Private Speech as Social Action

	Eric Hauser
journal or publication title	Language and Sociocultural Theory
volume	2
number	2
page range	119-138
year	2015-07-30
URL	http://id.nii.ac.jp/1438/00008891/

Private Speech as Social Action

Eric Hauser
University of Electro-Communications

Abstract

An important theoretical construct within the Vygotskian sociocultural approach to
second language learning is private speech. Within a conversation-analytic framework,
an agnostic stance is taken in this paper toward the possible intrapsychological
function(s) of private speech in order to 1) illustrate how private speech can be
identified within the details of talk-in-interaction and 2) how private speech can be
understood as social action. It is argued that attention to the details of how private
speech is produced is important in order to show how private speech has been identified
as such; that viewing private speech as social action allows for a more emic perspective;
and that, at least within interaction, private speech is social not just in origin, but each
time that it is produced.

Keywords: CA-SLA, private speech, social action, Vygotskian sociocultural theory

Within the Vygotskian sociocultural approach to second language (L2)
learning (Lantolf, 2000, 2011; Lantolf and Thorne, 2006), private speech is an
important theoretical construct. Drawing on Vygotsky's discussions of how children
appear to use language to mediate or regulate their behavior during the performance of
cognitively challenging tasks (Lantolf, 2000), researchers within this approach are
interested in whether and how L2 learners can use L2 private speech to regulate their
performance during such activities as picture-narration tasks (see Lantolf and Thorne,
2006, and McCafferty, 1994a, for extensive discussion). As stated by McCafferty
(1994a), 'a Vygotskian view of private speech affords a valuable window onto the
intrapersonal processes in which adult L2 learners engage in their efforts to self-regulate
in the face of the very complex process of learning a second language' (434). In
addition, learners may also use private speech to focus attention on, and manipulate L2
forms, which may facilitate internalization of those forms (Lantolf, 2003; Lantolf and
Thorne, 2006). Lantolf and Yáñez-Prieto (2003) and Ohta (2001), for instance, used
personal microphones to record the talk of students in university language (Spanish and
Japanese, respectively) classrooms that presumably was not audible to either other
students or the teacher and found that the learners used private speech in just this way.
Within the Vygotskian approach, then, private speech is argued to make the process of
L2 learning at least partially observable (Lantolf, 2003).

One important issue in research on private speech is whether and how it can
be distinguished from talk used for interpersonal communication, that is, social speech
(Donato, 2000; Wells, 1999). True to the Vygotskian perspective on the genesis of
intrapersonal psychological functions within interpersonal social interaction, private
speech is seen as having social origins (Alarcón-Rubio, Sánchez-Medina, and Winsler,
2013; Antón and DiCamilla, 1999; Damianova, Lucas, and Sullivan, 2012; DiCamilla
and Antón, 2004; Lantolf, 2000), which may make it appear in some cases similar to
social speech (Antón and DiCamilla, 1999; DiCamilla and Antón, 2004). In some
studies (e.g., Lantolf and Yáñez, 2003; Ohta, 2001), private speech appears to be
identified primarily on the basis of its supposed inaudibility to others. In other studies
(e.g., Lee, 2008; Stafford, 2013), all talk (and in the Lee study, gesture as well) seems to
be identified as private speech based on the fact that the participants were working

alone.[1] In other cases, which involve learners engaged in more interactive tasks, distinguishing private speech from social speech seems to be much more complicated. In McCafferty (1994b), for example, which involved a picture-narration task, three criteria were used to identify private speech, namely, that it was 'tangential to the narrative' (123), involved an 'effort to seek self-guidance' (123) and was 'concerned with mastering some task-relevant difficulty' (123). As a second example, while Antón and DiCamilla (1999) are not particularly clear about their criteria, DiCamilla and Antón (2004), a study which seems to draw on a larger version of the dataset used for the earlier study, state that they 'relied on an analysis of linguistic, paralinguistic, and discursive features of the utterances that more clearly reveal their *self-directed nature* in the contexts of the discourse' (41, emphasis in the original). Smith (2007) identified as private speech utterances which involved some sort of prosodic shift and 'were not explicitly social' (345). Moving to recent research on private speech that does not involve L2 learning, San Martín Martínez, Boada i Calbet, and Feigenbaum (2011) identified an utterance as private speech if it included 'audible verbal elements and/or silences that were not explicitly directed at another person and that indicated self-regulation' (222); Damianova, Lucas, and Sullivan (2012) used only a negative criterion to identify private speech as any utterance that could not be identified as social speech, based on linguistic and nonlinguistic cues; and similarly, Alarcón-Rubio, Sánchez-Medina, and Winsler (2013) identified as private speech 'any verbalization ... not explicitly addressed to another person' (103) based on such linguistic and nonlinguistic cues as pronoun reference, gaze, and volume.

As argued by Wells (1999) and Steinbach Kohler and Thorne (2011), though, and against the claims of Antón and DiCamilla (1999), Damianova, Lucas, and Sullivan (2012), DiCamilla and Antón (2004), and San Martín Martínez, Boada i Calbet, and Feigenbaum (2011), it may be that in social interaction, private and social speech cannot be clearly distinguished and that an instance of talk may be both simultaneously. From a more theoretical perspective, Larrain and Haye (2012) argue that all inner speech, of which private speech is considered a subtype, reflects its origins in social speech in being a kind of discourse which 'is not different in nature to "public" language' (17).[2] Goffman (1981), in a passage that I have not seen discussed in any of the work on L2 private speech, has this to say about the apparently private speech produced by the children that Vygotsky observed in his research:

> Given that Vygotsky's early work required an adult observer to be within listening distance, one could go on to suggest an additional interpretation, namely that for children the contingencies are so great in undertaking any task, and the likelihood so strong that they will be entirely discounted as reasonably intentioned persons if they fail (or indeed that they will be seen as just idling or fooling around anyway), that some voicing of what they are about is something they are always prepared to offer. An adult attempting to learn to skate might be equally self-talkative. (95--96)

That is, even talk which is clearly addressed to the self may nevertheless be targeted at others, to show for example that the speaker is 'reasonably intentioned,' and may thus involve a kind of social action.[3] That private speech can be viewed as social action is the central claim of this paper.

The concept of social action has a long history in sociology and is perhaps most strongly associated with Parsons (e.g., Parsons, 1937). There are various senses in which language use can be viewed as social action (Holtgraves, 2002), but the sense in which I use the concept draws on Ethnomethodology and Conversation Analysis (CA) (Garfinkel, 1967; Hayano, 2013; Heritage, 1984; Heritage and Stivers, 2013; Maynard, 2013; Pomerantz and Fehr, 1997; Sharrock and Button, 1991; Stivers, 2013; Wooffitt, 2005). This means, first, that social actions are designed for their context, both at the local sequential level (Heritage, 1984; Heritage and Stivers, 2013; Maynard, 2013;

Pomerantz and Fehr, 1997) and at the broader level of event, activity, and participants (Heritage, 1984; Pomerantz and Fehr, 1997). Much work in CA focuses on how social actions are responsive to prior actions and/or set up relevancies for subsequent actions, that is, how they are organized as parts of adjacency pairs (Hayano, 2013; Stivers, 2013). However, this is not the only way in which social actions implemented through talk can be designed for their local context. Second, social actions shape the context in which they occur, again at the local as well as the broader level (Heritage, 1984; Pomerantz and Fehr, 1997). Third, though they may flout social norms, the design of social actions shows participants' orientations to such norms (Garfinkel, 1967; Heritage, 1984; Wooffitt, 2005). Finally, social actions are designed to be recognizable (Garfinkel, 1967; Heritage, 1984; Sharrock and Button, 1991), though this should not be taken to mean that people do not perform actions which are intended to deceive, nor that participants' understanding of what an action is cannot be revised (Heritage, 1984). The recognizability of social actions has particular import for research on private speech, as this is relied on to identify instances of such speech.

That private speech during interactive tasks can have a social component has, of course, been recognized by those working within the Vygotskian theoretical framework. While this often seems limited to a recognition that it is produced in the presence of others (Lantolf, 2000), that 'a socially stimulating context is crucial for the elicitation of private speech' (Damianova, Lucas, and Sullivan, 2012: 447), or that it builds 'a sense of cooperation' (McCafferty, 1994a: 429) or 'collective externalization of metacognitive strategies' (McCafferty, 1994a: 430), it has also been recognized that social context may influence the choice of language for private speech (Lantolf, 2003). In addition, Smith (2007), while still arguing that private speech can be distinguished from social speech, provides examples to show 'that private speech can be responded to by others as if it were social in social interactive contexts' (348). At least one study has focused explicitly on the social aspects of private speech. Steinbach Kohler and Thorne (2011) argue that private speech cannot be dissociated 'from broader processes of social interaction' (70) and show how participants can use private speech 'as a resource for managing interaction' (77), how private speech can emerge as the result of some kind of misalignment among participants, and how private speech can index 'an agreement to disattend copresent parties' (86).

Similar to Steinbach Kohler and Thorne (2011), in this paper I attempt to show, through analysis of two examples, how talk produced by a relatively low-proficiency L2 learner of English, which is hearably self-directed and concerned with overcoming difficulty in producing an L2 utterance, and thus identifiable as private speech within the Vygotskian theoretical framework, can also be understood as a social action in a much more fundamental sense than has generally been recognized within this framework. It is not my intention to claim that private speech, including the private speech in my two examples, does not have a self-regulating function or play a role in the internalization of L2 forms. Rather, I adopt an agnostic stance toward the intrapersonal psychological function(s) of this talk, including whether talk can provide a window onto psychological processes. With such an agnostic stance, my analysis is not framed within sociocultural theory. Instead, I make use of CA to unpack the data. Below, after first describing the source of the data, the participants, and the transcription, I analyze how part of the talk in the two extracts is identifiable as private speech, in that it is self-directed talk, or self-talk (Goffman, 1981; Goodwin, 1987), as it will be referred to from hereon, and oriented toward overcoming difficulty with using the L2. I then show how the same talk is also identifiable as social action. My purpose in doing so is to criticize the Vygotskian view of private speech, but rather to engage with this view from a perspective that focuses on private speech as interpersonal rather than, or in addition to, intrapersonal, talk.

Data

The data are drawn from audiorecordings between myself (Eric) and a male adult first language (L1) user of Japanese, called Nori (pseudonym), that were made over a period of seven months. At the time of the first recording, Nori had just recently immigrated to Honolulu from Tokyo with his wife and two daughters, who were in elementary school. His English proficiency was fairly limited, though he was attending free adult school ESL classes. In order to provide him with opportunities to practice using English, Nori and Eric agreed to meet at Eric's apartment periodically to speak English. Nori also gave consent for these meetings to be recorded for research purposes. Eric was an L1 user of English and a proficient L2 user of Japanese, who had lived, studied, and worked in Japan for over seven years before coming to Honolulu for graduate school. Nori knew that Eric was proficient in Japanese and the two of them typically used Japanese whenever they conversed outside these meetings. Eric occasionally prepared some simple referential communication tasks (Yule, 1997) that took up a few minutes of meeting time, as in the first example below, and Nori sometimes brought things that he had questions about. However, most of the meeting time was used for unplanned conversation, as in the second example below. The data were collected in order to investigate L2 development longitudinally and I have used them that way in other work (e.g., Hauser, 2013a, b). However, for this paper, I will not be making any longitudinal comparisons. There is one final point about the collection of these data which is important to clarify. Though the data were collected in order to conduct longitudinal investigations, the data collection was also opportunistic. The data were not collected as part of a preplanned research project and the meetings did not involve, for example, attempts to elicit particular types of language. Nevertheless, such things as 'practicing English' and 'recording Nori's English for research' are features of the broader context, which may be treated as relevant by the participants.

The data have been transcribed based on conventions commonly used in CA (Jefferson, 2004). Japanese words have been transcribed in italics, with a morpheme-based translation on the line below. See the appendix for transcription conventions and symbols used to translate Japanese words. Through analysis of two excerpts from these data, I show, first, how some of Nori's talk can be identified as private speech. I then show how this private speech can be understood as social action.

Private Speech as Articulation of Explicit L2 Grammar Knowledge
Searching for the Superlative Form

In the first example, Nori and Eric are engaged in a referential communication task that involves Nori looking at a card that has a picture of something, such as an animal or a tool, and describing it so that Eric can guess what it is. Nori approaches this task as involving the giving of a series of hints. For the card that Nori is looking at during this part of the task, the object in the picture is a dinosaur.

Example 1

```
01 N: o:ld old animal.
02 (0.5)
03 E: [old animal?
04 N: [ah:
05 N: old animal. yeah.
06 (1.2)
07 N: yeah big. (.) big animal. .hh
08 E: mhm,
09→N: uh:: maybe uh::: (.) first uh (0.5) °big ↓bigger°
10→    (.) biggest. .h hhm .h [biggest      ] (0.2)=
11 E:                        [the biggest ] animal?
```

```
12 N: =manimal. .h n: .h (demo ne f:) (.) [no  :   :=
                        (but  IP)
13 E:                                       [>like a-<
14 N: = [n.
15 E:    [like an elephant?
16 (0.5)
17 N: ele- ↑no. (.) moh- ° m::° more- more- (.) .h (0.6)
18→   o:l- (0.4) old o(h)- (.) olde(h)st. .h oldest
19    animal.
20 E: thee oldest animal?
21 N: yeah_ .hhh (1.6) big ↓animal. ↑no .hh (0.6)
22    no livu=
23 E: =(here-) n.
24 N: ° n::°  not .hh (0.4) n::ot lih- ° n::°  live world.
25    ↑ah chigau na. ° ima n-°  (0.2)
              different IP      now
26 E: it doesn't live now?
27 (0.2)
28 N: yeah ↑not live now.
29 E: oh okay
```

Nori gives his first hint ('old old animal') in line 01. When Eric repeats this as 'old animal,' with rising intonation, in line 03, Nori confirms the repetition in line 05. However, Eric does not attempt to guess what the picture is of and Nori does not immediately give a second hint, leading to a fairly long gap in line 06. Nori then offers a second hint in line 07 ('big … big animal') and then starts to elaborate on this hint in line 09. The elaboration of this second hint is eventually articulated as 'biggest … animal' in lines 10 and 12. However, before he says this, Nori clearly encounters difficulty. He repeats 'uh,' often in elongated form, three times and pauses twice in line 09. As for his use of 'first' in this line, it can be understood retrospectively as the start of a translation of the Japanese '*ichiban ookii*' (literally, 'number one big'). After this 'first,' the third 'uh,' and a half-second pause, he then says quietly (noted in the transcript by the degree signs) 'big bigger,' with a noticeable drop in pitch on 'bigger' (noted by the down arrow). After a third brief pause, Nori then says 'biggest' with falling intonation. After some breathy sounds, he articulates 'biggest … animal' (with 'animal' sounding more like 'manimal'). However, it is not so simple, as 'biggest' is produced in overlap with Eric's 'the biggest' (line 11) and Eric says 'animal' during the pause in Nori's talk, which may have prompted Nori to add the word 'animal.'

When Nori says 'big bigger … biggest,' he is verbalizing grammatical knowledge of English which he possesses by virtue of being an L2 user of English, knowledge which he may have learned in secondary school English classes in Japan and/or in his ESL classes at adult school. Specifically, this is knowledge of the comparative and superlative forms of adjectives, which in some English language classrooms with a focus on grammar, such as those in Japanese secondary school, and in some commercial English language textbooks, such as the one Nori is using at adult school, is presented and taught as three-part lists (Jefferson, 1990); that is, as 'big, bigger, biggest' (or 'good, better, best,' 'interesting, more interesting, most interesting'). By verbalizing this particular three-part list, Nori is led, or rather, leads himself, to the form he wants, which is the superlative form. This verbalization is thus oriented to overcoming difficulty using the L2 and, indeed, leads to a successful solution of this difficulty. The articulation of this list of adjective forms also appears to be self-talk, as the first two items are produced with a notably lowered volume and the second item is produced with lowered pitch as well. The third item is produced with normal volume

and with falling intonation, which commonly indicates some sort of completion, so that this list is complete both structurally, in that it has three parts, and intonationally. Nori, though, does not orient to turn transition, as indicated by the audible breathy sounds, including two inbreaths, following the third item. Such an inbreath, when it follows one's own possibly complete talk, is indicative of intention and preparation to continue speaking and Nori does indeed continue following these sounds. The falling intonation and three-part structure can thus indicate that the list is complete, but not that Nori expects a response from Eric at this point. Eric also orients to this as self-talk as he does not respond to the first two items and only responds to the third after some delay. That is, he does not initially treat the falling intonation on the third item as an indication that Nori has completed his turn and that he should respond. For these reasons, Nori's 'big bigger … biggest' can be considered an example of private speech.

   Something very similar happens as the interaction continues. After Eric makes a guess in line 15 ('like an elephant'), marked as a candidate answer with rising intonation, Nori first rejects this in line 17 (see Hauser, 2013b, for what is termed *x-no negation* and its use by Nori for correction) and, eventually, elaborates on his earlier first hint by saying 'oldest animal' in lines 18 and 19. Again, though, he experiences difficulty as he tries to make this elaboration. In line 17, he makes four attempts to use 'more,' with three of these being cut off (indicated by the hyphens) and one being merely a softly articulated and elongated 'm' sound. After a brief pause, an inbreath, a longer pause, an attempt to say 'old,' and another pause, he apparently articulates another three-part list of adjective forms, 'old older … oldest,' though the second item is obscured by laughter (indicated by the letter *h* in parentheses). Once more, by articulating this three-part list of adjective forms, Nori leads himself to the superlative form, which he then uses in 'oldest animal,' to which Eric responds with 'the oldest animal' said with rising intonation. While the items that make up this list of adjective forms are not produced with noticeably lower volume, as with the earlier list, neither Nori nor Eric treat the structural completeness of the list or the falling intonation on 'oldest' as indicating that Nori has said something for which a response can be expected from Eric. Instead, Nori produces an inbreath before producing 'oldest animal' and Eric waits until Nori has articulated the utterance before responding. Once again, then, this three-part list of adjective forms can be understood as self-talk used to overcome some difficulty with the L2, and thus merits the status of private speech.

Searching for the Correct Pronoun

   In the second example, Nori and Eric have been talking about surfing and the fact that the waves have been high recently. Eric has commented that according to the surf report, the waves are smaller today.

Example 2

```
01  N:  n s'koshi.  ↑demo    [.hh
           little        but
02  E:                       [yeah little bit smaller
03      [today.
04  N:  [bah uh bu- but (0.9) .h ° uhm n n n°
05  E:  ah you went   [this morning?
06→N:                 [uh I-         I: ↓n. .hh ° n bo-
07→      boku ni ↑(haga n) ↓(oni) ↑I my ↓me° .hh ↑me .h (.)
           1SG    PP
08      ↑me:↓: (0.2) me think (.) big wave. h hh
09      [.h huh .hh              [y(h)es. huh
10  E:  [oh you think they're big [waves?
11  (0.9)
12  N:  .hh  [kowai.
               frightening
```

```
13 E:        [so they're bi- they're big for you.
14 N: n.
```

       In line 01, Nori produces in Japanese a very mitigated agreement that the waves are smaller, with an agreement token ('n'), the word '*sukoshi*' ('little,' used for amounts), and '*demo*' (used as a qualifier, literally 'but'). In line 02, Eric accepts the downgrading by reformulating it in English, with stress on 'little.' Nori then starts what appears to be a further mitigation of his agreement in line 04, but before he can bring it to completion, Eric displays a new understanding of Nori's surfing activities in line 05. In overlap with this, Nori first appears to continue what he started in line 04, saying 'uh I' as Eric says 'this morning,' and then producing a slightly elongated 'I.' However, he then responds to Eric's new understanding with a confirmation token ('n'). After an inbreath, he lowers the volume of his talk and utters what appears to be the first syllable of '*boku*,' a first-person reference form in Japanese, then '*boku ni*,' the same reference form followed by a postposition. This can be translated in various ways depending on the context, but includes 'to me' or 'for me.' He then produces a segment that is incomprehensible, but which appears to be in Japanese, followed by 'I my me.' With the shift up in pitch on the first item in this list and the shift down on the last, and the fact that it contains three parts, the sequence comes off as a complete unit. Nori then returns to normal volume as he repeats 'me,' repeats it again in elongated form at the start of line 08, and then, after a pause, uses it, saying 'me think … big wave.' In overlap with Nori's laughter, Eric reformulates this in line 10 with a candidate understanding, which Nori confirms, while continuing to laugh, in overlap with the end of the reformulation. In line 12, Nori evaluates his subjective experience of the waves in Japanese, in overlap with which Eric provides a second reformulation of 'me think … big wave' that Nori confirms in line 14.

       What Nori appears to be doing starting in line 04 is adding to the mitigation of his agreement that the waves have decreased in size. What this mitigation eventually comes to is that he thinks the waves are big or that they are big for him. However, he encounters different kinds of difficulty in trying to produce this mitigation in English, first difficulty in finding a way to continue after 'but' in line 04, then difficulty caused by the need to respond to Eric's display of new understanding in line 05, and finally difficulty in finding a way to continue after 'I' in line 06. He deals with this difficulty first by switching to Japanese. The first-person reference form that he produces, '*boku*,' can be translated into English as 'I' or 'me,' depending on the context, or as 'my,' if followed by a form of the genitive marker '*no*.' This is not the only first-person reference form in Japanese. There is a range of both first-person and second-person reference forms that index a range of social meanings, which do not change depending on case, and which, while generally translated into English with personal pronouns, function quite differently from English personal pronouns (Ono and Thompson, 2003). As he continues to deal with the trouble, he switches back into English. The three-part list of first-person pronoun forms that he produces is something that he almost certainly would have been taught in secondary school English class in Japan. Nori uses this list of forms to lead him to 'me,' which he is then able to use. Although 'me think … big wave' is ungrammatical, Nori is nevertheless successful in overcoming the difficulty of trying to further mitigate his agreement that the waves have gotten smaller. Additionally, the stretch of talk in Japanese and English from the end of line 06 to the first 'me' in line 07 comes off as self-talk. One reason for this is its lowered volume. Another reason is that, even though the raised pitch on 'I' and the lowered pitch on 'me' makes this list of forms sound like a complete unit, Nori follows this with an inbreath and repetition of 'me' with a higher pitch, both indicating that he is continuing and not treating 'I my me' as something for Eric to respond to. For his part, Eric does not attempt to respond to what Nori has said at a lower volume. He only responds after Nori has successfully produced the further mitigation of his agreement. This stretch of talk, then, in dealing

with difficulty and being self-talk, can be heard as private speech.

Private Speech as Social Action

Based on the two criteria that what Nori says is oriented to overcoming some kind of difficulty in using his L2 and that it is self-talk, many other cases of private speech can be found in the data. The examples above are particularly clear cases. Based on these examples, I will argue in this section that the production of private speech is, in addition to whatever else it may be, a social action.

Unshared Laughter as an Index of Difficulty

When people talk about or experience some kind of trouble or difficulty, one thing they may do to make light of the trouble is laugh (Glenn, 2003; Jefferson, 1984). While laughter is often shared, it may be inappropriate to join laughter that is being used in this way, as it would amount to making light of the other person's difficulty (Glenn, 2003; Jefferson, 1984). This type of laughter is thus termed unshared laughter. For instance, in Example 2, Nori laughs after he says 'me think … big wave,' first in overlap with Eric's reformulation of this and then as he confirms the reformulation. Eric does not join this laughter. One way of understanding this is that Nori is making light of his limited surfing skill, in that if the waves are big for him, this implies that they may not be big for other, more skilled surfers, so stating that the waves are big for him is an implicit admission of his limited surfing skill. By treating this as something to laugh at, a laughable (Glenn, 2003), he also shows that this is something that he is not overly concerned about.

In Example 1, Nori laughs in line 18 as he produces the possible three-part list of adjective forms, partially obscuring one of the forms and thus rendering ambiguous what he is actually trying to say. Again, Eric does not join this laughter. What Nori is targeting as the laughable is not completely clear, but it may be his use of this list of forms itself to help him find the superlative form, especially as he had just a few moments earlier done the same thing. Nori can this be heard as making light of his difficulty in finding the correct form and, by extension, of his limited proficiency in English. Through this unshared laughter that is part of private speech, he takes a public affective stance toward his English proficiency.

Construction of Interactional Space for Private Speech

When Nori produces private speech, he and Eric collaborate to construct space within the interaction for it to occur. If the interaction is thought of as moving from left to right, as is shown in the English-based transcript, the left boundary of the interactional space is constructed through resources use to initiate self-repair (Schegloff, Jefferson, and Sacks, 1977); that is, pauses (Example 1, lines 09 and 17), inbreaths (Example 1, line 17; Example 2, line 06), 'uh' (Example 1, line 09), cutoff words (Example 1, lines 17 and 18; Example 2, line 06), and elongated words (Example 1, lines 17 and 18; Example 2, line 06). For his part, Eric does not attempt to intervene when Nori initiates self-repair.

Nori constructs the right boundary by using an item from the three-part list in the private speech as part of a larger unit addressed to Eric. This is particularly clear in Example 1 with 'biggest animal' (lines 10 and 12) and 'oldest animal' (lines 18 and 19). It can also be seen in Example 2 with 'me think … big wave,' but the repetition of 'me' makes the right boundary somewhat fuzzy. While Eric refrains from interfering in or responding to the private speech, he collaborates in constructing the right boundary by responding to the larger unit addressed to him.

While Nori's private speech, as self-talk, is not responsive to anything in Eric's talk and does not make a response relevant, a place is nevertheless constructed for it within the local sequential organization of the interaction. It is both designed to fit the sequential context and to contribute to the creation of the local context for what comes after the right boundary.

Grammatical Knowledge Displayed as Three-Part Lists

Though this is not the only way in the complete dataset that Nori articulates his grammatical knowledge of English, each of the articulations of grammatical knowledge above is carried out as a three-part list. As shown by Jefferson (1990), three-part lists are not only very common in conversation (at least in English), but also have a normative element, in that when a specific third part is not available, it may be provided by a more general item (e.g., 'and stuff'). The construction of a list structured as three parts is something that participants can be seen to work at and achieve in interaction. An interesting possibility is that three-part lists may also provide a resource for organizing explicit L2 grammar knowledge. It is very likely that Nori has been taught, possibly multiple times, the particular lists used in the examples above as three-part lists to be memorized and their three-part structure may be an aid to remembering them. Moreover, even if these lists are memorized, their production is not something that just happens mechanically. Rather, Nori must reproduce the list each time he uses it, with the reproduction itself being a mnemonic strategy. In each case, he reproduces the list first as a three-part list and then chooses one item, in each case the final item, as the one to use. He thus shows at least a limited orientation within his private speech to the structural normativity of three-part lists.[4]

Display of Understanding of Source of Ungrammaticality

In Example 1, Nori's use of private speech twice leads him to choose the correct form for the superlative. However, in Example 2, Nori's private speech leads him to choose the accusative form of the first-person singular pronoun, which when serving as the subject of the verb 'think' results in the ungrammatical 'me think … big wave.' There are at least two, not necessarily incompatible, ways that his choice of 'me' is an understandable choice. First, it could be that as the last item in the three-part list of pronouns, it is the easiest one to choose. That is, the structure of the three-part list itself may have led him to choose the last item. Second, it could be his use of Japanese in his private speech that leads him to choose 'me.' His use of a first-person reference form ('*boku*') followed by a postposition ('*ni*') can be translated as 'for me,' 'to me,' or perhaps 'on me,' depending on the context. It is this second possibility that Eric appears to understand as the source of the choice of 'me.' In his second reformulation, Eric says 'they're big for you,' with stress on 'big' and 'you.' Uttered with falling intonation, it sounds like a candidate understanding (that is, a more or less tentative displayed understanding of what another has said) that Eric has more confidence in, compared with the first reformulation said with rising intonation. Also, while the nominative and accusative forms of the second-person pronoun in Standard English are identical, by reformulating what Nori has said in this way, Eric displays his understanding that what Nori was trying to say was something such as 'the waves are big for me.'

This second reformulation thus provides evidence that Eric has been paying attention to Nori's private speech, even though he does not treat it as something that he should respond to directly. This is similar to what was found in Smith (2007) and Steinbach Kohler and Thorne (2011). The former contains examples of a participant responding to what the author identified as private speech, which provides the basis for the author's argument that participants can treat private speech 'as if it were social in social interactive contexts' (348). In the latter, it is shown how the private speech of one participant related to a particular difficulty in the L2 can provide a resource for a different participant to identify what the difficulty is and offer assistance. Although as self-talk, private speech does not make a particular set of possible next actions relevant, it does contribute to the local context of the sequentially unfolding interaction.

Display of Trying to Use English

In the quote from Goffman above, he proposed that the children that Vygotsky observed were using talk to show the adult observer that they were doing the assigned task. In these meetings between Nori and Eric, they could much more easily converse in Japanese. Not only is Eric a proficient L2 user of Japanese with extended experience of living in Japan, something which Nori is aware of, but when they

converse at other times and places, they do so mostly in Japanese. The purpose of these meetings, though, is to provide Nori with the opportunity to practice speaking English, while what Eric receives for providing Nori with this opportunity is data for future research. It is not clear how well Nori understands the uses to which this data may be put, but he knows that Eric is a language teacher and presumably understands—especially as he has given informed consent—that the data will be used to investigate his learning of English.

      As discussed above, the activity that Nori and Eric are engaged in can thus be glossed as something like 'practicing English' and 'recording Nori's English for research.' Given the sort of activity, Nori is to some extent obligated to use English in these meetings. In fact, though, as can be seen in both examples, while Eric almost exclusively uses English, Nori uses a fair amount of Japanese. By using private speech to overcome difficulties with expressing himself in his L2 English, one thing that Nori accomplishes is to show that he is trying to use English. Goodwin (1987) demonstrates how self-talk provides resources for other participants to understand the behavior of the speaker, while at the same time the behavior of the speaker provides resources for understanding the meaning of the self-talk. In the examples above, Nori's private speech contributes to constituting the activity that he and Eric are engaged in as 'practicing English' and 'recording Nori's English for research,' while the activity provides the context for understanding Nori's private speech as showing that he is trying to use English.

      Some of the ways that the private speech in the examples above can be understood as social action, then, include 1) that laughter is used to index and make light of the difficulty that the private speech is being used to overcome, and through this, take a public affective stance; 2) that the participants collaboratively construct space for the private speech in the interaction; 3) that there is an orientation by Nori to the structural normativity of three-part lists; 4) that the private speech is paid attention to by Eric, which can influence what happens subsequently; and 5) that the private speech shows that Nori is trying to do what he is supposed to be doing, that is, use English, thus contributing to what they are doing as 'practicing English' and 'recording Nori's English for research.'

Concluding Remarks

      One thing that becomes apparent when reading research on private speech is that it is not always clear how some bit of talk has been identified as private speech. Even though the criteria are often clearly explained, how they are applied to identify particular instances is often not. While a few studies (e.g., DiCamilla and Antón, 2004; Steinbach Kohler and Thorne, 2011) provide examples and explain the basis for identifying the examples as private speech, it seems more common to provide examples without such explanation. In a few quantitative studies, all the data are numeric and no examples of private speech are provided at all. One result is that the reader must often accept, on the basis of such things as inter-rater reliability measures, the researcher's ability to apply the criteria. As shown above, by paying careful attention to the details of how talk is produced in interaction, it is possible to clearly demonstrate how the private speech has been identified by the researcher. This provides the reader with resources for assessing the validity of the researcher's claims, as opposed to the statistical reliability of the coding procedures.[5] A weakness of the data used for this paper is that it is only audio data. Some research on private speech has made use of video data and included gesture as part of private speech (McCafferty, 1998; Lee, 2008). With such data, careful attention to the details of gesture, gaze, and other aspects of embodied practice is as important as—perhaps even inseparable from—attention to the details of the talk.

      More importantly, by looking at private speech as social action, it becomes apparent that private speech is not just an etic category or a theoretical construct, but

also an action that the participants in interaction may orient to as what they are doing. Paying attention to the details of how private speech is produced and how it is (or rather, is not) responded to allows for a more emic perspective, in that it is through the details of their interaction that participants make visible, for one another and, incidentally, for the researcher, their understanding of what is happening. This lessens the risk of simply imposing the researcher's theoretical commitments onto the participants' talk (Schegloff, 1997). As discussed above, social actions are designed to be recognizable, though this should not be taken to preclude the possibility of deception and does not mean that participants' understanding of what an action is cannot be revised. When a participant produces private speech, it is produced so as to be recognizable as private speech, something that researchers rely on as they attempt to identify private speech.

I have argued elsewhere (Hauser, 2011, 2013b) against the use of so-called exogenous theories of learning within a CA approach to language learning (i.e., CA-SLA, see Kasper and Wagner, 2011). The primary target of this argument is the use of a different variety of sociocultural theory, situated learning theory (Lave and Wenger, 1991), in CA-SLA. However, the same basic argument applies to other theories of learning, such as Vygotskian sociocultural theory. As shown in Hauser (2013b), though, this does not mean that connections cannot be drawn between work in CA-SLA and work in other approaches. In this paper, I hope to have shown how connections may be drawn between work in CA-SLA and the Vygotskian approach to language learning (Lantolf, 2011), without, however, compromising the CA principle of avoiding *a priori* theorizing. As I see it, there are three such connections that can be drawn with regard to private speech. First, attention to the details of interaction and the provision of examples of how private speech has been identified can allow the reader to assess the validity of the researcher's claims. Second, by adopting an emic perspective, it is possible to see what the participants themselves are doing when they produce talk (and embodied action) that is recognizable as private speech. Third, private speech, at least when it is produced within interaction, is social not just in origin, but social each time that it is produced. There may thus be other reasons aside from task difficulty for the use of private speech.

Notes

1. This may lead, though, to an observer's paradox, in that the participants were being recorded and could be thought of as producing talk and gesture for the researcher. See the discussion below of Goffman's (1981) criticism of Vygotsky.
2. See also Wittgenstein (1953) for a critique of the notion of a private language distinct from public language.
3. Though they do not mention Goffman, recent experimental research in child development by McGonigle-Chalmers, Slater, and Smith (2013) has found that the amount of private speech produced by children is much greater when an observer is present than when an observer is absent, which supports Goffman's interpretation.
4. Within Vygotskian sociocultural theory, inner speech and private speech are described as elliptical. In these examples, though, the three-part lists are not elided.
5. See Hauser (2005) for a critique of the use of coding in research on interaction in Second Language Acquisition.

References

Alarcón-Rubio, D., Sánchez-Medina, J. A., and Winsler, A. (2013). Private speech in illiterate adults: Cognitive functions, task difficulty, and literacy. *Journal of Adult Development* 20: 100--111. doi: 10.1007/s10804-013-9161-y
Antón, M. and DiCamilla, F. J. (1999). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *The Modern Language Journal* 83: 233--247.
Damianova, M. K., Lucas, M., and Sullivan, G. B. (2012). Verbal mediation of problem solving in pre-primary and junior primary school children. *South African Journal of*

Psychology 42: 445--455. doi:10.1177/008124631204200316

DiCamilla, F. J. and Antón, M. (2004). Private speech: A study of language for thought in the collaborative interaction of language learners. *International Journal of Applied Linguistics* 14: 36--69.

Donato, R. (2000). Sociocultural contributions to understanding the foreign and second language classroom. In J. P. Lantolf (ed.) *Sociocultural Theory and Second Language Learning* 27--50. Oxford: Oxford University Press.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Cambridge: Polity Press.

Glenn, P. (2003). *Laughter in Interaction*. Cambridge: Cambridge University Press.

Goffman, E. (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Goodwin, C. (1987). Unilateral departure. In G. Button and J. R. E. Lee (eds.) *Talk and Social Organisation* 206--216. Clevedon, UK: Multilingual Matters.

Hauser (2005). Coding 'corrective recasts': The maintenance of meaning and more fundamental problems. *Applied Linguistics* 26: 293--316. doi: 10.1093/applin/ami010

Hauser, E. (2011). On the danger of exogenous theory in CA-for-SLA: A response to Hellermann and Cole (2009). *Applied Linguistics* 32: 348--352. doi: 10.1093/applin/amr015

Hauser, E. (2013a). Expanding resources for marking direct reported speech. In T. Greer, D. Tatsuki, and C Roever (eds.) *Pragmatics and Language Learning, Volume 13* 29--53. Honolulu: National Foreign Language Resource Center.

Hauser, E. (2013b). Stability and change in one adult's second language English negation. *Language Learning* 63: 463--498. doi: 10.1111/lang.12012

Hayano, K. (2013). Question design in conversation. In J. Sidnell and T. Stivers (eds.) *The Handbook of Conversation Analysis* 395--414. Malden, MA: Wiley Blackwell.

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Heritage, J. and Stivers, T. (2013). Conversation Analysis and sociology. In J. Sidnell and T. Stivers (eds.) *The Handbook of Conversation Analysis* 659--673. Malden, MA: Wiley Blackwell.

Holtgraves, T. M. (2002). *Language as Social Action: Social Psychology and Language Use*. Mahwah, NJ: Lawrence Erlbaum.

Jefferson, G. (1984). On the organization of laughter in talk about troubles. In J. M. Atkinson and J. Heritage (eds.) *Structures of Social Action: Studies in Conversation Analysis* 346--369. Cambridge: Cambridge University Press.

Jefferson, G. (1990). List construction as task and resource. In G. Psathas (ed.) *Interaction Competence* 63--92. Washington D.C.: University Press of America.

Jefferson, G. (2004). A sketch of some orderly aspects of overlap in natural conversation. In G. H. Lerner (ed.) *Conversation Analysis: Studies from the First Generation* 13--31. Amsterdam: John Benjamins.

Kasper, G. and Wagner, J. (2011). A conversation-analytic approach to second language acquisition. In D. Atkinson (ed.) *Alternative Approaches to Second Language Acquisition* 117--142. London: Routledge.

Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (ed.) *Sociocultural Theory and Second Language Learning* 1--26. Oxford: Oxford University Press.

Lantolf, J. P. (2003). Intrapersonal communication and internalization in the second language classroom. In. A Kozulin, V. S. Ageev, S. M. Miller, and B. Grindis (eds.) *Vygotsky's Educational Theory in Cultural Context* 349--370. Cambridge: Cambridge University Press.

Lantolf, J. P. (2011). The sociocultural approach to second language acquisition. In D. Atkinson (ed.) *Alternative Approaches to Second Language Acquisition* 24--47. London: Routledge.

Lantolf, J. P. and Thorne, S. L. (2006). *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.

Lantolf, J. P. and Yáñez-Prieto, M. –d. –C. (2003). Talking to yourself in Spanish: Intrapersonal communication and second language learning. *Hispania* 86: 97--109. http://www.jstor.org/stable/20062818

Larrain, A. and Haye, A. (2012). The discursive nature of inner speech. *Theory and Psychology* 22: 3--22. doi: 10.1177/0959354311423864

Lave, J. and Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation.* Cambridge: Cambridge University Press.

Lee, J. (2008). Gesture and private speech in second language acquisition. *Studies in Second Language Acquisition* 30: 169--190. doi: 10.1017/S0272263108080303

Maynard, D. W. (2013). Everyone and no one to turn to: Intellectual roots and contexts for Conversation Analysis. In J. Sidnell and T. Stivers (eds.) *The Handbook of Conversation Analysis* 11--31. Malden, MA: Wiley Blackwell.

McCafferty, S. G. (1994a). Adult second language learners use of private speech: A review of studies. *The Modern Language Journal* 78: 421--436.

McCafferty, S. G. (1994b). The use of private speech by adult ESL learners at different levels of proficiency. In J. P. Lantolf and G. Appel (eds.) *Vygotskian Approaches to Second Language Research* 117--134. Westport, CN: Ablex Publishing.

McCafferty, S. G. (1998). Nonverbal expression and L2 private speech. *Applied Linguistics* 19: 73--96.

McGonigle-Chalmers, M., Slater, H. and Smith, A. (2013). Rethinking private speech in preschoolers: The effects of social presence. *Developmental Psychology* 50: 829--836. doi: 10.1037/a0033909

Nguyen, H. t. and Kasper, G. (eds.) (2009). *Talk-in-Interaction: Multilingual Perspectives.* Honolulu, HI: National Foreign Language Resource Center.

Ohta, A. S. (2001). *Second Language Acquisition Processes in the Classroom: Learning Japanese.* Mahwah, NJ: Lawrence Erlbaum.

Ono, T. and Thompson, S. A. (2003). Japanese *(w)atashi/ore/boku* 'I': They're not just pronouns. *Cognitive Linguistics* 14: 321--347.

Parsons, T. (1937). *The Structure of Social Action.* New York: McGraw-Hill.

Pomerantz, A. and Fehr, B. J. (1997). Conversation Analysis: An approach to the study of social action as sense making practices. In T. A. van Dijk (ed.) *Discourse as Social Interaction* 64--91. London: Sage.

San Martín Martínez, C., Boada i Calbet, H., and Feigenbaum, P. (2011). Private and inner speech and the regulation of social speech communication. *Cognitive Development* 26: 214--229. doi: 10.1016/j.cogdev.2011.05.001

Schegloff, E. A. (1997). Whose text? Whose context? *Discourse and Society* 8: 165--187.

Schegloff, E. A., Jefferson, G. and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language* 53: 361--382.

Sharrock, W. and Button, G. (1991). The social actor: Social action in real time. In G. Button (ed.) *Ethnomethodology and the Human Sciences* 137--175. Cambridge: Cambridge University Press.

Smith, H. J. (2007). The social and private worlds of speech: Speech for inter- and intramental activity. *The Modern Language Journal* 91: 341--356. doi: 10.1111/j.1540-4781.2007.00584.x

Stafford, C. A. (2013). What's on your mind? How private speech mediates cognition during initial non-primary language learning. *Applied Linguistics* 34: 151--172. doi:10.1093/applin/ams039

Steinbach Kohler, F. and Thorne, S. L. (2011). The social life of self-directed talk: A sequential phenomenon? In J. K. Hall, J. Hellermann, and S. Pekarek Doehler (eds.) *L2 Interactional Competence and Development* 66--92. Bristol, UK: Multilingual Matters.

Stivers, T. (2013). Sequence organization. In J. Sidnell and T. Stivers (eds.) *The Handbook of Conversation Analysis* 191--209. Malden, MA: Wiley Blackwell.

Wells, G. (1999). Using L1 to master L2: A response to Antón and DiCamilla's

"Socio-cognitive functions of L1 collaborative interaction in the L2 classroom." *The Modern Language Journal* 83: 248--254.

Wittgenstein, L. (1953). *Philosophical Investigations*. Malden, MA: Blackwell.

Wooffitt, R. (2005). *Conversation Analysis and Discourse Analysis: A Comparative and Critical Introduction*. London: Sage Publications.

Yule, G. (1997). *Referential Communication Tasks*. Mahwah, NJ: Lawrence Erlbaum.

Appendix: Transcription Conventions and Translation Symbols

Transcription conventions (based on Jefferson, 2004):

.	falling intonation
?	rising (questioning) intonation
,	fall-rising (continuing) intonation
_	flat intonation
[	start of overlap
]	end of overlap (not always used)
=	continuation of a turn-at-talk across noncontiguous lines of transcript; also latching between turns-at-talk
:	elongation
-	cutoff
>x<	relatively fast talk
x	stress
(.)	audible silence of less than two tenths of a second
(0.5)	silence measured to closest tenth of a second
(bah)	best guess at mostly unintelligible talk
(xx)	unintelligible talk, best guess at number of syllables indicated by number of x characters
h	outbreath
.h	inbreath
(h)	laugh token within talk
°	start or end of relatively quiet talk
↑	shift in pitch up
↓	shift in pitch down
→	line of special interest for analysis

Translation symbols:

1SG	first-person singular reference form (Ono and Thompson, 2003)
IP	interactional particle (Nguyen and Kasper, 2009)
PP	postposition