〔Educational Study〕

Looking for life in scripted EFL dialogs: An experiment in conversation analysis, interaction, and learning

Don Carroll†

—— Contents ——

1. Introduction

Scripted dialogues might seem like an unlikely place to go looking for evidence of novice L2 interactional competence and co-construction. Certainly the talk to be analyzed in this paper is a far cry from the sorts of naturally occurring interactions that are typically the object of conversation analytic (CA) inquiry. In fact, this talk would seem to be the very antithesis of naturally occurring. For starters, this talk clearly fails the "dead social scientist test" that the talk should have happened,

† Professor Dr. (Communication Studies), Faculty of Arts and Humanities at Shikoku Gakuin University.

as it happened, whether or not there had been a researcher around to record it.(1) Nor is this talk the sort of "arranged free conversation" that has served as data for much of the CA and earlier SLA research on non-native speaker interaction (Long, 1981, 1983; Hosoda, 2000; Olsher, 2004, Gardner and Wagner, 2005; Seedhouse and Richards, 2005, Carroll, 2000, 2004, 2005). Neither is it "task talk" nor "classroom inter-action" (Seedhouse 1996, 2004). It isn't even "role-play" data in which participants are assigned roles and tasks but are at least partially free to come up with their own language within the confines of the role-play (Kasper and Dahl, 1991; Gass and Houck, 1999). These conversations are "scripted dialogues," pure and simple. And the activity in which the students and the teacher (alternatively "Callers" and "Call-receiver") are engaged might best be described as "performing a (telephone) dialogue."

In short, this talk is naturally occurring only in the minimal sense that it is actual recorded talk between human beings at some actual point in time.(2) In this one sense it is real and natural. So why analyze such ostensibly "unnatural" talk? Well, we might begin by recollecting an aphorism normally attributed to Einstein that "if we knew what we were doing, we wouldn't call it research, would we?" The initial challenge was to see if there were indeed observations worth making about such "canned" dialogue from a conversation analytic perspective. Were there interesting ways that the realities of conversational interaction, as well as novice L2 competencies, would reveal themselves in the micro-performance of a dialogue? Based on the findings of this preliminary study, I now believe there are.

The dialogue script itself is reasonably (but still not completely) true to life since it was scripted (by the author) taking into account CA research on the organization of English telephone calls (Schegloff, 1968, 1979, 2002a, 2002b, 2002c, 2004, Wong, 1984, 2000). Nevertheless, it remains

"pedagogic material." It presents a "normative" model for a telephone opening, a model that is accurate enough in its details but still grossly artificial in that it didn't emerge from the exigencies of some particular interaction between some particular parties engaged in living their real lives. Below is the dialogue as it appeared in the student's textbook.

((ring ring))[3]
Aki: Hello:↑
Ben: It's Ben.
Aki: Oh. Hi↑
Ben: Hi↓ How'z it goin.
Aki: OK.
Ben: (Hey) the reason I called is↑ [d'ya've a CD player?]
Aki: Yeah, why?
Ben: Could I borrow it?
Aki: Sure, no problem.
Ben: Thanks a million. I'll come by to pick it up.
Aki: OK.
Ben: OK.
Aki: [[Bye
Ben: [[Bye

The data examined in this paper is innovative for CA in one further respect: It is longitudinal. Brouwer and Wagner (2004:35) suggest that orthodox CA methodology must be expanded "...to be able to describe changes over time." For the most part longitudinal studies have not been the norm in CA research (cf. Wootton, 1997). However, Rasmussen and Wagner (2002) provided an early example of what "longitudinal CA" might look like in their study of a series of telephone openings by the same individual (a Danish company employee engaged in making

international telephone calls) which revealed how this particular individual streamlined ("routinized") his conversational openings over the span of these calls. Brouwer and Wagner (2004) argue that CA researchers interested in language learning will need to start collecting a different type of data, one which involves the same speaker (or speakers) in multiple interactions over time.

The data set for this investigation consisted of three sets of telephone recordings of the same scripted dialog made roughly one month apart. In each set there are six calls by the same six students.(4) The first set of recordings was made approximately one month into a first-year "talk-in-interaction" EFL course at a university in Japan (all students were Japanese). Prior to this recording, the students had performed this dialogue with their classmates in class at least 40 times.(5) The second set of recordings was made roughly a month later (and subsequent to a further 40 in-class performances).(6) The third and final set of recordings was made during the last week of the semester as part of a final conversation exam. In total, the students had performed this dialogue over one hundred times (with each other in class) prior to this final recording. Note: The rationale for having the students do this dialogue was not primarily to teach "telephone talk" but rather this dialogue served as a convenient vehicle for introducing students to a range of interactional practices including the relevancy of finely-coordinated speaker transition and the basics of adjacency pair organization. For most students, this was the first time they had ever spoken on the telephone in English. These data sets allow comparison among calls by different Callers within the same recording period and longitudinally for the same Caller over three different calls over the semester.

The full calls contain far more material that can be adequately dealt with in this paper. The focus will, therefore, be on the performance of

the opening of these dialogues. This is not entirely arbitrary. Past conversation analytic work on telephone conversations has shown that, far from consisting of fixed routines (sometimes unfortunately labeled "adjacency sequences" or "gambits"), conversational openings are exquisitely varied and a particularly rich site of interactional work. But even to cover everything that happened in the openings would be beyond the scope of this paper, so the discussion will be further restricted to only two aspects of the opening: 1) the Caller's response to the initial answer-to-summons and 2) the construction of the self-identification turn.

2. Openings as infinitely variable

Openings are all too often portrayed as immutable "boilerplate" routines to be run through in a rote or mechanical manner as a prelude to the "real talk." Nothing could be further from the truth. In fact, the corpus of 500 calls that comprised Schegloff's initial data set (1968) displayed a startling, perhaps even disconcerting, degree of variability. The openings found in EFL textbook dialogues are, in comparison, disappointingly one-dimensional as well as factually inaccurate (Wong, 2000). Just consider the subtle (yet sublime) differences in the ways that the following three openings unfold.

TG (Schegloff, 2002:251)

01 ((telephone rings))
02 Ava: H'llo:?
03 Bee: hHi:,
04 Ava: Hi:?
05 Bee: hHowuh you:?
06 Ava: Oka:::y?hh=

MDE (Schegloff, 2002:253)

01 ((telephone rings))
02 Marsha: Hello:?
03 Tony: Hi: Marsha?
04 Marsha: Yeah.
05 Tony: How are you.
06 Marsha: Fi::ne.

Joyce and Stan (Schegloff, 2002:258-259)

01 ((telephone rings))
02 ((receiver lifted))
03 J: Hullo:
04: S: pt Hi Joyce, it's Stan.
05 J: Hi Stan,=
05 S: -Hi can you hear me okay?<'cause the record player's on.
06 J: O:h yeah:h, I hear you fin[e.
07 [Okay good.

Impressionistically, these three openings might initially seem the same. They do share many commonalities. However, closer inspection reveals the subtle variations. Indeed, rather than viewing openings as boilerplate routines, it's perhaps better to conceive of each individual *opening-as-it-came-to-be-performed* as an observable precipitate of the act of real people living their lives in the real world.

Four central "motifs" in the organization of openings

Schegloff (1979) identified four types of sequential organization operative in openings: 1) *summons-answer* sequences, 2) *greeting-greeting* sequences, 3) *identification* sequences, and 4) *howareyou* sequences. But it would be inaccurate to think of these as "building blocks" or "minimal

units" to be stacked or shuffled. Instead each of these sequence types provides an overarching domain of organization, within which specific practices are ordered. They are more accurately motifs. And it is these organizational motifs that both explain the commonalities between openings while at the same time allowing for the subtle variability.

The first of these sequences, the summons-answer sequence, is most obviously involved in issues of availability. Schegloff (1979:337) suggests the label "pre-conversational" in that these sequences arrange for the "non-terminality" of the talk. Beyond this, the Call-receiver's choice of formats for doing the answer-to-summons displays a proposed framing for the call (e.g. as "casual" or as "business"). This first-talk-by-Call-receiver also initiates the work of mutual identification.

Perhaps the most ubiquitous answer-to-summons in English is "hello." The common misperception (both by the lay-public and second language learners) is that the Call-receiver is initiating a greeting sequence. Schegloff (1968), however, was able to argue, based on a corpus of 500 telephone call openings that the first-turn "hello::?" is better understood as a second-pair-part response to the ring of the telephone. This point is often lost on English language learners since textbook dialogues regularly omit any textual representation of the ringing (Wong, 1984, 2000). Also missing from textbook phone calls is any visual clue that "hello" in the opening of a call has a distinct and specialized pronunciation. The distinction between an answer-to-summons and a greeting is more obvious in Japanese where the former is usually realized as either "hai" or "moshi moshi" (as opposed to "konnichiwa").[7]

The distinction is also obvious in the following atypical telephone opening. In line 2, the Call-receiver performs a loud and gregarious first-pair-part greeting. When this unusual first turn is met with silence (line

3) D reconsiders the appropriateness of this and issues the normatively expectable answer-to-summons, which does occasion immediate uptake (though of an atypical sort).[8]

```
01    ((Telephone rings once - possibly picked up during first ring))
02    D: HI::!
03       (.)
04       Hello:?
05    L: It didn't even ring
06    D: Wha'? It didn't ring?
07    L: No.
08    D: Did you hear me say hello?
```

The relevance of no-gap, no-overlap speaker transition upon the completion of the answer-to-summons is not to be overlooked here. At issue here is recognition. The preferred action (in the CA sense of that term), in response to the voice sample, is a display of recognition, which should be done immediately. A gap following "hello::?" may be hearable by the Call-receiver as displaying a problem with recognition. A prolonged silence may even occasion a second elaborated issuance of "hello::::w?" or even result in the termination of the call. Typically openings happen rather quickly, lasting perhaps no more than five seconds.[9] But in these few seconds a great deal of vital interactional work is accomplished and what happens in these first few seconds of a call may well affect the trajectory of the first few minutes of talk or even the entire conversation. The openings in the scripted dialogues in this study lasted a bit longer, and in some cases much longer. The possible reasons for this will be considered in the next section.

3. Evidence of learning, but learning what?

The advantage of longitudinal data sets is that one can potentially witness learning over time. One domain of learning might be labeled "fluency" and the calls recorded at the end of the semester do indeed sound subjectively "more fluent." One objective measure of this fluency is the time taken to complete the first part of the call, running from the answer-to-summons by the Call-receiver to the end of the Caller's request. In five out of six sets of calls, we can see the times becoming shorter.

Time in seconds: First turn answer-to-summons→end of request			
Name	Call #1	Call #2	Call 3#
Moe	9.8	11.8	9.0 FASTER
Ema	19.3	28.0	13.0 FASTER
Kumiko	10.6	10.1	9.1 FASTER
Sayaka	12.0	10.5	10.0 FASTER
Yuko	11.5	11.4	11.9
Kazuhiro	10.4	9.4	9.5 FASTER

But what specifically results in these quicker performances? Is it just a matter of these novice L2 Callers learning to speak faster? That's part of the answer. But we shouldn't lose sight of the fact that two parties are involved here and that talk-in-interaction is always a joint accomplishment. In this sense we, the Caller and I (the "call-receiver"), are a team. The analysis here will focus on two aspects of these collaborations between Caller and Called in telephone openings: 1) the Caller's response to my answer-to-summons and 2) the construction of the Callers' self-identification turn.

4. The first turn and beyond

In the vast majority of telephone calls the first turn-at-talk, which we'll call T1, is occupied with an answer-to-summons by the Call-receiver.[10] This is most commonly realized in casual English telephone talk as a stylized "Hello::?" This one-word-turn carries an amazingly heavy interactional load. As the second-pair-part of the summons-answer sequence its most obvious job is to announce that the Caller is available for talk. Additionally, the choice of "hello" from a set of alternate T1 formats, for example, "Mr. Jones speaking," frames the call as non-institutional. But even beyond this, the saying of "hello::?" initiates the vital work of negotiating reciprocal identification by providing a voice sample. The preferred response here is an immediate no-gap display of recognition by the Caller in the second turn (T2) as in the following Japanese opening:

```
01        ((ring ring ring))
02    M: shibata de gozaimas:=
03    C: =oka::san [konban wa::]
04    M:           [ha::i      ] konban wa::
```

In line 2 (T1), the Call-receiver uses the semi-formal [family name]+[de gozaimasu] format common in Japan. In line 3, in perfect no-gap/no-overlap coordination with the completion of T1, the Caller displays recognition with a relational term followed by a greeting ("Good evening"). At the same time, the talk in T2 also provides a resource for the Call-receiver to possibly identify the Caller.[11] An immediate, no-gap display of recognition is the preferred outcome (Schegloff, 1979).

Yet, in each of the 18 individual calls in this study, the Call-receiver's

"hello::" in T1 is followed by a silence (a "gap" since it starts at a point of possible completion) and in several calls quite an extended silence, relative to typical silences in conversation.

Gap after Answer-to-summons ("Hello:?")[12]			
Name	Call #1	Call #2	Call #3
Moe	0.63/ 0.24 + .hhh	0.53 (slight click)	0.44
Ema	1.00	1.10	0.93
Kumiko	0.35 + .pt	1.19 (no click)	0.56/ 0.16 + click
Sayaka	0.65/ (slight breath)	0.44	0.36
Yuko	0.69 / 0.35 + .pt	0.82	0.94 (slight noise)
Kazuhiro	0.82	0.92	0.93

If the Caller does not provide a timely claim of recognition in T2, Caller-receivers may hear this as displaying trouble and this trouble can end up shaping not only the opening, but even the conversation as a whole (Schegloff, 2002). A gap at this point might, in some circumstances, suggest technical problems or, if the line is "hearably open," may warrant the Call-receiver making other sorts of interpretations, for example, that the other party is somehow incapacitated or "unwilling" to speak. In the following call from the data, an extended silence after the answer-to-summons is terminated when the unknown Caller hangs up.

```
01    ((telephone rings))
02    D: hello:?
03       (2.48)
04       ((telephone disconnects))
```

So what is the "trouble" being displayed in each of the 18 calls in this study? Is the silence unilaterally a fluency problem on the part of the

novice L2 Callers? According to the turn-taking system described by Sacks, Schegloff, and Jefferson (1974) next speakers monitor the moment-by-moment production of a turn-in-progress for clues as to when, in the immediate future, it is likely to reach a point of possible completion. That is, they "project" turn completion and orient to this point as the unmarked location for speaker transition. This suggests two possibilities. First, these novice L2 callers may not be engaged in monitoring the ongoing talk and projecting its incipient completion at all. They may simply be waiting for the prior speaker to stop before "gearing up" to speak, which on purely neurological grounds would result in a gap. One form of evidence that this is the case, for at least some of these speakers, are the various audible in-breaths, lip noises, clicks, and other preparatory sounds that preface their T2 identifications.

Kumiko - Call 1

```
01    ((telephone rings))
02    D: hello:?
03       (0.35)
04    K: .pt >a- it's kumiko nouchi. I'm in your talk-in-teraction=
05       =class.<
```

In the first-round calls, four of the six Callers (Moe, Kumiko, Sayaka, and Yuko) preface their turn-beginnings (Schegloff, 1996) with one of these sounds. In the second-round calls, there are only two instances of this phenomenon and in the third-round calls this appears to have been reduced to a single instance (the noise in Yuko's case might be an artifact of the recording). So here is evidence of learning of a sort. In fact, Moe and Sayaka (and arguably also Kumiko) display regular progress over the three calls towards minimizing this gap. On the other

hand, Ema, who is perhaps the least "fluent" of these novice L2 Callers, seems to make no progress in this respect.

A second possibility is that these novice L2 Callers are not yet socialized to the English practice of voice-only identification. In Japanese telephone conversations, the most common format in T1, in addition to "hai" (which does comparable work to English "hello::?"), includes a second element, a formal identification: [family name]+[desu/de gozaimasu] as in the following.[13]

```
01        ((ring ring ring))
02     S: hai shibata desu
03     M: minamoto des:
04     S: ara konbanwa::
```

These novice L2 Callers might, therefore, be withholding talk in the expectation of a formal identification. Seen from this perspective, the silence would be "mine" and not "theirs." Even though they know that it's not in the script, they may still have a hard time suppressing an orientation to their L1 practices to jump in immediately after "hello." Still, it remains an open question as to why not a single novice Caller was able to achieve a no-gap speaker transition at this point. It is certainly something that warrants some pedagogic attention.

5. Turn-construction of Caller identification in T2

Once the Call-receiver has established in T1 that he or she is available and provided a resource for identification (either through a voice sample or fuller format), it falls to the Caller to self-identify in the second-turn (T2) - as well as to display recognition (if necessary and/or possible).

Schegloff (1979:28-32) identified at least nine classes of T2 formats in his data:

1) Greeting terms
2) Answerer's, presumed answerer's, or intended answerer's name or address term, with quasi-interrogative intonation contours
3) Answerer's, presumed answerer's, or intended answerer's name or address term, with assertive, exclamatory, or terminal intonation contours
4) Question or noticing concerning answerer's state
5) "First topic" or "Reason for the call"
6) Request to speak to another ("switchboard" request)
7) Self-identification
8) Question re identity of answerer
9) Joke or joke version of above

The script provided to the Callers for this study uses the self-identification format as that format seemed most generically fitted to the context of calling one's teacher to make a request. Specifically, the students were told to identify themselves in the following manner: "It's [first name]+[family name]. I'm in your talk-in-interaction class." So when I received their telephone calls, I was oriented towards hearing this sort of full identification prior to responding.[14] (In retrospect, it might have been preferable to include in the script both the greeting + self-identification as this combination is found in many of the calls in Schegloff's corpus.) One might assume then that there would be no variation at all in terms of how these Callers constructed their T2 turns. Yet, this turned out not to be the case.

First, let's consider the variation between callers within a single set of

calls. In the following call from the first-round set, the Caller produces the expected format. And more importantly the two parts of the identification get produced as one unit, via a technique called a "rush through" (Schegloff, 1982). By constructing her turn in this manner she actively discourages speaker transition at the end of the first unit.

Kumiko - Call 1

```
01    ((telephone rings))
02    D: hello:?
03       (0.35)
04 -> K: .pt >a- it's kumiko nouchi.=I'm in your talk-in-teraction=
05       =class.<
06       (0.26)
07    D: oh, hi:
```

The next calls are similar but with one very crucial difference. They include a slight "beat of silence" (Sayaka) and/or a preparatory noise (Yuko, Moe) after the name and before adding the second turn construction unit. Some explanation is in order regarding what is meant here by a "beat of silence." Objectively, as measurable with sound editing software, there is a 0.22 second period of silence in the sound wave between the two parts of the identification. However, Sayaka is speaking in a very slow and measured fashion (as indicated by the outward pointing brackets) so from a participant's perspective the first word of the second part, in line 6, might be heard as coinciding with the next rhythmic beat.

Sayaka - Call 1

```
01    ((telephone rings))
02    D: hello:?
```

03 (0.65/ slight in-breath at mid point?)
04 S: <It's:: sayaka uchida.>
05 -> (0.22)
06 <I'm een your talk een interaction class.>
07 (0.10 maybe)
08 D: oh, hi:

Yuko - Call 1

01 ((telephone rings))
02 D: hello:?
03 (0.69/0.37 + .pt?)
04 Y: <it's yuko ono.>
05 -> (0.29/.hh?)
06 <I'm in your talk-in-interaction class.>
07 D: oh, hi:

Moe - Call 1

01 ((telephone rings))
02 D: hello:?
03 (0.63/0.24 + .hhh)
04 M: it's moe miyoshi.
05 -> (.hh)
06 I'm in your talk-een-interaction class
07 (0.08)
08 D: oh, hi:

These barely perceptible gaps might seem insignificant, but it's important to recognize that had I, as the teacher, not been oriented to the production of the full scripted format, I might have responded at that point (with "oh, hi") resulting in overlapping talkassuming, of

course, that I had achieved recognition from the name and voice. This is similar to what did happen in another call. In this case, I had allowed a delinquent student a "second chance" to call at a pre-specified time so I was waiting for this specific student to call and this is not inconsequential.

Akinori - Call 3 (not included in the corpus because he only called twice)

```
01     ((telephone rings))
02     D: Hello:?
03        (0.94)
04     A: <I:t's Akinori.>
05        (0.39)
06 -> D: Oh. [Hi.]
07 -> A:     [I::]'m in,
08        (0.43)
09     A: I'm in you:r talk-u (0.36)>interaction class<
10     D: Oh, YEAH! Yeah. I was waitin' for your call.
```

So it can be argued that in the case of the calls by Sayaka, Yuko, and Moe (above), there is at least some sense in which I co-participated in the construction of these T2 identification turns by withholding a display of recognition at a point of possible completion. In two of the calls, however, the novice Callers failed to produce the second part of the identification before I "gave up" and jumped in with a claim of recognition.

Ema - Call 1

```
01     ((telephone rings))
02     D: hello:?
03        (1.00)
```

```
04    E: <It's e:ma yamaue.>
05 ->    (1.08)
06    D: oh, hi:
```

Kazuhiro - Call 1

```
01    ((telephone rings))
02    D: hello:?
03       (0.82)
04    K: it's >Kazuhiro Takeuchi.<
05 ->    (1.39)
06    D: oh, hi:
```

Perhaps even more interesting than the variation among Callers in a given set of calls is the variation by individual Callers over the period of the three calls. In several cases, there is clear evidence that the Caller had learned to routinize this part of the opening. Consider, as one example, the manner in which Kazuhiro's T2 identifications played out in the other two calls. In the second call, he reduced his full name to first name only (as did most of the other Callers for some reason) and immediately launched into the second part of the identification.(15) Pacing issues, however, still led to my delay in producing a timely start-up on the completion of Kazuhiro's turn. By the third call, he had sorted out the pacing issues and I was able to achieve the desired (in this context) no-gap, no-overlap speaker transition.

Kazuhiro - Call 2

```
01    ((telephone rings))
02    D: hello:?
03       (0.92)
04 -> K: it's Kazuhiro. I'm ('n) your talk-in-interaction class.
```

05 -> (0.35)
06 D: oh, hi:

Kazuhiro - Call 3

01 ((telephone rings))
02 D: hello:?
03 (0.93)
04 -> K: >it's Kazuhiro I'm ('n) your talk in interaction class.<
05 -> D: oh, hi:

By the third round calls, five of the six novice Callers displayed clear evidence of routinization in their T2 identifications. Only Ema was still failing to provide the class identification but her production of her name identification was now much smoother than it had been in prior calls.

Another issue that had worked itself out by the third round of calls, and which accounts for some of the overall shortening of these perform-ances, is the coordination between the Caller's T2 identification and my subsequent claim of recognition. In four of the round-three calls, the Caller and I were able to achieve no gap transitions.

Kazuhiro - Call 3

04 K: >it's Kazuhiro I'm ('n) your talk in interaction class.<
05 D: oh, hi:

Moe - Call 3

04 M: it's moe.
05 (0.18/ slight click?)
06 I'm in your talk-een->interaction< class.
07 D: oh, hi:

Sayaka - Call 3

04 S: <It's: sayaka. I'm <een your talk een interaction class.>
05 D: oh, hi:

Yuko - Call 3

04 Y: <I'm yuko ono. I'm in your talk-in-interaction class.>
05 D: oh, hi:

This was not the case with Ema, who was still not producing the full format identification. In Kumiko's third call, the reason for the gap was less clear since she was deploying the full identification.

Kumiko - Call 3

01 ((telephone rings))
02 D: hello:?
03 (0.56/0.16 + click)
04 K: >it's kumiko. I'm in your talk-in-teraction class.<
05 -> (0.36)
06 D: .t oh, hi:

However, the production of Kumiko's T2 identification is compressed (her rendering of "talk-in-teraction" is missing a syllable) and somewhat faster than (her) normal, as indicated by the inward pointing arrows. In another study (Carroll, 2000) I reported that talk by novice L2 speakers produced with irregularities in the pacing makes precision next speaker start-up more difficult. That seems to have been the case here as well.

6. Conclusions

The most inescapable observation in all this is that there is a considerable amount of variation in what are ostensibly "fixed" dialogues. The words and grammar are immutable. The variation finds its way in via practices that are not usually considered within the scope of language teaching. In these calls, the variations centered on the negotiation and timing of speaker transitions. Many language teachers might not even consider this a problem. Others might wish to partition this off as "interactional competence" or as something "unteachable" in the classroom (I can assure you it is not). But more and more the view of language being painted by the wide range of researchers looking at social interaction and social construction, from Conversation Analysis to Sociocultural Theory, is one where language is, at its deepest levels, inseparable from language use and, ultimately, from the lifeworlds we inhabit and co-construct. If this view turns out to be valid, then the sorts of practices examined in this paper need to be treated as central to the phenomenon of language and language teachers may need to thoroughly re-examine what it means to "teach language."

A second point that we can take from this study is that the line between these novice Callers' competencies and my own is often hard to draw. As co-participants in the talk we manage to "talk life into" these scripted dialogues, each playing off the other, relying on one another. We saw how the silences after my response-to-summons might be heard as mine as much as theirs. We looked at how my actions (and inactions) may have affected the ways that they came to build "their" T2 identifications. And, if this is the case even in the first few turns in the performances of a tightly scripted dialogue, how much truer it must be in free-flowing novice L2 interactions. My research on face-to-face

novice-to-novice L2 interactions (Carroll, 2004, 2005 2006) strongly suggests that competence is co-constructed and shared across the parties to an interaction. And this fact has obvious consequences for how we define and assess "learner competence." In short, the view that learner competence is something that "resides within the skull" (to borrow a phrase from Goodwin, 1995) of an individual learner is debatable.

References

Brouwer, C. & Wagner, J. (2004). Developmental issues in second language conversation. In: *Journal of Applied Linguistics*, 11, 29-27.

Carroll, D. (2000). Precision timing in novice-to-novice L2 conversations. In: *Issues in Applied Linguistics*, 11(1), 67-110.

Carroll, D. (2004). Restarts in novice turn beginnings: Disfluencies or interactional achievements?. In: Gardner, R. & J. Wagner (Eds.) *Second language talk: Studies of native and non-native interaction* NEW York:cont:nuum (pp. 201-220).

Carroll, D. (2005). Vowel-marking as an interactional resource in Japanese novice ESL conversation. In: K. Richards & P. Seedhouse (Eds.), *Applying conversation analysis* London:Palgrave Macmillan (pp. 214-234).

Carroll, D. (2006). Co-constructing Competence: Turn Construction and Repair in Novice-to-Novice Second Language Interaction. Unpublished Ph.D. dissertation, University of York, York.

Gass, S. & Houck, N. (1999). *Interlanguage refusals: A cross-cultural study of Japanese-English. Studies on Language Acquisition*. Berlin: Mouton de Gruyter.

Goodwin, C. (1995). Co-Constructing meaning in conversations with an aphasic man. In: *Research On Language and Social Interaction*, 28, 233-260.

Heritage, J. M. (1984). A change-of-state token and aspects of its sequential placement. In: J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* Cambridge: Cambridge University Press. (pp. 299-345).

Long, M. H. (1981). Input, interaction and second language acquisition. In H.Winitz, (Ed.), *Native language and foreign language acquisition*. Annals of the New York Academy of Sciences, 379, 259-278.

Long, M. H. (1983). Native speaker/non native speaker conversation and negotiation of comprehensible input. In: *Applied Linguistics*, 4, 126-41.

Rasmussen, G., & Wagner, J. (2002). Language choice in international telephone conversations. In: K.K. Luke & T. Pavlidou (Eds.), *Telephone calls. Unity and diversity in conversational structure across languages ads cultures* (pp. 111-131). Amsterdam: John Benjamins

Sacks, H., Schegloff, E.A., &. Jefferson, G. (1974). A simplest systematics for the organization of turntaking in conversation. *Language*. 50(4), 696-735.

Schegloff, E. A. (1967). The first five seconds: The order of conversational openings. Unpublished Ph.D. dissertation, University of California, Berkeley.

Schegloff, E. A. (1968) Sequencing in conversational openings, In: *American Anthropologist,* 70, 1075-95

Schegloff, E. A. (1979) Identification and recognition in telephone conversation openings. In: G. Psathas, ed. *Everyday language: studies in ethnomethodology*. New York: Irvington (pp23-78).

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In: D. Tannen (Ed.), *Georgetown University roundtable on languages and linguistics* Washington DC: Georgetown University Press (pp. 71-93).

Schegloff, E. A. (2002a) Reflections on research on telephone conversation: Issues of cross-cultural scope and scholarly exchange, interactional import and consequences. In: Kang Kwong Luke, Theodossia-Soula Pavlidou (eds.) *Telephone Calls: Unity and diversity in conversational structure across languages and cultures*. Amsterdam: John Benjamins (pp. 249-81).

Schegloff, E. A. (2002b) Beginnings in the telephone. In: J.E. Katz, M. Aakhus, eds. *Perpetual contact: mobile communication, private talk, public performance*. Cambridge: Cambridge University Press (pp. 284-300).

Schegloff, Emanuel A. (2002c) Opening sequencing. In: J.E. Katz, M. Aakhus, eds. *Perpetual contact: mobile communication, private talk, public performance*. Cambridge: Cambridge University Press (pp. 326-85).

Schegloff, Emanuel A. (2004) Answering the telephone. In: Gene H. Lerner, ed. *Conversation Analysis: Studies from the first generation*. Amsterdam/Philadelphia: John Benjamins (pp. 63-107).

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. In: *Semiotica*, 8, 289-327.

Wilson, T. P., & Zimmerman, D.H. (1986). The structure of silence between turns in two-party conversation. In: *Discourse Processes*, 9 , 375-90

Wong, J. (1984). Using conversational analysis to evaluate telephone conversations in English as a second language textbooks. Unpublished master's thesis, University of California, Los Angeles.

Wong, J. (2000). "Applying" conversation analysis in applied linguistics: Evaluating English as a second language textbook dialogue. Paper delivered as part of an invited colloquium entitled "Conversation Analysis: A Methodological Resource for SLA in New Millennium," annual meeting of the Second Language Research Forum (SLRF), University of Wisconsin-Madison.

ENDNOTES

(1) In a message to the ETHNO list, Ian Hutchby refers to the use of this phrase by Jonathan Potter stating that: "what is recorded should have occurred that way had the researcher keeled over on the way to University that morning."

(2) It might be noted, though, that the fact that this talk did in fact occur, however artificially, is already more than can be said for the so-called imaginary or introspective data that have been the mainstay in some approaches to discourse analysis and pragmatics.

(3) Wong (2000) points out that EFL textbook representations of telephone calls regularly leave off any representation of the ringing, which represents the first pair part of the summons-reply adjacency pair.

(4) There were more students in the class, but not all students completed all three

calls.

(5) In order to be able to do this many practice performances it is necessary to have students work in pairs and to rotate these pairs quickly and efficiently. To do this I set of two rows of two side by side chairs (side by side to simulate the non-face-to-face nature of telephone talk). Students were given 30 seconds to run through the dialogue two times, changing Caller vs. Caller roles. When this time was up, I stopped any ongoing talk and had the students rotate such that the twostudents on the inside front moved to back of the opposite (inside) row and all other students on the inside row moved up on seat. In this way students eventuallyperformed this dialog with every member of the class. If there were an odd number of students, I placed "vacation chair" in the back of one of the rows that students would rotate through. Thus, with 10 students in the class, after one full rotation each student would have performed each role in the conversation 10 times with 10 different partners (over roughly 7-8 minutes), and run through the conversation as a whole 20 times.

(6) The students also had several opportunities to perform this dialogue with each other over the telephone using their cell telephones.

(7) It is worth noting, however, that in a small corpus of Japanese telephone calls that my CA students and I have collected over several years, very few are actually initiated with "moshi moshi." The most common answer-to-summons in Japanese telephone calls appears to be "hai, [family name] desu" or just "hai." Furthermore, the fact that moshi moshi may occur in other positions such as first-turn-by-Caller and may then be responded to with a further "moshi moshi" suggests that this practice involves more than merely responding to a telephone summons. One further complication in terms of terminology is that "greeting" is regularly translated into Japanese as "aisatsu," which is a much broader term including a wide range of fixed and semi-fixed expressions, some of which might only occur in leave-takings, or thanking.

(8) This call was received on an office extension telephone with a direct connection to D's home telephone-on the same internal telephone system. The reason for D's

use of "hi" is that D can be almost certain (at certain times of day) who is calling.

```
01    ((Telephone rings three times))
02    D: Hi
03    L: Hi it's me::
04    D: Yeah?
05    L: $It's an emergency$ I forget it's movie day.
06    D: Oh. that's ri::ght!
```

(9) This fact was captured in the title of Schegloff's Ph.D. dissertation: "The first five seconds: The order of conversational openings."

(10) Cell telephones and telephones with Caller ID are changing the ways that telephone openings are managed.

(11) Note that in this call, the Call-receiver's immediate production of "hai" ("yes") at the completion of "okaasan" claims recognition but doesn't actually display it, as the use of a name would have.

(12) These times are in seconds. The use of a 2-place decimal is not meant to imply that hundredths of a second are meaningful or even perceptually real to participants. Rather, this practice merely renders transparent that the timings have been measure using sound editing software.

(13) Some of my Japanese students have reported to me that this may be changing in recent years in response to the perceived danger of "giving too much information to strangers." So instead more calls are being answered with just "hai" or "moshi moshi."

(14) In order to ensure that I would be at home, the students were required to call within certain evening hours. This fact served as a resource for recognizing the likelihood that a given call was probably a student call. Such expectations can and do affect the ways that any given conversation may unfold.

(15) With the exception of Yuko each of the novice Callers dropped the family name from the second call onward. This may have been the result of the language socialization provided by practice in class with other novice Callers ("classmates") with whom use of the last name might have seemed superfluous.