Review article

Task-based instruction

Peter Skehan[1] King's College, London, UK

This article is organised in five main sections. First, the sub-area of task-based instruction is introduced and contextualised. Its origins within communicative language teaching and second language acquisition research are sketched, and the notion of a task in language learning is defined. There is also brief coverage of the different and sometimes contrasting groups who are interested in the use of tasks. The second section surveys research into tasks, covering the different perspectives (interactional, cognitive) which have been influential. Then a third section explores how performance on tasks has been measured, generally in terms of how complex the language used is, how accurate it is, and how fluent. There is also discussion of approaches to measuring interaction. A fourth section explores the pedagogic and interventionist dimension of the use of tasks. The article concludes with a survey of the various critiques of tasks that have been made in recent years.

Introduction

During the 1970s there were considerable moves within language teaching to embrace the communicative approach (Brumfit & Johnson 1979). As a result, a range of teaching activities gained prominence which emphasised the need for learners to focus on meaning and to convey information to one another (see Geddes & Sturtridge 1979; and Harmer (1983, 1st Edn.) for exemplifying work, and Harmer (2001, 3rd Edn.), Wesche & Skehan (2002) for more recent reviews.). At this time, the assumption seemed to be that it was not enough in language teaching to focus only on language structure, but that this needed to be accompanied by a concern to develop the capacity to express meanings (Widdowson 1978). The implications of these pedagogic developments were widespread, and influenced syllabus design (White 1988), methodology (Nunan 1989), assessment (Morrow 1977, 1979), and an early (and influential) proposal for the use of task-based approaches (Prabhu 1987), an innovation all the more remarkable for the setting in which it was introduced – state secondary school classes in Bangalore.

At the same time as these pedagogic developments, there was also considerable growth in the relatively new (research-based) field of second language acquisition, where a similar (but stronger) underlying claim was made – that naturalistic exposure to and use of language were prerequisites for development, but with some slightly different emphases – that naturalistic use has to come first and is necessary to drive forward interlanguage, i.e., the structural development of language. As Hatch (1978, p 404) put it:

> One learns how to do conversations, one learns how to interact verbally, and out of this interaction, syntactic structures are developed

During the 1970s and 1980s, the basic 'unit' which was shared by communicative approaches and second language acquisition researchers was the *communicative activity*, a relatively vague term, but easily interpretable (albeit perhaps differently) by researchers and teachers alike (Mitchell & Johnstone 1986). Interestingly, during the 1980s, two major developments occurred. First, the term 'communicative activity' became increasingly replaced by that of 'task', even though the actual events that were referred to were essentially the same (Rubdy 1998). Second, what might be termed strong and weak forms of a task-based approach to instruction emerged. In general, proponents of the weak position tend to assume that tasks are not the driving force for syllabus design; that the use of tasks is an adjunct to structure-based teaching; and that it may be possible to 'clothe' structures through tasks without compromise. This approach tends to assume an automatisation or practice view of learning (Bruton 2002). In contrast, those who take the stronger view of tasks have generally seen the engagement of acquisitional processes as central, although views on the conditions which engage such processing have changed. Initially, Krashen (1985) was very influential in arguing for both the necessity *and sufficiency* of input. However, evidence from a number of sources (e.g., the lack of sustained production development in immersion educated

Peter Skehan is Professor of Applied Linguistics at King's College, London. He has published widely in applied linguistics, particularly in the areas of second language acquisition and task-based instruction, individual differences in language learning, and language testing. With Pauline Foster he is currently researching the nature of fluency in task-based performance, as well as the relationship between quality of interaction and nature of linguistic performance.

[1] The author would like to thank Pauline Foster, Janet Hooper and four anonymous reviewers, who have all made very useful contributions to the present review.

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

children (Swain & Lapkin 1982; Harley & Swain 1984)) suggested that input alone is not sufficient.

A first reaction (cf. Hatch, above, and especially Long (1983, 1985a)) was to argue that *interaction* is crucial, as are the opportunities it provides for learners to receive personalised, well-timed feedback for areas of interlanguage which are problematic (Pica 1994). Hence the promotion of indices for the negotiation of meaning, such as comprehension checks, clarification requests, and confirmation checks, all regarded by one group of researchers as key indices of interactions in progress which would be supportive of acquisition. More recently the central generalisation about negotiation of meaning has been modified somewhat, again by Michael Long (Long & Robinson 1998), to argue that within the use of tasks, there needs to be a Focus-on-Form (FonF), such that, even though learners may be participating in interactions, with meaning as primary, there is some concern for form (form-in-general, rather than specific forms). As a result, the naturalness of communication is not compromised, but form, and potential for development, do have some priority. Long (1991) argues that without this, outcomes such as those from immersion education are likely to occur, with a lack of sustained development. We are currently in a period where the Focus-on-Form generalisation is widely accepted by almost all researchers working on task-based research, whether from Long's (1989) interactionist perspective, or those taking a more cognitive approach (Skehan 1998; Robinson 2001). (See Sheen (2002) for a dissenting view, based on his reading of relevant empirical work.)

The switch from a Negotiation of Meaning justification for the use of tasks to one linked instead to the concept of a Focus on Form is important since this latter term can function more naturally as an umbrella term for a number of contrasting approaches. Although a Negotiation of Meaning perspective is quite consistent with this latter construct, as is its more recent manifestation, recasting, (Doughty & Varela 1998; Doughty & Williams 1998a; Long, Inagaki & Ortega 1998), so are other approaches to tasks, including a general, cognitive, attention-driven perspective (Skehan 1998; Robinson 2001) as well, even, as some sociocultural approaches (Swain & Lapkin 2001). All of these assume that (a) interaction, in itself, is not enough, and (b) insinuation of a focus on form into interactions is vital. Approaches differ in how they try to achieve such a Focus on Form, as we will see below. Broadly, they vary in the reliance they place on feedback (cf. Doughty & Williams 1998b), attention allocation (Skehan 1998; Doughty 2001), interaction (Van Lier & Matsuo 2000) and output (Swain & Lapkin 2001). They are all consistent with current conceptualisations of the role of noticing, whether this is through input (Schmidt 1990) or noticing the gap in output (Swain 1995).

Another preliminary issue concerns who might be interested in the use of tasks. Essentially three major groups can be identified: researchers, testers, and teachers, and each group has different concerns when working with tasks. Researchers, other things being equal, are likely to see tasks as convenient or necessary means to explore theoretically-motivated questions. (The questions may themselves have pedagogic relevance, (or they may not), but that is not likely to be the immediate motivation for the stance that researchers take.) It is likely, in this case, that the emphasis will be on some kind of empirical study, and its level of validity, and the task that is chosen, and the way it is used will be secondary to the research questions which are being asked. It is therefore probable that the task will be fairly self-contained, and explored within a cross-sectional research design – the task is done, possibly under non-classroom conditions, the data are gathered, the researcher leaves and analyses these data. (Although see Bygate (1996, 2001) who shows that longitudinal designs are also appropriate, and Lynch and MacLean (2001) and Samuda (2001) for illuminating case study approaches.) Language testers are also likely to be interested in self-contained individual tasks. In their case, the motivations are to obtain data (a) arising from actual communication (b) that can be assessed (c) and which are standardised. Except that the overriding goals for testers will be to work with tasks of known qualities, and even, perhaps, to be able to choose tasks precisely because they do not intrude into the sort of performance which results, there is considerable similarity between testers' needs and those of researchers.

In contrast, a pedagogic approach, where tasks are investigated by teachers, will necessarily be classroom-based and not subject to the same degree of manipulation characteristic of a research study. Here, it is likely that the relevant time-span for the task will be extended, since teaching goals will go well beyond simply demonstrating some sort of experimental effect, and are likely to be integrated within some extended pedagogic sequence. Accordingly, one can (rather programmatically) distinguish between:

- the focussed task ⇒ of interest to (experimental) researchers and testers
- a pedagogic sequence ⇒ of interest to teachers and some researchers
- an entire class period ⇒ of interest to teachers and non-experimental researchers
- an extended sequence ⇒ of interest to longitudinal researchers, non-experimental researchers, and to teachers
- a syllabus, a textbook ⇒ of interest to course designers, materials writers

There is some degree of overlap here, in this rather idealised sequence. The purpose of the 'sequence' is

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

to bring out the different expectations of all these groups, and the different standards that are likely to apply. The sequence is characterised by a move from greater control, at the focussed task end, moving to greater classroom practicality. Similarly, in terms of purpose in using tasks, there is a movement from demonstrating theoretically-valued 'effects' of tasks, on the one hand, to enabling teachers and others involved in pedagogy to work with learners over extended periods to promote change in the learner's language system, on the other. Interestingly, the possibility of longitudinal research, which would draw researchers into exploring more extended time periods, has not been characteristic, perhaps because of the lack of control, expense, and also the more protracted manner in which findings would emerge.

Given this complex background as to how tasks have emerged as an interesting basis for acquisition, and for instruction and assessment, it is useful to offer a definition of the concept. The purpose in doing so is to clarify what might qualify as a task, and what is unlikely to, as well as provide links to theory and to different contexts for task use. Bygate, Skehan and Swain (2001, p. 11) offer a core definition of task, as follows:

> A task is an activity which requires learners to use language, with emphasis on meaning, to attain an objective.

Bygate *et al.* (2001) also provide extended definitions to relate more appropriately to research or to pedagogy, and within each of these, to learning, teaching, and testing. These extended definitions need not concern us here – essentially they attempt to be more specific in their scope, and to convey a little more clearly *why* tasks are used in slightly different contexts.

There have, though, been difficulties regarding the definitions of task. One source of difficulty concerns an additional feature, often used in task definitions – the relationship of the task to the real world. Long (1985b) relates this to things people do in everyday life, while others (e.g., Skehan 1998, Ellis, in press) prefer phrases such as 'bears a resemblance, direct or indirect, to the way language is used'. In so doing, they are preferring to emphasise the nature of the response by the learner (cf. Widdowson's (1979) 'genuineness') rather than a form of authenticity, defined only in relation to the real world occurrence of an activity. In other words, for Skehan and Ellis, it is the reaction of the learner or the research subject which is key, rather than the fact that a particular task may have, at some time, been used by native speakers (possibly in a different context). Guariento and Morley (2001) have also recently provided a discussion of task authenticity in broadly similar terms.

A second issue which arises from definitions concerns referential tasks. Critics of task-based approaches (Cook 2000; Seedhouse 1999; Block 2001) attack an exclusive concern with transactional and referential tasks, i.e., tasks which emphasise information transmission and exchange (Yule 1997). It should be clear from the above definition that there is no restriction implied which requires tasks to be of this sort. It may be that many task based studies are of this type, but there is no requirement that they should be so on definitional grounds, and indeed many research studies use non-referential tasks.

Now that we have explored the contexts which led to an interest in tasks, as well as the issue of task definitions, we turn to exploring different approaches to researching and using tasks in more detail.

Research oriented approaches to tasks

Given the vitality of the research into tasks, it is not surprising that a number of different approaches can be distinguished. We will explore four major approaches: a psycholinguistic approach to interaction; a social interactive approach; a cognitive perspective; and a concern for structure-focussed tasks. (See Yule (1997) and Ellis (2000) for other, complementary reviews of the task literature.) The section will conclude by examining individual variables in task research.

A mild warning is appropriate, though, at the outset. Research which has been conducted tends to be with adults (and some adolescents), generally at intermediate proficiency levels, and mostly with English as the target language. Many of the findings may well hold up with other ages, other proficiency levels, and other languages. But this cannot be assumed, and so until wider-ranging research appears, any claims which are made have to be limited in nature.

A psycholinguistic approach to interaction

This represents the first major research area to emerge into task-based instruction (see discussion above), and has been heavily influenced by Michael Long's (1983, 1989) proposals for the role of interaction, generally, and particularly with respect to the negotiation of meaning. Negotiation of meaning concerns the way learners encounter communicational difficulties while completing tasks, and how they do something about those difficulties. Long proposes that the interactional adjustments that learners make to address such difficulties serve to induce their interlocutors to modify the input they are providing. Even more significantly, the responses that learners receive when meaning is negotiated in this way delivers feedback to the learner at the most propitious moment. The feedback arises when meaning is problematic, and when the learner is thought to be most receptive (Pica 1994). In addition, it is likely to be personalised, since, as interactants try to take the communication

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

forward, what will happen naturally will be the provision of useful information on precisely the area of language that the learner is struggling with. In earlier work Long (1989) proposed that tasks which lead to beneficial negotiation of meaning of this sort are indexed by greater numbers of comprehension checks, clarification requests, and confirmation checks. These conversational moves are then taken to reflect the degree of usefulness of the interactions concerned. More recently, research in this tradition has moved to advocate a slightly different conversational feature, the *recast*, i.e., where an interlocutor rephrases something said by a non-native speaker, and so provides a model and feedback to the learner when the learner may be most open to such a contribution (Long, Inagaki & Ortega 1998). See Gass (2002) for a recent review of this work.

Research in this tradition has shown a number of effects of task characteristics and task conditions upon the incidence of negotiation of meaning and of recasts. Long (1989) proposed that convergent tasks (i.e., where participants have to agree on an answer) would produce more negotiation of meaning than divergent tasks (i.e., where no agreement is envisaged, as in a discussion). A study by Duff (1986) explored this possibility, but provided only weakly supportive results. Pica and Doughty (1985), though, showed that group and pair based interaction, where information exchange was required, provided more conversational modification than a teacher-fronted situation. Results such as these, reviewed in Pica (1994) and systematised in Pica *et al.* (1993) suggested that it is possible to design interactions in order to produce greater amounts of negotiation.

More recently, this interpretation of interaction research has switched to focus on recasts, i.e., repetitions of a learner's incorrect utterance, but with changes made in order to make it correct. The relevance, as with the Negotiation of Meaning studies, is the claim that tasks provide an effective frame for recasting to occur. It is proposed (Long, Inagaki & Ortega 1998) that this provides negative feedback which is helpful to the learner's development. Mackey, Gass and McDonough (2000) and Lyster and Ranta (1997), for example, report significant recasting in two classroom based studies, as do Doughty and Varela (1998), although in this case the situation was organised specifically to produce *intensive* recasting. Experimental studies, too, have explored recasts. Long *et al.* (1998) reported that recasts were more effective than correct models in bringing about short-term change in learner language. This recasting literature has been recently summarised by Nicholas, Lightbown and Spada (2001), who propose that recasting is more effective when a learner has already begun to use a particular language feature, and where the recasting helps the learner to make effective discrimination between alternatives. Nicholas *et al.* (2001) also propose that recasts are not equally effective in all areas of language.

To research naturally occurring recasts is to examine the reactive moves by learners that arise from conversation during tasks. Some related research has appeared recently which seems to accept a similar general framework, but instead takes a more interventionist approach. For example, Ellis *et al.* (2001) have researched the use of precasts, i.e., teacher or learner anticipations of trouble immediately ahead (a form of predicting gaps which are relevant). They argue that precasts are more likely than recasts to be followed by uptake, i.e., evidence that the 'feedback' which is provided is incorporated into learner utterances. Shehadeh (2001) makes similar claims for self-initiated recasting, where the learner seems aware, spontaneously, of a problem. Williams (2001) also argues for a stronger connection between learner-initiated repair and the likelihood of uptake. Finally, it is worth recalling a slightly earlier study. Noboyushi and Ellis (1993) argue that targetted intensive recasting is particularly effective at producing learning, i.e., evidence that feedback doesn't simply occur, but is noticed, recognised for its potential, and has an impact on the learner's developing interlanguage.

The Negotiation of Meaning/Recasting studies have, though, received some criticism. Aston (1986) has proposed that tasks which require a lot of negotiation of meaning irritate learners, with the result that any potential that negotiation might have to provide relevant feedback is not realised. Foster (1998) argues that negotiation of meaning does not occur in actual classrooms as frequently as might be expected from the lab-based studies. She also suggests that such negotiation of meaning as does occur is primarily lexical, and often accounted for by particular students who negotiate a great deal (sufficiently to generate statistical significance in the groups of which they are part), while most learners do not negotiate in this way at all. Lyster (1998; Lyster & Ranta 1997) criticises the more recent switch in emphasis towards recasts. He suggests that such recasts are not particularly frequent in classes; that the feedback that they provide is not noticed by learners; and that even when it is noticed it is not recognised as appropriate to the intended point of language, and even if it is so noticed, it is not incorporated in the learner's own speech.

A sociocultural approach to interaction

Although the negotiation of meaning literature is centrally concerned with interaction, it is a particular type of interaction, i.e., what happens when there is some level of conversational breakdown, and the impact of feedback at such points for the development of *form*. A contrasting view of interaction is that taken within sociocultural theory, where researchers

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

explore how learners co-construct meaning while engaging in interaction. In this case, there is no particular concern with negotiation of meaning, in the previous sense. In contrast, it is assumed that the interest in a task is to allow participants to shape it to their own ends and to build meanings collaboratively that are unpredictable and personal. An excellent recent overview of such an approach is provided in the chapters in Lantolf (2000).

In fact, there are different, often overlapping approaches within this general area. One concerns the way tasks are reinterpreted by learners, and made to respond to their individual needs and interests. Duff (1993), for example, (and see Coughlan & Duff (1994)) explored three tasks (a discussion, a picture description, and the retelling of a Cambodian folk story). She reports considerable variability in performance on the same task at different meetings, and it is clear that learner interpretation of task is a major influence on this. In effect, this evidence questions the feasibility of analysing task qualities in any sort of static way, and using such analysis as the basis for predicting consistent dimensions of task performance. It is clear that other research strategies, which tend to minimise such interpretations, need to address this issue.

A different approach, but still within sociocultural theory, is represented by Swain's work within the context of immersion education in Canada. There is a strong focus on structural development in this research, but the proposed key to such development concerns the potential of different tasks to enable language understanding to be scaffolded by participants, mutually, through interaction. For example, Swain and Lapkin (2001) compare dictogloss and jigsaw tasks for their capacity to engage learners in collaborative interaction. This research tradition is less concerned with demonstrating quantitative differences and instead tends to focus on language-related episodes (LREs) in learner interactions to show how each learner may contribute aspects of language structure that the other cannot, and, as a result, both learners in (say) pair work would benefit.

Still within sociocultural approaches, another perspective concerns the nature of interaction itself. Van Lier and Matsuo (2000), for example, explored whether interactions vary measurably in how symmetrical and collaborative they are. They examined whether interlocutors ratified one another's contributions, whether they responded to them and developed them, or in contrast, whether they failed to do these things, with one interlocutor dominating and controlling the interaction. They report clear differences between learners on these indices. Developing this approach, Nakahama, Tyler and Van Lier (2001) compared performances on a discussion and an information-gap task. The study points up the difference between a broader notion of interaction that arises naturally from sociocultural theory, and that connected with negotiation-of-meaning studies. Nakahama *et al.* (2001) show that, in contrast to claims for the superiority of information gap tasks made by the Negotiation of Meaning literature (Long 1989), in fact the discussion task, too, provides many learning opportunities, with this advantage arising from what might be termed the greater depth of the interaction, and facilitation for extended turns.

Cognitive perspectives

Task researchers who have taken a more cognitive perspective have focussed on the psychological processes typically engaged in when learners do tasks. They have explored three main areas: analyses of how attentional resources are used during task completion; the influence of task characteristics on performance; and the impact of different conditions under which tasks are completed. There are two contrasting approaches here, but with many similarities. Skehan (1998) proposes that attentional resources are limited, and that to attend to one aspect of performance (complexity of language, accuracy, fluency) may well mean that other dimensions suffer. Skehan and Foster (1997, 2001) argue for the existence of tradeoffs in performance, such that, typically, greater fluency may be accompanied by greater accuracy or greater complexity, but not both. Robinson (2001, in press), in contrast, advocates two propositions: (a) that attentional resources are not limited in the way Skehan and Foster (2001) argue, but instead learners can access multiple and non-competing attentional pools, and (b) that, following Givon (1985), complexity and accuracy in a task correlate, since they are each driven by the nature of functional linguistic demands of the task itself. So whereas Skehan and Foster argue for fluency being correlated with either complexity or accuracy (at best), Robinson argues that fluency contrasts with complexity and accuracy, which correlate with one another. The evidence is not decisive, with either of these interpretations of attention, but tends to support the limited capacity view. It is clear that more research is needed.

Both cognitive approaches explore how performance can be affected by task characteristics and task conditions. Regarding characteristics, the findings seem to be as follows:

Task Characteristic		**Influence upon performance and research basis**
• structured tasks, i.e. clear time line or macro-structure	⇒	clearly greater fluency, tendency towards greater accuracy (Foster & Skehan 1996; Skehan & Foster 1997, 1999)

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

- familiar information ⇒ greater fluency and greater accuracy (Foster & Skehan 1996, Skehan & Foster 1997)
- outcomes requiring justifications ⇒ justifications lead to markedly greater complexity of language (Skehan & Foster (1997))
- interactive vs monologic tasks ⇒ interactive tasks produce markedly more accuracy and complexity, monologic tasks more fluency (Foster & Skehan 1996, 1999; Skehan & Foster 1997, 1999).

This range of findings, and the generalisations they support, indicate that while task choice hardly guarantees focus on particular aspects of language, there is some predictability involved. This, in turn, could potentially link with more effective pedagogic decision-making.

Researchers have also explored the influence of the conditions under which tasks are done, or the manipulations in conditions available to experimenters. This has been one of the most active areas in task research. One line of investigation concerns the phases which are relevant to using tasks, i.e., what happens before the task, what during, and what after. Regarding the first of these, the bulk of the research has been on the role of pre-task planning. Following early research by Crookes (1989) and Ellis (1987), which argued for connections between planning and, respectively, complexity and fluency (Crookes) and accuracy (Ellis), a number of studies have tried to establish how task performance is influenced by planning. One can now offer the powerful and robust generalisation that almost all studies (e.g., Foster & Skehan 1996, 1999; Mehnert 1998; Ortega 1999; Skehan & Foster 1997, 1999) support a clear influence of planning on complexity and fluency: these performance features are almost always improved when there is planning. The situation with accuracy is not so clear. Some studies (e.g., Foster & Skehan 1996; Skehan & Foster 1997; Mehnert 1998) do suggest that accuracy is raised when there is planning, but other studies (e.g., Crookes 1989; Ortega 1999; Wigglesworth 1997, 2001) do not support this claim. It is certainly the case that any accuracy effect is smaller than the effects for other performance areas. Although it has been proposed that certain conditions are more likely to generate accuracy effects (e.g., not giving learners instructions; planning led by a teacher), using planning to dependably influence accuracy remains a challenge. It may be that on-line planning (Wendel 1997; Yuan & Ellis, in press), i.e., the capacity to regroup one's resources and plan while a task is being done, may be a more consistent correlate of this performance area.

Another area which has stimulated several studies is task repetition. An early study (Plough & Gass 1993) suggested that repetition may not be generally useful. Plough and Gass pointed to negative reactions on the part of learners required to re-do a task which had already been completed, and no particular advantages in terms of their performance. In contrast, Bygate (1996) has argued strongly for the value of task repetition. In a series of studies (Bygate 1996, 1999, 2001), he does not report dissatisfaction on the part of students completing narrative re-retellings. He also argues that task repetition has beneficial influences upon performance, with the repeated performance producing a more syntactic engagement as similar events in a narrative are retold, but with more complex syntax and a greater density of propositions. Following Levelt (1989), he suggests that the repeated performance enables greater focus on formulation and monitoring, as opposed to the demands of conceptualisation and effortful accessing of expressions during the first completion of the task (Bygate 1996). In addition, there is clear evidence that learners are able to access the original formulations that were used on the first telling (even after an interval of several weeks), and draw on this effectively on the second telling. Similarly Lynch and McLean (2001) report on a classroom activity in which medical specialists doing an English for Academic Purposes course had to repeat presentations on posters they had made to different 'visitors' to their poster. An issue concerned whether learners felt they were learning from the opportunity to repeat. It appeared to be the case that noticing useful language in ones interlocutor's contribution, or being able to improve one's own description of the poster varied from learner to learner, with the more advanced learners feeling more confident that they had made some improvement. Gass *et al.* (1999) also report improvement with video-based narrative retellings, as do Nemeth and Kormos (2001) using argumentation tasks. In their case, the repeated performance leads to greater focus on content and more use of supportive moves to underpin the arguments being advanced.

There has also been research into post-task effects. Skehan and Foster (1997, 2002) have shown that with interactive tasks, giving learners a post-task activity (such as the need to re-do a task publicly after the task has been done 'privately', or the requirement that learners transcribe one minute of their own task performance subsequent to the task itself) leads to significantly greater accuracy. In addition, the trend towards greater accuracy is stronger as time goes on – by the second and third weeks of a study, participants who have been required to engage in some sort of post-task activity showed significantly greater accuracy than in the first week of the study, suggesting that the channelling of attention selectively

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

towards accuracy is an influence which grows in impact.

Focussed tasks

Many supporters of the use of tasks in language instruction regard them as vehicles for learners themselves to set the agenda during interaction, and as a result, obtain feedback at points of interlanguage development relevant for them. From such a viewpoint, to design a task to predispose the use of a particular structural feature would be to defeat an important purpose for task use. Yet precisely this has been advocated by various groups of researchers. Loschky and Bley-Vroman (1993), for example, in exploring the relationship between tasks and particular language structures use the terms 'natural', 'efficient', and 'necessary', reflecting tasks which respectively (i) enable but do not predispose the use of a particular structure; (ii) are more effectively conducted if a particular structure is used; and (iii) force the use of a particular structure. They advocate the use of the third type of task, proposing that it offers the advantages of a task-based approach and also that of systematic, plannable instruction. It should be noted that the same criteria for 'taskhood' should be applied in this case too, so once again, the focus should be on meaning and outcome. Ellis (in press) also provides a justification for the use of such tasks, on the grounds that they are more effective in developing automatic processing and implicit learning, the former because such processing leads to a more sustained focus, and the latter because the greater structural clarity enables learning by induction to proceed more effectively.

Ellis (in press) distinguishes between three types of focussed tasks: structure based production tasks; comprehension tasks; and consciousness-raising tasks. In the first category, Fotos (1998) discussed work aimed at using tasks with larger classes, while Mackey (1999) reports on a task designed to elicit question forms, a task which was successful, but with some effort on the part of the learners, in leading to greater production of questions. A contrasting study of this type is that of Samuda (2001), who designed different tasks to induce (a) the use of modality, and (b) the 'there is/there are' form. She reports that the tasks were broadly successful in quite naturally eliciting the language forms artfully built into the input materials. She also focusses on the role of the teacher in such tasks, and reports that, in the case of the more complex area, modality, a significant amount of teacher intervention was needed to wean learners away from simpler expressions of modal meaning ('maybe', 'it is possible') and towards the use of modal verbs.

Ellis (2001; in press) also discusses comprehension-based tasks. These include input-enhancement tasks, such as those used in Doughty (1991), Jourdenais *et al.* (1995), and Trahey and White (1993). In all these tasks, some (selective) feature of the input is highlighted, on the assumption that implicit learning processes are thereby triggered. These studies do generally show positive results on performance. There are also input processing tasks, generally associated with the work of Van Patten (1996). In these, learners are given instruction on how to process input more effectively, in such a way that they will go beyond meaning-extraction, and also notice structural regularities in the input language. Results of such research are generally positive in terms of demonstrations of effects, but have received criticism because it is alleged that the effects are accounted for by other variables (see the exchange between Van Patten (in press) and De Keyser, Robinson, Salaberry & Harrington (in press)).

Finally, consciousness-raising tasks are those where (a) a specific feature of language itself is part of the task, and (b) the focus is on explicit learning. Of this task type, Newton (2001) explored how tasks may be used to promote vocabulary acquisition, while Fotos and Ellis (1991) used a focussed task to teach dative alternation, and reported that the indirect consciousness-raising approach was broadly as effective as conventional explicit instruction. Other studies and reviews (e.g., Fotos 1993, 1994; Sheen 1993) have not been so unequivocally positive. Such a task-type has also been linked with the issue of level of conscious awareness on the part of the learner. Leow (1998), for example, reports a correspondence between level of awareness and amount of learning.

Individual variables

An obviously relevant set of variables which has had relatively little attention in the literature concerns differences between learners. Few would argue that all learners respond to tasks in the same way – indeed a central factor with sociocultural theory is to enable the individual to interpret tasks in whatever way they think is appropriate. What is lacking is systematic research which probes just how learner differences may impact on tasks. Even so, there have been studies of gender (Gass & Varonis 1986), of proficiency level (Long & Porter 1985; Yule & McDonald 1990; Wigglesworth 2001) and of interlocutor familiarity (Plough & Gass 1993), but these have not tended to be regarded as central to the research tradition. More recently, though, there have been two studies in this area which indicate a deeper level reconceptualisation. Dornyei and Kormos (2001), in the context of researching motivation, in intact classrooms, distinguish between what they term high and low task attitude students. They report, from correlational evidence, that high task attitude students seem more affected by task manipulations than do low task attitude students. Mackey *et al.*(2002) have explored the relevance of differences in working

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

memory for the capacity to respond to recasting and feedback in tasks. They report a relationship between working memory capacity and the noticing of interlocutor feedback.

These two studies are important, in themselves and programmatically, because they indicate that group-based differences in task performance may disguise even greater differences for some students, and very small differences for others. In other words, if we could establish more clearly which individual differences are relevant for task performance, it might then be possible to increase the sensitivity of research designs by conducting condition-seeking studies (McLaughlin 1980), i.e., those which explore the consequences of *combinations* of conditions.

Measuring performance on tasks

One of the interesting features within task-based research has been the different ways in which researchers have operationalised and measured performance (Yule 1997; Skehan 1998; Ellis, in press). To a considerable extent, the different choices that investigators have made have reflected their theoretical positions. The following table clarifies this:

Cognitive approaches	Fluency Accuracy Complexity Lexical aspects of performance
Interaction: Negotiation of meaning	Negotiation of meaning measures • clarification requests • confirmation checks • comprehension checks • recasts
Interaction: Sociocultural theory	Measures of interactive involvement Measures of interactive symmetry

In addition, some investigators have explored the use of performance ratings (using conventional methods from language testing, e.g., Wigglesworth 1997, 2001) and other have also used participant ratings of task difficulty. Further, as Robinson (2001) points out, little use has been made of measures of outcome fulfilment (although see Yule & McDonald 1990; Yule & Powers 1994), or of dual-task research designs, where one might gauge degree of task interference, a strategy which has been widely used within cognitive psychology.

Although there are still areas of disagreement within task research, there has been significant progress in developing better performance measures. The complexity-accuracy-fluency dimensions of task performance have been justified both theoretically and empirically. Theoretically, the sequence implies the three stages of change in the underlying system (i.e., greater complexity, as more complex interlanguage systems are developed), acquisition of greater control over the emerging system (i.e., greater accuracy, as new interlanguage elements are used not simply haltingly and incorrectly, but instead with some reduction in error), and development of performance control, as elements are routinised and lexicalised (i.e., fluency) (Skehan 1998). In addition, Skehan and Foster (1997, 2001) provide evidence from factor analyses indicating how these different performance areas compete with one another for limited attentional resources, suggesting that each needs to be included in a study if any wide-ranging claims about performance are to be made.

A number of researchers have now used measures in each of these areas. One immediate issue is whether to use specific measures (e.g., targetting accuracy or complexity in a specific area, such as the article system, or Pienemann's developmental scale (1998)) or to use more generalised measures. Although earlier researchers (e.g., Crookes 1989) used specific measures, research in recent years has tended to take the latter option, and now a range of general measures are available. With respect to complexity, Foster, Tonkyn and Wigglesworth (2000) argue strongly for the use of what they term the AS-unit (Assessment of Speech). This, they argue, is more appropriate for a spoken language context, and provides a more appropriate measure than T-units (Hunt 1965) or c-units (Brock 1986), the first of which is more appropriate for written language and the second of which is not sufficiently clearly defined. Regarding accuracy, most researchers use a measure of percentage of error-free clauses, or errors per 100 words. In both these cases, generalised measures seem to provide more sensitive measures to detect treatment effects, even if the value of having specifically focussed predictions is thereby sacrificed.

Progress has been particularly clear with measures of fluency and lexical content. Regarding fluency, it is now increasingly accepted that finer grained analyses of fluency require separate measures of (a) silence (breakdown fluency), (b) reformulation, replacement, false starts, and repetition (repair fluency), (c) speech rate (e.g., words/syllables per minute), and (d) automatisation, through measures of length of run (Koponen & Riggenbach 2000). These seem to represent different sub-dimensions of fluency and are needed to give a comprehensive picture of performance in this area.

The measurement of lexical performance in tasks has, until recently, been beset with problems. The conventional measure which is available is the type-token ratio, i.e., the number of different words in a text (e.g., a narrative performance) divided by the total number of words in the text. Unfortunately, this measure is heavily influenced by text length. Given that most performances by second language learners

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

doing tasks are necessarily fairly short, this poses serious problems. Foster (2001) reports consistent correlations between type-token ratio and text length of around −.80, so it is clear that uncorrected type-token ratios are unsafe. Fortunately, Malvern and Richards (1997, 2002) have offered a statistic, D, which can be calculated (without contamination from text length) for texts of variable length. Their measure is available within the CHILDES Clan suite of programmes (MacWhinney 2000a, 2000b), and is likely to enable more effectively measures of lexical density in the language used within tasks.

Also problematic in the assessment of performance on tasks is the general notion of interaction. As we have seen, those persuaded of the importance of negotiation of meaning tend to rely on indices such as comprehension checks, recasts and the like. But it is clear that these are not the whole story in measuring interaction. It may be that in the future the sorts of measures developed by Van Lier and Matsuo (2000), which probe the extent to which interactants respond to one another's meanings, develop them and engage in reasonably balanced (symmetrical) interactions, will prove more effective at capturing quality of interaction.

Tasks and pedagogy

In general, task research has used cross-sectional research designs, generally of a quantitative nature, although increasingly with qualitative research components. This research may be done under laboratory conditions, or it may be located within intact classes. But in either case, the research 'intervention' has been brief, generally involving no more than two to three weeks, with probably less than one hour's intervention in each week. Researchers have attempted to use these more controlled conditions to chart the effects of task characteristics and the conditions under which tasks are done. The intention with such research is either to explore questions made significant by theory, or to establish claims which are generalisable to actual teaching situations (or both). In either case, though, the research does not reflect conventional extended pedagogic involvement, as would be regarded as typical by language teachers. In essence, this connects with the relative paucity of longitudinal research in second language acquisition.

Despite these qualifications, some claims can be made about the relevance of task research for pedagogy. Generalisations are emerging on the effects of different task characteristics on aspects of performance, e.g., the use of structured tasks to promote accuracy, or tasks based on familiar information to promote fluency, as reviewed earlier. Similarly, the research into task conditions, e.g., pre-task planning and post-task conditions, does connect, in a fairly direct manner, with teacher decision making and classroom practice, promoting complexity and fluency, in the first case, and accuracy, in the second. Teachers are routinely concerned with how to organise effective pre-task (cf. pre-reading, pre-listening) activities, and so planning, as an example of such an activity, could easily be incorporated within lesson planning, and can be assumed to be likely to produce the same sorts of effects that have been found in the research studies. Post-task activities are similarly essentially pedagogic in nature, and, when linked with research findings, an interesting basis for generalisation to classrooms.

But these applications of research findings do not really make sufficient connection with most classroom decision making. In that respect, one can go beyond applications of focussed studies, and identify other 'units' which concern teachers. These are:

- teaching sequences for an entire class (not just five minutes of task work)
- linked class teaching sequences
- project work
- extended periods (cf. coursebook series, syllabuses, schemes of work, etc.).

Proposals in each of these areas are less likely to be derived from research (which is not to say that relevant research is not desirable), but grounded in classroom experience.

The major alternatives regarding class and linked class sequences come from work by Samuda (2001) and Willis (1996). Samuda proposes a (long) class-oriented sequence which follows the stages (a) input data, (b) operations on input data, and (c) consolidation and reflection. In addition, she distinguishes between knowledge-constructing tasks, i.e., tasks intended to induce learners to pressure their interlanguage systems to develop new forms, and knowledge-activating tasks, i.e., tasks which do not push learners to develop new knowledge, but which instead promote salience, so that learners are more likely to mobilise language which they do know, but otherwise not be so likely to use. Finally, in terms of general principle, Samuda (2001) clarifies a role for the teacher within task-based work such that the teacher is present for long periods during the task. The teacher's role is then to 'lead from behind', i.e., to attempt to use the ways learners are engaging with the task and expressing meanings to provide relevant assistance with language form when learners themselves have created a 'need to mean'. In other words, the classroom methodology structures the way in which learners are induced to notice a gap in output (Swain 1995), and then the teacher supplies language relevant to the gap which has just been made salient.

An example of a knowledge-constructing task is what Samuda terms the 'things in pockets' task.

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

Learners are provided with some objects supposedly found in the pocket of a coat left in a plane. Their task is to speculate on the identity of the owner of the coat on the basis of these possessions. They are also provided with a matrix linking the objects with degrees of probability, and this constitutes the input data. The task targets modality, both at a simple level ('maybe', 'perhaps'), a slightly less elementary level ('it is possible'), and then the main focus for the task, the use of modal verbs. Learners are given such a focussed task because it creates an extended period in which a 'need to mean' has been contrived, and in which, as learners work on the task, the teacher can work with them, and introduce new forms, to attempt to induce the learners to allow these particular new forms to supplant the more restricted forms expressing modality. The task is completed with learners making a report, in the form of a poster, such that there is opportunity for consolidation and reflection on the forms which have just been introduced.

A contrasting approach to using tasks is advocated by Willis (1996), who proposes the following pedagogic sequence:

Pre-task

- introduction to the topic and task
- exposure to real language (tape recordings of native speakers completing the same task)
- use of texts and activities upon those texts

Task Cycle

- task
- planning
  - drafting and rehearsal
  - teacher assistance with language
- report

Language Focus

- analysis
- practice

There are similarities and differences here with the approach advocated by Samuda (2001). The most striking difference is that Willis' methodology for using tasks does not imply any pre-selection of form. It is intended that although a need to mean will be created, the language which is likely to be made salient in this way will be whatever language learners see as relevant. Although the pre-task may have an impact on what language learners perceive as relevant, essentially the role of the teacher is to react to whatever language emerges as important, and then help learners to address the gap which has been noted. It is also striking that, for Willis, planning follows rather than precedes the task. More properly, planning precedes the report phase, i.e., the phase by which learners should have done something about any gaps which have been noticed, and should be trying to mobilise resources so that the report phase is completed in a manner significantly better than the original task. It is only after the report phase that the lesson switches to an overt language focus. By this stage, in other words, the focus on language will have been 'earned' and made salient by all the preceding activity, and so consolidation and integration of new forms into an existing interlanguage structure can be accomplished more effectively. The approach, essentially, is consistent with Swain's (1995) proposals to encourage learners to 'notice the gap'. What are most important in Willis' approach are (a) input materials that push learners into 'primed' areas (cf. Loschky and Bley-Vroman's 'efficiency' condition), (b) learners are given considerable support by the teacher *after* they have noticed whatever gap may have emerged, and (c) language analysis, systematisation and consolidation take place *after* there has been some interlanguage restructuring.

Willis' (1996) proposals are for a methodology which is repeated, i.e., the same approach is used in different classes or short sequences of classes. But it is possible to consider larger 'units' of work, e.g., project work, coursebook series, and syllabus design. Unfortunately, at this level, there is no direct research to inform pedagogic decision making, although the approaches described here are consistent with descriptions of project work in Fried-Booth (1986) and Haines (1989). Little can be said save that research into task characteristics, conditions, and linked class teaching sequences is relevant, but would need to be applied with caution (Lightbown 2000). Even so, it is interesting to see that after something of a gap following the publication of the Cobuild series in the late 1980s (Willis & Willis 1988), one or two other major coursebook series are now incorporating tasks at more than a token level (see e.g., *Cutting Edge* (Cunningham & Moor 1999) and *Inside Out* (Kay & Jones (2001)). We will now have to wait and see in the next few years whether such approaches are practically and commercially successful, and make appropriate assumptions about the level of training on the part of the teachers who use them.

One interesting development, though, has been more at the methodological level. We saw with Samuda (2001) that it was important for the teacher in her study to work with students in developing clear form-meaning mappings that the learners would actually use (and that progress was hard-won, with some persistence required from the teacher). Two other researchers have explored how teachers can have an effective role within tasks, while they are running in the classroom, and without compromising the naturalness of the task. Lynch (1997) shows how teachers can avoid a heavy-handed approach to feedback by 'nudging' learners towards the correct use of forms, i.e., by a process of semi-intervention

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

(cf. Samuda's 'leading from behind'). Drawing on Swain (1995), Thornbury (1997) explores how teachers can exploit activities such as dictogloss and text reconstruction to induce learners to notice gaps and to restructure underlying interlanguage.

It is clear that one of the reasons why task-based instruction is not used more widely relates to the difficult role it implies for the teacher. If tasks and task completion are the driving force in class, teaching preparation is nothing like as exact. The teacher, in other words, has to be prepared for learners to take interactions in whatever direction they choose. Then the teacher has to be ready to provide the unpredictable help that will be required. This presupposes a broader type of readiness for almost anything to occur, compared to the more comfortable ability to prepare for the pre-ordained structure-of-the-day. Small wonder, then, that many teachers shy away from this approach.

Critiques of tasks

Over the last decade, there have been interesting critiques of a task-based approach to instruction, or of task-based research. Of course, some of the criticism comes from task researchers and is aimed at other task researchers! For example, cognitively oriented researchers do not, generally, work within a negotiation of meaning framework. Similarly, sociocultural theorists (e.g., Van Lier & Matsuo 2000) essentially doubt that the conceptualisation of interaction in other task-based approaches will enable acquisition to be accounted for.

But other critiques come from outside the task-based paradigm. Some critiques focus on just one aspect of a task-based approach, attack it, and seem to assume that as a result the entire task-based enterprise is doomed. For example, Sheen (1994) attacks tasks as if they were still input-dominated, and deriving their credentials from Krashen's (1985) work. He also criticises a small number of studies supportive of the claim that learners do not particularly learn errors from one another during tasks. Given the developments reported above, this critique needs to be widened and updated if it is to have any force. Similarly, one of the aspects of Seedhouse's critiques (1997, 1999) is that negotiation of meaning is inadequate as an account of the complexity of classroom interactional patterns – unproven and unprovable, as he claims. To say that accounts of interactional patterns are as yet inadequate and unproven in not contentious. To assert that a negotiation of meaning approach is unprov*able*, though, when little systematic evidence is presented that is relevant to this claim, is excessive, and hardly fair to the range of studies (unreviewed in his critique) which have been published within this approach. In any case, there is still the point that tasks are not simply vehicles for researchers to explore the relevance of negotiation of meaning – the vitality and diversity of task research indicates the limitations of Seedhouse's view.

Seedhouse's claim about the complexity of interactional patterns does chime, though, with a related set of critiques. A number of authors (Cook 2000; Block 2001; and see also Firth & Wagner (1997)) all claim that task-based research is limited by its excessive concern for referential tasks, and that what is missing from such research is the inevitable social dimension of language use. On occasions (Cook 2000; Block 2001) the critique is extended to include the quantitative approach to research that is used by many researchers. Clearly there is little that can be said to those who think that systematic, careful control of variables is impossible or meaningless in task-based research (as well, probably, as in other areas of applied linguistics). It is more relevant to accept that while there has been an emphasis on referential tasks in task research, the research has not been exclusively concerned with such task types, as Cook (2000) himself notes. Undoubtedly the emphasis reflects a belief that extraneous variables are easier to control with such tasks. One of the reasons for this is that researchers are aware of Seedhouse's point that working on tasks can be face-threatening, and that referential tasks are less likely to provoke problems in this regard.

But the crucial points here are first that most (but by no means all) task researchers are concerned with acquisition, and this does have a cognitive component: the assumption is that what is being acquired is an interlanguage system (Kasper 1997). There is no denial that there is a social reality, that the expression of meaning and communication are vital. But the research strategy adopted is that by focussing on acquisition, research will be viable and pay dividends. Second, a key issue is that a body of research is accumulating and that this body of research can point to generalisations, and findings, which seem to have emerged by following conventional research strategies. In other words, far from quantitative research in this area being impossible, a range of findings have been established which could not easily have been obtained in any other way. So it would appear that the participation of human beings in social encounters does have some predictable outcomes, and these could only be established through research with groups of subjects, with conventional statistical procedures, and through the methods researchers use to handle threats to internal and external validity (Porte, in press).

Conclusions

It is hoped that the present review has indicated that research into task-based instruction and learning, not to mention practical implementations of tasks within language teaching, demonstrate considerable vitality. There are a range of different approaches, often which

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

do not indicate complete agreement on all issues. The vitality of the field is almost certainly likely to benefit from this. In addition, researchers within the task-based area have made very considerable methodological progress within the last fifteen years or so. Standards of research have risen significantly. In particular, methods of measuring performance on tasks have expanded enormously, while showing signs of standardisation, where appropriate, to enable cross-study comparisons. There are significant new lines of research and application emerging on a regular basis, and it is to be hoped that over the next decade, the initial promise that the area has shown will be consolidated, both at a research and an application level.

References

Aston, G. (1986). Trouble-shooting in interaction with learners: The more the merrier? *Applied Linguistics*, **7**, 128–43.

Block, D. (2001). 'McCommunication': A problem in the frame for SLA. In D. Block & D. Cameron (Eds.), *Globalisation and language teaching*, Clevedon, Avon: Multilingual Matters.

Brock, C. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly*, **20**, 47–59.

Brumfit, C. & Johnson, K. (1979). *The communicative approach to language teaching*, Oxford University Press.

Bruton, A. (2002). From tasking purposes to purposing tasks. *English Language Teaching Journal*, **56** (3), 280–88.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.).

Bygate, M. (1999). Task as context for the framing, reframing, and unframing of language. *System*, **27**, 33–48.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In Bygate, Skehan & Swain (Eds.).

Bygate, M., Skehan, P. & Swain, M. (2001). (Eds.). *Researching pedagogic tasks: Second language learning, teaching, and testing*, London: Longman.

Cook, G. (2000). *Language play, language learning*, Oxford: Oxford University Press.

Coughlan, P. & Duff, P. (1994). Same task, different activities: Analysis of a second language acquisition task from an activity theory perspective. In J. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research*, Norwood, N. J.: Ablex.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, **11**, 367–83.

Cunningham, S. & Moor, P. (1999). *Cutting edge*, London: Longman.

DeKeyser, R., Robinson, P., Salaberry, R. & Harrington, M. (in press). What gets processed in processing instruction? A commentary on Bill Van Patten's 'update'. *Language Learning*.

Dornyei, Z. & Kormos, J. (2001). The role of individual and social variables in oral task performance. *Language Teaching Research*, **4**, 275–300.

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study on SL relativization. *Studies in Second Language Acquisition*, **13**, 431–69.

Doughty, C. (2001). Cognitive aspects of focus on form. In P. Robinson (Ed.).

Doughty, C. & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.).

Doughty, C. & Williams, J. (1998a). *Focus on form in classroom second language acquisition*, Cambridge: Cambridge University Press.

Doughty, C. & Williams, J. (1998b). Pedagogic choices in focus on form. In C. Doughty & J. Williams (Eds.).

Duff, P. (1986). Another look at interlanguage talk: Taking task to task. In R. Day (Ed.), *Talking to learn*, Rowley, Mass.: Newbury House.

Duff, P. (1993). Tasks and interlanguage performance: An SLA perspective. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice*, Clevedon, Avon: Multilingual Matters.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, **9**, 12–20.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, **4** (3), 193–220.

Ellis, R. (2001). Non-reciprocal tasks, comprehension, and second language acquisition. In Bygate, Skehan & Swain (Eds.).

Ellis, R. (in press). *Task-based language learning and teaching*, Oxford: Oxford University Press.

Ellis, R., Basturkmen, H. & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, **35** (3), 407–32.

Firth, A. & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal,* **81** (3), 285–300.

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, **19** (1), 1–23.

Foster, P. (2001). Lexical measures in task-based performance. Paper presented at the AAAL Conference, Vancouver, Canada.

Foster, P. & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, **18** (3), 299–324.

Foster, P. & Skehan, P. (1999). The effect of source of planning and focus of planning on task-based performance. *Language Teaching Research*, **3** (3), 185–214.

Foster, P., Tonkyn, A. & Wigglesworth, J. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, **21** (3), 354–75.

Fotos, S. (1993). Consciousness-raising and noticing through focus on form: Grammar task performance versus formal instruction. *Applied Linguistics*, **14** (4), 385–407.

Fotos, S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, **28** (2), 323–51.

Fotos, S. & Ellis, R. (1991). Communicating about grammar: A task-based approach. *TESOL Quarterly*, **25**, 605–28.

Fotos, S. (1998). Shifting the focus from forms to form in the EFL classroom. *English Language Teaching Journal*, **52** (4), 301–07.

Fried-Booth, D. (1986). *Project work*, Oxford: Oxford University Press.

Gass, S. (2002). Interactionist perspectives in Second Language Acquisition. In R. Kaplan (Ed.), *Handbook of applied linguistics*, Oxford: Oxford University Press.

Gass, S., Mackey, A., Alvarez, M. & Fernandez, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, **49** (4), 549–81.

Gass, S. & Varonis, E. (1986). Sex differences in NNS/NNS interactions. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition*, Rowley, Mass.: Newbury House.

Geddes, M. & Sturtridge, G. (1979). *Listening links*, London: Heinemann.

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

Givon, T. (1985). Function, structure, and language acquisition. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition*, Hillsdale, N. J.: Lawrence Erlbaum, 1005–1028.
Guariento, W. & Morley, J. (2001). Text and task authenticity in the EFL classroom. *English Language Teaching Journal*, **55** (4), 347–53.
Haines, S. (1989). *Projects for the EFL classroom*, London: Nelson.
Harley, B. & Swain, M. (1984). The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper & A. Howatt (Eds.), *Interlanguage*, Edinburgh: Edinburgh University Press, 291–311.
Harmer, J. (1983). *The practice of communicative language teaching* (1st Edn.), London: Longman.
Harmer, J. (2001). *The practice of communicative language teaching* (3rd Edn.), London: Longman.
Hatch, E. (1978). *Second language acquisition: A book of readings*, Rowley, Mass.: Newbury House.
Hunt, K. (1965). *Grammar structures written at three grade levels*, Champaign, Ill.: National Council of Teachers of English.
Jourdenais, R., Ota, M., Stauffer, S., Boyson, B. & Doughty, C. (1995). Does textual enhancement promote noticing? A think-aloud protocal analysis. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning*, Honolulu: University of Hawai'i Press, 183–216.
Kasper, G. (1997). 'A' stands for acquisition: A response to Firth and Wagner. *Modern Language Journal*, **81** (3), 307–12.
Kay, S. & Jones, V. (2001). *Inside out*, London: Heinemann.
Koponen, M. & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency*, Ann Arbor, Michigan: University of Michigan Press Robinson.
Krashen, S. (1985). *The Input Hypothesis*, London: Longman.
Lantolf, J. (2000). *Sociocultural theory and second language learning*, Oxford: Oxford University Press.
Leow, R. (1998). Towards operationalising the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis. *Applied Psycholinguistics*, **19**, 133–59.
Levelt, W. (1989). *Speaking: From intention to articulation*, Cambridge: Cambridge University Press.
Lightbown, P. ((2000). Classroom SLA research and second language teaching. *Applied Linguistics*, **21** (4), 431–62.
Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, **4** (2), 126–41.
Long, M. (1985a). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input and second language acquisition*, Rowley, Mass.: Newbury House.
Long, M. (1985b). A role for instruction in second language acquisition: Task-based language training. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition*, Clevedon, Avon: Multilingual Matters, 77–100.
Long, M. (1989). Task, group, and task-group interaction. *University of Hawai'i Working Papers in English as a Second Language*, **8** (2), 1–26.
Long, M. (1991). Focus on form: A design feature in language teaching methdology. In K. De Bot, R. Ginsbert & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective*, Amsterdam: John Benjamins, 39–52.
Long, M., Inagaki, S. & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, **82**, 357–71.
Long, M. & Porter, P. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, **19**, 207–28.
Long, M. & Robinson, P. (1998). Focus on form: theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom SLA*, Cambridge: Cambridge University Press.
Loschky, L. & Bley-Vroman, R. (1993). Grammar and task-based methodology. In Gass & Crookes (1993b).
Lynch, T. (1997). Nudge, nudge: Teacher interventions in task-based learner talk. *English Language Teaching Journal*, **51** (4), 317–25.
Lynch, T. (2001). Seeing what they meant: Transcribing as a route to noticing. *English Language Teaching Journal*, **55** (2), 124–32.
Lynch, T. & McLean, J. (2001). 'A case of exercising': Effects of immediate task repetition on learners' performance. In Bygate, Skehan & Swain (Eds.).
Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, **20**, 51–81.
Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, **19** (1), 37–66.
Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, **21** (4), 551–88.
Mackey, A., Gass, S. & McDonough, K. (2000). Do learners recognize implicit negative feedback as feedback? *Studies in Second Language Acquisition*, **22** (4), 471–97.
Mackey, A., Philp, J., Egi, T., Fujii, A. & Tatsumi, T. (in press). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences and instructed second language acquisition*, Amsterdam: Benjamins.
MacWhinney, B. (2000a). *The CHILDES Project: Tools for analysing talk: Volume 1: Transcription format and programs* (3rd Edn.). Mahway, N. J.: Erlbaum.
MacWhinney, B. (2000b). *The CHILDES Project: Tools for analysing talk: Volume 2: The Database* (3rd Edn.). Mahway, N. J.: Erlbaum.
Malvern, D. & Richards, B. J. (1997). A new measure of lexical diversity. In A. Ryan & A. Wray (Eds.), *Evolving models of language*, Clevedon, Avon: Multilingual Matters.
Malvern, D. & Richards, B. J. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, **19** (1), 85–104.
McLaughlin, B. (1980). Theory and research in second language learning: An emerging paradigm. *Language Learning*, **30**, 331–50.
Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, **20**, 83–108.
Mitchell, R. & Johnstone, R. (1986). The routinisation of 'communicative' methodology. In C. Brumfit (Ed.), *The practice of communicative teaching*, Oxford: Pergamon.
Morrow, K. (1977). *Techniques of evaluation for a notional syllabus*, Centre for Applied Language Studies: University of Reading.
Morrow, K. (1979). Communicative language testing: Revolution or evolution. In C. J. Brumfit & K. Johnson (Eds.).
Nakahama, Y., Tyler, A. & Van Lier, L. (2001). Negotiation of meaning in conversational and information gap activities: A comparative discourse analysis. *TESOL Quarterly*, **35** (3), 377–405.
Nemeth, N. & Kormos, J. (2001). Pragmatic aspects of task performance: The case of argumentation. *Language Teaching Research*, **5** (3), 213–40.
Newton, J. (2001). Options for vocabulary learning through communication tasks. *English Language Teaching Journal*, **55** (1), 30–37.

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press

Nicholas, H., Lightbown, P. & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, **51** (4), 719–58.

Nobuyushi, J. & Ellis, R. (1993). Focussed communication tasks and second language acquisition. *English Language Teaching Journal*, **47** (3), 203–10.

Nunan, D. (1989). *Designing tasks for the communicative classroom*, Cambridge: Cambridge University Press.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, **21**, 109–48.

Pica, T. & Doughty, C. (1985). Input and interaction in the communicative language classroom: A comparison of teacher-fronted and group activities. In S. Gass & C. Madden (Eds.), *Input and second language acquisition*, Rowley, Mass: Newbury House.

Pica, T. (1994). Research on negotiation: What does it reveal about second language learning, conditions, processes, outcomes? *Language Learning*, **44**, 493–527.

Pica, T., Kanagy, R. & Falodun, J. (1993). Choosing and using communicative tasks for second language instruction. In Gass & Crookes (1993b).

Pienemann, M. (1998). *Language processing and second language development: Processability theory*, Amsterdam: John Benjamins.

Plough, I. & Gass, S. (1993). Interlocutor and task familiarity: Effect on interactional structure. In Gass & Crookes (1993b).

Porte, G. (in press). *Appraising research in second language learning*, Amsterdam: John Benjamins.

Prabhu, N. S. (1987). *Second language pedagogy*, Oxford: Oxford University Press.

Robinson, P. (2001). Task complexity, cognitive load, and syllabus design. In P. Robinson (Ed.), *Cognition and second language instruction*, Cambridge: Cambridge University Press.

Robinson, P. (2002) Attention and memory during SLA. In C. Doughty & M. Long (Eds.), *Handbook of research in second language acquisition*, Oxford: Blackwell.

Rubdy, R. (1998). Key concepts in ELT: Task. *English Language Teaching Journal*, **52** (3), 264–65.

Samuda, V. (2001). Guiding relationships between form and meaning during task performance: The role of the teacher. In Bygate, Skehan & Swain (Eds.).

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, **11**, 17–46.

Seedhouse, P. (1997). Combining form and meaning. *English Language Teaching Journal*, **51** (4), 336–44.

Seedhouse, P. (1999). Task-based interaction. *English Language Teaching Journal*, **53** (3), 149–56.

Sheen, R. (1993). Double standards in research selection and evaluation: The case of comparison of methods, instruction, and task work. *RELC Journal*, **24** (2), 1–28.

Sheen, R. (1994). A critical analysis of the advocacy of the task-based syllabus. *TESOL Quarterly*, **28** (1), 127–51.

Sheen, R. (2002). 'Focus on form' and 'focus on forms'. *English Language Teaching Journal*, **56** (3), 303–05.

Shehadeh, A. (2001). Self and other initiated modified output during task-based interaction. *TESOL Quarterly*, **35** (3), 433–57.

Skehan, P. (1998). *A cognitive approach to language learning*, Oxford: Oxford University Press.

Skehan, P. & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task based learning. *Language Teaching Research*, **1** (3).

Skehan, P. & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, **49** (1), 93–120.

Skehan, P. & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.).

Skehan, P. & Foster, P. (2002, ms). The effects of post-task activities on the accuracy of language during task performance. King's College, London.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & Seidlehofer (Eds.), *Principle and practice in applied linguistics*, Oxford: Oxford University Press, 245–56.

Swain, M. & Lapkin, S. (1982). *Evaluating bilingual education: A Canadian case study*, Clevedon, Avon: Multilingual Matters.

Swain, M. & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In Bygate, Skehan & Swain (Eds.).

Thornbury, S. (1997). Reformulation and reconstruction: Tasks that promote 'noticing'. *English Language Teaching Journal*, **51** (4), 326–35.

Trahey, M. & White, L. (1993). Positive evidence in the second language classroom. *Studies in Second Language Acquisition*, **15** (2), 181–204.

Van Lier, L. & Matsuo, N. (2000). Varieties of conversational experience: Looking for learning opportunities. *Applied Language Learning*, **11**, (2), 265–87.

Van Patten, B. (1996). *Input processing and grammar instruction*, New York: Ablex.

Van Patten, B. (in press). Processing instruction: An update. *Language Learning*.

Wendel, J. (1997). *Planning and second language narrative production*, Unpublished doctoral dissertation, Temple University, Japan.

Wesche, M. & Skehan, P. (2002). Communicative teaching, content-based instruction, and task-based learning. In R. Kaplan (Ed.), *Handbook of applied linguistics*, Oxford: Oxford University Press.

White, R. (1988). *The ELT curriculum*, Oxford: Blackwell.

Widdowson, H. G. (1978). *Teaching language as communication*, Oxford: Oxford University Press.

Widdowson, H. G. (1979). *Explorations in applied linguistics, Vol.2*, Oxford: Oxford University Press.

Wigglesworth, J. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, **14** (1), 85–106.

Wigglesworth, J. (2001). Influences on performance in task-based oral assessments. In Bygate, Skehan & Swain (Eds.).

Williams, J. (2001). Learner-generated attention to form. *Language Learning*, **51**, 303–46.

Willis, J. (1996). *A framework for task-based learning*, London: Longman.

Willis, D. & Willis, J. (1988). *COBUILD Book 1*, London: Collins.

Yuan, F. & Ellis, R. (in press). The effects of pre-task planning and on-line planning on fluency, accuracy, and complexity in L2 oral production. *Applied Linguistics*, **24**, 1.

Yule, G. (1997). *Referential communication tasks*, New York: Erlbaum.

Yule, G. & McDonald, D. (1990). Resolving referential conflicts in L2 interaction: The effect of proficiency and interactive role. *Language Learning*, **40**, 539–56.

Yule, G. & Powers, M. (1994). Investigating the communicative outcomes of task-based interaction. *System*, **22** (1), 81–91.

https://doi.org/10.1017/S026144480200188X Published online by Cambridge University Press