

Chapter 7

Stimulated Recall

CASE STUDY: JONATHON RYAN
COMMENTARY: SUSAN GASS

CASE STUDY

Introduction

This chapter reports an investigation into areas of miscommunication in interactions between first and second language speakers of English at a New Zealand university. Specifically, it focuses on the way that problems may be triggered by the use of noun phrases that are referentially ambiguous or misleading in context. For example, from a referential perspective, the utterance ‘I saw him yesterday’ is only communicatively successful if the hearer correctly identifies which person (the *referent*) the speaker claims to have seen. Problems occur when the hearer either identifies the wrong referent (misidentification) or no referent (communication breakdown). The focus of the chapter is to report on how stimulated recall (SR) was used to identify such problems.

However, identifying miscommunication is inherently problematic. For example, many studies rely on the analysis of transcripts (e.g. Cook-Gumperz & Gumperz, 2002; Verdonik, 2010), but a serious methodological limitation of such approaches is that miscommunication often leaves no clear verbal trace, particularly when it goes unnoticed by both interlocutors. Unrecognized miscommunication is also an obvious limitation for studies involving self-reported data (e.g. Tzanne, 2000).

The approach of the present study was to access the hearer’s mental representation of the discourse and compare this to the speaker’s intended message. To achieve this aim, a film retelling task was used to elicit linguistic data, enabling strong grounds from which to identify the speaker’s intended meaning. This procedure was followed by a Stimulated Recall (SR) interview in which the hearers described how their mental model of the discourse developed. This approach potentially offers much richer insights into the hearer’s interpretation of learner speech than text-based approaches.

Methodological Focus

Stimulated Recall is a type of introspective research methodology, differing from Think Aloud (see Chapter 6) in that it involves the verbalisation of cognition retrospectively rather than concurrently. It is a method used to elicit qualitative data relating to the thought processes

associated with performing an action or participating in an event. To assist recall of these thought processes, a stimulus is used, such as a video recording of the activity. It is argued that such stimuli may enable a participant “to relive an original situation with vividness and accuracy” (Bloom, 1953:161).

For much of the 20th century such introspective methods were treated with suspicion (Ericsson, 2003). More recently, such methods have gained respectability through the theoretical basis provided by researchers such as Ericsson and Simon (1984), and the contribution of multiple researchers in developing best practice guidelines. Nevertheless, even supporters of SR express caution. There are, for instance, multiple pitfalls in implementing SR procedures (see Gass & Mackey, 2000:84-99), and so it is of some concern to Lyle that few studies “treat the procedures involved as unproblematic and few studies report the SR protocol in critical detail” (2003:861). In particular, Lyle warns that care is required to minimise the risk of SR data not accurately representing cognitive processes from the time of the original event, particularly in relation to processes such as re-ordering, reasoning, and ‘sanitization’.

Stimulated Recall has most often been used to explore aspects of cognition that lie behind the participants’ decisions and actions. For example, SR has been used to explore teacher cognition (Polio, *et al.*, 2006), learner cognition (Lam, 2008), language processing in translation (Dechert, 1987), and learner reflection and self-evaluation (Murray, 2010). Less frequently, SR procedures have been used to explore hearer responses to speech, such as Tyler’s (1995) study of the perceptions of conversational interactants, and Bloom’s (1953) study of learners’ thoughts during lectures and tutorials. The present study employs this methodology in an area that has not been explored in the studies cited: how interlocutors’ mental models of a narrative developed during discourse. In particular, it attempts to identify mismatches between the intended referential meaning of speakers and the interpretations of addressees.

The Study

The main research questions of the study concern the extent to which second language learners (SLL) are pragmatically competent in communicating referential speech acts in English, and the conditions under which referential miscommunication occurs.

I collected data from 60 participants, arranged in 30 dyads, in which each partner was assigned a role as either a speaker or a hearer. Some speakers were native English-speaking students and others were students for whom English was a second language; all the hearers were first language speakers (L1), and among these were teachers of English as a second language.

Linguistic data were elicited through a film retelling task adapted from Perdue (1984, 1993), in which an edited version of the silent Charlie Chaplin film *Modern Times* is used. The participants in each pair watched the first five minutes of the film together; the ‘hearer’ was then called away while his/her partner continued watching the movie. A few minutes later, the hearer returned and asked to be told what had happened in the film during their absence. This retelling task was video- and audio- recorded, and was immediately followed by a two-part stimulated recall interview with the hearer, using the video recording as recall stimulus. These interviews are the focus of the present chapter.

Procedural Steps

Prior to the interview, the hearers were informed that the main focus was on their thoughts at the time of the original interaction. They were asked to cast their minds back to when the narrative was being told, and comment on what they recalled of their understanding at that time. In the first part of the SR interview, the hearers were shown a video recording of the narrative interaction. The video recording was then periodically paused by either myself or the participant, and the participant reported their (original) understanding of the narrative or made other (often evaluative) comments. At this point, I was often, but not always, able to identify miscommunications that had occurred. Interestingly, I missed some miscommunications because my knowledge of the film influenced how I interpreted the hearers' comments (completing a circle of miscommunication).

In the second part of the interview, the hearers were shown Part Two of *Modern Times* and were asked to comment on anything that was different in the film to what they had pictured. At this stage I was able to confidently confirm or disconfirm earlier evidence of miscommunication, and also to identify further instances that had gone unnoticed.

The research protocols were adapted from those outlined in the literature, particularly those discussed by Gass and Mackey (2000). Prior to settling on specific protocols, I conducted 16 piloting sessions over six months, and spent considerable time reflecting on the efficacy of specific procedures. It became clear that substantial practice is required in conducting SR interviews: the techniques are subtle, and the data are easily compromised by inadequate technique. I found it particularly valuable to have 'critical friends' observe or participate in interviews and comment on the procedures. Many of the most useful comments came from critical friends with no direct connection to applied linguistics or the education setting but who brought skills from an unrelated field (e.g. a licensed professional counsellor). Even so, I noted further issues during formal data collection, forcing minor procedural modifications. Finally, I also commissioned a colleague to prepare and conduct a similar task with myself as the participant. This happened too late in the data collection process to influence the procedures I used, but it did strengthen my thinking in relation to best practice procedures and in relation to certain theoretical issues relating to the research topic. For future studies, I recommend that the researcher adopt the role of a participant in the final stages of piloting their data collection method.

Stimulated Recall Data Samples

Examples of the data are presented in Extracts 1 and 2 below, from an L1-L1 and a SLL-L1 interaction respectively (in each case the teller is labelled A, the hearer B, and the researcher Res). The left-hand column presents data from the narrative retelling task, while the right-hand column presents corresponding comments from the two-part SR interview. In relation to Extract 1, the hearer reveals in the first part of the SR interview that she interpreted all of the underlined pronouns to be referring to 'the fat woman'. However, as the hearer discovers while viewing the film in *SR part 2*, this had been a miscommunication.

Extract 1

<i>Narrative retelling</i>	<i>SR interview</i>
<p>A – Charlie falls on <u>the fat woman</u>, like two times and <u>she</u>'s all like like this (GESTURE)</p> <p>B – [LAUGHS]</p> <p>A – and then um <u>she</u> like gets up and looks real mad?, like <u>her</u> face?, and then suddenly like <u>she</u> pushes the policeman or something,</p>	<p><i>Part 1</i></p> <p>Res – who got up and looked mad?</p> <p>B – the fat woman</p> <p><i>Part 2</i></p> <p>B – ah, it WAS her [[<i>the banana girl</i>]], [LAUGHS] I thought the fat angry woman got angry</p>

In Extract 2, the hearer reveals that she had identified two possible candidates as the relevant referent, but was uncertain which one was intended. However, knowledge of the *Modern Times* narrative confirms that the speaker was referring to neither of these, but a third machine that had been only briefly sighted by the hearer.

Extract 2

<i>Narrative retelling</i>	<i>SR interview</i>
<p>A – so, you know at the last part, we see together, when they brought the machine?</p>	<p>B – when she said the machine, I thought of two machines obviously*</p> <p>B – there was the conveyor belt,</p> <p>Res – ah, and that other one</p> <p>B – and the other big machine where the guy talked to him and he changed the speeds</p>

*Note: Some intervening comments have been omitted

Methodological Implications

A number of methodological issues arose during piloting, and this section focuses on how these were addressed. These relate to time lapse, memory or re-interpretation, use of video stimulus, timing of recall prompts, and the formulation of recall questions.

A critical issue when using SR is how much time elapses between the activity and the interview, as the greater the delay, the greater the potential for memory decay (Gass & Mackey, 2000). One of the decisions I faced was whether to closely examine the video recording prior to the SR. However, it seemed that the potential benefits of doing so might be outweighed by issues of reliability stemming from memory loss. In addition, because I was present in the room during the retelling task, I was able to mentally note salient issues at that time. I therefore decided to conduct the interviews immediately after the retelling task, with perhaps two or three minutes of transition time in setting up the video playback. On reflection, this approach appears to have been the right choice, although I became aware during transcription of certain issues that I would like to have probed further. However, these missed opportunities appeared relatively few, and were minimised by my training as a researcher through the piloting stage: I became much more aware of how to elicit full and appropriate information, and of potential ambiguity in the data and the hearers' responses.

Secondly, it is essential in SR that participants comment on memories rather than on a present interpretation. The interviewer can guide interviewees by using questions with past tense verbs (Gass & Mackey, 2000:92-93) and adverbial time markers. However, past tense verbs

appear sometimes to be overlooked by hearers (and are, after all, not always indicators of past time), and a subtle, yet important issue may be the placement of adverbial time markers at the beginning of an utterance rather than at the end. For example, the placement of the adverbial at the end of the question in Sentence 1) below may prompt the hearer to begin formulating an answer relating to their present interpretation before realizing that the question relates to a memory. This could compromise the accuracy of the recall. In contrast, the second sentence seems to clearly establish prior to any further mental processing that a memory is being sought.

- 1) What did you think she think she meant when she first said it?
- 2) When she said first said that, what did you think she meant?

A third procedural issue relates to the effective use of the video stimulus. Although it is usually argued that video is the most powerful prompt of memory recall, it has also been argued that video may “produce a much more foreign stimulus than audiotapes” (Yinger, 1986:271). One consideration is that observing oneself on video may reposition the interviewee in relation to the event. For example, studies of teacher cognition employing SR techniques face a problem in that participants view themselves from a perspective that is very different from their original experience of the actual event. Initially, I filmed both the speaker and the hearer, assuming that the hearers’ physical responses would provide important recall stimulus. However, I found that hearers often became distracted by seeing themselves on video, sometimes focusing more on their appearance than on recalling their understanding of the narrative. In response, I altered the direction of the camera to capture only the speaker. This modification appeared more effective in re-creating the experience of the original interaction.

A fourth procedural issue is the timing of recall prompts. During piloting, I frequently paused the video immediately after the use of a referring expression (RE) to discover which referent the hearer had identified. Underlying this decision was an (erroneous) assumption that hearers immediately identify the referents of REs. However, even before recognising this error, I found these pauses to be unnatural as they interrupted the flow of discourse. Furthermore, the hearers often seemed unsure and would ask to view the clip again, even when they felt there had been no miscommunication during the original interaction. It soon became clear that some REs are resolved not immediately, but at the end of a clause or tone unit (see Kehler, 2002). In response, I began pausing only at natural discourse boundaries, such as complete syntactic units, tone units, and episode boundaries.

This issue of timing of recall prompts relates to a fifth procedural issue: the focus of recall questions. Initially, after the use of a RE, I would directly ask the hearer which referent they had identified. This type of question is illustrated in Sentence 3) below. However, I found it more effective to ask the participant to describe a more general picture of their understanding of the narrative at that point. This approach enabled the discourse to be segmented into larger units that did not disrupt the discourse flow, and often also clarified in one utterance how the hearer resolved a number of REs. Further, it appeared unnatural to frequently ask questions about the identity of referents, and seemed to slightly unnerve the participants, perhaps feeling more like an interrogation than a discussion. I therefore preferred to use question types such as those illustrated in Sentences 4) and 5) below. Single questions like these were able to elicit the hearer’s understanding of a larger stretch of discourse, and in so doing, reasonable inferences were able to be made of how the hearer interpreted multiple REs.

- 3) When she said ‘that big guy’, who did you think she meant?
- 4) What was your understanding of the film at this point?
- 5) What was your mental picture of the film at this point?

A crucial question raised in the literature (e.g. Gass & Mackey, 2000:20-24) relates to the type of cognitive experience that can be researched using SR techniques. It is widely assumed, for example, that declarative knowledge is available for recall, while procedural knowledge is not. The assumption of mainstream approaches to literacy is that narrative comprehension does result in declarative knowledge: from very early stages of learning to read, learners are asked to declare aspects of their understanding of narratives. The evidence for this assumption appears strong, and in this respect, miscommunication appears to be a suitable topic for investigation with SR. It is important to note, however, that the processes involved in comprehending narratives are mostly procedural, and other types of analysis are required to identify what triggers miscommunication.

This simple declarative view of comprehension-outcomes aligns with traditional approaches to reference, which hold that in fully successful communication hearers resolve all references. Yoshioka, for example, states that “[i]n order to construct intelligible discourse, it is essential that the identities of referents are *made clear at all times*” [emphasis added] (2008:236). With this in mind, in the initial stages of the data collection I assumed that the hearers should be able to identify the referents of every RE. I further assumed that any doubt or ambiguity represented at least communicative strain, if not communicative breakdown. I persisted with this view through the piloting stage, and it undoubtedly influenced the types of questions I asked, as well as the significance I placed on certain unresolved references. However the following comment changed my thinking:

Extract 3

<i>Narrative retelling</i>	<i>SR interview</i>
A – and in the next part it shows Charlie and that other guy still doing . . whatever’s [LAUGHS] happening on the conveyor belt,	Res – when she said Charlie and that other guy, did you know who she meant?*
	B – oh nah yeah, I was thinking it was this guy, or that guy, but I didn’t really care, I was like ‘anyway, next part of the story’

*Note: Some intervening comments have been omitted

In this extract, the hearer indicates that she had not, in fact, attempted to precisely resolve the reference, despite her obvious cooperation in interpreting the overall narrative. What appears to happen is that, under some circumstances, the speaker will signal, and the hearer will tolerate, some referential ambiguity (see Ryan, 2012, for discussion of degrees of referentiality). An important methodological implication is that a direct question from the researcher (as in Extract 3 and Sentence 3), may actually prompt the hearer to settle on an interpretation of the reference *during* the SR interview. That is, the researcher’s question may prompt a resolution when no such resolution actually occurred at the time of the original interaction. To counter this problem in subsequent interviews, direct questions about reference resolution were used more sparingly, and with greater caution.

The issue illustrated in Extract 3 also highlights the importance of minimising participant anxiety. It was only after re-evaluating my theoretical stance on reference (and introducing the concept of degrees of referentiality) that I was able to recognise that my question in Extract 3 was a leading one. However, this insight may have gone unnoticed if the hearer had not been sufficiently confident, relaxed, and perhaps strong-willed: these factors may have enabled her to assert what she recalled of the interaction, rather than be led by the question.

The importance of avoiding leading questions is well known, and is one of the key recommendations by Gass and Mackey (2000). However, when checking my understanding of the participants' utterances, I sometimes found an inherent tension between not leading the participant, and needing to present myself as a cooperative listener. To confirm or disconfirm miscommunication, I often needed to check my understanding of what the interviewee said. Ideally, checking would be done with an open question (*who stole the bread?*), but to repeatedly do this may have appeared either uncooperative or perhaps signalled problems with the interviewee's recount. I therefore settled on a balance between open and confirming questions (e.g. *so the girl stole the bread?*), reserving the latter for cases where I was reasonably confident in my interpretation. This use of leading questions for confirmation checking is supported by Kvale (1996:158).

In studies involving SR, attention must also be given to the concept of face, and in particular, face threats. This issue becomes relevant in many applied linguistics studies where participants may feel that they need to justify their actions or use of language, particularly where they perceive a threat to their status as teaching professionals, or competent language users. The risk is that participants report not on their recall, but provide a rationalisation or justification of their behaviour, especially if they feel defensive. In the present study, the most salient face threats are those created by miscommunication. As Tzanne (2000) argues, miscommunication poses a face threat to the speaker's need to be viewed as a competent communicator, and also to the hearer's need to be viewed as an intelligent, cooperative listener. Although I did not interview the speakers (where face appears most at risk), there were three interviews in which I felt that the hearers became slightly defensive over some questions relating to miscommunication, and on some occasions appeared reluctant to acknowledge some of the problems. In one such case the participant was a teacher working at a school where I had previously been the academic manager, and she may have felt (despite the pre-interview assurances) that there was an evaluative aspect to the interview (see also the discussion in Chapter 4). Although I developed procedures to minimise anxiety, Bloom (1953:162) notes that the rapport established during the interview is critical to eliciting the interviewees' more private thoughts. In hindsight, it may have been better to spend longer developing rapport between myself and the participants.

Reflection

In this study, I used SR to address a fundamental problem in miscommunication research: how to confidently identify miscommunications that go unnoticed by the interactants. While there can be no certainty that all miscommunications were identified, it is clear that a large proportion of those that I identified would have remained undetected in a text analysis approach, or any other approach which did not seek to compare directly the addressee's interpretation with the original narrative (see also Ryan & Barnard, 2009).

As I have argued, there appear to be strong reasons to believe that the interpretation of oral narratives results in declarative knowledge that can reside in long-term memory and be available for accurate recall. There are also, however, fundamental limitations in the use of introspective methodologies, including SR. In the present study, for example, despite doing the stimulated recall immediately after the narrative task, the delay between a participant's initial experience of the narrative and the recall means that some memories were lost. This problem appears inevitable when information is transferred from short- to long-term memory (Ericsson & Simon, 1984). There are also a number of other issues of validity, such as whether participants are recalling memories or responding directly to the stimulus, and whether introduction of the stimulus alters the original memory. In addition, there are validity issues specific to this study, such as the effect of face concerns on participants' willingness to identify miscommunication. These concerns cannot be dismissed, but rather, need to be managed through maintaining 'best practice' procedures, such as those summarized by Lyle (2003:865-866).

Turning now to issues of teacher cognition, I initially intended to include an exploration of how miscommunications involving teachers (as hearers) compared to those involving non-teachers. As Gass and Varonis (1984) demonstrate, familiarity with learner language assists in making sense of SLL discourse, and James (1998:211) has noted that if an error type occurs frequently, then interlocutors have "to learn to accommodate it, and to make adjustments in one's readings." A consequence could be that teachers become oblivious to the potential for certain persistent learner errors to trigger miscommunications with less experienced interlocutors. Ultimately, due to time and space constraints, such issues were not pursued in this study, but SR techniques are likely to prove useful in exploring them in future research.

In conclusion, when used with due caution, SR seems an appropriate method to elicit data in miscommunication research and, more generally, in research concerning the interpretation of discourse. It could also be used to research learner success in comprehending L1 oral and written language, and to explore aspects of teacher cognition in relation to the interpretation of interlanguage. However, it is important to remember that SR provides access only to the conscious outcomes of comprehension, not to the cognitive processes involved. The actual triggers of miscommunication must be inferred within a relevant linguistic, socio-cultural, or cognitive framework.

COMMENTARY

Stimulated recall is a methodological tool that is part of a broader set of methodologies known as verbal reports (Bowles, 2010). Verbal reports themselves cover a range of elicitation types, including those which occur concurrently with an event (think alouds) as well as those which occur after an event. Stimulated recalls fall into the latter category and in the past decade have become a frequent way of understanding second language learners' cognitive processes while carrying out a task, most frequently an oral interactive task.

An early discussion of verbal reports can be found in works of Ericsson and Simon (1984, 1993). In their work, they categorize reports according to the following characteristics: 1) temporal characteristics (concurrent or retrospective) and 2) whether actual thoughts are tapped or whether additional information is being provided by a participant. Following these two parameters, stimulated recall is retrospective and requires thoughts about a prior task.

Stimulated recall is, by definition, a complement to other data since other data are used as a stimulus for the recall. Simply put, a stimulus from a task is used as the basis for asking participants about their thoughts during that task. Production (and even receptive) data (e.g., reading) are not sufficiently rich to allow researchers to understand learners' concurrent thought processes. Stimulated recalls fill this gap.

With some production data, for example, writing data, one can more easily obtain data on what learners are doing during production. These are known as think aloud reports. With oral production or interaction data, this is clearly not possible since one cannot simultaneously participate in an oral task and verbally provide thoughts about the task. In other words, it is virtually impossible to gather concurrent reports during oral interaction (but see Chapter 6 in this volume where Jinrui Li reports on the use of think aloud procedures during the marking of assignments).

The most common scenario for the use of stimulated recall as a methodological tool in second language research (see Gass & Mackey, 2000 for a fuller description) is for learners to complete an oral interactive task. The stimulus for the stimulated recall is the audio or video recording of the task itself. The researcher and the learner meet following the interaction and watch/listen to the recording of the task. The tape is stopped at strategic points (e.g., when there is a communication breakdown) and the learner is asked to comment on what they were thinking about at the time of the task. The stimulus, thus, serves as a reminder of the event.

Like any research tool, stimulated recalls must be done carefully, or, as described more fully in Gass and Mackey (2000), one can easily find oneself with contaminated data. In other words, if we are aiming to understand what learners are thinking as they are producing language, we must be certain to as great an extent as possible that the thought processes that are being verbalized truly reflect the thought processes at the time of the original task. The purpose of having a stimulus is to trigger actual memories of what someone was thinking about. This minimises the well-known problem of veridicality (see Bowles, 2010 for further discussion of this issue). All too often, recall comments slide into comments about what learners are thinking about when seeing the video/listening to the audio rather than what they were thinking about

during the original task production. Another way of maximising validity of the tool is through timing. The closer the recall is to the event the more likely it is that the recall itself will not be influenced by memory decay. A final issue related to validity of data has to do with how questions are asked. A question such as “What were you thinking about when you said/wrote x?” is appropriate because it asks about thoughts at the time of doing. A question such as “What are you thinking?” or even “What were you thinking?” is not appropriate because it can often lead learners into producing their thoughts at the time of the recall, even they those thoughts may be about the original event. In such cases, one might get a response like “I was thinking that I wish I had said X”. A researcher is then left not knowing whether that is what the learner was thinking after the fact or whether that is what the learner was thinking at the time he or she was involved in the interaction.

Comments on Ryan’s case study

With this discussion as backdrop, I turn to Jonathon Ryan’s report which is a thoughtful commentary on the use of stimulated recall. And, it breaks new ground in the area of oral production data as he investigates specific referential information. His approach departs from data that I and others have collected which have attempted to delve into thought processes involving learners’ specific use of language and reactions to corrective feedback (e.g., Mackey *et al.*, 2000) or reactions to the benefits of captions (Winke, *et al.*, 2010), or even thoughts about rating oral speech samples, as is common in testing research (e.g., Winke, *et al.*, 2011), or even with teacher data (Polio, *et al.*, 2006).

Jonathon brings into the mix of research questions addressed by stimulated recall questions of miscommunication, or lack of understanding during listening. This is clearly not an area that has been explored using this methodology. In his report, the issues discussed relate to the methodology without much discussion of non-native/native interaction so it was difficult for me to see in this short reflective piece how second language issues played a role, although it should be noted that the reader is referred to his dissertation. This difficulty was further compounded because, as far as I could tell, the stimulated recalls were done with native speakers as the participants. Notwithstanding, Jonathon’s paper confirms difficulties already known to be a problem with stimulated recalls and brings in others that are not as commonly dealt with. He is to be commended for the careful and thoughtful modifications made along the way in response to difficulties encountered as he was conducting his research. Important also is his commentary on the role of stimulated recall in the study of teacher cognition. One final comment relates to his methodology, although not related to stimulated recall directly: the telling of the end of a story to someone who had seen only the first part of a video is an innovative and realistic twist on the story-telling methodology used by many to elicit data. It sets the scene for a more natural data-collection methodology.

Problems revealed: Old problems and new problems

I have dealt with the problems of validity of stimulated recalls and Jonathon acknowledges many of them as well. In Ericsson and Simon’s model, the authors refer to different types of verbalisations. Verbalisations with the least amount of reactivity are those that do not have metacognitive information. They also discuss verbalisations that have additional

information (metacognitive). These they claim may slow processing and cause changes in cognitive processing (see Bowles, 2010). It is not clear whether Jonathon differentiated between these two types of verbalisations. For example, he says “The video recording was then periodically paused by either myself or the participant, and the participant reported their (original) understanding of the narrative or made other (*often evaluative*) comments” (emphasis mine). If Simon and Ericsson are correct, these latter would not be accurate reflections of cognitive processes. Further, we do not know whether these responses were or were not included in the final analysis.

Decay, as noted above, and as recognized by Jonathon, is of critical importance when dealing with veridicality of stimulated recalls. He used this concept as a guide to set up the recalls to occur within 2-3 minutes following the actual event. From his description, however, it was not clear how long the original interactive session was. If the sessions were long in duration, then the recall was itself quite a distance from the original event, despite the 2-3 minute lapse in time between the end of the original interaction event and the beginning of the recall session. This however cannot be helped other than making the interaction event short.

In describing his thought processes as he was setting up his experiment, he noted that some participants were bothered by seeing themselves on the video (personally, this has never been the case in my research, or, at least, I was not aware of it). As a result, he faced the camera away from the hearers to avoid this discomfort. As with all research, choices have to be made and this one solved one of the problems, but it results in the inability to question just those moments of lack of understanding by being able to ask “You look confused here. What were you thinking?”

His careful thinking about the questions to be asked was excellent. For example, Sentences 1 and 2 are impressively thought out. Contextualizing the question by putting the adverbial first shows the care that was put into this research and is an indication of how subtle differences may influence one’s research tool.

There were other issues raised that can be characterised as new issues in that they are not often discussed in studies that use stimulated recalls. One such issue is that of face. While this is not a new issue in the pragmatics literature or the wider applied linguistics literature, it is not often discussed in the context of recalls. Jonathon’s sensitivity to this issue comes from the fact that some of his participants were teachers who may have felt that their professional credibility was at stake. As he points out, “participants may feel that they need to justify their actions or use of language, particularly where they perceive a threat to their status as teaching professionals...” It is not clear how this aspect affects their recalls, but it is a reminder that recalls are influenced by more than just a recollection of what a participant was thinking at the time of the event.

Yet another outcome of his detailed and thoughtful examination of his procedures, of previous literature, and of his data was the recognition that referring expressions may not be resolved until “the end of a clause or tone unit”. This knowledge led him to not stop the tape until there was a “natural discourse” boundary. Doing so earlier would have been tantamount to asking participants to reveal incomplete cognitive processes. This is an interesting and novel contribution.

In conclusion, Jonathon Ryan has added yet another dimension to the use of stimulated recalls. He has taken the reader on a well-reasoned journey and shown that meticulous piloting and constant thinking about a research tool can lead to beneficial and important changes. Finally,

he shows how a research tool, in this case stimulated recall, can be made to do what it is intended to do.

Reflective Questions

1. Susan Gass referred to the two types of verbalisation identified by Ericsson and Simon. How clearly can these be differentiated in a stimulated recall session?
2. What linguistic and other difficulties are likely to occur when conducting SR sessions with participants in their second language?
3. Reducing the length of time between the event and the subsequent SR session is critical, as Susan Gass says, to reduce memory decay. However, a delay of a few minutes reduces the amount of time that the researcher has available for pre-SR reflection or preliminary analysis of the recently-received data. How do you think this dilemma might be resolved?
4. Video-recordings are commonly used to stimulate a participant's recall of events. What other means can be used to stimulate recall?
5. For what sort of activities do you think SR procedures are desirable, or even necessary?
6. Jonathon makes the point that, despite extensive piloting, he made changes to his procedure once his fieldwork was underway. To what extent do you think this reduces the reliability of his study?
7. Make a plan of the technological (e.g. recording) and logistical (e.g. time and place) issues required to carry out a SR session.
8. Replicate Jonathon's narrative task with one or two colleagues. What issues of miscommunication arise? You can download excerpts from the *Modern Times* film on youtube.
9. Draft a letter of information to potential participants about a hypothetical research project you intend to carry out. Include a paragraph explaining why and how you want them to take part in a stimulated recall session.

References

- Bloom, B. (1953) Thought processes in lectures and discussions. *Journal of General Education* 7, 160-169.
- Bowles, M. (2010) *The think-aloud controversy in second language research*. New York: Routledge.
- Cook-Gumperz, J. and Gumperz, J. J. (2002) Narrative accounts in gatekeeping interviews: Intercultural differences of common misunderstandings? *Language and Intercultural Communication* 2 (1), 25-36.
- Dechert, H. W. (1987) Analysing language processing through verbal protocols. In C. Færch & G. Kasper (eds.) *Introspection in second language research* (pp. 96-112). Clevedon, England: Multilingual Matters.
- Ericsson, K. A. (2003) Valid and non-reactive verbalization of thought during performance of tasks: Towards a solution to the central problems of introspection as a source of scientific data. In A. Jack & A. Roepstorff (eds.) *Trusting the subject?: The use of introspective evidence in cognitive science* (Vol. 1, pp. 1-18). Exeter, England: Imprint Academic.

- Ericsson, K. A. and Simon, H. A. (1984) *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.
- Gass, S. and Varonis, E. M. (1984) The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning* 34 (1), 65-87.
- Gass, S. M. and Mackey, A. (2000) *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- James, C. (1998) *Errors in language learning and use: Exploring error analysis*. London: Longman.
- Kehler, A. (2002) *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- Kvale, S. (1996) *InterViews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage.
- Lam, W. Y. K. (2008) Metacognitive strategy use: Accessing ESL learner's inner voices via stimulated recall. *Innovation in Language Learning and Teaching* 2 (3), 207-217.
- Lyle, J. (2003) Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal* 29 (6), 861-878.
- Mackey, A., Gass, S. and McDonough, K. (2000) How do learners perceive interactional feedback? *Studies in Second Language Acquisition* 22 (4), 471-497.
- Murray, J. (2010, July) *Politeness and face in professional speaking tests*. Paper presented at the 18th International conference on pragmatics and language learning, Kobe University, Japan.
- Perdue, C. (ed.) (1984) *Second language acquisition by adult immigrants: A field manual*. Rowley, MA: Newbury House.
- Perdue, C. (ed.) (1993) *Adult language acquisition: Cross-linguistic perspectives* (Vol. 1 Field Methods). Cambridge, England: Cambridge University Press.
- Polio, C., Gass, S. and Chapin, L. (2006) Using stimulated recall to investigate native speaker perceptions in native-nonnative speaker interactions. *Studies in Second Language Acquisition* 28 (2), 237-267.
- Ryan, J. (2010, July) *The miscommunication of referential speech acts in ESL*. Paper presented at the 18th International conference on pragmatics and language learning, Kobe University, Japan.
- Ryan, J. (2012) Acts of reference and the miscommunication of referents by first and second language speakers of English. PhD Thesis, University of Waikato.
- Ryan, J. and Barnard, R. (2009) 'Who do you mean?' Investigating miscommunication in paired interactions. *The TESOLANZ Journal* 17, 44-62.
- Tyler, A. (1995) The coconstruction of cross-cultural miscommunication: Conflicts in perception, negotiation, and enactment of participant role and status. *Studies in Second Language Acquisition* 17 (2), 129-152.
- Tzanne, A. (2000) *Talking at cross-purposes: The dynamics of miscommunication*. Amsterdam: John Benjamins.
- Verdonik, D. (2010) Between understanding and misunderstanding. *Journal of Pragmatics* 42 (5), 1364-1379.
- Winke, P., Gass, S. and Myford, C. (2011) *The relationship between raters prior language study and the evaluation of foreign language speech samples, TOEFL iBT Research Report*. Princeton, NJ: Educational Testing Service.

- Winke, P., Gass, S. and Sydorenko, T. (2010) The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology* 14 (1), 65-86.
- Yinger, R. J. (1986) Examining thought in action: A theoretical and methodological critique of research on interactive teaching. *Teaching & Teacher Education* 3 (2), 263-282.
- Yoshioka, K. (2008) Gesture and information structure in first and second language. *Gesture* 8 (2), 236-255.