answering turn, that he asked this question so that he could develop his own views on this subject, not because he did not know the meaning of **death**. The final comment-

\(^\text{15}\)Three other definitions that display a structure similar to that of the definition in Excerpt 8 were found in the three classes, though not all occurred in the same environment as the one in Excerpt 8.
ing turn of the second student confirmed the rhetorical nature of this definition, because he indicated that he understood that the first learner was not in fact asking for help in understanding the meaning of death.

On the basis of these procedures, a total of 82 attempted and/or completed spoken definitions have been identified in the three classes that constitute the current database in the larger corpus. The eight definitions cited in the data section are representative of the range of definition types used in the database as a whole.

ANALYSIS

The first question we might ask ourselves is what conversational resources L10 draws on in Excerpts 1–8, first to understand what coral means, and ultimately to explain this word to her fellow students. The definitions in Excerpts 1–7 are not planned. They are locally occasioned (i.e., they are constructed in real time by interlocutors who use the resource of turn-taking to develop a definition sequence cooperatively). And they are also done as repairs. Following Schegloff et al. (1977) and other writers, repair consists of any work participants engage in to clear up conversational trouble as it

---

16Preliminary counts of the number of attempted and/or completed spoken definitions instantiated in the data indicated that a total of 80 definitions had been produced in the three classes, of which 79 were associated with target words or phrases that occurred in the original source readings. Thus, only one definition focused on a word or phrase that had been locally occasioned by the talk itself. This occurred in Class 3 during a short period of time when the composition of the groups was being reconstituted and two students briefly engaged in light-hearted banter, which included a word that one of the participants did not understand. Because the preponderance of definitions done on words or phrases from the source readings across the three classes was surprising, the data were checked again a number of times to make sure that this figure was accurate. These checks identified two more definitions of words that were locally occasioned by the talk itself in Class 3, but which themselves defined material offered as a paraphrase of a phrase from one of the source readings. This lapsed distribution is probably caused by the way in which the tasks were set up by the teachers, who focused the learners' attention on the importance of understanding the source readings and thereby unconsciously directed the learners to organize their talk around these readings.

17Note that this does not mean that these definitions are unplanned, because the recipient design of a conversation ensures that all talk is constructed so as to be relevant to what has just been said and to what a conversationalist demonstrably believes his/her interlocutor(s) to know.

18I do not mean to imply that all spoken definitions must occur as repairs in all conversational contexts in order to count as definitions. What I am claiming is that repair is so relevant to achieving spoken definitions that, at least, in the present database from the three classes, all definitions were achieved as repairs. This may be a function of the way in which the tasks were set up.


occurs. This work usually involves a switch in the focus of the discourse and is often signaled by the presence of such lexical devices as well, you know, I mean and by repetitions and/or reordering of syntactic units; furthermore, nonlexical devices such as silences, pauses, false starts, cut-offs, lengthenings, emphatic stress, hesitation markers such as uh or uhm, gestures, and other paralinguistic phenomena are also frequently present. Finally, participants may modulate repairs by laughing or by using modal verbs or question forms to make themselves sound more tentative, and therefore less challenging, to their interlocutors. The greater the number of these microlevel signals of repair found in a stretch of talk, the stronger the evidence that participants are repairing the interaction.

Notice that if we look at the first 24 lines of Excerpt 5, we can find several of the technical signals of repair mentioned earlier. More specifically, there is a switch in the focus of the discourse at line 1 (where L10 is reading from the article to herself) from reading for information to asking about the meaning of coral at line 3. Interrum silences occur at lines 2 and 4; interturn pauses are found at lines 10, 12, 19, and 21; and intratum pauses occur at lines 7, 9, 11, 13, and 24. We also find multiple instances of various hesitation markers at lines 7, 8, 11, 13, and 18; cut-offs at lines 6, 11, and 22; and syllable lengthenings at lines 7, 11, 13, and 18. Furthermore, L9 constructs her turns at lines 5, 20, and 23 as questions. Finally, notice that repetition occurs not only at lines 13, 14, and 17 but also across excerpts in L10's repeated requests for an explanation of what coral means. Thus, there can be no doubt that these participants are repairing their talk in this excerpt (as in the other excerpts cited).

But how can we demonstrate that repaired conversations are actually a resource for language learning? More specifically, what evidence can we use to show that the conversational resources of turn-taking and repair not only help L10 to understand what coral means but also promote some kind of learning, as opposed to mere short-term, localized understanding?

First, let us discard the kind of evidence that we should not rely on to answer these questions. Notice that L10 repeatedly uses the change-of-state token oh in Excerpts 2 and 3 and actually closes these sequences with the combined tokens oh ok at lines 25 (in Excerpt 2) and 18 (in Excerpt 3).

20Repair is also analyzed in terms of who initiates and who completes a repair, with participants orienting to preference rules that vary according to the turn-taking system in operation at any given time (see the references given in footnote 14). However, there is no structural evidence that by orienting to a particular preference rule, learners are provided with superior opportunities for language learning. Thus, no analysis of the preference structure of definitions is offered in this chapter.

21Notice incidentally that Heritage (1984) demonstrated that one of the environments in which such change-of-state tokens occur is in the context of other-initiated repair, as in these two extracts. The use of such change-of-state tokens provides the type of independent corroborating evidence called for by Jacobs (1986; 1987), evidence that repair is a mechanism that is relevant to spoken definitions.
Because the teacher does not pursue this matter any further, we have reason to believe that he or she interprets L10's use of these tokens (particularly at line 25 of excerpt 2, where she says, "Oh ok oh I see thank you") as evidence that L10 has understood what coral means. But such a reading (on the part of either the teacher or the analyst) of what L10 means by these final tokens would be erroneous, because L10 subsequently initiates two more sequences on the same word. I suggest that in Excerpt 2, L10 is only indicating that she understands that corals are found at the bottom of the sea, and in Excerpt 3, that a definition of what coral means must be provided in the subsequent whole-class activity. But for the reasons already outlined, L10 cannot have fully understood what coral means by the end of Excerpt 3. These "appropriate" answers therefore do not constitute sufficient evidence of understanding (see also Hawkins, 1985, on the difficulty of interpreting learners' appropriate answers).

What evidence, then, should we rely on to demonstrate localized understanding and perhaps even learning? The most important evidence that we can point to in this regard occurs in the marked turns of excerpt 5 (lines 18, 23–30, 32, 35, 38–39). At line 23, in response to L10's question at line 18 asking whether coral is food for fish, L9 replies that coral is like stone. This information seems to trigger a breakthrough in understanding for L10. More specifically, L10 vehemently asserts at line 24 that she has understood (which indicates at least that L10 is rather confident that she has indeed understood this word and is willing to expose herself to a potential loss of face if she turns out to be wrong). At line 25, she independently provides the extra information that coral is very beautiful, which L9 corroborates at line 27. At lines 27, 29, and 30, L9 provides more descriptive information about coral, which L10 overlaps with further assertions that she has understood, at lines 28 and 32. In addition, in the last part of her turn at line 32 and also at line 35, L10 again goes further by providing a translation of this word into Chinese. Because L9 is not a Chinese speaker, she does not understand these translations; consequently, L10 translates the Chinese term back into English by saying "coral" at line 38. The correctness of this translation is corroborated by L11 (who is also a Chinese speaker), who overlaps L10 at line 39 by also saying "coral." We may therefore conclude that L10 has indeed understood what coral means.22

22 Notice that no attempt was made to use self-report procedures or thinking-aloud protocols to obtain the kind of triangulated data (i.e., where the investigator checks his or her interpretations against those of the teacher and student) used by Hawkins (1985) and advocated by Cohen and Olstain (this volume, chapter 8) and Selinker and Douglas (this volume, chapter 6). The reason for this is that ethnomethodology discounts the use of ethnographic data unless there is internal evidence in the conversational data themselves warranting their introduction (Schegloff, 1987). Once a text has been produced, the analyst is said to be as well placed to analyze the data as the individual(s) who first produced the text. In this regard, the production of other texts by the original participants, as they explain or comment on what they really meant in the primary text, only serves to confuse the issue.

However, notice that the analysis offered in this chapter is internally triangulated insofar as three different sets of publicly available evidence in the conversations of the participants are analyzed. As pointed out by Jacobs (1986, 1987), it is the sum of these converging pieces of evidence that allows us to conclude that L10 has understood what coral means. Thus, although there are some rare instances in which ethnographic triangulation might be the most convenient means of understanding what participants meant (see the strawberry example cited by van Lieshout (1988), where the highly idiosyncratic use of this word can be understood only by having access to a previous conversation or, failing this, by asking participants what they meant), such external triangulation is unnecessary in this case because sufficient text/internal evidence exists to make the argument stick.
part statements (Abelson, 1967) like the following: An A is a B, which does C. Thus, coral would fit in the A part of this statement, the information classifying coral as a type of fossil found at the bottom of the sea would fit in the B part of the statement, and the information about coral being an important habitat or home for fish would fit into the C part of the statement. This more formal and complex definition also provides internal textual evidence that L10 has actually learned the meaning of coral.

CONCLUSION

I have demonstrated two things in the preceding paragraphs. In the first part of this chapter, I showed that the nomothetic characterization of descriptive studies as limited and non-theory-generating need not be unquestioningly accepted. Interpretive explanations, inspired by hermeneutics, of ordinary data have considerable theoretical power, which may be generalized beyond the single events on which they are based. And in the second part, I applied CA to explain how the mechanisms of turn-taking and repair may be used as resources for successful second-language learning.24

If these proposals for respecifying SLA studies have any merit, our notions of what constitutes the most valuable kind of research will have to be reassessed. This does not mean that I believe that research in the nomothetic tradition has no value. As I indicated at the beginning of this chapter, I acknowledge that qualitative and quantitative research represent complementary paths to new knowledge about the phenomenon of SLA and that the results of nomothetic research are in many ways impressive. But it does mean that researchers will have to reevaluate whether the current dominance of the nomothetic tradition is entirely beneficial to the field. At a time when epistemology and theory construction are high on the SLA research agenda, I believe that an ethnomethodological respecification of the process of SLA research is of the essence because it simultaneously problematizes the way in which we try to theorize SLA and provides a distinctive methodology for analyzing this complex phenomenon. Thus, in my opinion, the value of ethnomethodology lies in the fact that it provides the field with opportunities for critically reevaluating its dominant epistemology and also for reassessing the criteria by which it measures its level of success as a scientific endeavor. These are matters that a maturing discipline cannot afford to ignore.

24Although I have confined myself here to analyzing how the mechanisms of turn-taking and repair are used to learn new lexis, there is no reason why this approach might not also be used to analyze the acquisition of phonology or syntax. In this regard, see Hatch (1978) and Hawkins (1988) for a discourse-analytic account of the effect of scaffolded talk on the development of syntax-in-conversation. This research could easily be reanalyzed in conversation-analytic terms.

ACKNOWLEDGMENTS

I would like to thank Susan Gonzo, Ron Cowan, Barbara O’Keefe, and John Schumann for their comments on earlier drafts of this chapter. The errors it contains are of course my own.

APPENDIX: Transcription Conventions:
Adapted from van Lier (1988)

T
L1, L2
L
L3?
LL
/yes//yah//ok//
//huh?///oh///

= 24

a) turn continues at the next identical symbol
b) if inserted at the end of one speaker’s turn and the beginning of the next speaker’s adjacent turn, it indicates that there is no gap at all between the two turns

(+)(+)(1)

pauses; (+) = a pause of between .1 and .5 seconds; (+ +) = a pause of between .6 and 9 seconds; and (1) (2) (3) = pauses of 1, 2, or 3 seconds, respectively

?!
ok. now. well.
so, the next thing

er, the:::

emphasis
SYLVIA
°the next thing
... (radio)

((coughs))

((unintelligible))

teacher
identified learner
unidentified learner
probably learner 3 (L3)
several or all learners simultaneously
overlapping or simultaneous listening responses, brief comments, etc., by two, three, or an unspecified number of learners

rising intonation, not necessarily a question
strong emphasis with falling intonation
a period indicates falling (final) intonation
a comma indicates low-rising intonation, suggesting continuation
one or more colons indicate lengthening of the preceding sound
italic type indicates marked stress
capitals indicate increased volume
degree sign indicates decreased volume
single parentheses indicate unclear or probable item
double parentheses indicate comments about the transcript and note nonverbal actions
indicates a stretch of talk that is unintelligible to the analyst
no-
Peter
[s:m]
<hhh>

in-drawn breath
exhaled breath
laughter tokens

REFERENCES


