Center for Second Language Classroom Research
Social Science Research Institute
University of Hawai'i at Manoa

Technical Report No. 3

Task classification: a cross-disciplinary review

Graham Crookes

May, 1986
Technical Reports
of
The Center for Second Language Classroom Research
Social Science Research Institute, University of Hawaii at Manoa

The opinions expressed in Technical Reports of the Center for Second Language Classroom Research are those of the authors, and do not necessarily reflect the views of the Social Science Research Institute, of the University of Hawaii, or of the funding agency. This report was published by the Social Science Research Institute, University of Hawaii at Manoa. All rights reserved. No parts of this Report may be reproduced in any form or by any means without the prior written permission of the Director, Social Science Research Institute, University of Hawaii at Manoa.

Additional copies of this and other Technical Reports may be obtained for $2.00 each, which includes printing and postage, by writing to the Director, Center for Second Language Classroom Research, Social Science Research Institute, c/o Moore Hall 570, 1890 East-West Road, University of Hawaii at Manoa, Honolulu, Hawaii 96822, U. S. A. Checks should be made out to the Research Corporation of the University of Hawaii.


The Center for Second Language Classroom Research

University of Hawaii at Manoa, Summer 1986

The Center for Second Language Classroom Research (CSLCR) was established at the University of Hawaii at Manoa in the fall of 1983. Administratively part of the University's Social Science Research Institute (SSRI), it is a joint venture of SSRI and the Department of English as a Second Language (ESL).

SSRI provides a 50% time position for the Center's Director, (currently Dr. Craig Chaudron), who is a faculty member in the Department of ESL, plus administrative and technical support. Funded by outside grants and contracts, additional CSLCR staff are drawn from the faculty and students of the Departments of ESL, Linguistics, and Educational Psychology.

The work of the CSLCR includes research, curriculum development and training projects in the general area of second language (SL) education. This includes basic and applied research on SL teaching and learning, on education through the medium of a second language, and on classrooms where second dialects are present (e.g. Hawaiian Creole English). English and other second languages are included in this work. The Director of the Center coordinates all research projects and actively pursues new projects and collaborations with other agencies.

Acknowledgments

For their help in the preparation of this Report, the author gratefully acknowledges (in alphabetical order) Craig Chaudron, Dennis Day, Michael Harrington, Mike Long, Jack Richards, Ted Rodgers and Charlie Sato.
Since the middle 1970s (at least), the general ESL pedagogical repertoire has contained a number of tasks which focus on dyadic and small group communication (e.g. Allwright 1976, Geddes & McAlpine 1978, Byrne 1979). They mimic to some extent the natural interaction between native speaker and non-native speaker -- a major area of investigation in the developing SLA field of the 1970s. These tasks are often referred to as 'information-gap activities' or 'communication tasks' (usually without any attempt to define the terms). They play a major role in courses constructed according to the tenets of the "communicative movement" (Littlewood 1981: viii), or Communicative Language Teaching (CLT). In addition, as a result of their relevance to the forms of naturalistic second language acquisition (SLA), some of these tasks have been found to be useful for experimental research into SLA (e.g. Long 1980, Pica & Doughty 1985, Long & Porter 1985), and studies of this sort are increasing in number.

For example, a typical task of this sort was used by Buckley, Samuda and Bruton (1978: 134) in a study of group work (and subsequently in a study of error correction, Bruton & Samuda 1980), "Spot the Difference":

[for this task, each person is given a picture, depicting the same scene, but each one differs from the others in two small details...To solve the problem...the group has to identify these differences without looking at any other picture, apart from the one distributed to each individual at the beginning.(134)

CLT has made extensive use of tasks. Unfortunately, the design of tasks, like other aspects of materials design, has not usually been regarded as an activity which could be informed by reference to empirical research into the processes of SL learning. Littlewood (1981) organizes a survey of what he terms "communicative activities" purely according to surface characteristics such as their social or functional aspects, rather than their psycholinguistic characteristics or utility for SL learning. Little reference is made to work which would justify on empirical grounds the intuitively appealing use of these tasks to facilitate second language learning.

Littlewood separates the activities into "functional communicative activities" (FCAs) and "social interaction activities" (SIAs). In FCAs, "the teacher structures the situation so that learners have to overcome an information gap or solve a problem", and "the only aim is to exchange meanings successfully in order to complete a task or solve a problem" (21). In SIAs, there is a constraint that language chosen be socially appropriate, unlike in FCAs.

Littlewood suggests (without evidence) that because of the nature of the FCAs,

there is a wide range of communicative functions that
are unlikely to occur (e.g. greeting, inviting) ... the situations ... bear little outward resemblance to those [the learners] will encounter outside the classroom ... the learners' social role is unclear and generally irrelevant. (39)

FCAs may (it is said) be turned into SIAs through being made more realistic -- requiring learners to "simulate the social roles involved in the interaction" (43). Other SIAs are
1) classroom management activities in the L2;
2) content teaching in the L2;
3) the "conversation session", (for which the instructor must be co-communicator, and must "provide materials which are capable of sustaining interaction without the teacher's presence"), and
4) role-playing and simulation.

It should be obvious that the SIAs are in fact the bulk of language learning activity in a second language. As Littlewood indicates, they are not distinct from "FCAs": a student's proficiency will determine whether or not the social appropriacy of language is being coped with. Littlewood's survey may, then, be taken as preliminary evidence that a large part of language behavior in the SL classroom can be considered as a series of tasks, or "communicative activities".

1.2 Content area use of 'task'

The preceding section refers to tasks as most commonly thought of in current SL education. This is a rather narrow view, however, as many non-SL educators and researchers have also made use of the term. The task as a unit of analysis is, of course, one which is not only applicable to language classroom activities, but also to the content classroom and the 'real world'. Doyle (1983) provides an analysis of schoolwork in terms of tasks, (which have less in the way of communication or cooperation than those described in the preceding section). He also provides a psychologically principled system for grouping (academic) tasks.

Doyle (1983: 160) states

the concept of 'work' provides a useful metaphor for approaching the study of what students do in school.

Consequently he views the curriculum as "a collection of academic tasks" (cf. Doyle 1979, 1980), and indicates that

[the term "task" focuses attention on three aspects of students' work: (a) the products students are to formulate ...; (b) the operations that are to be used to generate the product...; and (c) the "givens" or resources available to students while they are generating a product. (1983: 161)
Tasks influence learners by directing attention to particular aspects of content and specifying ways of processing information to the extent that

the nature of exploratory behavior with respect to any stimulus configuration is modulated by the tasks in which the subject is involved at the time of encounter. (Nunnally & Lemond 1973: 79)

Students learn

what a task leads them to do ... acquire information (facts, concepts, principles and solutions) ...[and] practice operations (memorizing, classifying, inferring, analyzing) used to obtain or produce the information demanded by the task. (Doyle 1983: 162)

Doyle goes on (162-3) to draw on a body of work in psychology concerned with the analysis of cognitive processes underlying human aptitude and performance (Curtis & Glaser 1981, Greeno 1976, Resnick 1976). This work relates to the cognitive components of academic tasks (Anderson, Spiro & Montague 1977, Calfee 1981). Within this domain, he cites in particular the work of Greeno (1976) and Merrill & Boutwell (1973) as the basis for his specification of "general categories of cognitive operations that are involved in task accomplishment", e.g.

1 memory tasks in which students are expected to recognize or reproduce information previously encountered...
2 procedural or routine tasks in which students are expected to apply a standardized and predictable formula or algorithm...
3 comprehension or understanding tasks in which students...are expected to (a) recognize transformed ...versions of information previously encountered, (b) apply procedures to new problems... or (c) draw inferences from previously encountered information or procedures...
4 opinion tasks in which students are expected to state a preference for something...(162-3)

He notes that memory, procedural and comprehension processing may interfere with each other in accomplishing a given task.

Doyle's discussion of the characteristics of tasks in context (i.e. the classroom) acknowledges that an additional important part of the student's 'work' are those activities which are necessary for students to participate with others and to decode teacher expectations and instructions.

It is of interest that a recent substantial work on the nature of classroom content learning, produced independently of Doyle, uses almost the same basic unit of analysis. Mohan (1986) is an exploration of how the learning of a second language can be coordinated with the learning of content (a topic of particular concern to Canadian educational researchers, to content teachers
with students of limited English proficiency (LEP), as well as to SL teachers generally). While observing that "current views of language teaching can broadly be termed 'communicative'", Mohan (1986: v) points out that language teaching lacks a satisfactory understanding of the context for communicative discourse. He recognizes that attention has been directed to such contexts in their broadest sense as "culturally recognized social activities in which language plays a role" (ibid.). (He equates the latter with speech events.) However, he suggests that there is not yet an adequate general model of the relationship of activities to discourse.

Mohan argues that the structure of classroom instruction which is most helpful to the non-native speaker is one which reflects the structure of human activities. Following Dearden (1968), Mohan states that any human activity has two aspects -- action and theoretical understanding:

> [t]he specific, practical side we term an action situation, and the general, theoretical side we term background knowledge. (1986:42)

Thus, to take a classroom example, solving an algebraic equation is an action situation which requires (or leads to the learning of) background knowledge, *viz.* "rules for symbol manipulation and proof procedures" (44), the two parts together constituting, for Mohan, an 'activity': "a broad integrating idea relevant to all teaching and learning" (45).

It should be clear from this description that if Mohan's 'activity' is placed alongside Doyle's version of 'task', the two are identical except that Mohan has explicitly separated out the role of background knowledge. Both researchers argue for the extensive relevance and usefulness of their concepts, and for many purposes, the terms 'task' and 'activity' as used by these researchers could be used interchangeably (and Mohan indeed cites other writers who use the term 'task', without apparently feeling the need to make an explanatory comment.)

1.3 Equivalence of content and SL uses of 'task'

A further source of evidence for the utility of 'task' (as a conceptual unit) comes from a major review of research into teacher planning and decision-making (Shavelson & Stern 1981). Shavelson & Stern observe that while most teachers are trained to plan instruction in the orthodox manner by specifying objectives, determining students' entry behavior, choosing and sequencing activities to result in attainment of objectives, and evaluating outcomes, this is not in fact what teachers do in practice.

Research on teacher planning has found that the instructional activity is the basic unit of planning (Clark & Yinger, 1979; Peterson et al., 1978; Smith and Sendlebach, 1979; Yinger, 1977; Zahorik, 1975). (477)

What is the reason for the "mismatch" between the prescribed
model and what teachers actually do? Shavelson & Stern suggest that the demands that classroom instruction place on the teacher have the result that decisions about activities become of primary importance, to the possible exclusion of other considerations. (To prevent terminological confusion, Shavelson & Stern advise following Doyle's 1980 usage, which they equate with 'activity' as used by the researchers quoted immediately above.)

The findings cited by Shavelson & Stern are independently supported by a study of Swaffar, Arens & Morgan (1982) in SL classrooms. In attempting to account for the failure of comparative methods studies to show "clear, lasting superiority... in terms of student performance", Swaffar et al. investigated the assumption that classroom practices actually conform to given teaching methods, philosophies or approaches. Structured interviews of teachers conducting classes which utilized two supposedly different methods showed that there was no clear distinction between the classroom practices used across groups. The authors observed that

[Methodological labels assigned to teaching activities are, in themselves, not informative, because they refer to a pool of classroom practices which are universally used.](31)

What differences there are merely refer to the priorities assigned to tasks:

defining methodologies in terms of characteristic activities has led to distinctions which are ... not real ... not confirmable in classroom practice.(32)

In conclusion they state:

any analysis of methodologies needs to commence in terms of task, order [of tasks] and learning strategies. This is the way we, as foreign language teachers, interpret the pragmatics of the classroom.(32)

The convergence of independent research which has thus been found is good evidence of the potential strength and utility of 'task' as a major unit of analysis throughout all educational research and design.
2. Effects of task characteristics on language use and language learning

The area of research which is often referred to now as Second Language Acquisition (SLA) grew out of work done in the 1970s concerned with naturalistic SL learning. Since it was often assumed that all that was necessary for SL learning was comprehensible verbal interaction between native speaker (NS) and non-native speaker (NNS), the object of interest was the dyadic, informal NS-NNS 'free' conversation. Many early SLA studies merely placed NSs with NNSs and recorded the discourse which followed, which was usually not oriented to a specific objective. (For review see Long 1981.) However, obviously not all 'natural' language acquisition situations are undirected -- indeed, many of the beginner's encounters with the target language (TL) in situ are specifically directed towards achieving particular goals: making a purchase, obtaining directions, etc. At least part of the intent of later SLA investigations has been to identify the most productive aspects of dyadic interaction in the TL. These investigations have therefore involved experimental conditions of not only informal 'free' conversation, but also discourse engaged in with a particular end in mind. Investigators have tended to see the linguistic environment as a resource which, if comprehensible, can in some way be learned from. Additionally, however, it has been tentatively suggested that learners' SL development is partly affected by their opportunities to participate in L2 discourse (the "output hypothesis", Swain 1985). If this is so, the structure and demands of the task with regard to learners' production will also be relevant.

The following section reviews studies grouped as they relate to particular task characteristics, and subsidiary factors. Obviously, there are a great many criteria which could be used to organize the complex domain of tasks. In the present case, given that the utility of tasks for promoting second language learning is a major concern, the principal criterion for task classification must be psycholinguistic, with other psychologically relevant factors coming close behind. A psycholinguistically motivated task characteristic would be one which can be shown to affect the nature of language produced in performing a task in ways which are relevant to SL processing and SL learning. (As relatively little work has been done compared to the complexity of the domain, first language communication research and language production studies are also included where results have general implications for future SL studies.)

2.1 Information-transfer: one-way and two-way tasks

2.1.1 Background The psycholinguistically-motivated task characteristic to which most work has so far been devoted concerns information transfer. In some tasks involving two or more participants, it is necessary for different information held by each participant to be shared verbally, in order for the task
to be successfully completed. Such "two-way tasks" are contrasted with others in which verbal information transfer is also necessary for task completion, but where the information is allocated solely to one party, who is required to convey it to the other party, the role of the latter being mainly passive. The work of Long (1980) has demonstrated the comparative advantage that two-way tasks have over one-way tasks in terms of the amount of interactional modification they necessitate for successful task completion (for NS-NNS dyads). Long compared the discourse of 16 NS-NNS dyads on a variety of tasks with that of NS-NS dyads. Significant modifications were made in the discourse structure of NS speech, notably in interactional features (repetitions, expansions, confirmation checks, etc.) so that NS utterances were made more comprehensible than would otherwise have been the case. When two-way tasks were compared with one-way tasks, differences between NS-NS and NS-NNS conversation were greater on seven out of nine measures of interaction. In two-way tasks, it may be argued that NSs are obliged to make greater efforts to ensure as much as possible of their discourse was comprehensible to their NNS pairs, so as to permit the dyads to complete the tasks. In that such modifications increase comprehensibility, they also increase the likelihood that the verbal interaction can be learned from (providing task difficulty is controlled for, or an independent measure of success is provided; see Pica, Doughty & Young 1986).

This line of research was subsequently extended to NNS-NNS interaction. Varonis & Gass (1985) demonstrated that more negotiation of meaning takes place in informal NNS-NNS discourse than in similar NS-NNS discourse. Studies by Doughty & Pica (1984, Pica & Doughty 1985) constitute support for Long's findings for one-way versus two-ways tasks with respect to NNS-NNS dyads. In Pica & Doughty (1985) the discourse of teacher-fronted discussions was compared with that of small group (n=4) discussions, using two typical ESL communication activities focusing on decision-making and values clarification. (116)

The teacher-fronted groupings were given biographical data on five families living in the twenty-first century and then [were] asked to select the family most eligible to adopt a child. (Pica & Doughty, in press: 3)

The small groups "had to reach a consensus as to which one of six potential recipients would benefit most from a heart transplant". However, since in each condition the relevant information was equally shared, completion of the task did not require that every member participate equally, nor that utterances be maximally comprehensible to all members. Pica & Doughty report that utterances which contained interactional adjustments constituted only 6% of the total, in the small group condition.
A subsequent study (Pica & Doughty, in press) repeated the design given above, this time with two-way tasks. Participants were each provided with information concerning a small number of flowers in a plot. Each participant received information concerning different flowers, and had to exchange information with others so as to arrive at a specification of the complete plan of the flower-bed. Under these conditions, interactionally-adjusted utterances accounted for 24% of the total in the small group condition. (Other aspects of these experiments concerning group membership and size are discussed below.)

2.1.2 Shared assumptions Gass & Varonis (1985) investigated a "subset" (158) of Long's interactional modifications in one- and two-way tasks. Their one-way task required one member of each pair to describe a picture to the other member, without letting it be seen, in such a way that the picture could be drawn by the member not describing. The two-way task was a "jigsaw" listening task, in which participants listened separately to tapes of an interview between a detective and two of four characters suspected of a robbery, then came together to pool their information to determine which suspect was the actual robber. In the discourse produced, Gass & Varonis examined

nonunderstanding routines ... define[d] as those exchanges in which there is some overt indication that understanding between participants has not been complete. (1985: 151)

They refer to such exchanges as examples of "unaccepted input" and found no difference in the number of indicators of unaccepted input between one- and two-way tasks. They argue that the one-way/two-way cut is not the only factor of importance, and that the amount of shared information background is a factor too. Since there is more shared background in the two-way task, they state, it provides less chance of partial or complete communication breakdowns: "the greater the shared set of assumptions, the less need for negotiation". [2]

Varonis & Gass' result is consistent with that of Gaies (1982). This study was a partial replication of Long (1981), with one difference being that Gaies' NS-NNS pairs were acquainted with one another, whereas Long's subjects had not been. Long had found a significant difference in the degree of topic maintenance (ratio of topic-continuing moves to topic-nominating moves) in conversation, comparing NS-NNS dyads with NS-NS dyads. Gaies did not find such a difference. He suggested that as a result of participants' familiarity with each other, they brought greater "shared knowledge" to the conversational interaction. Consequently, there was less likelihood of complete communication breakdowns, and resultant 'dropping' of topics.

2.1.3 Discourse topic Crookes and Rulon (1985), using part of the data base analysed in Long (1980), compared the discourse of NS-NNS dyads engaged in free conversation with that which accompanied the performance of two other two-way tasks, and looked at changes in interlanguages (ILs) in the immediately
following discourse. The latter two tasks were of the type loosely called 'problem-solving' (see below for a caveat concerning the use of this term). In one, partners had to agree on which of four items was the 'Odd Man Out' and in the other, on the differences between two similar pictures ('Spot the Difference'). A major concern of the study was the extent to which NSs provided "feedback" to NNSs -- that is, were there occasions when, following a non-target-like usage in the NNS's IL, the NS partner provided a target language-like equivalent? (Such occasions might be considered specially salient cases of comprehensible input.) It was found that provision of such feedback occurred significantly more often in the discourse accompanying the problem-solving tasks than in free conversations. However, in only one of these two tasks (the Odd Man Out task) could more change in the ILs be observed than in the free conversation condition. Thus while both provided a more favorable environment for language learning than did free conversation, it seemed possible that the nature of one of them, and perhaps of the structure of the discourse it produced, was more desirable than the other.

In addition, it had been hypothesized that differences between conversation on the problem-solving tasks and free conversation were also due to the extent that discourse topics were continued rather than dropped. It seemed likely that if the same linguistic material were used repeatedly in the course of a conversation because the task entailed the recycling of discourse topics, such a conversation would be potentially more useful to the NNS than one in which many items occurred once only. However, of the two 'problem-solving' tasks, the one which had long stretches of discourse on the same topic (Spot the Difference) was the one which had less measurable effect on IL. A closer examination of the linguistic content of sections of discourse showed that the same linguistic material could appear in different topic segments, and contrariwise, material in a single topic segment could, in some circumstances, be quite varied and non-repetitive. So selecting tasks according to whether they were likely to generate long stretches of discourse on the same topic would not appear to be a desirable procedure.[3]

It is suggestive that of the two problem-solving tasks, the one which was less productive of immediately observable IL development (i.e. destabilization) was that in which the task provided visual support to both members of the dyad. Even though the pictures used were not identical, they were versions of the same picture differing only in certain limited features. It could be argued that the situation is analogous to the Gass & Varonis study cited above, with regard to "shared assumptions".

That visual support can affect discourse is demonstrated by Nurss & Hough (1985), who investigated young (L1) children's production of oral narrative on different tasks.[4] The initial task was simply to tell an original story, without visual stimulus. Subsequent tasks required children to tell a story after reviewing a wordless picture story book; to respond orally
to a single picture with implied action by telling its story; to dictate a story which applied to a single picture with implied action; and to follow and give directions to make a product. Significant differences favoring the no-picture situation were found:

in summary...telling an original story elicited a greater quantity of language and somewhat more mature language structures than the other tasks, although each task yielded slightly different structures.(283)

2.1.4 Convergence Another task-characteristic derives from the work of Duff (1986). Duff examined the discourse produced by NNS dyads doing 'problem-solving' and 'discussion' tasks (her terms). On the basis of the type of interaction produced and the "focus of one dyad member relative to another" (148) she argues for regarding this sort of discourse as either 'convergent' or 'divergent'. For Duff "a problem-solving task" is one in which "the shared goal of learners is to reach a mutually acceptable solution" (150). It is thus an example of a "convergent task type". Duff states that tasks such as 'discussion' may be set up so that pairs take (or are given) different positions on an issue and are asked to defend them and attack that of their partner. Such a task involves "opposite or independent goals for each member", and thus falls into Duff's "divergent task type" category (150).

Duff indicates that in both cases, success requires speakers to "acknowledge and incorporate the other's output" (151), and found that convergent tasks (such as problem-solving) will produce more negotiation than divergent ones, because both parties ...have vested interest or responsibility in the ultimate decision or solution that is reached.(151)

Specifically, her results show that convergent activities lead to frequent exchange of turns and more communication units, whereas divergent activities lead to longer turns of greater syntactic complexity. (The latter type of activity may be desirable if attention is being directed to 'output'.) Summarizing, Duff states that

the extended discourse (long turns) in [discussion] reduces opportunities for negotiation of input ... coupled with the greater syntactic complexity of [discussion], this reduces ... the amount of comprehensible input available.(170)

By comparison, convergent activities (as manifested by problem-solving activities in this experiment) produce more questions and shorter turns, and thus, it may be assumed, more comprehensible input is available in the discourse which accompanies their performance.

Although Duff's work suggests the possibility of a connection between task characteristics and overall syntactic
complexity of (oral) discourse generated, attempts to predict specific sentence-level units of discourse from task-type may be unsuccessful, as an exploratory study (Butler-Wall 1983) suggests. In this instance, NS dyads performed a range of tasks, grouped by the experimenter into five categories, for which psycholinguistic motivation was not explicitly provided. The categories were: narrative (e.g. 'tell a fairy tale'); descriptive (e.g. 'describe your family members'); expository (e.g. 'give directions to your home'); speculative (e.g. 'discuss the future of the world in ten years'); and cooperative (e.g. 'play Odd Man Out'). The resulting discourse was analysed in terms of a range of syntactic structures: right and left dislocation, topicalization, clefting and pseudoclefting, adverb preposing, extraposition, deletion under identity, complex nounphrase shift, passive, and 'there' insertion. Passive and 'there' insertion were found to appear in many of the tasks at above a criterion level of 5% of T-units. However, clear connections between task and frequency of use of a particular structure did not appear.

2.1.5 Descriptive terms for tasks The terms 'problem-solving' and 'discussion' are already in use to refer to tasks, though whether the tasks these labels have been applied to are really different, in terms of the language they produce or their potential for facilitating L2 learning has yet to be established. Furthermore, exactly what is meant by these terms may differ from one user to another. An early L1 study is suggestive in this respect. Morris (1966) found that group tasks could be differentiated in terms of their effects on two types of group behavior: "production-related" and "process-related". "The type of task...significantly affected the distribution of almost 60% of the group activity" (55\(^1\)), with a task which might superficially be characterized as Discussion (anything of the form 'reach a consensus on the issue of ...') generating a high amount of 'process' activity, and a Production task ('write a story about-' being low on this dimension. Problem-Solving (e.g. 'develop a plan to ...') tasks were found to generate approximately equal amounts of each type of behavior. (But note the way Morris uses the term 'problem-solving'.) Morris refers to a "two-dimensional behavioral space", one dimension of which reflects group behavior oriented towards the production of the task product ("structuring an answer, producing specific solutions"), the other dimension representing the processes (such as "clarifying" and "explaining") involved in doing the task. All three of his tasks can be located at different positions within the behavioral space. This approach provides a way of demonstrating empirically the differences between tasks, or of validating a task category system, (depending on the nature of the behavioral dimensions selected). Unfortunately, small group work is "noteworthy for its diversity of task settings ... that have only infrequently been replicated" (Mabry 1985:80).

2.2 Aspects of task complexity

The widely-used information-processing model of human cognition incorporates the conception of limited processing
capacity. The attentional demands of a task with multiple components, or its overall complexity, can produce cognitive strain, with detrimental effects on performance (overall, or in one area). Attempts to address this aspect of task have until very recently been mainly in L1 studies, and in areas not specifically related to language. Relevant L2 studies have focused on language production and elicitation, rather than communication.

2.2.1 Small group work on task complexity Segal (1982: 334), in (L1) work on small group behavior, provides some useful comments on "task complexity". She notes

[t]he operational definition of task complexity has varied with studies and has not been standardized. Two dimensions of the variable are (1) the number of goals, and (2) the number of paths to these goals (Fiedler, 1967; Shaw, 1976)...

A third dimension contributing to the complexity of the task is the amount of specific, unequivocal information given the decision maker (Driver and Streufert, 1969). These dimensions are differentiated from those of difficulty (Shaw, 1976) and uncertainty (Perrow, 1967; Thompson, 1967), dimensions that appear to be a function of a specific unit's resources rather than inherent in the task environment.

Segal reports that task complexity affects the sequence and extent of recycling of topical units in discussions. Task complexity, in terms of number of goals and the number of paths to them, may therefore be a relevant factor in task design and selection.

2.2.2 Non-language work on task complexity Indicators of task complexity which come from non-group research do not appear to go against those mentioned above. For example, in Hogarth's (1975) examination of decision-making tasks, complexity depends on the number of attributes per alternative and the number of common characteristics of the alternatives being considered (though it is not clear which of these factors is more important, or whether a multiplicative relation between them may exist: Onken, Hastie & Revelle 1985: 15). With regard to cognitive strain, increases in task complexity beyond processing capacity will result in not only more errors, but less deliberation.

Other indications of task complexity may be derived from the domain of 'task analysis':

the study of performance on complex tasks, such as those encountered in school, to reveal the psychological processes involved.(Resnick & Ford 1978: 378)

An understanding of the relative complexity of tasks is, however, not a principal aim of these investigations, so much as the extraction from a complex task of elements or subprocesses which can themselves be individually taught. Resnick & Ford distinguish between rational and empirical analyses of tasks, and point out
that the sub-procedures used by those who perform a task are often not those suggested by a purely logical or rational analysis of the task — a point to be borne in mind by SL investigators of this area.

2.2.3 Effect on sentence-level phenomena Ability to produce discourse (and the quality of that produced) depends partially on the demands of the subject matter — its complexity and familiarity. Obviously there is a trade-off: a topic of (in some sense) 'high' complexity may nevertheless be handled fluently if it is one the speaker often handles, and one of low complexity may not be handled well if it is one the speaker rarely deals with. Holmes (1984: 117) notes that in earlier studies of language production (Ford 1982, Ford & Holmes 1978) subjects had been required to talk on a topic of personal interest. As a result, the counter-intuitive conclusion had been arrived at that sentences are planned relatively independently ...

the detailed planning occurs basic clause by basic clause. (117)

He comments that this conclusion may result from the fact that research has been done principally on speakers who do not have any chance to plan or practice. Consequently, in Holmes (1984) the nature of the experimental task was altered (by comparison with other related studies). Subjects were required to continue a conventional story, the first few sentences of which they were provided with. Detailed analysis of pauses and other phenomena observed in the subjects' output indicate that under these conditions subjects [can] construct more highly integrated utterances, which have been largely thought out and organized before expression ... compared with the relatively unplanned discourse of spontaneous speech. (129)

Holmes also cites a comparable study (Goldman-Eisler & Cohen 1970) which demonstrates that prepared speeches were more syntactically complex (in terms of presence of passives) than the discourse of spontaneous discussion. Thus any study which considers syntactic features or the overall syntactic complexity of discourse may need to control for topic familiarity and preparation. (See also Ochs 1979 and Crystal 1980 on planned versus unplanned discourse.)

2.2.4 L2 studies At least since the 1970s, SL researchers have been making use of a variety of elicitation tasks to obtain data (e.g. Naiman 1974, and for recent instances see e.g. Wagner 1983, Dechert 1983.) However, the failure of some investigators to consider the effects of elicitation task on the nature of the language elicited has led to the conclusions of a number of prominent studies (e.g. Dulay & Burt 1974, Bailey, Madden & Krashen 1974) being identified as potentially artificial (Larsen-Freeman 1975, 1978). Other studies (e.g. LoCoco 1976) have been weakened by failures to report details concerning
elicitation task-characteristics. (See Hood & Schieffelin 1978, Hulstijn & Hulstijn 1984, Chaudron 1985 and Hull 1986.) Tarone (1979), drawing on Labov (1969), has called for a much fuller reporting of task, in addition to interlocutors, surroundings and topic, than has generally been given. In her own work, she has given specific attention to the relation between NNSs' language and elicitation tasks used. In Tarone (1983) she reviewed a variety of studies and concluded that IL can be viewed as a continuum of styles, which are defined in terms of "the amount of attention paid to language form" (152). She predicted that when more attention is directed to language form, learners' interlanguage (IL) will be more influenced by target-language (TL) forms, and when less attention is paid to form, IL will show less TL influence.

These predictions, as they apply to IL phonology, were examined in Sato (1985). It was suggested (183) that styles may be "concretized" as communicative tasks (e.g. "spontaneous conversation, oral reading, structured interview, elicited intuition, elicited imitation, etc."). Sato tested Tarone's prediction that acquisition of TL forms should take place first in speech produced on more formal tasks (Sato 1985: 183).

In this case-study of a Vietnamese boy acquiring English, the 'tasks' used were free conversation, reading of continuous text, and elicited imitation. Of Sato's three tasks, the "formal" task (in Tarone's terms) is reading of continuous text, and it is this one, according to Tarone, where evidence of learning should first appear. But in fact, when word-final consonants were considered, it was in the "vernacular style" task: conversation, that (in two of four samples) there was a greater incidence of TL-like production. Sato concluded that it is not just attention paid to language form, but also the attentional demands of other aspects of the task: "recall and encoding of rhetorical structure, lexical items, clause sequencing etc." which must additionally be taken into account. An earlier investigation (Oyama 1976) into the IL phonology of Italian immigrants to the US also compared performance in casual speech and on an oral reading, and found more TL-like pronunciation on the informal speech sample than on the reading.

A follow-up study (Tarone 1984) failed to confirm Tarone's predictions: IL is most influenced by TL in performance on tasks which elicit informal speech (the vernacular speech style), despite the fact that this is the style in which least attention is supposed to be paid to language form. (Tarone explained the result by reference to discoursal phenomena.) Finally, Parish, Tarone & Taghavi (1986) recognize the fact that this line of research is seriously threatened by its attempts to apply sociolinguistic concepts to a psycholinguistic problem. Parish et al. present a reanalysis of Tarone's (1984) data, and conclude that task attentional demand is the principal predictor of accuracy levels of IL forms, following (though not acknowledging) Sato's critique.
Alternative attempts to explain IL variability are to be found in Ellis (1984, 1985), who places considerable importance on possibility of planning: tasks which permit the utilization of planned discourse will allow the production of IL discourse with different characteristics to those which do not. The concept of automaticity is also invoked (equated with ease of access to linguistic knowledge, as in Bialystok 1982:183). However, until investigators take to heart Sato’s comments cited above, and make detailed analyses of the full range of simultaneous demands made on the cognitive system by the task of interest, post hoc explanations which just focus on speech style, or the prevalence of planned discourse are likely to be unsatisfying.

2.3 Moderator variables

2.3.1 Group composition A factor which has been involved in a number of investigations is the composition of the 'group': specifically, whether the participants include a teacher or not.

Long, Adams, McLean and Castaños (1976) reported that NNS participants in a dyadic discussion task utilized a wider variety of forms of communicative language (pedagogic moves, social skills and rhetorical acts) than NNSs in a larger, teacher-fronted group engaged in discussing the same question. The discussion concerned "similarities and differences between man and other animals" (139). The authors also suggested (following Barnes 1973) that the pressures of a large group situation might cause students' utterances to be briefer and less complex, because of the more public nature of the setting -- Barnes' "audience effect". They argued that this would encourage the establishment of an impoverished linguistic environment, with negative implications for learning.

Several recent studies by Pica and Doughty (1985, in press) have also addressed this question. These compared teacher-fronted and NNS group decision-making discussions, hypothesizing that there would be more interactional modification in the latter format. This was not the case for their one-way task, but when a two-way task was used, significantly greater interactional modifications did occur for the group (NNSs, n=4). Although Pica & Doughty (in press) make it clear throughout that they were comparing complete classes plus a teacher with small groups minus a teacher, they observe (12) that "as the number of interlocutors decreased, the amount of negotiation increased significantly ... p<0.001)". It should be pointed out that despite this, size and composition are confounded here.

Rulon & McCreary (1985) also compared teacher-fronted classes with small groups and report findings similar to Pica & Doughty's first study. Their task was to complete an outline by listing and discussing the advantages of the colonists and the British during the American Revolution. (1985: 8)
This too was a task whose structure did not require all participants to take an equal part, or be obliged to fully negotiate input so as to complete the task. One or two fluent or knowledgable students could have dominated the discourse; less able students could have opted out, to some extent. Rulon & McCreary found no significant differences in length of speech unit nor syntactic complexity between small group (n=3) and teacher-fronted discussions, though they recognize problems with their measure of speech unit (words/communication unit). Clarification and coverage of content in small groups was equivalent or better than in the teacher-fronted class, however.

2.3.2 Group size In a study concerning group testing of oral SL proficiency, the elementary observation that participants in larger groups speak less has been documented (Liski & Puntanen 1983: 241).

Antony (1986) describes the discourse of NNS groups of size two to five performing "task-based consensus activities" (for which no definition is supplied). With regard to size, he observes

[group size does not seem greatly to affect the number of wpm [words per minute] of the group as a whole, so smaller groups likely generate more wpm per student. Larger groups, however, seem to introduce new ideas more quickly, more simultaneous starts, and more brief overlaps. So while smaller groups may provide more practice in speaking, larger groups may well provide more valuable input.(1986: 5)

Antony observes that "nearly all the language produced" can be classified as either metacommunicative, (reg procedure and organization / language problems), or "task-solving". He notes that "[f]or language students a certain amount of this metacommunication deals with correctness of form in the target language (1986: 9)". [5] However, Antony states that "the vast bulk" of his groups' language is task-solving in nature, and goes on to develop a functional analysis of the language used. Since only one task was used, a knowledge of the relationship between task type and functional characteristics of language generated (if any) has yet to be determined.

2.3.3 Output mode Gass & Varonis (1985) compared their own study with that of Doughty & Pica (1984), and noted that comparison of results in the two studies should take into account the differing modes of output. The former study contains a one-way task whose product requires pencil-and-paper work (drawing), whereas the latter contains a task whose output involves object manipulation (as also in Pica, Doughty & Young 1985). This variable, "medium of response/psychomotor demands" has been largely ignored in L2 studies of this kind so far. (But see Shortreed 1985; and for non-SL work, see McGrath 1984, below.)
2.3.4 Sex of interlocutor Gass & Varonis (1985) also consider the variable 'sex':

in general, men signal unaccepted input more often than women. This suggests that initially women, at least in interactions with men, feel less confident in indicating a lack of understanding...[or] that they understood more and did not need to negotiate meaning. (159-60)

Recent work in (L1) small group behavior suggests that task content can favor the participation or success of one sex over the other, but this is not necessarily the case when content is selected to be neutral. All-male groups appear to differ from all-female groups in terms of that functional quality of language which may be used to establish dominance patterns, but differences between male and female behavior decreases when the group observed is mixed (Mabry 1985). Male-only groups appear to produce more problem-solutions, female-only groups more creative solutions (Wood, Polek & Aiken 1985).

2.3.5 Personality/individual ability SL studies which attempt to predict discourse characteristics from task characteristics do so on the assumption that individual difference variables are not strong enough to eliminate such effects. If this is so, it is at least partly a function of subject selection from normal, adult, language-oriented groups. However, 'ability to communicate' is a factor which could influence such results one way or another, and it has so far been taken for granted or ignored. This is probably because communicative competence is seen principally in SL research as a function of SL control, rather than a personality variable. Within communication studies, however, it is seen as an individual difference variable which affects comprehension and overall communicative ability -- see e.g. Beatty & Payne (1984), Wiseman & Abe (1986), Hale (1980); and compare Allwright's remark (1976: 10) that

a good communicat[or] with a poor command of English may be better able to cope with communication problems in English than much more fluent speakers who are not good communicators.
3. Current language syllabus types and the task-based syllabus

The second language research which has been reviewed in the preceding section has not existed in a vacuum. Its long-term intent is to improve SL instruction. Such instruction has not, of course, been held in abeyance while results of SLA research are awaited. There has been a certain amount of cross-fertilization and some investigators have given attention both to small scale laboratory experiments into the characteristics of SL tasks, and to task-related aspects of syllabus design. Arguably, a concern with communication, and the role of "communicative activities" both in the classroom and as a basis for naturalistic acquisition of second languages is the basis for both a substantial portion of the work reviewed in the preceding section, and for some of that to be examined here.

Given the extensive utility of the concept 'task' outlined in Section 1, it is perhaps not surprising that attempts would be made to use it as the basis for language syllabus design. The principal reason for such attempts has been the mismatch between what is known about the processes of language acquisition, and the nature of language itself as a means of communication, on the one hand, and the most common forms of language syllabus, on the other. In this section, attention will be directed first to the inadequacies of existing second language syllabi. Subsequently, a number of counter-proposals based on tasks will be examined. As a major function of a syllabus is selection and sequencing, an underlying concern here is with what are appropriate units, and how they may be grouped or sequenced. (The lack of information available with regard to the latter point is the justification for Section 4.)

3.1 Background

Although in practice the structural (or grammatical) syllabus is probably the most widely used SL syllabus throughout the world, this is not because of its universal acclaim. The last decade or so has seen it criticized by proponents of alternative syllabi -- whose own proposals have in turn been criticized. Current positions on appropriate syllabus forms vary widely.

The standard works of Halliday, McIntosh & Strevens (1964), and Mackey (1965) contain what became the orthodox approach to SL language syllabus design concerning selection and grading. There was early opposition to this line from Newmark (1964, 1966) Newmark & Reibel (1968), Reibel (1969), and Macnamara (1973), all of whom argued against the selection and sequential arrangement of "language data". However, unlike more prominent syllabi or Methods both then in existence and to come (Richards 1984a), these proposals had no international institutional backing, and thus few practical manifestations (but see Newmark 1971 and Allwright 1976).
A subsequent, more substantial change in syllabus design was provided by the widely-known notional (often 'notional/functional') approach.

[A] grammatical syllabus is one which accords highest priority to grammatical criteria and sees the structure of language teaching as being principally provided by an ordered sequence of grammatical categories ('structures'); a notional syllabus would seek to change the balance of priorities by placing emphasis on the meanings expressed or the functions performed through language. (Wilkins 1981: 83)

The initial speculations of the principal exponent of notional syllabi (Wilkins 1972) were subsequently tempered to a limited set of circumstances (Wilkins 1974), and criticized by others who had at first welcomed them (e.g. Widdowson 1971, 1978). In a later paper, Wilkins drew attention to some misunderstandings concerning the relative position of notions and structures. He pointed out that a range of emphases is possible according to the needs of the learners for whom the course is designed -- from a principally grammatical syllabus all the way to a fully notional syllabus, and that while he had argued (Wilkins 1976) against some aspects of a grammatical syllabus,

[t]he notion that an individual can develop anything other than a rudimentary communicative ability without an extensive mastery of the grammatical system is absurd. (Wilkins 1981: 85)

Thus (with hindsight) notional/functional and structural (or grammatical) syllabi are not seen as necessarily separate by Wilkins, nor is one taken as superior to the other.

At present, as in the 60s, a substantial number of applied linguists concerned with curriculum accept the idea that language course organization should be based principally on formal analyses of language structure. Of these, one group seems prepared to accept a structural syllabus at least under some circumstances (Brumfit 1984a, Widdowson 1984). Another group (Yalden 1984, Allen 1984) has a conception of the syllabus as including content, pragmatic, methodological and structural elements. Within this section of applied linguists, then, there is disagreement as to the units out of which a language syllabus should be constructed. Other applied linguists reject the idea of a fixed language syllabus, whatever its linguistic units (Breen 1984, Candlin 1984a, Long 1985).

3.2 Criticisms of structural syllabi

There are a number of arguments against both purely structural syllabi, and variants with notions, et cetera added in. Wilkins (1972), referring to purely structural forms,
identified (1) negative effects on motivation for those learners who need to be able to communicate as soon as possible; (2) inefficiency, in the idea that the whole of the grammatical system is to be taught when not all learners would need it all; (3) the subordination of meaning to form; and (4) the artificiality of grouping structures according to form rather than communicative content, with resultant negative pedagogical effect.

Notional aspects of the type of syllabus under consideration have been criticized on psychological grounds, by some who would probably prefer a purely structural syllabus. Thus, Brumfit (1981: 91, 92) states "principles of organisation must be answerable to a view of how language is learnt", but observes that the principal proponent of notional syllabi does not (in Wilkins 1976, 1981) address himself to learning theory. Brumfit continues

until we have some way of saying 'X is a notion and Y is not, and we can test them in the following ways', we are talking about a vacuous concept.

Paulston (1981) also criticizes Wilkins' work for paying no attention to a theory of language acquisition, and for being based on logico-deductive reasoning, not empirical evidence. It is certainly the case that no consideration was ever given to the psychological reality of the units of notional-functional syllabi until recently (see Cook 1985).

Similar arguments can also be advanced against the structural component of a language syllabus -- in particular, Brumfit's remark above applies here, too. In this type of syllabus, the specifications of the language to be learned are derived from a formal analysis of the language, largely unrelated to and unmotivated by a view about how language is learned. Many of the units of such formal grammatical analyses become problematic when approached from a psychological or language processing point of view (Clahsen 1980, Pienemann 1985, Ford 1982: 797). There is, in addition, extensive evidence against the conception that learners move in a simple step-wise fashion to control of a second language (see e.g. Schumann 1978, Felix 1981, Bahns 1983, Gass 1983.)

From a pedagogical point of view, it may be argued (as does Long 1985) that because structural syllabi are based on formal analyses of language (target language (TL) forms, not interlanguage (IL) forms), they result in lessons whose 'point' is neither the activities the learner will be called on to perform subsequently in the 'real world', nor the units in which the L2 learner learns the language.

Besides selection, a major function that a syllabus performs is sequencing. However, as Long (1985) notes, sequencing decisions related to linguistic syllabi (whether overtly linguistic: notional-functional/structural, or covertly
linguistic: situational) are based (sometimes) on needs analysis, but more often on materials writers' intuitions. Even if sequencing decisions were based on explicit criteria concerning relative grammatical complexity, they would vary according to the linguistic theory used. Of course, not only does no complete grammar exist (even for English), no current major syntactic theory has been shown to have psychological reality. (See Kenny 1980, Ortony 1975, Pullum 1986.) In practice, L2 syllabus sequencing decisions are not based on proven psychological complexity or learning difficulty,[6] nor on empirically-supported judgments concerning valency, for either structures or notions. Schinnerer-Erben (1981: 11), having reviewed a substantial body of second language related work concerning sequencing, states that

the criteria which are commonly used to establish traditional sequences are rather feeble. Difficulty is not easily defined and it is of questionable value. Frequency/utility is also difficult to establish and has not been proven helpful in the learning process. And, natural sequences do not really exist in sufficient detail to be used as the basis for a precise order, nor have they been shown to facilitate learning in a second language situation.

However, she does not rule out the importance of organization completely. After considering some aspects of (mainly Ausubellian) cognitive learning theory related, however, to the learning of content, she concludes

there is considerable evidence to show that organization does facilitate learning, [but]...we cannot at present use the criteria available to effectively organize the material.(25)

3.3 Alternatives to structural syllabi

The alternatives to analytic syllabi are based on the assumption that language (like other psychomotor skills, compare Welford 1968, Levelt 1978) is something that may be acquired holistically, rather than synthetically. This position is stated by Newmark (1966), who asserts that

if the task of learning to speak English were additive and linear ... it is difficult to see how anyone could learn English.(77)

He maintains that "language is learned a whole act at a time, rather than an assemblage of constituent skills". Newmark conceived of a language course of an 'analytic' type, where large 'chunks' of language are presented and learned: the presentation possibly on videotape, the learning through observation, and practice through extensive use of roleplay. Reibel (1969) argued for the irrelevance of sequencing to what was known about language learning. Both identified a psychological basis for their position in Bandura's Social Learning Theory (Bandura & Walters 1963, Bandura 1977), a highly respectable, well-developed
and researched theory, which has been inexplicably ignored by second language researchers (though not L1 researchers: see Rosenthal & Zimmerman 1978).

Of current language courses which do not use synthetic syllabi, some are based on Methods such as Counseling-Learning, the Natural Approach, etc., which are accurately described as "responses to innovations at the level of instructional theory" (Richards 1984b: 22), and are not responses to the inadequacies of the synthetic syllabus. However, leaving these to one side (along with the brief pioneering work of Allwright and Newmark), there are some recent attempts to address this problem which should be considered: Candlin (1984a), Long (1985), the Malaysian Language Syllabus (Kementerian Pelajaran Malaysia 1975; also Samah 1984) and the Bangalore project (Prabhu 1983). Of these, the first two are theoretical arguments and partial descriptions concerning a language syllabus of this type based on 'tasks'; the third is a practical example with little in the way of theoretical rationale; and the status of the fourth is problematic.

The larger part of Candlin (1984a) provides a non-language specific critique of the synthetic syllabus, detailing the passive, alienated conception of the learner and authoritarian, all-knowing picture of the teacher it presupposes. For examples of the desirable alternative Candlin draws on the work of others, in particular Long (1985). (See also Candlin 1984b).

For Candlin, a task is

one of a set of sequencable, differentiatable and problem-posing activities which involve learners in some self-reliant selection among a range of variably available cognitive and communicative strategies applied to existing or acquired knowledge in the exploration and attainment of a variety of pre-specified or emergent goals via a range of procedures, desirably interdependently with other learners in some social milieu. (Candlin 1984b)

Long (1985) sees syllabus design in the broader context of program design (as do Brumfit, Allen and others, by comparison with the narrower conceptions of Wilkins and Widdowson). A major part of current problems, he indicates, is the mismatch between choices made in each area of program design. Even then, the current options in syllabus design and teaching method have no basis in SLA theory and research.

Long proposes instead that 'task' be taken as the basic unit of the syllabus, and defines task straightforwardly as

a piece of work undertaken for oneself or for other, freely or for some reward...by 'task' is meant the hundred and one things people do in everyday life, at work, at play, and in between. (Long 1985: 89)

He points out that it can serve to organize all four major stages
of syllabus design:
1) identifying learners' needs
2) defining syllabus content
3) organizing language acquisition opportunities
4) measuring student achievement.

Candlin, too, finds this point worthy of emphasis, saying

[a] first step towards achieving a syllabus of this second type is to attempt this accommodation by making purposes, content, methods (or, better, learning experiences), and evaluation, interdependent. (Candlin 1984a: 9)

Long points out that in designing a task-based syllabus, stage 1 is simplified by the fact that occupational analyses already exist (Dictionary of Occupational Titles, [DOT], U.S. Dept. of Labor 1965, 1977). Target tasks thus identified would need to be grouped by type, and pedagogical tasks could be derived from such task types. Pedagogical tasks in turn could be selected and sequenced to form a task syllabus. This aspect of Long's work suggests an orientation towards VESL/ESP syllabi. (See below, Section 4, for analyses of employment-related task characteristics which should feed into this approach.)

Whereas selection of tasks is confined, for Long, to "ensuring task types are adequately represented", grading is more of a problem. Variety, pace and duration are to be considered, but obviously, difficulty is a prime consideration. The latter, Long suggests, is determined by
1) number of steps needed
2) number of parties involved
3) presupposed knowledge
4) intellectual challenge
5) spatio-temporal displacement.

Because of the preliminary nature of the work, no support for this (intuitively appealing) set of determinants of difficulty is advanced, nor is any attempt made to define them in more operational terms.

The Malaysian Language Syllabus (Kementerian Pelajaran Malaysia 1975) and the Bangalore Procedural Syllabus (Prabhu 1983) are attempts to apply some unorthodox forms of syllabus within the regular secondary education systems of two third world countries. In neither case has implementation been as smooth or as well-informed as might ideally have been wished.

The Malaysian case obviously antedates most of the theoretical work discussed so far. Original implementation was influenced by the British 'communicative approach' to language teaching, though it has recently been reinterpreted by some as an early attempt at a task-based syllabus. Following governmental specification (in 1975) of three simple program goals, 24 general objectives were listed by program developers, and the political, business and educational establishments were asked to add, delete or modify the list. The revised list was taken to constitute a
needs analysis (Richards 1984b), and used as the basis for constructing English language "products". (For further details, see Rodgers 1984, Samah 1984.) An example of an objective is "follow and understand a talk on specific topics"; an example of a "classroom procedure" is

A foreign visitor has expressed interest in a poster showing local tourist attractions. Describe the attractions shown on the poster. (Malaysian English Language Syllabus 1975, cited in Richards 1984b: 36)

The Bangalore project began in 1979. Early influences were similar to those of the Malaysian communicative syllabus. Its major proponent states

[there is...no syllabus in terms of vocabulary or structure, [and] no preselection of language items...The basis of each lesson is a problem or a task...Tasks consist of such things as: interpreting a schedule, a map or a set of rules; giving a set of directions...; drawing pictures from instruction...(Prabhu 1983: 3)

By comparison with the suggestions of Long and Candlin, other aspects of classroom structure are traditional. There is considerable emphasis on "receptive language", and classes are "teacher-centered". Group work is apparently discouraged "because of the fear that learner-learner interaction will promote 'pidginization'" (Brumfit 1984b: 236-7)[8], and there is little or no student-student communication.

The major problem[9] with the Bangalore project is that no rationale is provided for either the selection or the grading of the 'tasks', let alone a coherent theoretical basis. Almost as serious, and a point which even sympathetic commentators such as Brumfit have noted, is the lack of hard information about the success of the project. The recent work of Beretta (1986, Beretta & Davies 1985) attempts to redress the situation. Beretta & Davies (1985) report results from a post hoc comparison of intact groups, where it is specifically admitted that three out of four "experimental" groups had "better qualified, more highly motivated teachers", and were accustomed to being observed and treated as "guinea-pigs" (123). The results are interpreted as being not unfavorable to the Bangalore project. (See also Beretta 1986.)

In brief, neither of these two programs can be ignored, but as yet their documentation, in terms of reports, evaluations, syllabi, etc., is insufficient evidence on which to base conclusions as to the efficacy of task-based syllabi.
4. Non-SL research concerning task organization

Some partial answers to the questions raised in seeking a thorough approach to sequencing and selection for a language task syllabus are perhaps to be found in work done in the areas of a) human performance research and b) small group research. The former may be of particular relevance to the development of task syllabi with a VESL/ESP orientation, though it also addresses general procedures for the development of task taxonomies. The latter provides what appears to be the least inadequate descriptor system for tasks currently in existence.

4.1 Human performance research

4.1.1 Task classification procedures in human performance research

Human performance research (Dunnette & Fleishman 1982: xi) investigates human traits, abilities and performance levels with a view to matching the individual to the task. This area is related almost exclusively to psychomotor tasks, most of which require little or no communication. These tasks are thus somewhat distant from those most commonly used in language research, but the procedures used to classify them should apply generally to task taxonomy design.

Of two major subdomains within human performance analysis, that concerning procedures for development of task taxonomies has the most potential for task syllabus design. The second subdomain investigates individual difference factors as they affect task performance. This might be of more interest to the language researcher who makes use of tasks, rather than the syllabus designer. It should be noted that (with a few exceptions, e.g. Gass & Varonis 1985), SL researchers have tended to assume that subjects' individual characteristics other than those relating to language (and cultural) experience were irrelevant to task performance, since the actual outcome of the task has usually not been an issue. But if, for example, task outcome is to be used as a measure of comprehension (Pica et al. 1985), individual factors become more significant. Of the two areas, less work has been done on the development of task taxonomies than on developing taxonomies of person characteristics re jobs (Dunnette 1982: 4).

Fleishman (1978) describes work related to developing systems for classifying tasks that would improve our predictions and generalizations about factors affecting human performance...[W]hat we need is a learning and performance theory which ascribes [to] task dimensions a central role.(1978: 1007)

He refers to two strands of work concerning "ability requirements", and "task characteristic descriptions"). The first classifies and sequences tasks by the abilities they demand of an operator. According to Fleishman,
a limited number [of abilities] (nine or ten in the psychomotor and nine in the physical proficiency areas) seemed to account for most of the variance in several hundred tasks. (1978: 1009)

However, these tasks, for example, "multilimb coordination ability, response orientation, rate control", are principally psychomotor in nature.

The second area of relevance, task characteristic description, assumes tasks can be described and differentiated in terms of the intrinsic objective properties they possess (such as "goals, input stimuli, procedures, responses and stimulus-response relationships").

In order to differentiate among tasks, the components of a task were treated as categories within which to describe task characteristics or descriptors (e.g. number of procedural steps, variability of stimulus locations, number of output units). (1013)

Fleishman and colleagues have been able to predict performance on tasks in terms of ability requirements (characteristics of the operator) necessary for tasks varying in complexity on a number of factors.

Fleishman and Quaintance (1984: 66ff.) bring together an extensive quantity of material concerned with the problem of classifying human tasks. They state that there are four main issues in developing a classificatory system:

1 the set of descriptors must be selected with regard to the theoretical and conceptual base;
2 tasks must be classified as reliably as possible;
3 qualitative or quantitative classification should be chosen: classification may represent different types of things, or differences in degree of common dimensions;
4 criteria for evaluation of the classificatory system must be specified.

Concerning 1, Fleishman & Quaintance state:

[initially, the relevant concepts should be reviewed and translated into well-defined hypotheses bearing on the nature of the attributes that should be considered relevant within the classificatory system. (66)]

After the domain has been specified,

some decisions must be made as to exactly what attributes and descriptors are selected for evaluation. (66)

They indicate there are three lines of attack. Either (a) "evaluate all relevant descriptors", or (b) a random sampling of them, or (c)
attempt to specify significant attributes and
descriptors in the sense that they are likely to
differentiate the relevant classes and/or are of some
practical concern within the context of the classification
attempt.(66)

Concerning 2: Fleishman & Quaintance draw attention to the
interrelatedness of taxonomic issues:

the purpose for which classification is attempted has
implications for the definitions of tasks...(which]
determine ... the appropriate conceptual bases for task
description and differentiation.(68)

They also emphasize that the descriptors serving to classify the
tasks should ideally be operationally defined, and at least
capable of objective identification. This may be compared to the
SL situation -- at present, for example, there is disagreement
concerning one of the task characteristics being used in SL
research, with some referring to 'one-way' versus 'two-way'
tasks, others to tasks in which information transfer is
'optional' or 'required', and yet others considering positions on
a continuum between one-way and two-way tasks. This would appear
to be the sort of problem, though in a different domain, that
Fleishman and colleagues have begun to address.

Fleishman & Quaintance's third and fourth points are
obviously applicable to the development of a task taxonomy for
SLA/language syllabi, though little if any SL task-related
research addresses these points.[10] SL investigations
addressing these points will be necessary for a fully worked-out
task syllabus design program.

4.1.2 Human performance research concerning communicative tasks
The main body of the work reviewed in Fleishman & Quaintance
(1984) considers human abilities with respect to tasks, which are
often those that an operative in an industrial or military
setting would perform alone, or at least without extensive
communication with others. However, a small set of studies
concern team performance, for which communication is required.
The initial objective (413) was
to define the domain of team performance by focusing on
performance dimensions that make effective, synchronized
work possible.

Nieva, Fleishman & Rieck (1978) proposed four major categories of
team performance functions:
1) orientation functions (processes for generating and
distributing information necessary to task accomplishment to
team members);
2) organizational functions (processes necessary for members
to coordinate actions necessary for task performance);
3) adaptation functions (processes occurring as team members
adapt their performance to each other to complete the task);
4) motivational functions (defining team objectives and "energizing the group" to complete the task).

Attempts were made (Shiflett, Price & Schemmer 1981, and Shiflett, Eisner & Inn 1981) to validate this taxonomy. Thirteen activities carried out in teams by Army motor crew and bridge-construction Platoons were videorecorded. The tapes were then viewed by "naive judges" who had been given "a short training experience". Fleishman & Quaintance give an example concerning just one aspect of one of the four dimensions measured: "information exchange about member resources or constraints". Tasks varied from "[a] very simple task requiring little exchange of information about resources", e.g. "loading troops onto a truck", to "[a] very complex task structure requiring [a] high degree of information exchange about internal resources", e.g. "10km march on cloudy night in dense terrain; no loss of equipment or personnel". Judges were required to rate the tasks they observed in terms of the tasks' demand for exchange of information concerning resources and constraints, and in terms of the level of information exchange that actually occurred. In most categories, it is reported that subjects were able to "detect reliably" the functions. The lack of detailed reportage and apparent unconcern for careful measurement of rater agreement etc. is unsatisfactory. However, this study is an attempt to attack problems of a parallel nature to those faced by language task users.

4.1.3 Job-related task classification Other research in the human performance domain is summarized by Peterson & Bownas (1982), who provide a review of "the major existing task taxonomic systems". (See also Farina 1973.) This research may be linked to Long's suggestion concerning use of the Dictionary of Occupational Titles for derivation of pedagogical tasks from job descriptions, in that if there are a limited number of descriptors which can identify most jobs, it may be possible to make use of these in the construction of pedagogic tasks for a VESL/ESP task syllabus. The pioneer attempt at developing a job-related task taxonomy was that of Fine (1963): the 'Functional Job Analysis (FJA)', which formed the basis of early editions of the DOT. Peterson & Bownas argue that

[t]he FJA model classes jobs as a whole, rather than investigating task structures within jobs...it seems unlikely that these rather broad constructs will offer much value in determining specific job or task requirements.(55)

Subsequent investigations concerned the 'Position Analysis Questionnaire (PAQ)' (McCormick 1976, McCormick, Jeanneret and Meecham 1972).

The PAQ evolved from an intitial attempt to determine rationally all activities engaged in by workers. The resulting instrument focuses most heavily on work behavior in blue-collar skilled and semi-skilled manual jobs. (Peterson & Bownas 1982: 56)
McCormick, Jeanneret & Meacham (1972) used the 183 item PAQ to investigate 536 jobs, and, following factor analysis of the responses arrived at 5 general factors and 27 factors specific to topical subsections of the PAQ. Subsequent studies have used larger and more representative job samples, and arrived at slightly increased numbers of factors. Dunnette & Fleishman summarize:

depending on the level of generality or specificity desired, it appears that any given job may be described with the Position Analysis Questionnaire according to 12 quite general categories (e.g. clerical tasks, supervision tasks, physical tasks, etc.) or according to 32 much more specific categories (e.g. decision-making tasks, physical coordination tasks, etc.). (1982: 4)

However, commenting on the balance of the PAQ, Peterson & Bownas (57) observe that "most of the PAQ's discriminatory power is centered on blue-collar positions". Some of the factors derived from it would be inapplicable not only to white-collar work, but also to the sort of tasks used in SL research and syllabi. Obviously, any attempt to make use of instruments like the PAQ would need to consider modifications, or indeed whether it was necessary to use the entire instrument.

4.2 Small group research

Small group research can be divided into three major areas (McGrath 1984: xiii): studies investigating "groups as vehicles for delivering social influence"; those dealing with "groups as structures for patterning social interaction"; and "groups as task performance systems". It is obviously the last section of small group research which is of interest here. However,

there has really been very little study devoted to the ... analysis of task differences, in a systematic way, that takes into account how those differences affect group task performance. (53)

McGrath summarizes and integrates the existing work in this area (e.g. Carter, Haythorn & Howell 1950, Shaw 1973, McGrath & Altman 1966, Hackman 1968, Steiner 1972, and Laughlin 1980 inter alia) to provide the currently most fully developed task classification scheme for small group work. He notes

Ideally, the categories of such a classification scheme should be (a) mutually exclusive...(b) collectively exhaustive...(c) logically related to one another...[and] (d) useful. (60)

The McGrath system provides four major categories: Generate, Choose, Negotiate & Execute. McGrath argues that these are the four general performance processes involved in group activities. An activity's status with respect to each of these categories

30
enables it to be assigned to one of eight major types:

- **Generate**
  1. planning tasks
  2. creativity tasks: such as brainstorming
- **Choose**
  3. intellective tasks: solving problems which have a correct answer
  4. decision-making tasks: where arriving at a consensus is required
- **Negotiate**
  5. cognitive conflict tasks: viewpoint conflicts to be resolved
  6. mixed-motive tasks: negotiation/bargaining/coalition-forming tasks
- **Execute**
  7. contests/battles/competitive tasks
  8. performances/psychomotor tasks.

McGrath states that "the eight types can accommodate virtually all tasks used in group research", and points out that tasks can be additionally classified in terms of their cognitive, affective and conative aspects, with types 1 & 2 being in some sense maximally cognitive, type 5 most affective, and type 8 most behavioral; and can be further classified in terms of their need for or tendency to promote "cooperation or facilitative compliance", or "conflict or contrient interdependence".

The system appears to provide a useful preliminary framework for locating almost all tasks mentioned so far. As such it may help in organizing current and future SL task-related investigations, and might serve as a possible nucleus for a syllabus-oriented task classification system. It is not necessary to conceive of this system as unitary nor complete, however. Obviously, none of the categories are psycholinguistically motivated, but it may be that additional descriptors which are, such as 'information-structure', can be added. The system gains support from its adherence to basic principles of classification system design (exclusivity, exhaustivity, etc). Whilst based on extensive previous research, it has not, however, been developed through the rigorous procedures outlined by Fleishman & Quaintance, nor has any attempt been made to validate it empirically.
5. Summary

It has been shown that the category 'task', as used by researchers generally, is widely applicable and has psychological reality. Much, if not most, of human purposeful activity, whether in employment or in the classroom, can be seen as a series of tasks -- some having a communicative aspect, others not.

Section 1 reviewed the use of 'task' as a psychologically real term of analysis, both within second language studies and in more general educational research.

Section 2 provided a sketch of what little is known concerning the interaction of task characteristics and language. Though task characteristics cannot predict the precise nature of language generated, there are early indications that some affect the language that task participants use. They thus affect opportunities for language to be produced ('output'), the linguistic environment ('input'), and opportunities for it to be learned from ('intake'). This suggests that ideally, both syllabus designers and SL researchers would be able to select tasks whose potential for language learning (given particular learner variables) would be known. At present, however, there is disagreement even about information-transfer, one of the few task characteristics so far identified as significant in terms of language generated.

Section 3 further explored the applications of 'task', and recapped the arguments against existing syllabi and those in favor of task-based process syllabi. Organizational considerations suggest that selection and sequencing are still necessary, but this section concluded by indicating that no specific answers to questions concerning task selection and sequencing have yet been found.

Consequently, Section 4 considered research into human performance and small group work. Researchers in the first of these areas have preliminary methodology for the development of task taxonomies, which may be of particular relevance to ESP/ESL task syllabi. Work in the second area has suggested a taxonomy of potential relevance to SL communicative task classification.

It is hoped that this review has demonstrated the multivariate nature of tasks. For the second language researcher, a knowledge of which variables are the most potentially confounding, in addition to those of specific interest in any particular investigation, should permit rational, principled selection of experimental tasks.

For the syllabus designer, the principal problems have been stated by Long: selection and sequencing. The complex nature of tasks makes it obvious that there can be no mere linear sequencing (or univariate classification). Language, in its fullest sense, also cannot be reduced to a "linearly ordered
content domain" (Wilson & Brock, 1985). Theoretically motivated and/or empirically based decisions must be taken concerning the most appropriate groupings of tasks. Then within groups, perhaps a reduced number of variables can be applied to rank tasks within groups, or alternatively to specify what factors can be manipulated to arrive at task variations of differing ranks. Looking at the problem from the perspective of an individual task -- ideally, it should be possible to identify the characteristics of the task which would make it appropriate for a particular learner-group at a particular stage in a sequence, etc. Each task would be identifiable in terms of a set of variables, and each could thus be conceived of as occupying a position in a multidimensional grid.

On the basis of the relatively small number of studies surveyed so far -- few of which have specifically addressed themselves to the topic at hand with regard to language, it should be clear that possibly only a small part of such a multidimensional grid has yet been revealed. Other variables affecting performance, and other significant task characteristics are yet to be discovered. At the present time, some general research questions are as follows:

**Attention**
What is the nature of the interaction between attention to discourse and change in IL? In what way (if at all) do task attentional demands interfere with learning? Are task attentional demands a function of task complexity?

**Complexity**
What are the most important components of task complexity? Do the putative variables which alter task ranking within a type really result in language which is different?

**Task-characteristics**
What criteria should be used to distinguish tasks? Is one-way/two-way a dichotomy or a continuum, and are there other information-structure factors of importance, e.g. visual support? Are there other aspects of the task or discourse which have positive or negative effects on the utility of the linguistic environment, such as Duff’s 'convergence'? What is the role of shared information background in this respect? Do different tasks permit the application of different learning strategies and thus result in more or less learning?

**Input versus Output**
Is quality of linguistic environment a restricting factor on learning? Is quality of output in task discourse as important as quality of input, i.e. what is the balance between intake and practice in contributing to IL development? Do some tasks permit more of one than the other, and should this be capitalized on?

**Syllabus-design and classroom utilization issues**
What procedures should be used to derive pedagogical tasks from job descriptions? Given that there is to be no language-based sequencing, how necessary is task-complexity-based sequencing? Is there an "intrinsic motivation" factor in (some) tasks? Does it outweigh their lack of relevance? How important is task content in VESL/ESP situations?
This is, in fact, a very undelimited set of questions. No one experiment, however well-conceived or complex, could possibly be able to answer them all. At such an early stage, experimental work is beginning to tackle one or two small corners of the problem. Undoubtedly, work will go on at a number of sites, relatively independently. This review will have succeeded in its aims if it reminds researchers to take into consideration the breadth of issues bearing upon the problem, and in so doing, develop work which can be cumulative in nature, through constructing investigations which are compatible with, and build upon, those of others.

Notes

1 For specific studies, see Grosz (1982), Levinson (1979), and Hutchinson (1978).

2 It should be mentioned, however, that there is no necessary connection between non-understanding and negotiation, nor between negotiation and acquisition.

3 A related aspect of discoursal quality would be cohesiveness. L2 work has yet to consider this, but for L1 work see Friedlander, Thibodeau, Nichols, Tuoker & Snyder (1985).

4 Since this study involved not only L1 subjects, but subjects who were young children (12 each from grades k to 3, i.e. approximate ages 5 to 8), results must be taken as no more than intriguing, but a replication for adult L2 subjects would appear to be called for.

5 It may be wondered whether the presence of metacommunicative segments of task-based discourse (or even 'free' discourse) would be an indicator of the potential of such discourse for SL learning. The answer may relate to the overall purpose of the interaction as perceived by the NNS. (See also Faerch 1985.)


7 See Miller, Treiman, Cairn & Roos (1980), for a critique of the validity of DOT individual measures.

8 See Long & Porter (1985) for a review which allays such fears.

9 From a practical point of view, it has also been severely hampered by the fact that there has been no institutional or research support, with some instructors participating on an unpaid, voluntary basis. The project will cease in 1987 when its
instigator, N. S. Prabhu, moves from Madras. (E. Ramani, personal communication.)

10 S. Ross (p.c.) is currently conducting SL task research which makes use of tasks classified on differences of degree on a common dimension; some such (unreported) attempts are made in the Bangalore project (A. Beretta, p.c.).
REFERENCES


Brumfit, C. J. 1984a. Function and structure of a state school syllabus for learners of second or foreign languages with heterogeneous needs. ETJ Documents 118, 75-82.


Buckley, P., V. Samuda & A. Bruton. 1978. Sensitizing the learner to group work - a method. Practical Papers in Language Education 1, 123-152.


Candlin, C. N. 1984b. Implementing process curricula in ESL: the issue of 'task'. Paper presented at the 18th. annual TESOL Convention colloquium on syllabus/curriculum design, Houston TX.


Hutchinson, T. 1978. The practical demonstration: an analysis of the effect on discourse structure of a visual display. *Practical Papers in English Language Education* 1, 1-42.


Labov, W. 1969. The study of language in its social context. 
*Studium Generale* 23, 30-87.


