Learning talk analysis

NUMA MARKEE AND MI-SUK SEO

Abstract

Since the beginning, second language acquisition (SLA) studies have been predominantly cognitive in their theoretical assumptions and programmatic agendas. This is still largely true today. In this paper, we set out our proposals for learning talk analysis (LTA). LTA synthesizes insights from linguistic philosophy, ethnomethodology, conversation analysis, discursive psychology, and the discourse hypothesis in SLA. LTA points to behavioral, process-oriented accounts of mind, cognition, affect, language, and language learning that are agnostic about a priori theoretical claims that such traditionally psychological constructs underlie SLA. Instead, LTA treats these constructs as observable, socially distributed interactional practices. While an ethnomethodological re-specification of SLA studies is a key agenda item of LTA, LTA is also concerned to foster an on-going conversation with all SLA researchers. The paper defines LTA, discusses how the various intellectual traditions it invokes form a coherent whole, provides a sustained, empirical exemplification of how LTA works, and suggests possible areas for future collaboration between behavioral and cognitive SLA researchers.

1. Introduction

Since the beginning, second language acquisition (SLA) studies have been predominantly cognitive in their theoretical assumptions and programmatic agendas (see Selinker 1972, who constitutes an early foundational text for SLA studies). This is still largely true today (see Long and Doughty 2003). However, foundational texts often contain the seeds of interpretations that are diametrically opposed to understandings that, over time, gain conventional currency. For example, let us briefly consider two interpretations of Plato’s account of how Socrates demonstrated to a man named Meno that an uneducated slave he
Numa Markee and Mi-Suk Seo

owned knew Pythagoras' theorem, without having ever been taught this theorem.

Edwards (1997) notes that the Meno text is normally cited as a demonstration of innate knowledge, and as a foundation for Cartesian notions of mind and cognition, and thereby as a key philosophical resource for, e.g., Chomskyan theories of universal grammar, nativism, and mind. However, Edwards shows how the Meno text can also be used to show that “the nature of experience, understanding, and events are accomplished discursively, in how they are described and narrated” (Edwards 1997: 37).

In this paper, we develop our proposals for learning talk analysis (LTA). LTA synthesizes insights from linguistic philosophy (LP), ethnomethodology (ET), conversation analysis (CA), discursive psychology (DP), and the discourse hypothesis (DH) in SLA. LTA points to behavioral, process-oriented accounts of mind, cognition, affect, language, and language learning that are agnostic about a priori theoretical claims that such traditionally psychological constructs underlie SLA. Instead, LTA respecifies these constructs as observable, socially distributed interactional practices. We further argue that if LTA can work up behavioral explanations of how participants do learning that are independent of cognitive explanations, we may then ask: “what is left for cognitive approaches to explain?” We believe that the answer to this question is: “much less than we have historically assumed.”

This answer poses a challenge to all SLA researchers. For behaviorally oriented SLA researchers who are either inspired by or sympathetic to Firth and Wagner’s (1997) call for a respecification of SLA studies (and we include ourselves in this camp), the challenge is this: LTA proposes that there is no need to invoke a priori theories of learning to anchor behavioral accounts of SLA. To do so is to put the cart before the horse. The point of such research is to describe language learning behavior in its own terms, and if any theory eventually emerges from such accounts, it will be a process oriented, member’s theory (Markee 2008).

This position not only suggests that the emerging tendency among some conversation analysts to incorporate various language learning theories into their work as an a priori point of departure (see, for example, Brouwer and Wagner 2004; Hellerman 2007, 2008, this issue; Mondada and Pekarek-Doehler 2004) is premature. But this position also differentiates LTA from sociocultural1 and language socialization theories of SLA (see Lantolf and Thorne 2006; Watson-Gegeo 2004; Watson-Gegeo and Nielsen 2003). We acknowledge the considerable value of the on-going collaboration between CA and sociocultural theory

1. See Edwards (1997: 45–48) for an excellent, nuanced discussion of the similarities and differences between sociocultural theory and DP.
Learning talk analysis

(see Hall 2007; Hellerman 2008), but we also argue that DP is probably an even better partner for CA than sociocultural theory is. This is because CA and DP share the same ethnomethodological epistemology, whereas sociocultural theory does not. Thus, LTA does not assert a priori that intra-mental cognition emerges from social interaction between experts and novices. Rather, it treats this question as an empirical issue. Furthermore, from a methodological standpoint, LTA is self-consciously more cautious than sociocultural theory and shies away from the notion that cognitive processes can be “observed directly in the linguistic interactions that arise among speakers as they participate in problem-solving tasks” (Donato and Lantolf 1990: 85). As our analysis of Fragment (2) shows (see page 52 and Note 6), demonstrating that cognition can be directly observed in interaction is a difficult, perhaps impossible, task.

For cognitive researchers, the challenge is even more complex. As already noted, LTA proposes a major respecification of SLA studies, an idea that has not always been well received in the past (see Gass 1998; Kasper 1997; Long 1997). However, as we show through our empirical work in this paper, what is different about LTA from previous proposals for CA-for-SLA is that it can generate research questions – such as “what do process-oriented language learning rules look like?” – that emerge in the first instance from theoretically unmotivated analyses of language learning behavior. We can then pass these grounded findings on to cognitive SLA for further study (see pages 57–59). Collaboration of this kind would be of interest to cognitive researchers who, whether they agree with LTA’s agnostic stance on cognition or not, are nonetheless open to behavioral accounts of SLA (e.g., Ellis 2003; Larsen-Freeman 2007). We therefore mean to engage all SLA colleagues in a potentially fruitful dialog on these and other important issues.

2. Cognitive perspectives on mind, cognition and language learning

Cognitive SLA’s theories of mind, cognition, language, and language learning are widely known. Consequently, we will review these positions in the briefest terms. Basically, all these constructs are predominantly understood from the perspective of individual, cognitive psychology. In the nativist metaphor of Chomskyan linguistics, at least, language is a formal system of abstract rules, and syntax is an autonomous module of knowledge (competence) that is prewired at birth in the mind-brain of all normal human beings (Chomsky 1965). This language acquisition device (LAD) is the mechanism that enables L1 learners to acquire their native language quickly and efficiently, despite being exposed to input in the environment that is corrupt and underspecified. Universal grammar (UG) has also had an important impact on SLA studies, although
there is considerable ambiguity as to whether (and, if so, when and how) UG becomes available to SL learners (Bley-Vroman 1989).

Other cognitive theories of SLA take different positions from those outlined above. For example, connectionism views language learning as a process of distributed processing, during which the individual’s mind-brain constructs and strengthens ever-expanding networks of connections of different weights among non-representational neural nodes (McLelland, Rumelhart et al. 1986). It is these rule-like networks that do the work of language learning in this model, rather than the LAD posited by UG. In yet another cognitive model of SLA, computational limitations on short-term memory constrain the individual mind-brain’s capacity to process and learn new language (Skehan 1998). In this account, language is described in terms of generative rules that resemble those of Chomskyan universal grammar. These rules are computationally expensive in terms of short-term memory, and are invoked when there is a need to produce novel, complex, or accurate speech. In addition, there are exemplar-based rules, which arrange speech in pre-assembled packets of language. These rules have the advantage of being computationally cheap in terms of short-term memory, and are invoked when there is a need for rapid, fluent talk.

Thus, cognitive SLA is not monolithic and we can expect different answers from supporters of these different theories to the following three questions: what is the exact nature of language rules? Is language learning qualitatively different from other forms of learning? And what is the precise relationship that exists between language and mind? However, all three theories agree that mind is an individual construct. Furthermore (leaving connectionism aside for the moment), they agree that language is a formal object that resides in the individual mind-brain, whose rules are best described independently of social context. And last but not least, they unanimously claim that SLA theorizing and research are psycholinguistic enterprises.

Cognitive metaphors of SLA have obviously been productive during the last 30 years. However, we believe that, even in their own terms, their intellectual scope is unnecessarily narrow. The cognitive scientist Edwin Hutchins (1995) argues that cognition and learning are constructs that go beyond the individual. More specifically, Hutchins suggests that the relevant unit of analysis is not the individual, but whole ecological systems, such as, for example, a navigation team and the instruments it uses to get a naval vessel from A to B. In this distributed version of cognitive science, individuals are members of larger ecosystems of contributing agents and technologies. This position contrasts sharply with the highly individualistic version of cognitive science that is still the norm in cognitive SLA (see, for example, Long and Doughty 2003). We maintain that this individualistic perspective is excessively restrictive or, worse still, simply out of date.
3. Learning talk analysis

We first define LTA and other key terms that are used within this definition. We then unpack the constituent elements of this definition and explain how these elements are interrelated as a coherent intellectual whole. Next, we illustrate how LTA works empirically by working up an analysis of language learning behavior. Finally, we sketch out some possible areas of collaboration between LTA and cognitive research.

3.1. Definitions of key terms

As already noted, LTA draws to some extent on the work of linguistic philosophers interested in ordinary language (e.g. Wittgenstein). However, LTA’s main source of intellectual inspiration derives from ET and its offshoots: CA and DP. Within SLA, LTA also draws on the DH. LTA may be defined as follows:

1. LTA is programmatically concerned with developing fine-grained, emic accounts of naturally occurring language learning behavior that occurs in real time. However, written cultural artifacts that participants observably talk into relevance may also be analyzed (sources: CA, DP).

2. Language learning behavior is one of the many public actions of everyday life that members routinely achieve in and through talk (sources: LP, ET, CA, DP, and DH).

3. LTA is agnostic as to whether cognitive and affective constructs, structures and processes underlie language learning behavior. Instead, LTA invokes the insight that the natural habitat of grammar (however specified) is talk-in-interaction (source: CA). By extension, the same is true for mind, cognition and affect (sources: CA, DP and DH). Thus, LTA is concerned with how members routinely:
   a. Co-construct the grammar of interaction as an emergent phenomenon (sources: CA, DH)
   b. Use psychological metaphors during talk (source: DP)
   c. Manage cognitive and affective topics during talk (sources: CA and DP)
   d. Formulate so-called psychological scripts (such as a restaurant script, or a conversation partners script) in and through talk (sources: DP and DH).

ET proposes that all competent members of society have at their disposal a methodological resource for analyzing and making sense of everyday life: the observable behaviors which participants engaged in any activity routinely

---

2. Language learning behavior is not an a priori, etic construct. More specifically, language learning behavior involves the social construction of talk by members who observably orient to learning work as the principal business of the current interaction.
display to each other to understand the world around them. By extension, this methodological resource is also available to professional analysts because they too are members of society (Garfinkel 1967). This methodology entails analysts adopting an agnostic position of “ethnomethodological indifference” (Garfinkel and Sacks 1970: 63) toward etic (i.e., researcher-centric) explanations of human behavior as a basic guiding principle for their work. To recycle the example of the Meno text, ET does not necessarily deny the existence of mind, cognition, affect, etc. as psychological constructs. But ET is critical of the methodology used (here, the Socratic method) to demonstrate innateness. As we have already seen, the Meno text can be used just as convincingly to demonstrate how practical, common sense reasoning works as everyday practice.

CA, which is the most important offshoot of ET, is an approach to the analysis of talk-in-interaction (i.e., ordinary conversation and institutional talk; see Schegloff 1987) that focuses primarily on how the normative rules that describe turn taking, sequence organization and repair practices shape members’ co-construction of talk in real time (Schegloff 2007). More specifically, speakers orient to a set of recursive turn taking practices (Sacks, Schegloff and Jefferson 1974: 704) which specify that, whenever an initial transition relevance place (TRP) of an initial turn constructional unit (TCU) is reached, and a “current speaker selects next” technique is used, then the selected speaker has the right to the floor. If a “current speaker selects next” technique is not used, then the current speaker may, but need not, continue, in which case the first person that starts talking takes over the floor. If none of these choices are exercised, the rule set re-applies at the next TRP until a transfer of speakership is achieved.

These practices involve speakers parsing the emerging grammatical structure (however this is specified) of TCUs, and use this on-line information to project when they may start their own turns-at-talk. Thus, in a rather literal sense, participating in interaction involves “doing grammar on the fly.” These turn taking practices also interact with an independent set of repair practices, which serve to fix any emerging trouble that may occur during talk (Schegloff, Jefferson and Sacks 1977). More specifically, these repair practices are described in terms of who initiates and who completes a repair. So, repairs may be self-initiated and self-completed, or they may be self-initiated and other-completed. Conversely, they may be other-initiated and self-completed, or other-initiated and other-completed. In addition, repairs are analyzed in terms of their sequential position in relation to the original trouble source. Thus, repairs may occur in first, second, third, or (more rarely) in fourth position. Going back to our “do-

---

3. Note that this methodology has now been extended to SLA data. With such data, extra care must be taken to ground the analyses in the observable behaviors being studied.
ing grammar on the fly” analogy, these repair practices may also change the unfolding grammatical shape of sentences. For example, speakers use repair to: 1) convert sentences into subordinate clauses; 2) transform questions into assertions; and 3) change *wh*-questions into *yes-no* questions (Schegloff 1979). Finally, as Schegloff (1996) argues, the natural habitat of grammar is interaction. Such a grammar is “produced piece by piece, incrementally, through a series of ‘turns so far.’ These features support the openness of talk-in-progress to considerations of interactional import and reactivity, recipient design, moment-to-moment recalibration, reorganization and recompletion, and to interactional co-construction” (1996: 55–56).

CA has also explored topics such as whether, and if so, how, CA might contribute to the development of a theory of mind (Schegloff 2006a) and other cognitive issues. Whereas such matters have been a secondary interest for CA, DP has always been programmatically concerned with issues of how cognition and affect are done in and through talk (see the collections edited by te Molder and Potter 2005 and van Dijk 2006 for contrasting CA and DP perspectives on these issues). DP is closely allied to CA epistemologically, and has come over the years to use its analytic techniques as its principal methodological tools. Nonetheless, DP has a different research agenda from CA. As Edwards (2006) notes:

DP approaches the topics of cognition, mental states and psychological characteristics as matters under active management in talk and text. The start point is everyday discourse considered as a domain of social practice. Mostly, it is talk-in-interaction (Schegloff 1987), but written text is analysed too. The key to DP is that it is primarily a way of analyzing talk and text. It does not start with psychological questions, and does not offer a rival theory of mind. Not does it deny the reality and importance of subjective experience. Rather, DP rejects the assumption that discourse is the *product* or *expression* of thoughts or intentional states lying beneath it. Instead, mental states, knowledge, thoughts, feelings, and the nature of the external world, figure as talk’s topics, assumptions and concerns (Edwards 2006: 41, emphasis in the original).

It is in these varied senses that the natural habitat of grammar, mind, cognition, and affect is talk.

As we have seen, CA, DP and ET contribute foundational insights to LTA. However, none of these disciplines has anything to say about language learning. For this, we turn to the DH, which states that “one learns how to do conversation; one learns how to interact verbally, and out of this interaction, syntactic structures are developed” (Hatch 1978: 404). To update the DH, the turn taking, repair and sequence organizations described above form the bedrock of discourse-based language learning behaviors in SLA. Furthermore, by invoking the DH as an important source of ideas for LTA, we set up a dialog between
the DH and DP on how “psycholinguistic” scripts are actually done by participants.

3.2. Unpacking the constituent elements of LTA

Wittgenstein’s ideas about language are divided into two distinct, diametrically opposed bodies of work. In his early work (Wittgenstein 1961), he employed mathematics to analyze language in a highly abstract form. But in his later work (Wittgenstein 1953, 1981), he repudiates these ideas. Rejecting the notion that language is the property of a socially marooned individual, Wittgenstein argues that language is a public, not a private, matter. Furthermore, when he says that “words are deeds” (Wittgenstein 1980: 46), he is arguing that meaning in ordinary language is constructed in and through use.

These ideas tie in closely with ET’s interest in documenting how mundane action is achieved as behavior. Speaking to this issue, Garfinkel (1963: 90) famously asserts, “... there is no reason to look under the skull since nothing of interest is to be found there but brains.” This statement has set off a fierce debate within ET, CA and DP as to whether these disciplines should be programmatically anti-cognitive or merely non-cognitive. Writers such as Button, Coulter, Lee and Sharrock (1995), Coulter (1979), and Lynch and Bogen (2005) argue in favor of a strong anti-cognitive stance, which involves “investigating nominally ‘cognitive’ themes without trading in mentalistic notions of cognition” (Lynch 2006: 95). So, for Coulter (1999, 2005), any concession to cognitivism gives away too much, and he criticizes both CA and DP writers for being too accommodating to cognitive theories of mind (see Potter and Edwards 2003, for a rebuttal of this criticism). In contrast, CA and DP adopt a non-cognitive – or, perhaps more accurately – variously agnostic stances toward the matter of how cognition and affect may be manifested in talk-in-interaction.

Where does LTA situate itself in relation to these debates? We acknowledge that, as Kitzenger (2006) notes, there is currently some tension between CA and DP concerning what precisely can and cannot legitimately be said about cognition by CA and DP practitioners (see, for example, the criticism of Drew (2005) by Potter (2006)). However, we embrace this tension, because it forces LTA to be methodologically self-conscious about how it makes its arguments. This is why we talk about language learning behavior rather than language

---

4. The apparent discrepancy in the publication dates of Wittgenstein’s early and later work is accounted for by the fact that his work was translated into English at different times after his death.
To explain what we mean by this phrase, we update Markee’s (2008) description of how specific types of actions cluster together to achieve language learning behaviors that have very specific characteristics (the material in square brackets represents updated material). Language learning behaviors are massively achieved as:

- Repair sequences that may contain initial statements of non-comprehension, and/or emphatic assertions of understanding (these verbal behaviors are often accompanied by smiling, clapping, and embodied actions such as thinking gestures and pointing to information in written texts);
- Changes of epistemic state, including the use of tokens such as “oh” (Heritage 1984, 2005) in ordinary conversation or other speech exchange systems that are closely related to ordinary conversation;
- Participants independently volunteering new information that connects the learning object to practices or knowledge that are already part of their interactional repertoires;
- Translation from one language to another.
- When participants observably use combinations of such actions and also produce new language (which may, at the same time, contain talk that recycles verbatim quotations of material lifted from previous turns: see Markee 1994, 2000), participants achieve an interactionally based “transparency of understanding” (LeBaron and Koschman 2003: 121). That is, participants have observably displayed that they have learned new language in the short and/or long term.

This empirically derived definition of language learning behavior allows us to align ourselves with the agnostic stances of CA/DP discussed previously. We reject the potential criticism that, in so doing, we are giving away the intellectual store before it has even been built. Instead, we rely on a strict interpretation of CA’s emic principles and methods (see Markee and Kasper 2004) to provide LTA with the epistemological and methodological rigor that is required to achieve two programmatic goals: (1) respecifying SLA studies along behavioral lines (Firth and Wagner, 1997; Kasper 2006, 2008, this issue; Lafford 2007; Markee 1994, 2000, 2008; Mori 2002, 2004; Mori and Hasegawa, this issue), while (2) retaining the ability to engage the broader SLA community in a productive dialog.

LTA is also agnostic about having to choose between “acquisition” and “participation” metaphors of learning (Sfard 1998). CA/DP is a powerful resource for showing how participants construct both metaphors of learning. For exam-

5. Note that we are not making an a priori claim that language learning occurs only in and through language learning behavior. As an anonymous reviewer pointed out to us, language learning can occur through “self-study (e.g. memorization), private speech, and other forms of ‘non-talk’ (or non-observably talk) sorts of language learning behaviors [sic].”
ple, Fragment (1) below shows how SJ uses a “putting new knowledge in the head” gesture in line 43 – which invokes an individual acquisition metaphor – as a resource for making a summative claim of learning a rule of sentence level grammar. This claim of understanding is accompanied by laughing and smiling behaviors in lines 42 and 43, which suggest that both participants are pleased to have reached an apparently satisfactory conclusion.

(1) (Individual acquisition and collaborative participation metaphors / TJ&ST / 2001)

38 TJ: [we don’ say many. many food many water= 
39 SJ: =much time?
40 TJ: right. much or a lot of.
41 SJ: °yeah.°
42 TJ: [yeah. hm heh heh heh heh
43 → [((ST smiles and taps his head with finger))]

Even from the limited data we have shown here (see Fragment (2) for the complete version of this fragment), it is also clear that SJ’s claim of understanding in line 43 is the result of collaborative work between SJ and TJ in lines 38–42. This talk invokes an equally member-relevant social participation metaphor of learning. Consequently, these two metaphors need not be mutually exclusive. Along the same lines, Fragment (1) also shows that language learning behavior is socially distributed activity that happens in the inter-subjective space between participants. But SJ’s pointing gesture and smiling behavior in line 43 show that he has undergone an individually relevant “change of state” (Heritage 1984). This public change of state further “does” achieving a pedagogically desirable outcome in this institutional speech exchange system. For these reasons, we have no a priori opinion as to whether mind, cognition, affect, and language learning are individually or socially distributed behaviors. Without contradiction, the data suggest they are both.

Finally, how can LTA synthesize insights from the DH and DP to help us rethink our understanding of psychological scripts? Hatch (1992) and Edwards (1994, 1995, 1997) both use the classic example of a restaurant script to explain...
Learning talk analysis

how these work. The psycholinguistic approach to script theory argues that we organize our expectations about what is supposed to happen in particular situations (such as going to a restaurant) according to a set of instructions that specify not only what will happen but also the order in which specific actions associated with a given script will happen. For example, diners in the USA who go to an upscale restaurant in a large city might expect to:

1. Be greeted by a host or hostess, whose duties include asking the guests how many people are in the party, seating them, and distributing menus;
2. Be greeted by a server, who explains what chef’s specials might be available, and who then takes drink orders;
3. After an appropriate period of time has elapsed, during which guests consider what they want to eat and make small talk, the server serves the drinks, and then takes the dinner order;
4. After another appropriate period of time has elapsed, the server brings the courses as appropriate (for example, appetizers, salad, and then the main course);
5. As guests finish each course, the server or other staff whose job it is to bus tables will clear plates;
6. Once the meal is finished, the server presents dessert menus, and cycles 4–5 are repeated;
7. Once the meal is finished, the server presents the check;
8. The server takes the payment from the guests, and thanks them;
9. It is customary for guests who go to restaurants in large cities to tip the server between 15–20 % of the total bill. Smaller tips may be appropriate in smaller cities.

Minor variations are theoretically possible within scripts, and these may be updated in light of new information. So, restaurant scripts in France differ from American scripts, and these differences must be learned. While such idealized scripts may sound familiar, the way in which restaurant scripts actually work in ordinary conversation often provides countless exceptions to these rules (see Edwards 1997: 147–149). For this reason, as Potter and te Molder (2005: 31) note in their commentary on Edwards’ work, instead of proceeding from “Script W (which refers to ordered and orderly features of the real world) to Script PC (which refers to features of an individual’s perceptions and cognition) to Script D (which refers to the way events are described as orderly, or to departures from the standard order),” DP reverses the canonical order of psychological script theory to Script D → Script PC → Script W. This is the procedure we adopt in the empirical example of LTA in action that follows below.
3.3. An empirical example of LTA in action

Our data include three fragments of language learning behavior that are taken from ESL tutor-tutee interactions. All three fragments involve the same two individuals. Our overall purpose is to develop a sequentially based CA of language learning behavior and its possible consequences. In so doing, we develop a Script D account of ESL tutor-tutee talk. As we work up these cognitively agnostic analyses of learning talk, we discuss what LTA can and cannot say about mind, cognition, affect, and language learning as individually and socially distributed processes.

Let us begin with what we can say. In Fragment (2), we identify a conversational environment in which cognitive work of various kinds is likely to occur: a zone of interactional transition or ZIT (Markee 2004). Misunderstanding is one of the common phenomena that occur in a ZIT that is constructed when participants switch from one speech exchange system to another. In our data, the tutee, SJ, sometimes misunderstands the actions of TJ, his tutor, because they engage in two different activities (i.e., ordinary conversation and pedagogical activity) on a moment-by-moment basis during the same speech event. As a ZIT is constructed, SJ sometimes mistakes TJ’s pedagogical action for a conversational response. More specifically, in Fragment (2) (which is the full version of Fragment (1)), SJ has been talking about an internet-based English program that he has recently been using to improve his listening comprehension.

(2) (many time / TJ& SJ / 2001)

01 SJ: tch! because h uh: (1.2) t! diffi- u:::m
02 → yeah many ti- I- I (1.0) I need (.) many time,
03 → (1.0) I need many time to hhh hear e:n I:
04 (.) exercise thru pronounce.
05 TJ: → mm [hm, [a lot of time.
06 → [((TJ raises his index finger and
07 [points toward SJ repeatedly with
08 [eye gaze on SJ)]

Nuna Markee and Mi-Suk Seo
Learning talk analysis 49

09 SJ: []° so:
10 SJ: → yeah. many time. a:nd=
11 TJ: → a lo: t of.
12 SJ: hm?
13 TJ: → a lo: t of time.
14 (1.2)
15 SJ: (uh s-) pronounce?
16 TJ: → you need a lo: t of time.
17 SJ: [(uh)
18 TJ: → [not many (. ) time.
19 (0.8)
20 SJ: one uh one hour (0.2) more than one hours
21 one hour.=
22 TJ: → right. so: a lo: t.
23 (1.0)
24 TJ: a lo: t of (. ) time.
25 SJ: you- you mea:n I must do (0.2) two hour to
26 three hour[rs]?
27 TJ: → (no: no no no.
28 → ih heh you said many. I need many,
29 SJ: mm [hm
30 TJ: → time? you should say I need a lo: t of (. )
31 time.
32 SJ: → [eh hyeah. a lo: t of ti[me.
33 → [(SJ snaps his fingers)]
In lines 1 through 3, SJ is trying to explain that it takes a lot of time for him to understand each exercise segment in the program. In responding to SJ’s turn, TJ produces a response token in line 5, acknowledging SJ’s prior talk, and then quickly adds another increment which corrects SJ’s linguistic error. In so doing, TJ constructs a ZIT during which he moves from an orientation to ordinary conversation to pedagogical talk. TJ marks the boundary between these two speech exchange systems with his pointing gesture described in lines 6 through 8 and the accompanying frame grab. TJ’s gesture, and the way he also lifts his head back at the same moment, seem to say: “pay attention to what I am saying.”

In terms of ordinary conversation, TJ’s immediately following error correction in the last part of line 5 is unusual, because his response token “mm hm”...
looks similar to what Schegloff (1981) calls a “continuer.” In ordinary conversation, continuers are found in contexts where the recipient passes up the opportunity to initiate any action, including repair, and other-correction is usually done right after the trouble source turn is finished. However, in the context of a ZIT, repair is sequentially relevant as an institutional, specifically pedagogical, response to SJ’s error. But SJ does not immediately understand that TJ has switched from ordinary conversation to pedagogical talk. Seemingly misled by TJ’s initial response token, SJ first overlaps the second part of TJ’s “mm [hm,” by saying “[so::” in line 7. This shows that, at this point in the talk, SJ is still orienting to the norms of ordinary conversation. In line 10, SJ treats TJ’s slightly transformed error correction as a repetition of a prior utterance with downward intonation. This prosodic contour is sometimes used to indicate the recipient’s acknowledgement, confirmation, agreement, or alignment with prior talk (Schegloff 1997). More specifically, in line 10, SJ first produces a response token “yeah,” which may be understood as projecting a more explicit acceptance of correction, such as a repetition of TJ’s correction. However, after this response token, SJ repeats his prior problematic utterance (“many time.”), thereby misaligning with the sequential implicativeness of TJ’s prior turn. In sum, in the sequential context of this ZIT, the two participants display divergent understandings of what action is being performed through TJ’s utterance, “a lot of time.” (i.e., error correction vs. acknowledgement), and thus which speech exchange system they are orienting to as being locally relevant at that specific moment in the interaction.

In line 11, in response to SJ’s misunderstanding, TJ starts to repair his talk, but this time he repeats the main part of the trouble source in line 5 by stressing and stretching the sound to make it clearer, thereby directing SJ to notice the correction. However, in line 12, SJ initiates repair (“hm?”) in a form that is similar to what Drew (1997) calls “open class repair initiators”, possibly indicating he is confused about TJ’s intended action. In line 13, TJ repeats the trouble source again (a common reaction to a non-specified repair initiator), but this time, he adds a noun component, “time”, producing a full noun phrase as he did in line 5. By responding to SJ’s repair initiation with a repetition of his prior correction (lines 5 and 10) instead of directly clarifying his intended action, TJ seems to be giving SJ one more chance to notice the correction himself. In line 14, however, the pause of 1.2 seconds and the conversely noticeable absence of a change of state token such as “oh” (Heritage 1984) suggest that SJ is still struggling to figure out what TJ is doing. Finally, in line 15, SJ says “uh s-pronounce?” which may represent an attempt on SJ’s part to reinstate the course of action that TJ interrupted in line 5.

In lines 16 and 18, TJ deletes the topical relevance of SJ’s prior turn by continuing to provide grammatical feedback (“you need a lot of time.” and “not many time.” respectively) on SJ’s problematic talk in line 3. However, the
way TJ constructs this alternative rendering seems further to contribute to SJ’s confusion. That is, instead of saying, “You should say, I need a lot of time, not many time,” he says, “You need a lot of time,” as if he were advising SJ to spend a lot of time practicing pronunciation (i.e., more than SJ had been). SJ may thus be hearing TJ to be saying that “a lot” is greater than “many” in English. In any case, SJ displays his understanding in lines 20 and 21 that TJ is talking about the amount of time for practicing pronunciation, not the linguistic form. But TJ then adds to SJ’s confusion when, in line 22, he accepts SJ’s candidate understanding (“right.”), while also repeating the trouble source, signaling that the problem has still not been resolved. In lines 25 and 26, SJ proffers the candidate understanding “you- you mean I must do (0.2) two hour to three hou[r]s?”

Finally realizing that the communication is going seriously awry, TJ directly disconfirms SJ’s candidate understanding in lines 25 and 26 by saying “no: no no no.” in line 27. As Stivers (2004) observes, multiple sayings are used in sequential contexts where a participant tries to stop an in-progress course of action that the prior speaker has insisted on pursuing. Thus, TJ’s multiple “no’s” indicate his determination to stop SJ’s persistently misaligned actions. TJ prefaces his talk in line 28 with two laughter tokens, which acknowledge his recognition of the source of SJ’s misunderstanding (i.e., TJ’s own pedagogical interventions). TJ then explicitly clarifies his intended action in the original trouble source turn in line 5, thereby correcting all of SJ’s misaligned candidate understandings. In line 32, with an “oh-prefaced response” (Heritage 1984) that is accompanied by a snap of his fingers, SJ displays a change in cognitive state: from not-understanding to a claimed understanding of TJ’s intended action, and to a claimed understanding of his own incorrect linguistic performance. Note also the quality of SJ’s turn in line 39: a latched “=much time?”. The haste in SJ’s delivery suggests that he is eager to demonstrate his understanding to TJ. Furthermore, through his gesture in line 43 (tapping his head with finger), SJ makes the claim that he has either: 1) understood what the complex repair sequences in this fragment were all about; or 2) made an error that he should not have made because he already knew the grammar rule. The first interpretation assumes that the rule is new to SJ, while the second assumes that this talk is a review of a rule he already knows.

Before we continue our analysis, note how difficult it is to identify when important cognitive moment(s) occur in talk-in-interaction. As we have already mentioned, there are important disagreements within ET, CA and DP about whether to be atheistic or merely agnostic about the theoretical status of cog-

---

6. This is the core of our critique of Donato and Lantolf (1990).
Learning talk analysis

nition. But even if we situate ourselves on the agnostic end of this continuum, as Heritage (2005: 194) points out:

Oh merely enacts ‘changes of state’, but whether the change of state enacted involves a change of attention, memory orientation or knowledge is left to be inferred from the context in which the oh is produced. Similarly, whether the knowledge accrued in a particular informing that is acknowledged with oh is significantly new . . . or merely incrementally confirmatory . . . is likewise inferred, though intonational and other resources may be used to discriminate the weight and unexpectedness of the information involved . . . (Heritage, 2005: 194).7

We therefore refer readers back to our previous observation that language learning behaviors potentially cluster together and suggest that any claim that a significant change of state has occurred should preferably be buttressed by multiple, independent tokens of evidence such as the ones listed in our definition of language learning behavior.

So far, we have seen how an interactionally achieved display of socially shared cognition that is designed to address SJ’s emerging language learning needs as an individual is co-constructed. But can we use such data as a source of empirical evidence to substantiate the claim that SJ may be making progress in language learning? Fragment (3), which occurs about 60 seconds after Fragment (2), provides an observable clue. In this fragment, SJ is trying to explain why he wants to have tutoring sessions with TJ more often.

(3) (much time / TJ& SJ / 2001)

01 SJ: the:: reason uh: I::: (0.2) "um uh:::.hh I
02 want to: (0.2) study with you, or another
03 ti:me.hhh (0.5) the:: Amy- ((name changed))
04 uh::: (0.5) uh uh-bis- uh:::.hh the reason I
05 want to: study with you,=
06 TJ: =mm h[m,
07 SJ: [(I:) study:: more times,
08 TJ: yeah.
09 SJ: → mhm:::. if I (1.5) spea:k many time,
10 → [um (. ) much time,
11 → [(SJ points to TJ with left index finger and
12 → [then smiles at TJ])

7. We are grateful to an anonymous reviewer who encouraged us to underline the fact that on-going tensions exist within ET, CA, and DP on this issue.
In line 9, as he provides an account for his request, SJ uses the ungrammatical form, “many time” again, which TJ had problematized and corrected in Fragment (2). However, this time, SJ quickly self-correction his linguistic error. The change from other-correction to self-correction not only enacts a change in the preference structure of SJ’s talk (self-correction is preferred over other-correction), it also suggests (if we may explicitly invoke an individualistic, acquisition metaphor of SLA at this moment in the analysis) that SJ’s behavior indicates an increased level of individual control over the target form. At the

---

8. As Junko Mori (personal communication, August 8, 2008) points out “what is important socially is that [SJ] does explicitly display his learning to [TJ], so the display is socially motivated.”
same time (if we may just as quickly switch to a more participatory metaphor of learning,) note how, right after he makes the error in lines 10 through 12, SJ points to TJ and self-corrects with a smile while gazing at TJ. Through his gesture, SJ invokes their shared interactional history regarding the target item and portrays his correction or improvement to be a consequence of the collaborative work they engaged in during Fragment (2). Notice also that TJ joins SJ’s laughter in line 13 with laughter of his own in line 14. By documenting how these clusters of language learning behaviors in lines 9–14 interact, we are not only developing an analysis of remembering as social behavior that is agnostic about the status of memory as an individual cognitive state, we are also beginning to amass a range of independent, convergent evidence that SJ’s formulation of the countable/non-countable rule is observably undergoing some kind of change.

If we may briefly make an etic comment about the talk-so-far, although SJ’s form is still not perfect considering the pragmatic content that he seems to be trying to convey (i.e., “If I spend a lot of time speaking English, I will learn English better”), it seems that SJ is indeed making progress in choosing the right form between two contrastive forms: “many” vs. “much.” However, he does not use the correct form “a lot of” that was provided by TJ in Fragment (2), but the form “much.” This may be due to the fact that in line 40 of Fragment (2), TJ says “right. much or a lot of.” This may have suggested to SJ that there is no semantic or pragmatic difference between the two expressions.

Let us now move on to Fragment (4), which includes another instance of short-term language learning behavior. Fragment (4) occurs about 2 minutes after Fragment (3). In this fragment, SJ, is trying to apply the linguistic knowledge interactionally occasioned in the prior repair sequences to a similar context.

(4) (some time / TJ&SJ / 2001)

01 SJ: → uh:: now- v’yeah v’ and now (0.2) tch! I have
02 → some [times.
03 (0.2) / (SJ moves his left hand in the air in
04 [front of his torso])
05 (0.2) / ((SJ continues hand gesture))
06 SJ: → um I h- I ha- I ma- I can make some times uh
07 → uh: [some time?
08 → [((SJ puts his left hand in the air in
09 [front of his torso, furrowing his
10 [eyebrows and gazing at TJ])}
In lines 1 and 2, while describing the changed situation in his availability for tutoring sessions, SJ uses the plural form “times” to describe an amount of time. In lines 6 and 7, after several attempts at self-repair, he completes his turn, but repeats the same error. However, he quickly problematizes his own utterance and produces an alternative form. He invites TJ to confirm the accuracy of this alternative form by using upward intonation that is choreographed with appropriate bodily conduct, as shown in lines 8 through 10. Subsequently, TJ confirms SJ’s candidate formulation by repeating it with downward intonation. In line 17, SJ reproduces his original sentence in an accurate form one more time.

Let us now look at SJ’s attempt at self-repair in this fragment from the more global perspective of the behaviors that are instantiated in Fragments (2)–(4). In Fragments (2) and (3), SJ and TJ have engaged in a range of increasingly complex, collaboratively achieved language learning behaviors, during which SJ observably begins to make distinctions about how expressions that describe a large amount of time, such as “a lot of time” and “much time” work in English. Now, in Fragment (4), he is faced with a situation in which he wants to describe a smaller amount of time. Although he and TJ did not explicitly talk about “time” vs. “times” in Fragments (2) and (3), in Fragment (4), SJ seems to be trying to: 1) develop a grammar rule (however technically specified) which incorporates evidence about how the countable/non-countable rule for “a lot of time” and “much time” works; and 2) extend the scope of this rule to a new learning object “some time/some times.” This new rule seems to be concerned with general as opposed to specific amounts of time.
3.4. Some possible areas of collaboration between LTA and cognitive research

In this section, we explore some possible areas of collaboration between LTA and cognitive research. In addressing our remarks to cognitive researchers, we do not mean to imply that there is no basis for collaboration with our behaviorally-oriented colleagues. Far from it; we anticipate plenty of discussion with these colleagues, and welcome their comments and critiques of LTA. But the goal of starting a dialog with cognitive researchers is likely much more difficult to achieve. And if such a discussion does not ensue (whether as a result of our or others’ efforts), behavioral researchers (including ourselves) and cognitive writers are likely to speak past each other, if we speak to each other at all (Larsen-Freeman 2007).

To summarize the argument so far, we have shown that LTA can develop highly detailed, emic accounts of how SJ and TJ do an ESL tutor-tutee Script D. As they enact this script, they use the resources of turn taking, repair, and sequence organization to co-construct cumulative language learning behaviors in the short term. To the extent that these normative, members’ rules provide an accurate account of how these participants accomplish everyday actions, it seems that these practices have a high degree of psychological reality for members (Potter and te Molder 2003). As SJ and TJ construct the language learning behaviors that we have analyzed, they not only display an orientation to grammar as an emergent interactional phenomenon, but also to grammar as the specific topic of their talk, as when they try to construct a grammatical rule that accounts for countable/non-countable nouns. Finally, note that these behavioral accounts are independent of cognitive accounts of SLA, and are agnostic as to whether cognition underlies language learning in the first place.

Now, let us examine what LTA cannot say on the basis of these analyses, as this enables us to return to the question we asked at the beginning of this paper: “what is left for cognitive approaches to explain?” We do not claim, for example, to have demonstrated that SJ has figured out how the grammar rules (however technically represented) that he and TJ invoke during the course of their talk actually work (he most likely has not). Nor do we claim that we have evidence of long-term learning (though see Hellerman (2007, 2008), Markee (2008), and Young and Miller (2004) for preliminary demonstrations of how longitudinal work in CA-influenced SLA studies might be carried out). But most importantly, if we are to go beyond our natural constituency of behaviorally-oriented SLA researchers and engage our cognitive colleagues in a worthwhile discussion, we again need to point out – without any apologies – that LTA cannot say how grammatical rules or rule-like processes are represented in the individual mind (assuming that such a construct exists on an a priori basis). That being said, we believe that the analyses we have presented
in this paper may be used by colleagues who are interested in construction grammar, learning constructions, associative learning, emergentism, and connectionism (see for example, Ellis 2003, who provides a comprehensive review of these matters).

If the purpose of a collaboration such as this is ultimately to develop a comprehensive theory of mind (ToM) (see, for example, Enfield and Levinson 2006) – and note that this represents a particularly ambitious level of collaboration, and that more modest proposals for collaboration may be more appropriate, particularly in the beginning stages – then it must be accepted that the study of interaction as described in this paper is foundational to such an enterprise.

Speaking to this issue, Schegloff (2006a) proposes three candidate universals about interaction that would have to inform a ToM:

1. Whether in turn-taking, repair or sequence organization, the default form is the minimal form;
2. The concept of “nextness,” or next-prior positioning (for example, next turn, next elements or increments in TCUs, adjacency pair sequences, and the placement of first position repairs) is crucial to understanding the organization of talk;
3. There is a preference for progressivity, so that any behaviors that interrupt progressivity (for example, pauses, cut-offs, in-breaths, and other signals of incipient repair, particularly other-initiated repair) are analyzable for their potential import as resources that members use to produce and recognize what is going on.

Similarly, Schegloff (2006b: 154) suggests that “any model of processing for interaction should be designed for ‘multiple passes’ . . . and that turns out to mean ‘multiple passes’ for each order of organization [i.e., turn taking, repair and sequence organization] that is inescapably implicated whenever ‘talk in interaction’ – actual or potential – is the state the participants find themselves in.”

We would like to build on these ideas to suggest that, even if our cognitive colleagues limit themselves to a concern with the acquisition of syntax, semantics, morphology, phonology, lexis, etc. – which LTA does not – they will have to develop rules that incorporate findings such as these into their formal representations of such rules. So, for example, the countable/non-countable rule invoked by TJ and SJ in our data would minimally have to include a process-oriented component that includes Schegloff’s candidate universals of interaction and the notion of multiple passes in order to adequately describe the kind of (cognitive?) work that participants do as they attempt to get something (e.g., language) right both in the moment and over time.
4. Conclusion

In this paper, we have proposed a methodology for analyzing language learning behavior that draws on Wittgenstein’s later work, ET, CA, DP, and the DH in SLA. This methodology, which we have called by the (hopefully) theoretically neutral term “learning talk analysis”, is agnostic as to whether cognition and affect, etc. underlie language learning, or whether these traditionally cognitive constructs are just manifested in or through talk-in-interaction. Whatever the truth(s) of this matter, we have shown that LTA is capable of developing complex analyses of language learning behavior that are independent of cognitive models of learning, and yet have a great deal of psychological reality for participants. We have further elaborated on the circumstances under which LTA seeks to engage behaviorally oriented and cognitive colleagues in a theoretical and empirical dialog that may potentially transcend the current stale mate that still largely exists between these two important constituencies in SLA, even if (or, just as likely, when) we may disagree with each other. To put this more colloquially, we have kicked the hornet’s nest; let the arguments begin!

University of Illinois at Urbana-Champaign
<nppm@uiuc.edu>
California State University, Sacramento
<mseo@csus.edu>

Appendix: Transcription conventions

[ ] A left bracket indicates a point of overlap onset.
= When two lines by different speakers are connected by equal signs, it indicates a latching. When two lines by the same speaker are connected by equal signs, then it indicates a single continuous utterance with no break or pause.
(0.5) Numbers in parentheses indicate silence, represented in tenths of a second.
(.) A dot in parenthesis indicates a micro-pause; less than 0.2 second.
. The period indicates a falling or final intonation contour.
? A question mark indicates rising intonation.
, A comma indicates continuing intonation.
! An exclamation mark indicates strong emphasis with falling intonation
↑ Arrows indicate shifts into especially high pitch.
word Underlining indicates stressed syllables
:: Colons indicate the prolongation or stretching of the sound just preceding them.
○○ The degree signs indicate the talk between them is soft or quiet.
- A hyphen after a word or part of a word indicates a cut-off.
⟩⟨ The combination of more than and less than symbols indicates a rushed talk.
hhh This sign is used for hearable aspiration (e.g., breathing, laughter).
.hh This sign is used for inhalation.
(( )) Double parentheses are used to indicate transcriber’s descriptions.
( ) Single parentheses indicate uncertainty on the transcriber’s part.

References


Numa Markee and Mi-Suk Seo


