Lang. Teach. (2009), **42:1**, 109–130 © Cambridge University Press
doi:10.1017/S0261444808005442

Replication Studies

Negotiated interaction in the L2 classroom

Johannes Eckerth King's College London, UK
johannes.eckerth@kcl.ac.uk

The present paper reports on an approximate replication of Foster's (1998) study on the negotiation of meaning. Foster investigated the interactional adjustments produced by L2 English learners working on different types of language learning tasks in a classroom setting. The replication study duplicates the methods of data collection and data analysis of the original study, but alters the target language (L2 German) and adds a stimulated recall methodology. The results of the replication study partially confirm Foster's results, and introduce some further differentiated findings. It is concluded that the original study's concern with the transferability of laboratory findings to classroom settings should be investigated in greater detail.

1. Introduction

Ever since the basic assumption endorsed in the seminal work of Hatch in the late 1970s that learners learn the structure of a language through interaction rather than learning grammar in order to interact (Hatch 1978), the relationship between interaction and acquisition has been one of the core issues in second language acquisition (SLA) research. A central claim resulting from this research is that, though interaction may not be strictly necessary, it nevertheless constitutes the primary means by which language learners obtain data for language learning, not only because interaction is how most learners receive input, but also because the input obtained through interaction is more conducive to acquisition than input received in other ways. Within SLA research, input obtained via interaction has been conceptualized and researched in terms of 'comprehensible input' (Krashen 1981), 'negotiation of meaning' (Long 1983), and 'comprehensible output' (Swain 1985, 1995, 2005). The interrelatedness of these three notions is concisely articulated in Long's revised version of the Interaction Hypothesis:

> I would like to suggest that negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways. (Long 1996: 451–452)

Given the assumed acquisitional potential of negotiated interaction, the question arises as to how to best create a learning environment in which negotiation of meaning and interactional adjustments can occur. From the middle of the 1980s onwards, a large number of studies

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

have been conducted which observe language learners working on specially designed tasks. These 'referential communication tasks', as they have been called (see Yule 1997 and Pica, Kang & Sauro 2006 for an overview of their use in L2 research), have been classified along several task variables (see Pica, Kanagy & Falodun 1993) such as information distribution (information shared among vs. split between interactants) and information flow between interactants (one-way vs. two-way). Typically, then, such tasks have been investigated as to how they induced L2 production, the negotiation of meaning, and/or the modification of output (see Pica 1994; Gass 1997; Gass & Mackey 2006, 2007; Mackey 2007 for research overviews).

Although this strand of research is primarily concerned with the cognitive and interactive conditions of language acquisition, such studies on negotiated interaction also claim to be directly or indirectly relevant to SLA and L2 pedagogy. Tasks, although they were developed for research, are claimed to be of immediate applicability as a teaching tool in the L2 classroom (e.g. Long & Porter 1985; Long & Crookes 1987; Pica et al. 1996; Pica 2005; Eckerth 2006). However, these studies have typically not been conducted in classrooms:

> The bulk of task-based research in SLA has taken place under laboratory conditions rather than in actual classrooms. A question of some interest to SLA researchers interested in the pedagogical application of their research is whether the results they have obtained can be replicated under normal classroom conditions. Task evaluations conducted by teachers, then, can serve as a way of testing the transferability of research findings. (Ellis 1997: 230)

What Ellis stated over ten years ago still reflects the current tendency (see Gass, Mackey & Ross-Feldman 2005). Given the immense productivity of task-based research – see the research overviews in book length (Ellis 2003; Long forthcoming) as well as several edited volumes (Bygate, Skehan & Swain 2001; Leaver & Willis 2004; Müller-Hartmann & Schocker-von Ditfurth 2004; Edwards & Willis 2005; Ellis 2005; van den Branden 2006; García Mayo 2007; van den Branden, van Gorp & Verhelst 2007; Eckerth & Siekmann 2008; Samuda & Bygate 2008), and the ever-growing popularity of task-based language pedagogy (see Lee 2000; Nunan 2004; Swan 2005), the question of the transferability of research results to the language classroom is a pressing one. The issue at hand is not the external validity or even the legitimacy of experimental and laboratory L2 research per se (see Hulstijn 1997 for the possibilities and limitations of experimental SLA research). Rather, the critical question is whether an approach to L2 pedagogy which claims to be firmly founded in and closely related to research findings is not at risk of losing some of its credibility, if this research does not attempt to demonstrate its validity in laboratory AND in classroom settings. One way to do this is to attempt to reproduce the results of a laboratory study in the L2 classroom, an approach chosen by Foster (1998). Another approach is to use the same tasks in both settings and in one and the same investigation, as recently exemplified by Gass et al. (2005). As the Gass et al. (2005) study was in part a reaction to Foster (1998), in the following I shall discuss the latter before focusing in the main part of the paper on the former.

Gass et al. (2005) conducted a comparative same-methodology/different-settings investigation of learners' task-based interaction. Three different tasks were administered in the two settings. The subjects were 74 native speakers of English who were enrolled in third-semester university-level Spanish courses. In one of the three dyadic tasks, information exchange was optional – on the basis of shared information about several universities, pairs

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

of students had to decide which university they would recommend a friend to choose. In the other two tasks, information exchange was obligatory – the tasks involved spotting differences in two similar pictures, one of which was held by each student, and drawing a driving route on a shared map on the basis of different pieces of information held by each of the students in the pair. Intact classes were randomly assigned to either classroom or laboratory settings. Whereas 44 students completed all three tasks during their regular class time with their regular instructor, 30 students from the other two classes accomplished the tasks in a laboratory setting with a research assistant. Learners' interactions in both settings were audio-taped and coded for negotiation of meaning, language-related episodes, and recasts. The two guiding research questions were: (i) In what way does task-based interaction in the classroom compare to task-based interaction in a laboratory setting? and (ii) How do different tasks influence learners' interactions in both settings? In order to answer these questions, two-factorial repeated-measures analyses of variance (ANOVAs) with the factors 'task' and 'setting' and the dependent variable 'learners' interactions' were conducted. Overall results showed no significant interaction between the setting and task factors, and no significant differences between the two settings. Significant differences, however, were reported between the different tasks in terms of all three measurements, i.e. amount of meaning negotiation, language-related episodes, and recasts. Thus, the authors conclude, although the setting had no influence on the frequency of the three interactional patterns under investigation, the type of task did (Gass et al. 2005: 596–597).

The Gass et al. study was partly motivated by Foster's (1998) claim that her own results 'call into question the typicality of previous research into the incidence of negotiation of meaning and the justification therefore for constructing an SLA theory upon it' (Foster 1998: 19 as cited in Gass et al. 2005: 578). The indirect reference here is to Pica et al. (1989). Foster argues explicitly against the use of inferential statistical analyses of the quantitative results employed in the Pica et al. (1989) study; first because 'complex statistical computations obscure what is happening at an individual level' (Foster 1998: 8), and second because the purpose of her own investigation 'was not to test a hypothesis but to observe individual students' classroom performance' (ibid.). Whereas Pica et al. inferred from their results gained in a laboratory setting that negotiation of meaning and output modification were 'alive and well' (Pica et al. 1989: 83), Foster concludes from her own study 'that uncoached negotiation for meaning is not "alive and well" in the classroom, and, given the minute number of syntactically modified utterances, is much too fragile to bear the weight of the SLA theory that researchers like Pica have built upon it' (Foster 1998: 19).

Given the inconsistent empirical results achieved by Gass et al. (2005) and Foster (1998), the author of the present paper decided to conduct a replication of Foster's (1998) original study. This decision was motivated *inter alia* by the huge impact of Foster's study in the field. Not only does Foster's 1998 paper rank among the 50 most frequently read articles that have been published in the refereed journal in question since its first issue in 1980 (see www.applij.oxfordjournals.org), but it has also been frequently cited and discussed in major books (e.g. Skehan 1998; Ellis 2003) and papers (e.g. Gass et al. 2005; Seedhouse 2005; Swan 2005).

The goal of the present replication study was twofold. First, it was set up to confirm or otherwise the RESULTS of Foster's (1998) study. Second, it attempted to shed more light on

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

the validity of Foster's INTERPRETATION of these results. In particular, the replication sought to address the question whether it is warranted to consider Foster's results as a function of the classroom setting in which the data was collected. It furthermore attempted to explore whether, in so far as Foster's findings diverge from results achieved in the laboratory, Foster's results evidence a lack of transferability of experimental research findings. The distinction between the results and the conclusions derived from those results was informed by the findings of previous and current research into task-based interaction.

With respect to RESULTS, a great number of studies investigating students' interactions when working on information-gap tasks in a controlled environment have reported high frequencies of meaning negotiation. Among them have been studies that included monitored sessions outside the classroom (Long 1985; Pica & Doughty 1987; Pica et al. 1989, 1991, 1996), special researcher visits to actual classrooms (Pica 1991), and breaks during regular class meetings (Doughty & Pica 1986). All these studies used different kinds of information-gap tasks and were conducted at different places with differing participants by different researchers. Thus, it is noteworthy that all of these studies report similar overall frequencies of meaning negotiation. These results, then, can be taken to show the robustness of meaning negotiation based on information-gap tasks across participants, target languages, and level of L2 proficiency. Given such a robustness of the interactional pattern across experimental settings, it appears to be highly appropriate to investigate task-based meaning negotiation under non-experimental classroom conditions.

With regard to the INTERPRETATION of these results, at least three issues have to be considered, the first referring to research settings, and the other two related to research methodology. First, as has been pointed out by Gass et al. (2005: 579), 'claims about research in classroom contexts require us to decide what is meant by a classroom'. Obviously, types of classrooms – status and function of the L2, intensity, characteristics, goals, purposes of instruction, etc. – can differ to a large degree, as can different laboratory settings and the perception of these settings by different learners. In other words, if the participants believe that a certain research IS simply part of their teaching program, then it is a classroom happening in the relevant sense. Rather than assuming characteristics of a certain setting, such an assumption should be verified against learners' perceptions of the setting. Second, there is growing empirical evidence that learners' task-based interactional behavior cannot be explained exclusively by the variable 'task type'. Using transcript-based and individualized post-hoc tests, some studies (Swain & Lapkin 1995, 1998, 2001; Williams 2001; Loewen 2002, 2005; Eckerth 2008a) have shown how learners exploit learning opportunities beyond the pedagogic focus of the task, thus adapting the task to their individual learning needs and goals. In the same vein, several studies have shown how different learning activities can result from the same task when completed by different learners (Coughlan & Duff 1994; Gourlay 2005) and how these learning activities might be contingent on learners' perceptions of the task (Slimani-Rolls 2005). In light of these findings, task-based research is slowly shifting from the conceptualization of the L2 learner as a mere 'task executioner' to a 'task interpreter' (Breen 1987; Ellis 2003; Mondada & Pekarek Doehler 2004; Lantolf & Thorne 2006; Eckerth 2008b). Third, there is a growing awareness in the field that a straightforward identification of classes of interactional moves on the part of the researcher may or may not conform with the communicative intentions of the interactants. Thus, whereas a speaker's repetition of an

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

interlocutor's previous utterance with rising intonation has typically been labeled 'clarification request' (e.g. Long & Porter 1985; Pica et al. 1989; Pica 1994; Gass 1997; Gass et al. 2005), it is not always clear whether the speaker's intention actually had been to clarify a lack of understanding (see Firth & Wagner 1997; Hauser 2005).

For these reasons, and in terms of data collection, a stimulated recall methodology was added to the present replication study. On the one hand, it was assumed that this procedure would not interfere with the main data collection instrument, i.e. the administration and observation of different tasks in the classroom (see below). On the other hand, it was expected that the triangulation of interactional data with the interactants' accounts of their own performance could help to better understand and interpret the results of Foster's original study. It was hypothesized, for instance, that it was not the difference in setting per se which caused the incongruent results in the studies of Foster (1998) and Gass et al. (2005), but rather the learners' perception of these settings (issue 1 above) as well as differences in research methodology (issues 2 and 3 above). Investigating students' perceptions of their task work was judged as a suitable means to bridge both interpretative issues, the research setting issue and the research methodology issue. With respect to the latter, it was expected that students' accounts of their task performance would provide a check of the researcher's interpretative accuracy. For instance, Foster suggests that students may perceive the laboratory as a more formal setting, where the emphasis is rather on accuracy, and thus may apply a 'check and clarify' strategy (Foster 1998: 19). Accordingly, the replication study attempts to explore whether students may perceive the classroom as a more relaxed setting, thus focusing rather on fluency and applying a 'let it pass' strategy (Firth 1996: 243). Consequently, empirical evidence supporting such an assumption of differing student perceptions could explain some of the frequency differences in input negotiation and output modification between the two studies. Therefore, the approximate replication reported in the remainder of this paper closely duplicates the methods of data collection and data analysis of Foster's original study, while altering the target language (L2 German).

2. The original research study

Foster's original study was set up to see '"what the student in the classroom does" with the negotiation of meaning' (Foster 1998: 5). For this purpose, lower intermediate ESL learners in an actual classroom were observed while they were working in small groups and in pairs on different language learning tasks. The study sought to investigate to what extent the learners would produce a) talk in general, b) comprehensible input, and c) modified output, and whether the variables 'task type' (optional vs. required information exchange) and 'participant structure' (group vs. pair work) would affect a), b), and c).

3. The replication study

The replication study closely followed the research procedures adopted in the original study. As will be specified in the following sections, the relevant parameters such as participants

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

(3.1), setting (3.2), tasks (3.3), data collection procedures (3.4) and data coding (3.5) were identical or closely comparable to those in the original research, whereas the stimulated recall methodology (3.6) has been added to the research design.

3.1 The participants

While the original study was conducted with ESL learners at a municipal college in Great Britain, participants in the replication study were students in an L2 German university course. The class met four times a week for approximately three hours a day as compared to three times a week for two hours a day in Foster's study. With the exception of some L1 speakers of various Slavic languages, the students came from similar L1 backgrounds as in the original study (Arabic, French, Korean, Spanish) and from a similar range of ages (20–42 years compared to 17–41 years) with an average age of 23 years (as compared to 21). As in Foster's study, they were of lower intermediate L2 proficiency and had been assigned to the class based on a written test and an interview. While the ESL learners consisted of typical 'part-time learners in colleges throughout Britain' (Foster 1998: 5), the L2 German learners can be characterized as typical college and university applicants in Germany. As in the original study, the class had been running for eight weeks at the time of data collection. Due to erratic attendance, 12 out of the 16 students in the class were observed in this study, most of them female, as was the case in the original study.

3.2 The setting

In line with the original study, the setting of the classroom was preserved as far as possible. Therefore, all recordings were made during three scheduled lessons. In the case of the replication study, these recordings were part of a more extended sequence of data collection for the purpose of a larger research project (see Eckerth 2008c). At the time of the recordings the students had become fairly accustomed to the presence of the researcher and the recording devices. The researcher had been introduced and occasionally acted as a co-teacher, unlike in Foster's study, where researcher and teacher were identical. As in the original study, on each of the data collection occasions the students were asked to do one task which was presented as part of the normal class routine. They all did the task at the same time and in the same room, as is normal classroom practice.

3.3 The tasks

In Foster's study, the two variables 'task type' (optional vs. required information exchange) and 'participant structure' (small group vs. pair work) were systematically combined, exemplified by four tasks with varying 'task foci' (three linguistically unfocused tasks, one linguistically focused consciousness-raising task). The replication study did not vary the participation pattern, but rather had all tasks completed in pairs of learners. Such a departure from the

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

original study has been motivated by two factors. First, research on the effect of participant structure on task-based interactional behavior is fairly consistent in showing a more balanced distribution of turns and negotiation moves when compared to small groups (e.g. Pica 1987, 1994; Gass 1997). Second, a particular interest of the replication study was to investigate Foster's claim that her results were due to the nature of the classroom as opposed to a laboratory setting (Foster 1989: 19). Thus, the replication study had to be relatable to relevant research that addressed the transferability issue either directly (Gass et al. 2005) or indirectly (Slimani-Rolls 2005). As both of these studies exclusively employed dyadic tasks, the use of dyadic tasks was felt to enhance the potential of relating the results of the replication study to the current discussion. However, the replica maintained the variation and comparison of 'task type' and 'task focus' of the original study.

Three tasks were chosen for the replication study, two linguistically unfocused referential tasks and one linguistically focused consciousness-raising task. In contrast to Foster's study, the tasks were not taken from actual textbooks, but had been developed by the researcher. In the case of the grammar-based task, for instance, this decision was guided by the interest of shifting the focus of the task more towards meaning and meaning negotiation, as opposed to a rather mechanical manipulation of L2 forms which is typical of Foster's (1998: 6) grammar-based task. With respect to the two referential communication tasks, one was identical in structure to Foster's 'consensus task' and the other was a variation of her 'picture difference' task (Foster 1998: 6). As picture difference tasks had been frequently used in the classroom under investigation, such a variation has been introduced in order to keep students motivated and interested. Thus, as the tasks developed for the purpose of the replication study closely corresponded to Foster's tasks as well as to typical textbook tasks of the respective type, such a procedure was not expected to affect the comparability of the two studies. In order to ensure their authenticity and their face value as judged by the students, the tasks were didactically embedded in the curriculum through extensive discussion with the teacher cooperating in the investigation. Both the researcher and the teacher had the impression that the tasks were perceived by the learners as part of the regular language study program. The tasks recorded for the replica (see Appendices A–C in the electronic supplement) were as follows:

1. Consciousness-raising task 'text repair'

   Students were required to work on a written L2 text (e.g. a farewell letter) with missing grammatical information. Completing a text repair task involved inserting items such as articles, reflexive pronouns, or verbal complements, and inflecting pronouns and verbs appropriately. Students working on this text in pairs had to produce a written version of the text that is grammatically correct, semantically coherent, and pragmatically appropriate. Though this task focuses on specific L2 features (reflexive prepositional verbs), it does not confront learners with isolated, decontextualized linguistic forms, thus allowing them to make use of their own language resources. However, as there is no obligation to exchange information, this task is, like the 'grammar-based task' in Foster's (1998: 6) original study, classified as optional information exchange.

2. Assembly task 'women'

   Students held sets of numbered pictures showing different women. Through the detailed description of a particular woman by one learner, the interlocutor had to identify this

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

woman on his/her own worksheet. When in doubt about the woman's identity, the interlocutor was supposed to inquire about specific features of the person in question. This could only be done by students sharing individually-held information and is classified – like the 'picture difference' task in Foster's (1998: 6) original study – as required information exchange task.

3. Decision-making task 'desert'

   Students were provided with one worksheet showing a list of 16 equipment items available for a planned desert trip. They were required to argue for and to agree on a selection of the most useful eight items to be taken on the trip. Once they had reached a consensus on which items to select, they received a further piece of information which set them a new problem to solve. They again had to reach a consensus, which entailed further problems and choices until the task ended. As all the information was available to both students, this task is, as the 'consensus' task in Foster's (1998: 6) original study, classified as optional information exchange task.

3.4 Data collection

As in the original study, recordings were made at weekly intervals during scheduled classes. The tasks were all part of the normal lesson plan and were not presented as being in any way 'special'. In contrast to the original study, an external microphone was connected to the tape recorder, ensuring high sound quality. No tapes had to be discarded from the data analysis due to unintelligible speech, as reported by Foster (1998: 7). This resulted in five recorded dyads per task, in contrast to the three dyads and groups per task in Foster's study. Foster transcribed the first five minutes of three tasks and the first ten minutes of the remaining task, halving the scores from the ten-minute recording to enable comparison. In the replication study, the first ten minutes of each interaction were transcribed and coded for data analysis. Thus, while replicating the original data collection procedure, the amount of data transcribed and coded was larger both in terms of number of interacting dyads and amount of interactional time. The decision to quantitatively alter the original study's data collection procedure was taken in the interest of assembling a larger, that is, more representative corpus (2175 vs. 918 c-units in the original study; see below) as well as reducing potential biases from the recording of untypical 'warming-up' sequences (Aston 1986: 132 as cited in Gass et al. 2005: 580). As in Foster's study, some learners were observed for all three tasks, but, because of erratic attendance, some were not. Table 1 shows the composition of learner dyads and their distribution across tasks.

3.5 Transcription and coding

Identical with the original study (Foster 1998: 7), the 'c-unit' was used to measure students' speech production. Foster, Tonkyn & Wigglesworth (2000) have suggested the 'A-unit' as better suited for the measurement of spoken language and have, rather atypically, attempted to operationalize the unit in a detailed way. However, in order to ensure comparability, the replication study used the same coding procedures as the original investigation. Thus,

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Table 1 Recorded learner dyads across tasks.

Learner dyad	Consciousness-raising task	Referential tasks	
		optional info-exchange	required info-exchange
Nino & Macit	×	×	×
Esin & Tamuna	×	×	×
Olga & Yusuf	×	×	×
Prerana & Ismail	×		
Sanie & Odeta	×		
Nancy & Ismail		×	×
Sanie & Fatih		×	×

all data was counted for c-units, defined as 'utterances, for example, words, phrases, and sentences, grammatical and ungrammatical, which provide referential or pragmatic meaning to NS–NNS interaction' (Foster 1998: 8, referring to Pica et al. 1989 and Brock 1986). Thus, whereas false starts would be excluded from analysis, elliptical constructions, which abound in oral language and are meaningful and pragmatically sufficient, would be counted. Along with the original study, negotiated interaction was coded for CLARIFICATION REQUESTS (a request for further information from an interlocutor about a previous utterance), COMPREHENSION CHECKS (the speaker's query of the interlocutor(s) as to whether or not they have understood the previous speaker utterance(s), and CONFIRMATION CHECK (the speaker's query as to whether or not the speaker's (expressed) understanding of the interlocutor's meaning is correct) (Foster 1998: 8, referring to definitions from Chaudron 1988: 45). Also in accordance with the original study and in order to measure modified output, the transcripts were coded using the definitions from Pica et al. (1989: 88) for SEMANTIC MODIFICATION (through synonym, paraphrase, or example), MORPHOLOGICAL MODIFICATION (through addition, substitution, or deletion of inflectional morphemes and/or functors), PHONOLOGICAL MODIFICATION, and SYNTACTIC MODIFICATION (through embedding and elaboration in clauses).

3.6 Stimulated recall

For reasons stated earlier in this paper, a stimulated recall methodology was included in the design of the replication study. Data was collected in the form of interviews with pairs of learners who had jointly worked on a task the day before. After class, students were accompanied to the researcher's office, where the purpose of the interview was explained to them. They were assured that the interview would have no effect on their course grades and that all information would be held confidential (see section 3.6 of the e-supplement for more details).

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

4. Results

As the purpose of the investigation was to observe task-based language learner performance in an actual L2 classroom rather than to test pre-specified hypotheses, both the original and the replication study left the coded data in the form of simple totals and percentages. However, although Foster (1998: 8) explicitly states her intention to 'observe individual students' classroom performance', no effort has been made to keep the composition of the groups and dyads constant. As Foster (1998: 8) claims, this was not possible 'without compromising the naturalistic setting of the study'. In the present study, however, it was generally the same students working together on the different tasks. This procedure was not a condition of the research project imposed onto the classroom, but rather accorded with the established and common classroom practice of working with one's neighbor. It allowed for a performance comparison across learners as well as across tasks for that subset of students who completed all three tasks.

In line with the structure of Foster's paper, the data will be presented according to the research questions posed earlier: language production (Tables 2 and 3 in the e-supplement, section 4.1), comprehensible input (Tables 4 and 5 in the e-supplement, section 4.2) and modified output (Tables 6 and 7 in the e-supplement, section 4.3). Within each of these three categories, the first table specifies the performance of the dyads, and the second table breaks these figures down as to individual learner performance. Assigning negotiation moves to individual students was not based on the assumption that the initiators of an input-negotiation move or an output-modification move would benefit more (or less) from this move than their interlocutors. Rather than the assumption of unequal benefits in terms of language acquisition, the guiding interest was to see in what way certain tasks varied in stimulating a more or less equal distribution of speech production, negotiation and modification within dyads.

4.1 Language production

Language production was analyzed as a function of task type and individual student. First, c-units were calculated to measure the amount of language produced by each dyad (Table 2 in the e-supplement). A cross-task comparison reveals clearly less speech production for the referential task with required information exchange as compared to the referential task with optional information exchange. As can be seen in Table 2 in the e-supplement, this does not only hold true in terms of the total of c-units produced by all dyads, but mostly also in terms of individual dyads. Such a result is in conflict with Foster's scores, which show the same ratio for the groups, but the reverse for the dyads. Furthermore, compared across dyads and consistent with Foster's study (1998: 9), the overall amount of language production varies widely. For example, in the consciousness-raising task, the c-units produced by different dyads vary between 171 and 89, and in the referential task with required information exchange, between 171 and 96. A closer look at the performance of different dyads also reveals that in both types of referential tasks it is always the same dyad who talks most (Esin–Tamuna)

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

or least (Nancy–Ismail). Also instructive in this regard is Table 3 in the e-supplement, which reports the individual variation of speech production within the dyads.

In contrast to a small-group constellation, a dyad makes it hard to completely drop out of the verbal interaction. Accordingly, there are no instances when a learner does not produce any or hardly any language, as reported by Foster (1998: 10). However, the individual students in a dyad contribute varying proportions of the amount of language produced by the dyad as a whole. In contrast to the student Yusuf, who produces less than half the amount of language his interlocutor does (49%; Olga & Yusuf, Table 3 in the e-supplement), the distribution of c-units between Nino and Macit or Nancy and Ismail is rather balanced (92% for both dyads; Table 3 in the e-supplement). Again, a comparison across learners is informing. Looking at the dyads which have carried out two or three different tasks clearly shows that if one learner verbally dominates his partner, this learner dominates not only in one, but in all tasks.[1] The final section will consider in what way these results reflect the relationship between particular task features and individual task performers.

4.2 Comprehensible input

Negotiation of meaning was measured by calculating the number of negotiation moves (confirmation checks, clarification requests, comprehension checks; examples are provided in the e-supplement, section 4.2, Transcripts 1–3) across tasks and dyads. Table 4 in the e-supplement shows both the total number of negotiation moves and their percentages when compared to the number of c-units produced by a dyad. As a whole, the corpus includes 123 negotiation moves with all the students producing at least a few negotiation moves. Foster's study shows a lot of students not negotiating at all, particularly when working in a group rather than a dyad (Foster 1998: 13). Thus, the factor of participation pattern seems to have an impact not only on the amount of language production but also on the amount of meaning negotiation. When it comes to the impact of task type, in line with Foster's results, negotiation was relatively more frequent in the two referential tasks. Negotiation was also the most frequent in the task that required information exchange. This applies bothto the average percentages of all dyads as it does for the single dyads. This picture changes somewhat when looking at the relative distribution of negotiation moves within the dyads (Table 5 in the e-supplement).

Although one might assume that the split distribution of information among both learners would result in a mutual responsibility for information exchange and therefore in a balanced share of meaning negotiation, this is not the case. When comparing the ratio of negotiation moves produced by the two learners of a dyad, the two tasks with shared information distribution and optional information exchange produce higher average percentages (44% and 28%, Table 5 in the e-supplement). However, a comparison across learners reveals that if learners dominate their interlocutors in terms of negotiation moves, they do this not only in one task, but in all tasks. This finding is consistent with the pattern of verbal dominance with regard to the distribution of language production within dyads (Table 3 in the e-supplement).

[1] 'Verbal domination' has been operationalized purely quantitatively as producing a higher amount of the linguistic unit under investigation (c-units, input negotiation moves, output modification moves) than one's interlocutor.

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Table 2 Absolute and relative overall frequency of negotiated input and modified output.

Study	Total speech production, c-units	Negotiated input moves	Modified output moves	Ratio c-units : neg. input	Ratio neg. input : mod. output
	n	n	n	%	%
Foster	918	87	20	9.5	23
Eckerth	2175	123	32	5.7	26

4.3 Modified output

Modified output, that is, utterances which were semantically, morphologically, syntactically, or phonologically altered in response to a negotiation move, were calculated across tasks and dyads (see the e-supplement, section 4.3, Transcripts 4–7). Table 6 in the e-supplement shows both the absolute and relative number of negotiation moves produced by a dyad.

In relation to the overall amount of language production, the extent of output modification is rather limited, as it is also in Foster's data. A total of 123 negotiation moves (Table 4 in the e-supplement) is followed by an output modification in 32 cases. As Foster discovered (Foster 1998: 15), no consistent pattern can be found when comparing tasks with optional as opposed to tasks with required information exchange. More revealing is the comparison across learners. Almost half of all output modifications are produced by one dyad, Esin & Tamuna. Furthermore, relating output modification to input negotiation, the students in this dyad are also found to be the most active input negotiators (Table 4 in the e-supplement). As with language production and input negotiation, modified output has also been measured with respect to the relative distribution of output modifications within the dyads (see Table 7 in the e-supplement)

The dyad-internal distribution of modified output (Table 7 in the e-supplement) is consistent with the distribution of negotiation moves (Table 5 in the e-supplement) in two ways. First, a comparison across tasks shows again the two tasks with optional information exchange to produce a more balanced distribution of output modification. Second, a comparison across learners reveals that if learners dominate their interlocutors in terms of output modification, they do this not only in one task, but in all tasks. Again, this is consistent with the distributional pattern of language production (Table 3 in the e-supplement) and input negotiation (Table 5 in the e-supplement) within dyads. Table 2 summarizes the overall frequency of meaning negotiation and output modification in Foster's (1998) original and in the replication study.

5. Discussion

As the main purpose of a replication study is to better understand the results, outcomes, and implications of the replicated study, it will be helpful to summarize Foster's (1998: 17ff.) findings briefly. First, based on the total scores for groups and dyads in relation to task type,

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Foster states (i) no recognizable pattern relating task type to language production, (ii) the most consistent occurrence of meaning negotiation during pair-work on tasks with required information exchange, and (iii) the most frequent occurrence of output modification when working in a dyad setting independent of task type. Second, based on the individual scores, Foster reports that (i) the amount of meaning negotiation and output modification per student was rather low, and (ii) the range of scores between individual students was extremely wide. Third, she explains these results with reference to (i) the potentially painstaking, frustrating, and face-threatening nature of extensive negotiation and modification, resulting in a 'pretend and hope' rather than in a 'check and clarify' strategy, (ii) the classroom environment as opposed to a laboratory setting with its tightly designed information-gap tasks, and (iii) she calls for tasks which do allow and encourage the reflection and negotiation of form, rather than absorbing learners' cognitive resources by the need to exchange information most efficiently.

The present study did not compare group and dyad settings. However, in comparison to Foster's study, it found a more consistent distribution of language production (Foster: Table 2 and 3, pages 9–10; present paper: Table 2 and 3 in the e-supplement) and a more consistent occurrence of meaning negotiation (Foster: Table 4, page 12, and Tables 5 and 6, page 13; present paper: Tables 4 and 5 in the e-supplement). This outcome indirectly supports the assumption that a dyadic setting is both more conducive to a balanced distribution of speech and negotiation. In a pair, as opposed to a small group, it is much harder for a single student to retreat or to drop out of the communication.

Neither the original nor the replication study could confirm the overriding effect of task type (required vs. optional information exchange) on the amount of language production and meaning negotiation established by former studies (e.g. Pica & Doughty 1985; Doughty & Pica 1986). Foster explains her results as a consequence of the learners' adaptation of the tasks, resulting in a strategy that 'could reduce some information exchange tasks to a format whereby the side holding the information need only answer yes or no to the informed guesses of the other side' (Foster 1998: 11). With regard to dyadic task completion, such task adaptation strategies are unequivocally confirmed by the findings of the replication study (see Transcript 11 in the e-supplement).

The evidence of task adaptation points to the necessity of investigating individual learners' actual task-based interactions more closely. 'Across the data as a whole', Foster (1998: 18) summarizes her results, 'the range in individual scores is so wide, and the lack of participation by so many students is so striking as to make any statistics based on group totals very misleading'. With respect to the dyadic tasks, this outcome is consistent with the results of the present replication study. It was consistently the same dyad (Esin & Tamuna) who talked most (Table 2 in the e-supplement), negotiated most (Table 4 in the e-supplement), and modified most (Table 6 in the e-supplement), rather independently of task-type. Furthermore, the replication study allowed not only for the comparison of dyads, but also for the comparison of individual students across tasks. The present study shows that if one learner dominated one dyadic exchange with respect to one of the three investigated parameters – language production, negotiation of input, modification of output – s/he would dominate all exchanges with respect to this very parameter (Tables 3, 5 and 7 in the e-supplement). Moreover, there

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

is a strong tendency for the learner who verbally dominates with respect to one parameter, to dominate with respect to all three parameters (Tables 3, 5 and 7 in the e-supplement).

Variation of meaning-negotiation across dyads and individual learners is also reported by Slimani-Rolls (2005). While 60% of the negotiation sequences in her study come from only 25% of the learners, another 25% of the learners do not contribute at all (ibid.: 202). Gass et al. (2005), however, do not show asymmetrical distributional patterns of meaning negotiation across dyads (ibid.: 598).[2] Neither variation of meaning negotiation between single learners and dyads nor variation between settings is confirmed by the results reported by Gass et al. (2005). The question of individual variance and the related question of transferability of experimental research findings to classroom settings deserve therefore more extended discussion. It may well be the case that the consistency of Gass et al.'s results across individuals, across dyads, and across settings is a function of the characteristics of the settings investigated, and the research methodology, i.e. the instruments of measurement used.

As has been mentioned earlier, Foster (1998) accounts for her main results – comparatively low frequency of interactional adjustments combined with marked variation across subjects – by referring to the very nature of task-based learner–learner interaction in the classroom, as opposed to task-work in the laboratory. She considers that task-based learner–learner interaction in a classroom might be seen by the students as a rather informal activity taking place in the sheltered intimacy of a pair or a group, and with the focus on practicing the language and maintaining the smooth flow of conversation rather than on accuracy, self-monitoring, and learning. Thus, if students regard pair work as a 'light-hearted and informal part of class, rather than a pedagogical activity specifically designed to promote SLA', Foster (1998: 19) claims, 'we cannot be surprised if they are relaxed enough about communication problems to let them pass, thereby missing opportunities to gain comprehensible input and to create modified output'. However, Foster (1998) provides no specific evidence to support her claim that it was the classroom environment which stimulated the learners in her study 'to let them [communication problems] pass'. In the present study, combined and related evidence from task completion and stimulated recall data (see e-supplement, Transcripts 12–15) exemplifies how students may deal with communication problems during task-work in the L2 classroom. Task-based learner–learner interaction appears not just as the verbal equivalent of a cognitive language learning activity; it is at the same time a communicative event and a social process that is mediated by socio-affective variables. At times, learners might even decide to 'react to social motives at the expense of their own pedagogical advancement, to preserve their social relationships' (Slimani-Rolls 2005: 208). As research has shown, comprehension can take place without much input processing (Ellis 1994: 92). Top–down processing of existing linguistic knowledge (Faerch & Kasper 1986; Ellis 1994), guessing from the linguistic context (Frantzen 2003), guessing based on knowledge of social appropriateness (Hymes 1972), pretending comprehension and hoping for more clues from the ongoing conversation (Hawkins 1985; Firth 1996) are all strategies that contributeto the understanding of one's interlocutors' utterances. Rather than following a 'the more the

[2] As Gass et al. (2005) report variance between single dyads only in the case of meaning negotiation (in one out of the three tasks used), and although being of vital interest in current interaction research, language-related episodes (Swain & Lapkin 1998; Williams 2001; Eckerth 2008a) and recasts (Loewen & Philip 2006; Long 2007) will not be considered in this paper.

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

merrier' principle (Aston 1986; Allwright & Bailey 1991: 145), students' moderate use of requests for clarification and confirmation or checks of one's interlocutor's comprehension appear to reflect the social nature of the 'interactional architecture' (Seedhouse 2004) of the L2 classroom.

In what way, then, do the results from this replication and other work on meaning negotiation in the L2 classroom (Foster 1998; Slimani-Rolls 2005) relate to the findings of Gass et al.'s (2005) carefully designed comparative study? As mentioned earlier in this paper, results of the ANOVAs showed no significant differences between the two settings. However, significant differences were reported between the different tasks in terms of all three measurements, that is, amount of meaning negotiation, language-related episodes, and recasts. Thus, the authors conclude that whereas the setting did not have any influence on the frequency of the three interactional patterns under investigation, the type of task had (Gass et al. 2005: 596–597).

As Gass et al. (2005) used different measures (absolute frequency counts rather than relative to the amount of speech production), it is hard to directly compare their quantitative results with the outcomes of Foster's original study and the present replication study. However, this is hardly a satisfying answer. After all, if, as Gass et al. (2005) convincingly demonstrate, participants do the same thing whether working in the classroom or in the laboratory, then why care about the students' perceptions, interpretations, and motivations? One possible answer could be that we indeed do not have to care if we could know for sure that they were doing the same thing. In other words, what may be coded and counted as the same interactional pattern on a discourse level might turn out to reflect quite different cognitive operations and discourse functions.

Since the comparison of task-based interaction in an actual classroom versus a laboratory setting is at the very center of the Gass et al. (2005) study, the authors take great care to describe the conditions under which the data has been collected (ibid.: 583–584). However, while the reported data collection procedures clearly contribute to the internal validity of the study, they might at the same time threaten its external, or, more precisely, ecological validity.[3] 'Ecological validity', as Bronfenbrenner (1979: 29) defines the term, 'refers to the extent to which the environment experienced by the subjects in a scientific investigation has the properties it is supposed or assumed to have by the investigator'. To what degree this has been the case in the Gass et al. (2005) study, however, is hard to assess. For one thing, it is rather surprising that all three tasks, which were different in detail but similar in design, were administered in uninterrupted succession and without any variation of the social pattern in class, that is, dyadic learner–learner interaction. For another, the reader is further surprised to learn that these three-task sessions hardly varied in length, both in the classroom and the laboratory setting, with all of them lasting between 50 and 55 minutes. Based on the study reported in this paper, and on abundant episodic evidence from language teachers, it is hard to imagine that all learners accomplished all three tasks in nearly the same time. One wonders, then, how exactly the teacher in the classroom setting was 'keeping

[3] Far from discrediting experimental research, it is generally acknowledged in SLA research that studies implemented under controlled conditions lack the ecological validity that research during actual class time can offer (DeKeyser 2003; Doughty 2003).

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

the students on task where necessary' (Gass et al. 2005: 584), and how much interference was required to do so. Therefore, while the classroom may have looked like a classroom, it may not have been perceived as such by the students. As the present study shows that students' task perceptions can influence their task performance, such an account of learners' interpretations of their own task work and interactional behavior could have helped with the appraisal of the investigation's ecological validity.[4]

6. Conclusion

Growing empirical evidence demonstrates that tasks of the kind commonly used in SLA research 'are not just performed but rather are interpreted, resulting in activity that is 'constructed' by the participants in accordance with their particular motives and goals' (Ellis 2003: 187). Using transcript-based and individualized post-hoc tests, some studies show in what ways learners exploit tasks beyond the pedagogical focus of the task, thus adapting the task to their individual learning needs and goals (see Swain & Lapkin 1995, 1998, 2001; Williams 2001; Nabei & Swain 2002; Loewen 2002, 2005; Adams 2007; Eckerth 2008a). Moreover, studies such as Kumaravadivelu (1991), Coughlan & Duff (1994), Gourlay (2005) and Harris (2005) have revealed different activities emerging from the same task when completed by different learners. Thus, what has been called the 'permeability of tasks' (Mondada & Pekarek Doehler 2004: 512) or their 'interactional reconfiguration' (ibid.: 510) points to two major challenges in task-based research.

First, there seem to be good reasons for bringing together qualitative analyses of contextualized task-based interactions with quantifications of isolated linguistic units. Rather than perpetuating unproductive 'paradigm wars' (Edge & Richards 1998: 335), such a multi-methodological perspective seeks to connect different approaches in order 'to link, but not reduce, one perspective to another' (Wertsch 1998 as cited in Ellis & Larsen-Freeman 2006: 578). Such an integrated approach might be ONE way to inform us about potential gaps between 'intended' and 'actual' pedagogy (Seedhouse 2005) and to improve the content validity of our research constructs. If supported by further research, evidence of learners' task adaptation, I suggest, does NOT question the validity of interactional adjustments and its significance for second language acquisition (see Gass 2003; Doughty 2004; Pica et al. 2006; Gass & Mackey 2006, 2007; Mackey 2007). It also does 'NOT prevent tasks and their range of influence from being researchable at the level of PROBABLE outcomes', as Skehan & Foster suggest (Skehan & Foster 1999: 117; my emphasis). However, they conclude, this does not imply that 'task characteristics might have a DETERMINISTIC impact on performance' (ibid.; my emphasis).

[4] Further support for the assumption of the influence of the environment on subjects' interactional behavior comes from research such as Oliver (2000) and Nicholas, Lightbown & Spada (2001) on negotiation and feedback in actual classrooms as compared to laboratory settings. Nicholas et al. (2001: 749), for example, conclude that 'recasts appear to provide more useful input to learners in the laboratory setting than in the classroom setting'. However, both the study conducted by Oliver (2000) and the research reviewed by Nicholas et al. (2001) include classroom settings that consisted of teacher-fronted lessons as opposed to task-based pair work in the laboratory settings. Therefore, it is hard to know whether the differing interactional patterns were due to the different settings, the different tasks, or both.

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Second, studies on task adaptation sketch a picture of the language learner as an intentional and reflective practitioner as well as a social agent, rather than a data processor with purely cognitively defined 'data needs' (Pica 1998: 11). Unfortunately, there are still rather few accounts of how learners perceive tasks and how they position themselves within a task. Apart from the work that has been done within a sociocultural approach to language learning (Roebuck 2000; Lantolf & Thorne 2006), an integration of the learners' perceptions and perspectives has been attempted (e.g. Nakahama, Tyler & van Lier. 2001; Mackey 2002, 2006; Nabei & Swain 2002; Swain & Lapkin 2002; Lindgren 2002; Adams 2003; Egi 2004; Ewald 2004; Carpenter et al. 2006; Polio, Gass & Chapin 2006; Roehr 2006; Watanabe & Swain 2007, Tavakoli forthcoming) and should be pursued in psycholinguistically oriented SLA.

Combining a stimulated recall methodology with an analysis of task performance, the present study tried to show in what way learners' task perceptions can influence their task-based interactions. This study also attempted to contribute to our understanding of task-based L2 classroom learning as being perceived by the learners as a cognitive activity, a communicative event, and a social process. However, as the empirical base of these and related findings is still limited, all conclusions have to be provisional in nature. Finally, they call for more classroom-based interventionist studies (Brumfit & Mitchell 1990: 12). Such investigations require not only a sound understanding of research methods but also a thorough knowledge of the learning–teaching environment in which any interventions and research instruments are introduced. Only then will we arrive at research designs which do not result in 'disrupting the class or compromising the data' (Foster 1998: 21).

Note

This paper is accompanied by supplementary material on the Journal's website, journals.cambridge.org/LTA.

Acknowledgements

I would like to thank Willis Edmondson and Tucker Childs as well as the anonymous reviewers for their insightful suggestions and feedback on an earlier version of this paper. Needless to say that all remaining shortcomings are my own responsibility. Finally, I would like to thank the editorial board of *Language Teaching*, above all Graeme Porte, for suggesting an e-supplement to the printed paper. I believe that providing additional space for the publication of relevant data can significantly increase the transparency and thus the credibility of qualitative, transcript-based L2 research.

References

Adams, R. (2003). L2 output, noticing, and reformulations: Implications for IL development. *Language Teaching Research* 7.3, 347–376.

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (ed.), *Conversational interaction in second language acquisition*. Oxford: Oxford University Press, 29–51.
Allwright, D. & K. Bailey (1991). *Focus on the language classroom*. Cambridge: Cambridge University Press.
Aston, G. (1986). Trouble-shooting in interaction with learners: The more the merrier? *Applied Linguistics* 7.2, 128–143.
Breen, M. (1987). Learner contributions to task design. In C. Candlin & D. Murphy (eds.), *Language learning tasks*. Englewood Cliffs, NJ: Prentice Hall International, 23–46.
Brock, C. (1986). The effect of referential questions on ESL classroom discourse. *TESOL Quarterly* 20.1, 47–58.
Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
Brumfit, C. & R. Mitchell (1990). The language classroom as a focus of research. In C. Brumfit & R. Mitchell (eds.), *Research in the language classroom*. London: Modern English Publications, 3–15.
Bygate, M., P. Skehan & M. Swain (eds.) (2001). *Researching pedagogic tasks: Second language learning, teaching and testing*. London: Longman.
Carpenter, H., K. Jeon, D. MacGregor & A. Mackey (2006). Learners' interpretation of recasts. *Studies in Second Language Acquisition* 28.2, 209–236.
Chaudron, C. (1988). *Second language classrooms*. Cambridge: Cambridge University Press.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edn.). Hillsdale, NJ: Lawrence Erlbaum.
Coughlan, P. & P. Duff (1994). Same task, different activities: Analysis of SLA task from an activity theory perspective. In J. Lantolf & G. Appel (eds.), *Vygotskian approaches to second language research*. Norwood, NJ: Ablex, 173–193.
DeKeyser, R. (2003). Implicit and explicit learning. In Doughty & Long (eds.), 313–348.
Doughty, C. (2003). Instructed SLA: Conditions, compensation and enhancement. In Doughty & Long (eds.), 256–310.
Doughty, C. (2004). Effects of instruction on learning a second language: A critique of instructed SLA research. In B. VanPatten, J. Williams, S. Rott, S. & M. Overstreet (eds.), *Form–meaning connections in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum, 181–202.
Doughty, C. & M. Long (eds.) (2003). *The handbook of second language acquisition*. Oxford: Blackwell.
Doughty, C. & T. Pica (1986). 'Information gap' tasks: Do they facilitate second language acquisition? *TESOL Quarterly* 20.2, 305–325.
Eckerth, J. (2006). Three theses on the pedagogical relevance of second language acquisition research. *Dil Dergisi Turkish Language Journal* 130.1, 18–36.
Eckerth, J. (2008a). Investigating consciousness-raising tasks: Pedagogically-targeted and non-targeted learning gains. *International Journal of Applied Linguistics* 18.2, 119–145.
Eckerth, J. (2008b). Task-based language learning and teaching — old wine in new bottles? In Eckerth & Siekmann (eds.), 13–46.
Eckerth, J. (2008c). Task-based learner–learner interaction: Investigating learning opportunities, learning processes, and learning outcomes. In Eckerth & Siekmann (eds.), 89–118.
Eckerth, J. & S. Siekmann (eds.) (2008). *Research on task-based language learning and teaching: Theoretical, methodological and pedagogical perspectives*. New York: Peter Lang.
Edge, J. & K. Richards (1998). May I see your warrants, please? Justifying outcomes in qualitative research. *Applied Linguistics* 19.3, 334–356.
Edwards, C. & J. Willis (eds.) (2005). *Teachers exploring tasks in English language teaching*. Houndmills: Palgrave MacMillan.
Egi, T. (2004). Verbal reports, noticing, and SLA research. *Language Awareness* 13.4, 243–264.
Ellis, N. & D. Larsen-Freeman (2006). Language emergence: Implications for Applied Linguistics. *Applied Linguistics* 27.4, 558–589.
Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (ed.), *Implicit and explicit learning of languages*. London: Academic Press, 79–114.
Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.
Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
Ellis, R. (ed.) (2005). *Planning and task performance in a second language*. Amsterdam: John Benjamins.
Ewald, J. (2004). A classroom forum on small group work: L2 learners see, and change, themselves. *Language Awareness* 13.3, 163–179.

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Faerch, C. & G. Kasper (1986). The role of comprehension in second-language learning. *Applied Linguistics* 7.3, 184–199.
Firth, A. (1996). The discursive accomplishment of normality. On 'Lingua Franca' English and conversation analysis. *Journal of Pragmatics* 26.2, 237–260.
Firth, A. & J. Wagner (1997). On discourse, communication, and (some) fundamental concepts in SLA. *The 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., A. Tonkyn & G. Wigglesworth (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics* 21.3, 354–375.
Frantzen, D. (2003). Factors affecting how second language Spanish students derive meaning from context. *The Modern Language Journal* 87.2, 168–199.
García Mayo, M. (ed.) (2007). *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters.
Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.
Gass, S. (2003). Input and interaction. In Doughty & Long (eds.), 224–255.
Gass, S. & A. Mackey (2006). Input, interaction and output: An overview. *AILA Review* 19, 3–17.
Gass, S. & A. Mackey (2007). Input, interaction and output in second language acquisition. In B. VanPatten & J. Williams (eds.), *Theories of second language acquisition*. Mahwah, NJ: Lawrence Erlbaum, 175–199.
Gass, S., A. Mackey & L. Ross-Feldman (2005). Task-based interactions in classroom and laboratory settings. *Language Learning* 55.4, 575–611.
Gass, S. & C. Madden (eds.) (1985). *Input in second language acquisition*. Rowley, MA: Newbury House.
Gourlay, L. (2005). Directions and indirect action: Learner adaptation of a classroom task. *ELT Journal* 59.3, 209–216.
Harris, K. (2005). Same activity, different focus. *Focus on Basics* 8.1, 7–10.
Hatch, E. (1978). Discourse analysis and second language acquisition. In E. Hatch (ed.), *Second language acquisition*. Rowley, MA: Newbury House, 401–435.
Hauser, E. (2005). Coding 'corrective recasts': The maintenance of meaning and more fundamental problems. *Applied Linguistics* 26.3, 293–316.
Hawkins, B. (1985). Is an 'appropriate response' always so appropriate? In Gass & Madden (eds.), 162–178.
Hulstijn, J. (1997). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition* 19.2, 131–143.
Hymes, D. (1972). Models of the interaction of language and social life. In J. Gumperz & D. Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*. New York: Holt, Rinehardt & Winston, 35–71.
Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
Kumaravadivelu, B. (1991). Language-learning tasks: Teacher intention and learner interpretation. *ELT Journal* 45.2, 98–107.
Lantolf, J. & S. Thorne (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
Leaver, B. & J. Willis (eds.) (2004). *Task-based instruction in foreign language education*. Washington, DC: Georgetown University Press.
Lee, J. (2000). *Tasks and communication in language classrooms*. Boston, MA: McGraw-Hill.
Lindgren, E. (2002). The effect of stimulated recall on 14-year-olds' L1 Swedish and EFL writing and revision. *Language Teaching Research* 6.3, 267–268.
Loewen, S. (2002). The occurrence and effectiveness of incidental focus on form in meaning-focused ESL lessons. Ph.D. dissertation, University of Auckland.
Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition* 27.3, 361–386.
Loewen, S. & J. Philip (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *The Modern Language Journal* 90.4, 536–556.
Long, M. (1983). Native-speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics* 4.2, 126–141.
Long, M. (1985). Input and second language acquisition theory. In Gass & Madden (eds.), 377–393.
Long, M. (1988). Instructed interlanguage development. In L. Beebe (ed.), *Issues in second language acquisition: Multiple perspectives*. Rowley, MA: Newbury House, 115–141.

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (eds.), *Handbook of second language acquisition*. San Diego, CA: Academic Press, 413–468.
Long, M. (2007). Recasts: The story so far. In M. Long (ed.), *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum, 75–116.
Long, M. (forthcoming)). *Task-based language teaching*. Oxford: Blackwell.
Long, M. & G. Crookes (1987). Intervention points in second language classroom processes. In B. Das (ed.), *Patterns of classroom interaction*. Singapore: SEAMEO Regional Language Centre, 137–152.
Long, M. & P. Porter (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly* 19.2, 207–228.
Mackey, A. (2002). Beyond production: Learners' perceptions about interactional processes. *International Journal of Educational Research* 37.3/4, 379–394.
Mackey, A. (2006). Epilogue. From introspections, brain scans, and memory tests to the role of social context: Advancing research on interaction and learning. *Studies in Second Language Acquisition* 28.2, 369–379.
Mackey, A. (2007). Interaction as practice. In R. DeKeyser (ed.), *Practice in a second language*. Cambridge: Cambridge University Press, 85–110.
Mondada, L. & S. Pekarek Doehler (2004). Second language acquisition as situated practice: Task accomplishment in the French second language classroom. *The Modern Language Journal* 88.4, 501–518.
Müller-Hartmann, A. & M. Schocker-von Ditfurth (eds.) (2004). *Task-based language learning and teaching*. Tübingen: Narr.
Nabei, T. & M. Swain (2002). Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness* 11.1, 43–63.
Nakahama, Y., A. Tyler & L. van Lier (2001). Negotiation of meaning in conversational and information gap activities: A comparative discourse analysis. *TESOL Quarterly* 35.3, 377–405.
Nicholas, H., P. Lightbown & N. Spada (2001). Recasts as feedback to language learners. *Language Learning* 51.4, 719–758.
Nitta, R. & S. Gardner (2005). Consciousness-raising and practice in ELT coursebooks. *ELT Journal* 59.1, 3–13.
Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.
Oliver, R. (2000). Age differences in negotiation and feedback in classroom and pairwork. *Language Learning* 50.1, 119–151.
Pica, T. (1987). Second language acquisition, social interaction, and the classroom. *Applied Linguistics* 8.1, 3–21.
Pica, T. (1991). Classroom interaction, participation, and negotiation: Redefining relationships. *System* 19.4, 437–452.
Pica, T. (1994). Research on negotiation: What does it reveal about learning conditions, processes, outcomes? *Language Learning* 44.3, 493–527.
Pica, T. (1998). Second language learning through interaction: Multiple perspectives. In V. Regan (ed.), *Contemporary approaches to second language acquisition in social context*. Dublin: University College Dublin Press, 9–31.
Pica, T. (2005). Classroom learning, teaching, and research: A task-based perspective. *The Modern Language Journal* 89.3, 339–352.
Pica, T. & C. Doughty (1985). The role of group work in classroom second language acquisition. *Studies in Second Language Acquisition* 7.2, 233–248.
Pica, T. & C. Doughty (1987). The impact of interaction on comprehension. *TESOL Quarterly* 21.4, 737–758.
Pica, T., L. Holliday, N. Lewis, D. Berducci & J. Newman (1991). Language learning through interaction: What role does gender play? *Studies in Second Language Acquisition* 13.3, 343–376.
Pica, T., L. Holliday, N. Lewis & L. Morgenthaler (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition* 11.1, 63–90.
Pica, T., R. Kanagy & J. Falodun (1993). Choosing and using communication tools for second language instruction. In G. Crookes & S. Gass (eds.), *Tasks and language learning: Integrating theory and practice*. Clevedon: Multilingual Matters, 9–34.
Pica, T., H. Kang & S. Sauro (2006). Information gap tasks: Their multiple roles and contributions to interaction research methodology. *Studies in Second Language Acquisition* 28.2, 301–338.

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Pica, T., F. Lincoln-Porter, D. Paninios & J. Linnell (1996). Language learners' interaction: How does it address the input, output and feedback needs of L2 learners? *TESOL Quarterly* 30.1, 59–84.
Polio, C., S. Gass & L. Chapin (2006). Using stimulated recall to investigate native speaker perceptions in native–nonnative speaker interaction. *Studies in Second Language Acquisition* 28.2, 237–267.
Roebuck, R. (2000). Subjects speak out: How learners position themselves in a psycholinguistic task. In J. Lantolf (ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press, 79–95.
Roehr, K. (2006). Metalinguistic knowledge in L2 performance: A verbal protocol analysis. *Language Awareness* 15.3, 180–198.
Samuda, V. & M. Bygate (2008) *Tasks in second language learning*. Houndsmill: Palgrave Macmillan.
Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. Malden, MA: Blackwell.
Seedhouse, P. (2005): 'Task' as a research construct. *Language Learning* 55.3, 533–570.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
Skehan, P. & P. Foster (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning* 49.1, 93–120.
Slimani-Rolls, A. (2005). Rethinking task-based language learning: What we can learn from the learners. *Language Teaching Research* 9.2, 195–218.
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass & Madden (eds.), 235–253.
Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (eds.), *Principle & practice in applied linguistics*. Oxford: Oxford University Press, 125–144.
Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (ed.), *Handbook of research in second language teaching and learning*. Mahwah, NJ: Lawrence Erlbaum, 471–483.
Swain, M. & S. Lapkin (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics* 16.3, 371–391.
Swain, M. & S. Lapkin (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal* 82.3, 320–337.
Swain, M. & S. Lapkin (2001). Focus on form through collaborative dialogue: Exploring task effects. In Bygate et al. (eds.), 99–118.
Swain, M. & S. Lapkin (2002). Talking it through: Two French immersion learners' responses to reformulation. *International Journal of Educational Research* 37.3/4, 285–304.
Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics* 26.3, 376–401.
Tavakoli, P. (forthcoming). Investigating task difficulty: Learners' and teachers' perceptions. *International Journal of Applied Linguistics*.
Van den Branden, K. (ed.) (2006). *Task-based language education: From theory to practice*. Cambridge: Cambridge University Press.
Van Den Branden, K., K. van Gorp & M. Verhelst (eds.) (2007). *Tasks in action: Task-based language education from a classroom-based perspective*. Newcastle: Cambridge Scholars Publishing.
Watanabe, Y. & M. Swain (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research* 11.2, 121–142.
Wertsch, J. (1998). *Mind as action*. Oxford: Oxford University Press.
Williams, J. (2001). The effectiveness of spontaneous attention to form. *System* 29.3, 325–340.
Yule, G. (1997). *Referential communication tasks*. Mahwah, NJ: Lawrence Erlbaum.

JOHANNES ECKERTH is Senior Lecturer at King's College London. Prior to this position, he has taught and conducted research on L2 German and L2 English at Applied Linguistics departments in The Netherlands, Germany, Turkey and the USA. His research interests focus on interfaces between SLA research and L2 pedagogy. He has widely published on SLA and L2 classroom interaction and has recently co-edited (with Sabine Siekmann) a volume entitled *Task-based language learning and teaching: Theoretical, methodological, and pedagogical perspectives* (Peter Lang, 2008).

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press

Comment from the author of the original study

I often thought that I would replicate my 1998 study, having been told on so many occasions that my results were the random outcome of poor data-gathering. It was my first foray into classroom research and I was dealing with recording in a very noisy place with equipment that appears ante-diluvian now. I was not surprised that the pervasive background noise (a normal feature of communicative classrooms) made some of my recordings unusable, but I was never convinced that somewhere in all that untranscribable chatter were the missing negotiations of meaning. If, ten years later, I were doing that replication, the recording equipment would be completely different, but the authentic classroom setting I was so keen to preserve would be the same. A learner in a classroom is aware that a task with a partner is as much a social interaction as a pedagogic exercise and crystalline clarity is neither a necessary nor desirable goal.

As in my study, the participants in this replication are in their scheduled class, and they are doing nothing other than normal class work. Apart from using only dyads (which makes sense given the way groups allow participants to remain silent), the procedures and analyses are all very similar. And gratifyingly, as summarised in Table 2 here, the numbers come out rather similar also. More importantly, the author has improved on my study. He has been able to get considerably more data than I did, and this strengthens his conclusions. He has also added a stimulated recall procedure which demonstrates what I was only guessing at, that learners will put up with partial understanding in order to keep the task interaction moving forward. This paper lays out together the transcript evidence and the recall evidence in a neat demonstration of how learners will smooth over the bumps rather than make explicit their lack of complete understanding.

In seeking an explanation for the results in Gass et al. (2005), which show no difference in the incidence of negotiation of meaning in classroom or laboratory settings, the author very plausibly suggests that their classroom setting might not have appeared much like a classroom to the participants, who had to do three tasks in seamless progression. But the way to test properly the effect of setting on L2 performances is to use a within-participants design comparing task interaction in an environment that very clearly is a classroom and one that very clearly is not, and this has not been done. For me, though, it is not so much whether negotiations of meaning happen in one setting or another, or in one task type or another; rather it is whether their overwhelmingly lexical nature makes them as valuable as is claimed.

Dr. Pauline Foster
Reader in Applied Linguistics, St. Mary's University College
fosterp@smuc.ac.uk

https://doi.org/10.1017/S0261444808005442 Published online by Cambridge University Press