Introduction

The issue discussed in this paper is whether task variation affects the degree of monitoring in the interlanguage of the second language learner. We will begin with a discussion of task variation and go on to consider theoretical issues relating to monitoring. There will then be a critical review of some recent research on task variation in interlanguage and this will be followed by a look at methodological considerations. Finally, the major results of the research and possible future directions will be discussed.

Task variation

Dulay, Burt and Krashen (1982:247) present a dichotomous view of task. They posit that tasks can be focused either on communication or on linguistic form. The former they term "natural communication tasks" (NCTs) and the latter "linguistic manipulation tasks" (LMTs). Dulay et al feel that data generated by NCTs permit researchers to make statements concerning the learner's "normally developing and subconscious grammar", whereas LMTs permit statements concerning the learner's "metalinguistic awareness, that is the conscious knowledge and manipulation of the rules and forms of a language". In their view, these two types of linguistic knowledge do not seem to be directly related and the tasks tapping each provide quite different results.
Hyltenstam (1983:58), Long and Sato (1984:4) and Tarone (1982:71) do not divide tasks into two distinct types. Hyltenstam distinguishes between "observational data", obtained from "more or less" spontaneous speech and writing, and "experimental data" obtained from at least eight task types:

1. elicited production; for example, Berko tests, BSM, SLOPE, guided composition;
2. manipulation of linguistic material; for example, sentence combining;
3. intuition, grammaticality judgement tests;
4. introspection;
5. cloze procedure;
6. imitation;
7. dictation or partial dictation;
8. translation.

However, he does not suggest that learners handling these tasks are involved in one or other of two distinct processes. Tarone questions Dulay, Burt and Krashen's (1982) assumption that tasks can be plotted dichotomously and posits that they can be plotted along a continuum, ranging from spontaneous speech to the most controlled activities such as discrete-point grammar tests.

Whatever theoretical standpoint one takes, it is crucial to control for task in any study that purports to make claims about how much or how little monitoring one task involves, relative to any other. Before dealing with these methodological considerations, however, this paper will look at the theoretical issues underpinning these views on task and at some recent research.

Theoretical issues

all have varying definitions of monitoring. Krashen's "Monitor Theory" is dichotomous: what is "acquired" is subconscious knowledge of the target language, whereas what is "learnt" is conscious knowledge and cannot become "acquired". Krashen further claims that Monitoring (capital M) can only operate on the conscious grammar and, even then, only under certain conditions. Originally, there were only two conditions, time and focus on form; however, after a number of studies reported findings that strongly suggested other conditions pertained, Krashen further narrowed down the circumstances in which his Monitor might be used. His more recent conditions are that learners must know the rule and be at a relatively high proficiency level. In addition, he claims there are wide individual differences ("over-users", "optimal users" and "under-users" of the Monitor) and that, in any case, only certain rules can be Monitored. These are rules that are "simple" and "easy to learn" (like the third person present simple tense "___s" morpheme), yet tend to be acquired late (if at all).

Bulstijn (1982:13) does not distinguish between Monitoring in Krashen's sense and monitoring in general because he feels it is hard to operationalize them in empirical research and that focus on form may well be due to either or both. Morrison and Low (1983:229-231) also argue for a broader definition than Krashen's. They define monitoring as a critical faculty (involving not only syntax, but phonology, lexis, truth and social propriety) and a creative faculty (enabling inferencing). In other words, monitoring could just as well focus on communication as on form.
Whereas Krashen's model is a dichotomy, Tarone (1979:183) postulates that the degree of attention learners pay to their speech can be shown on a continuum along which learners "style-shift" (1982:78). One end of the continuum represents a "superordinate" style, which is non-systematic and permeable not only to other linguistic systems such as socially prestigious variants of the native language but also to target language variants (Beebe 1980). The other end of the continuum, on the other hand, represents a "vernacular" style, which is systematic and impermeable. Tarone feels that Krashen's position is difficult to prove and that it is therefore more useful to regard monitoring (small m) as "paying attention to speech", which is a much broader interpretation.

A review of recent empirical research.

The "morpheme studies" took approximately nine English morphemes (the number varies slightly from study to study) and devised obligatory contexts for them. (For a thorough review of Supplied in Obligatory Context (SOC) analysis see Long and Sato 1984). The Bi-Lingual Syntax Measure (BSM), written for child learners by Burt, Dulay and Hernandez, was used in many of these studies (for a review, see Burt and Dulay, 1980) and the results almost invariably revealed what Burt and Dulay call the "natural order" for acquisition of these morphemes. They concluded that this order showed use of the Monitor (in Krashen's sense) was not taking place; only an "unnatural order", they felt, would indicate its use.

There was, however, no task variation in these studies and
it was precisely this issue that Larsen-Freeman (1975) addressed. She designed a study to determine whether the sequence of acquisition of grammatical morphemes for second language learners previously found in the "morpheme studies" would appear in tasks other than those requiring speech production. She used five different tasks (which she called "reading, writing, listening, imitating and speaking tasks") to collect data from 24 instructed adult learners. For the speaking task, she deliberately used the BSM, even though it was designed for children, to compare the results with those of the other tasks. Indeed, the results showed that, of the five tasks, the BSM was the most efficient at reproducing the "natural order" of acquisition of English morphemes. The other tasks showed some consistent orders for some morphemes, but there was not sufficient consistency to claim the "natural order".

Unfortunately, Larsen-Freeman used the same story for both the reading and writing tasks. This makes it impossible to know whether it was the written mode that produced a simple ordering of morpheme acquisition or whether it was because the same story was used for both tasks. In other words, the study failed to control a variable that could easily have been controlled. In addition, it is arguable whether either of these two tasks should be classified as primarily eliciting reading or writing. In fact they could probably both be more accurately described as discrete-point grammar tasks focusing on ten English morphemes, one using a multiple-choice format and the other blank-filling. This is important since these misnomers could lead to confusion
when comparing the results of this study with those of others that have genuine reading or writing tasks.

Indeed, all four of Larsen-Freeman's tasks in this study could be regarded as variations on a theme of discrete-point grammar task, where items are supplied in obligatory contexts (SOC). If we follow this interpretation, her results could be said to indicate that it would take a discrete-point task to promote sufficient focus on form to bring out the conscious Monitor (Krashen) and hence an "unnatural order". Krashen, Sferlazza, Feldman and Pathman (1976:150) suggest that Larsen-Freeman's tasks allowed more response time than the BSM. They do not substantiate their claim, however, so their suggestion cannot be validated.

Even though Larsen-Freeman's tasks may not have involved exactly what their names suggest, the study was one of the first to aim at controlling for task in interlanguage performance and its results forced researchers to be far more circumspect in making claims. It opened up hitherto uncharted areas; two important examples of such areas are, first, the effect of language mode on monitoring and, second, language other than the nine highly researched morphemes.

Houck, Robertson and Krashen (1978) replicated Larsen-Freeman's (1975) "writing" task on twenty-two "intermediate" learners and got similar results, which they interpreted as an "unnatural order" that showed use of the conscious "Monitor". Unfortunately, however, this study did not juxtapose tasks, so no inter-task comparison can be made. Nevertheless, Houck, Robertson and Krashen (1978) report on a study that did involve
task variation. Spoken language was elicited through pictures "in the style of the BSM". Their procedure was to tape the resulting interviews, get the students first to transcribe their own interviews and then to correct their own transcripts. A "natural order" was found for both the pre-corrected and corrected transcripts with only a 3% improvement between the two. This meant that Krashen's two original conditions to bring out the Monitor (time and focus on form) failed to do so.

It is unfortunate that the replication of Larsen-Freeman's writing task and the two transcription tasks were not carried out on the same subjects; as it is, there are fewer data on task variation than there would otherwise have been. One wonders, too, whether the subjects could really be relied on to transcribe their own L2 data accurately; additionally, there must have been a temptation for them to "monitor" during transcription rather than afterwards. This might explain the small margin of improvement between the two versions of the transcripts. Houck, Robertson and Krashen (1978) do not address these crucial procedural and methodological matters.

Fathman's Second Language Oral Production English (SLOPE) test, although designed for children (like the BSM), was used in a number of studies on adults. The test consists of a series of pictures and accompanying questions that create obligatory contexts for target items. Fathman (1975b:34) admits that these items of morphology and syntax were selected partly because they can easily be written for obligatory contexts. Krashen, Sferlazza, Feldman and Fathman (1976:150) ran the SLOPE test on
sixty-six instructed and uninstructed adult learners and found a "natural order". They subsequently reran the test on the subjects, but required written responses. The only result they give from this is that the third person present simple singular "___s" morpheme jumped from 20th place to 8th place (out of 20) in the rank order. Krashen, Sferlazza, Feldman and Fathman (1976) do not mention the serious possibility of a practice effect (which could have been eliminated by running the test in the written mode on half of the subjects first); this result, then, cannot be taken as a serious attempt at finding effects for task variation in interlanguage.

Fuller (1978), reported in Krashen (1981), also ran a written and oral version of the SLOPE test on the same subjects. Despite the extra time allowed for the written version, no significant differences were found, a "natural order" being reported in both. There are neither enough details of this study nor of the SLOPE test (which is unavailable) to make further comment.

Krashen, Butler, Birnbaum and Robertson (1978) report on two studies. In one of these, they aimed at finding out if the amount of "Monitoring" (to be judged by how "natural" the morpheme order was) varied between two different writing tasks. The fifty-eight instructed subjects were asked to write a "fast" and an "edited" version of a composition. A 6% increase in accuracy was found between the fast and edited versions, but it did not occur consistently for any items except the third person "___s" morpheme, nor for all subjects, and it was not enough to change the rank order. Even "edited" versions of written tasks
do not seem to bring out the "Monitor". Krashen, Butler, Birnbaum and Robertson conclude that their subjects did not focus on form, even though they had been instructed to do so; instead they focused on meaning. This is a crucial point and will be taken up again under methodological considerations.

Lo Coco (1976) compared three methods of second language data collection, in effect three tasks. Her subjects were twenty-eight university students taking an elementary Spanish course; all were native speakers of English. The three tasks were free composition, picture description and translation. There was a range of obligatory contexts, from none in the first task, to almost total control in the third. Lo Coco says that the results seem to indicate that the translation task caused the subjects to focus on form marginally more than the other two tasks. This may well be so, yet this study fails to show it. At no point in her report does the researcher say whether the picture description or the translation tasks were in the spoken or written mode. In addition, her categories for errors tend to be so wide as to be meaningless: all errors involving verbs are categorized together. The results of this study cannot be compared with those of any other study nor should they be regarded as having any validity.

Probably the most thorough and extensive study relevant to this review was conducted by Tarone (1985). There were twenty subjects, all of them adult instructed learners. The tasks were a written grammar test, an oral narration, and an oral interview. A fourth task was dropped as it was too difficult for the
subjects. One of her hypotheses was that the interlanguage style least influenced by the target language would be the vernacular style (that produced in the narrative task) and the style most influenced by the target language would be the careful style (that produced on the grammar test). Surprisingly, this was not upheld except for the third person "___s" morpheme. Learner performance on direct object pronouns and articles actually improved as tasks required less attention to form. Tarone (1985) speculates that this inverse relationship between the degree of attention to language form required by a task and the cohesiveness of the discourse may in fact be fairly common. She suggests that articles and direct object pronouns are important for textual cohesion, whereas third person "___s" is not.

There are a number of studies that report on phonological task variation in interlanguage performance. Long and Sato (1984) mention four such studies and report the general finding that the more controlled the task, the higher the proportion of target language phonological variants elicited. Space here permits the citation of only one of these studies. Beebe (1980:436) administered four tasks all in an "interview setting", (a conversation, the reading of a passage, the reading of twenty-five isolated words and a listening perception test). There were nine subjects, all adult Thais living in New York. Beebe neither reports their proficiency level nor whether they were instructed learners. These are flaws in reporting; the study itself appears to have been carefully conducted. The researcher, however, admits (1980:437) that for the "interview conversation", the number of tokens for each phonological variable varied with the
speaker, whereas the word-listing task was more uniform. This is an inherent methodological problem for this type of task (and will be dealt with in the next section). Statistics were simply percentages of correct forms that had been transcribed and scored by the same person. The findings (1980:443) suggest that monitoring operated on a continuum, increasing in formal contexts (reading a word-list) and decreasing in less formal contexts (interview conversation).

Methodological considerations

Tarone (1979:181) likens interlanguage to a chameleon in that it "varies with the subtlest shifts of situation" and says that scrupulous data-gathering methods are therefore crucial in enabling us to make empirically based claims from our studies. Seeking to systematize second language data gathering, she took Labov's (1969) methodological axioms (style-shifting, attention, vernacular, formality and good data) that show the "observer's paradox":

"if we get good recorded data, we get bad data in the sense that the speaker has focused attention on speech and style-shifted away from the vernacular, which is the most systematic interlanguage style and therefore what we want to study" (Tarone 1979:188).

There is no easy answer to this, but clarity in reporting procedures in studies is obviously paramount. Tarone (1979:189) calls for explicit detail on the following: task, interlocutors, physical surroundings and topic. Indeed, if all SLA researchers reported as fully and honestly as Tarone (1985), the field would be considerably further ahead than it is. Tarone goes on to suggest that certain procedures be followed; for example,
if the "vernacular" is required, the surroundings should be as informal as possible. In addition, she stresses that only through individual longitudinal studies of interlanguage in informal contexts will we ever be in a position to realize the systematic nature of interlanguage. She concludes by saying that elicitation devices and research in the classroom should not be abandoned. There are limitations to this method of data gathering, but it is needed to enable a distinction to be made between classroom interlanguage and that used elsewhere. Ultimately this is important for language teaching since the former is likely to contain more native as well as target language variants (and therefore be less systematic) than the latter. We might thereby gain insights as to how and when target language variants move from the classroom style to the vernacular or unmonitored style.

Long and Sato (1984) identify four crucial methodological "intrinsically related" issues in interlanguage studies. First, researchers should focus on process, not just product to obtain their "findings". However, many of the studies reviewed in this paper used SOC analysis, which focuses only on product. Consequently, for the paradigm "go-went-goad-went" (for the acquisition of English irregular past tense forms) for instance, a subject would be deemed to have acquired the form, whether (s)he were at (transitional) stage two or (target) stage four of the process. Instead, Target-Like Use (TLU) analysis could be used to prevent this flaw. (For a review of TLU analysis, see Long and Sato 1984).
Long and Sato also stress the need to treat function, not just form. Some elicitation procedures may show that a learner can manipulate a form, but that does not in itself indicate whether or not (s)he has mastered its function. SOC analysis, though frequently used in studies, equates form with function, yet evidence is urgently needed on whether monitoring focuses on meaning and communication, rather than on redundant morphemes such as third person "___s", (Krashen, Butler, Birnbaum and Robertson, 1978). Third, they call for multiple levels of linguistic analysis. Researchers should go beyond the much studied nine or ten morphemes. It would be useful, for instance, to have information on some of the factors mentioned by Morrison and Low (1983) such as lexis, truth and social propriety, as well as on the areas raised by Tarone's (1985) speculations about discourse cohesion. In addition, such areas as discourse and collocational accent could be studied. Lastly, Long and Sato mention the need to consider the context of the learner's performance, in other words whether his/her use of language is appropriate in the situation in which it is used. SOC analysis fails to detect this as well.

To conclude this section on methodological considerations, it is essential for researchers to control as many variables as possible to enable us to isolate that variability of interlanguage data which is related to the acquisition process. The sheer quantity of potential variables in our studies is alarming: variables between and within individuals and cultures, age, instruction, proficiency level, second or foreign language environment, task, interlanguage itself and so on. Tarone feels
that, at the very least, we should aim at controlling task within studies, gather data on a wider range of tasks and ensure that data are comparable across studies.

Major results

The preceding section on methodological considerations makes it clear that research is still at far too rudimentary a stage to permit us to put tasks on a continuum with regard to the amount of monitoring they induce. Just when it appeared that one task at least (discrete point grammar) could be placed at the monitoring end of any such continuum, Tarone (1985) reported her highly significant findings suggesting that style-shifting in interlanguage is more complex than previously predicted. It looks as if SLA research to date has only led us to the point where we are beginning to ask the right questions.

Future directions

The section on methodological considerations inevitably touched on future directions by outlining hitherto unfilled gaps in terms of the data we collect and how we analyze them. This section, then, deals with future directions other than those already mentioned. Tarone's (1985) findings obviously need to be followed up. In particular, we should reassess Krashen's notion of "easy to learn, simple grammar" in the light of her ideas about discourse cohesion.

In addition, the findings of studies not reviewed in this paper (because they do not look at task variation) should be followed up with studies that control for task. Seliger (1979),
for example, found no relationship between oral performance on the indefinite article and being able to state the "a/an" rule. It is possible that such a relationship would have been found if a written task had been used in addition to the oral one. Abraham's (1984) finding could also be followed up. She found evidence that the third person "___s" is acquired attached to particular verbs (rather than in isolation) and that phonetic environment, and perhaps frequency, influence the order in which the morpheme is attached to verbs. This finding reinforces Tarone's point about researchers giving the fullest possible information about their studies. In other words, if all the studies looking at the third person "___s" morpheme had given the fullest information about which verbs were elicited and with which verbs subjects supplied the morpheme, we could go back to those studies and retrospectively see if there was a pattern across studies.

Hopefully, with the benefit of hindsight, SLA researchers will now be able to run studies that genuinely address these still unanswered questions.
BIBLIOGRAPHY


