

2 RESEARCH ON INTERACTION BETWEEN LANGUAGE LEARNERS

This chapter provides an overview of three established perspectives in research on interaction between language learners. This includes the negotiation of meaning perspective, a broadly construed cognitive perspective, as well as the sociocultural perspective.

The first section of the chapter covers the negotiation of meaning perspective, the second section the cognitive perspective, and the third section the sociocultural perspective. These sections provide a brief background, a review of representative research and findings, as well as a critique of the relevant perspectives. A final section summarises implications for the development of a method for representing and analysing the dynamics of learner interaction.

2.1 Negotiation of Meaning Perspective

This section is organised into three sub-sections. The first sub-section gives a brief background to the negotiation of meaning perspective. This is followed by a review of some representative research on negotiation of meaning between language learners. The final sub-section contains a critique of the negotiation of meaning research, and an assessment of any implications for the aims of the present research.

2.1.1 Background to the Negotiation of Meaning Perspective

Negotiation of meaning refers to how communication breakdowns are overcome through the use of rhetorical devices such as clarification and confirmation requests, comprehension checks, as well as repetitions, repairs, corrections and completions. It was first suggested as a psycholinguistic rationale for understanding interaction between language learners by Long and Porter (1986). However, it is in essence an extension of observations made of native speaker/non-native speaker (NS/NNS) interaction, in non-pedagogical settings, or what is known as the study of 'foreigner talk' (cf. Long, 1981, 1983). The psycholinguistic rationale for negotiation of meaning is that it makes input received from a native-speaker (NS) more comprehensible for the non-native-speaker (NNS). In this respect, the rationale is an extension of Krashen's comprehensible input hypothesis (cf. Krashen, 1982). That is, *if comprehensible input facilitates language acquisition, and if negotiation of meaning makes input more comprehensible, then negotiation of meaning facilitates language acquisition.* Furthermore, since negotiation of meaning actively involves the NNS, through her use of the mentioned rhetorical devices, the resulting conversational adjustments, in relation to this particular learner's current level of second language competence, are uniquely appropriate for language acquisition to occur. Finally, this emphasis on language *acquisition*, through

comprehensible input, can be traced back to the influence of Chomskian linguistic theory, and the notion that human beings have an innate faculty for acquiring language through a largely unconscious process (cf. Chomsky, 1959).

2.1.2 Research on Negotiation of Meaning in Learner Interaction

Early studies on the negotiation of meaning between language learners served to extend this perspective from the original application to foreigner talk, or NS/NNS interaction. However, these early studies did not necessarily focus on pedagogical contexts.

As is customary in these studies, the review of research, which appears in the following, will continue to use the terms native speaker (abbreviated to NS), and non-native speaker (abbreviated to NNS).

One example of this early research, which sought to extend the concept of negotiation of meaning to learner interaction, was a study by Porter (1986). Porter compared NS/NNS and NNS/NNS interaction on a whole range of variables, including clarification and confirmation requests, comprehension checks, communication strategies (such as verification of meaning), definition requests and indications of lexical uncertainty. As measured by these variables, Porter found that NNS/NNS interaction showed many of the same negotiation of meaning patterns as NS/NNS interaction. Another study with a similar aim as that of Porter, by Varonis and Gass (1985), instead likened negotiation of meaning to vertical *pushdowns* in the horizontal progression of a conversation. Such pushdowns, therefore, contained the clarification and confirmation requests, and the comprehension checks. The researchers found that pushdowns occurred with greater frequency in NNS/NNS interaction than in both NS/NS and NS/NNS interaction. The researchers concluded from this that NNS/NNS interaction involves more negotiation of meaning than both NS/NNS and NS/NS interaction. Furthermore, the authors speculated that the greater frequency of pushdowns in NNS/NNS interaction was due to the learners' 'shared incompetence' in the second language, and that this shared incompetence served to break down the social constraints on repair-patterns normally observed in conversations between native speakers (cf. Schegloff, Jefferson, & Sacks, 1977). Finally, within their sample of NNS/NNS dyads, the number of pushdowns was related to how much the learners had in common. For example, two learners that did not share L1 background, or did not have the same level of proficiency in the L2, would negotiate more.

Although predating the psycholinguistic rationale underlying the negotiation of meaning perspective, findings by Schwartz (1980) may usefully be added to the above body of evidence. Schwartz compared NNS/NNS interaction with findings based on NS/NS baseline

data. Schwartz' was interested in whether NNS/NNS dyads would show the same patterns of repair as NS/NS dyads. He concluded that,

the second language learners [...] gave the speaker of a trouble source repeated chances to repair his [sic] own speech [...] but] when the trouble source involved was a matter of incompetence in syntax, lexicon, or phonology, the other speaker also made repairs. (1980, p. 151-152)

This again corroborates the findings of Varonis and Gass (1985) about the possible effects of a 'shared incompetence' in an L2.

The above early studies paved the way for research on negotiation of meaning in the language classroom. A first round of such classroom-based research compared negotiation of meaning in teacher-fronted and small group-work conditions. A representative set of such studies, by Pica and Doughty (Pica & Doughty, 1985a, 1985b) found that the amount of negotiation of meaning was very low indeed in their classroom setting, and that uncritically generalizing findings from non-classroom NS/NNS or NNS/NNS interaction, to the language classroom, was probably unjustified. However, Pica and Doughty speculated that the so-called decision-making tasks they had used for their small group-work condition might explain the lack of negotiation of meaning. This set the stage for a spate of research into the differential effect of different task types on negotiation of meaning between language learners. Hence, from this point forward the negotiation of meaning research may usefully be seen as part of a larger task-based perspective.

Three studies that are representative of this early task-based research were all conducted in pre-academic ESL settings in the US (Doughty & Pica, 1986; Duff, 1986; Gass & Varonis, 1985). In reviewing these studies it is important to note how the different researchers defined their task types. Doughty and Pica (1986) distinguish between tasks that require information exchange, and those that do not. Hence, their contrast is between tasks with *optional* versus *required exchange* of information. By contrast, Gass and Varonis (1985) differentiate between *one* and *two-way* tasks, where a one-way task refers to information flowing in one direction only (e.g., describing a picture from behind a screen), and a two-way task refers to information flowing both ways (i.e., providing each participant in a dyad with some of the necessary information). Note also that in terms of Doughty and Pica's distinction, both of Gass and Varonis' task types require information exchange. Finally, Duff (1986) distinguishes between *convergent* tasks, which require arriving at a consensus, and *divergent* tasks, which allow disagreement as an outcome.

The method and findings of these studies includes:

- Doughty and Pica (1986) added up clarification and confirmation requests, comprehension checks, as well as repetitions, into an overall modified interaction

variable. They found significant differences in this overall variable in favour of the required exchange tasks. Furthermore, since the researchers found the repetition measure to be “puzzling at best to analyze” (p. 317), they subtracted all repetitions to see whether it would change the result. However, the result remained the same.

- Gass and Varonis (1985) counted the number of pushdowns (as in their earlier study reviewed in the above; cf. Varonis & Gass, 1985), but found no significant differences between one-way and two-way tasks.
- Duff (1986) counted clarification and confirmation requests, comprehension checks, collaborative checks (in which explicit feedback about agreement or disagreement was sought), as well as a number of different types of questions (all as separate categories). The study found significantly more confirmation checks, as well as referential and subject-related questions in the convergent task. However, in the remaining variables there were no significant differences, and, consequently, the overall effect on the negotiation of meaning process was somewhat ambiguous.

An additional study of task effects on negotiation of meaning was from an adult EFL setting in Sri Lanka (Brown, 1991). Brown made distinctions between task types along three continua. These continua included 1) tight/loose, with a ‘tight’ task allowing learners less freedom in deciding what to do, 2) closed/open, with ‘closed’ tasks having definite questions and answers that had to be covered, and 3) procedural/interpretive, where a ‘procedural’ task was defined as ‘getting things done’, without the need for any individual interpretations. Brown counted repetitions, prompts, rephrasings, as well as repairs (including clarification and definition requests, and comprehension checks), but found no significant differences between tasks classified differently on his three continua.

Finally, while not explicitly task-based, research by Oliver (1998) set out to extend the research on negotiation of meaning, which invariably had been conducted with adult participants, to an Australian primary ESL setting. The pupils were between 8 and 13 years old, were paired in NS/NS, NS/NNS and NNS/NNS dyads (the NNS/NNS dyads did not share a common L1), and the pupils were given two different tasks to do. Oliver counted clarification and confirmation requests, comprehension checks, as well as self and other repetitions. The young learners in Oliver’s study, including the NNS/NNS dyads, showed that they were quite able to negotiate for meaning, but at levels of frequency somewhat below that of adult learners. However, in Oliver’s young learner data, the interactional configuration (NS/NS, NS/NNS or NNS/NNS) did not seem to matter for how much negotiation of meaning there was, as measured by any of the variables. This contrasted with the findings of Schwartz (1980) and Varonis & Gass (1985) (cf. discussion of these studies in the above). Oliver used the lack of any differences between interactional configurations in her young learner data as

evidence that “the pattern of exchange that occurs may be influenced in some way by the developmental level of the participants” (1998, p. 379). In other words, it may be that children worry less about any shared competence, or incompetence, in a language, and worry less about social constraints on patterns of repair.

2.1.3 Critique and Implications for the Research Aims

The above review of studies shows that there is some evidence that negotiation of meaning varies across task types, age of learners, as well learners’ backgrounds. However, the fact that these findings are based on statistical analyses of frequency data means that they do not contribute in any direct manner to the research aim of accounting for the dynamics of learner interaction. Nevertheless, a closer look at a few of the potential limitations of the negotiation of meaning research reveals some more indirect implications.

One limitation of the negotiation of meaning research is indicated by the related critiques offered by Hawkins (1985) and Aston (1986). Hawkins’ critique is built on what she sees as the implicit assumption that negotiation of meaning constitutes *appropriate responses* to learners’ non-comprehension. That is, there is an implicit assumption that so long as negotiation of meaning takes place, comprehension, and thereby also language acquisition, will follow. To back up her critique, Hawkins conducted a study where she used a stimulated recall design to explore whether comprehension was in fact an outcome of negotiation of meaning. Independent coding of the transcribed learner interaction data and the protocols from the stimulated recall interviews found that only 44% of negotiation of meaning sequences were claimed by the learners to result in comprehension. Aston’s (1986) critique focused on the difference between *formal* and *substantive* understanding. According to Aston, formal understanding is nothing more than the “performance of a ritual of understanding or agreement” (1986, p. 139), while substantive understanding involves an actual convergence of the interlocutors ‘worlds’. Aston concluded that “in order to show that a greater frequency of negotiation entails more input being made comprehensible, it would be necessary to have information on the frequency with which such negotiation achieves substantive understanding” (1986, p. 134).

The critiques by Hawkins and Aston point to the potentially important limitation of simple cumulative, or abstract, measures in research on learner interaction. That is, even though there may be a sound psycholinguistic rationale for the coding of a certain feature of interaction, such as negotiation of meaning, the critiques suggest that it is potentially rewarding to take an in-depth look at what the feature that is coded actually does in the learner interaction. Likewise, a dynamical account of learner interaction may be at a certain

level of abstraction, and might benefit from a more in-depth look at what the features that are coded actually do in the learner interaction.

A related limitation of the negotiation of meaning research are the many different ways of defining task types, as well as the overlap between these. With the possible exception of required exchange of information tasks (cf. Doughty & Pica, 1986; as well as discussion in sub-section 2.1.2), the findings about any task effects are quite ambiguous. The fact that all the definitions of task types are all formulated in abstract terms, and not linked to what the learners are doing in any meaningful sense, may be a possible clue for explaining the ambiguous findings. The present research does not purport to identify how different task types affect learner interaction. However, a tentative implication for the present research may be that any representation and analysis of learner interaction may benefit from making a close link to what it is, in meaningful terms, that the learners are doing.

A third limitation of the negotiation of meaning research relates to the selection of participants, data and settings. One feature of this research is that all the studies reported above, except the study by Oliver (1998), were conducted with adult participants. Furthermore, the great majority of the studies involved participants from University students enrolled in pre-academic language programmes (but cf. Brown, 1991). Another feature is that some of the studies have coded only limited segments of the available data for negotiation of meaning (sometimes in order to derive standardised samples for comparison). For example, the study by Varonis and Gass (1985) used only the first five minutes of each conversation they recorded. Finally, most of the studies reviewed above were conducted in settings, or with tasks, controlled by researchers, and not the learners' ordinary teachers. Recent studies in more unperturbed classroom settings have made striking observations about the greater variability and inconsistency of their 'ecological valid' data on learner interaction (e.g., Foster, 1993, 1998; Jacob, Rottenberg, Patrick & Wheeler, 1996). Foster goes as far as to say that "some of the current claims in Second Language Acquisition research are of academic rather than practical interest since the researchers have lost sight of the world inhabited by language teachers and learners" (1998, p. 21).

The observations made by Foster (1993, 1998) and Jacob et al. (1996) indicate that there are indeed unpredictable situational dynamics involved in learner interaction in language classrooms. This, in itself, validates the aims set for the present research. Furthermore, given the nature of the disconnection between research and practice, which according to Foster can be seen in the negotiation of meaning research, unpredictable situational dynamics may be hidden by any, or all, of the following factors: 1) relying on data from a limited population of learners; 2) coding and analysing only part of a conversation, activity or task; 3) controlling data collection too closely.

A final observation that can be made of the negotiation of meaning research is the lack of later studies replicating the findings reported above. The reason for this may lie in changes the negotiation of meaning perspective underwent in the late 1980s, and the 1990s. These changes may be traced back to Swain's (1985, 1995) output hypothesis. This hypothesis was motivated by research findings made in French immersion classrooms in Canada, where the English native language background students did not seem to progress beyond a certain level of productive language proficiency in French. In particular, the students did not seem to achieve the same syntactical accuracy as their French native language peers (cf. Cummins & Swain, 1986). Swain has argued that these learners need to process language syntactically if they are to progress towards a more target-like use of French. Moreover, the output hypothesis claims that learners will only process language syntactically if a teacher (or task) *pushes* learners to produce accurate, appropriate and coherent output (Swain, 1985).

Swain's suggestions seem to have resulted in the development of two strands of research on learner interaction. One strand has a combined focus on negotiation of meaning and on what is called modified output. These strands largely remain true to the roots of the negotiation of meaning perspective, i.e., the notion of language acquisition through an innate language faculty. However, this strand has not resulted in as much research on learner interaction (but cf. García Mayo & Pica, 2000; Pica, Lincoln-Porter, Paninos & Linnell, 1996; Shehadeh, 2001). The other strand of research draws on a later refinement of Swain's output hypothesis. This hypothesis suggests three functions of output that can be used as evidence for learners processing language syntactically. These functions include a) noticing gaps in one's language resources, b) testing hypotheses about how to use the language, and c) talking about language (as in metatalk) (cf. Swain, 1995). This strand has produced a great deal of research on learner interaction. However, through its focus on in-depth analyses of single episodes of interaction, it has at the same time developed close commonalities with the sociocultural perspective. Consequently, this strand will be discussed in the third section of this chapter, which covers the sociocultural perspective.

2.2 Cognitive Perspective

This section discusses research on interaction between language learners within a broader cognitive perspective. The discussion begins with a sub-section that gives a brief background to the perspective. A sub-section that reviews research on cognitive processes in learner interaction follows this. The final sub-section contains a critique of the cognitive perspective, and an assessment of possible implications for the aims of the present research.

2.2.1 Background to the Cognitive Perspective

The background to this broader perspective is not as clear as is the case for the negotiation of meaning research. Early work within this perspective tends to focus on different cumulative measures of language use, and only implicitly appeals to cognitive processes in interpreting results. Later work, however, puts a more explicit emphasis on the cognitive processing involved in the production of language. In this respect, the perspective draws on early information-processing accounts of language learning (e.g., McLaughlin, Rossman & McLeod, 1983), as well as more recent developments in psycholinguistics (cf. Skehan, 1998). There is also some overlap with what is known as a skills-based view of language and language learning (cf. Johnson, 1996). Nevertheless, all the research, which is reported under the heading of the cognitive perspective, has in common a focus on the cognitive processes involved in language use. Moreover, research within this perspective accepts that cognitive processes can be both conscious and unconscious, and that it is possible to direct these cognitive processes through careful consideration of task types and conditions (Skehan, 1996). Hence, the research within this cognitive perspective, as it is presented here, is explicitly task-based.

2.2.2 Research on Cognitive Processes in Learner Interaction

A study by Tong-Fredericks and colleagues (Tong-Fredericks, 1984) is illustrative of the early research within the cognitive perspective, including a focus on language use, and providing only implicit references to the cognitive processing of this language use. This study undertook a comparison of three different communication tasks: 1) a problem-solving task, 2) a role-play task, and 3) an authentic interaction task (which was similar to the role-play task - except that the learners 'acted' as themselves). The researchers compared the number of turns per minute (t/m), the number of self-corrections per minute, and the speed of speaking in words per minute (wpm). The most interesting results for turns per minute were for the problem-solving task, which averaged 13 t/m (as compared to 9 and 11 t/m for the role-play and authentic tasks). Tong-Fredericks argues that the greater number of turns per minute in the problem-solving task may be associated with interruption pressure. That is, Tong-Fredericks hypothesised that the other task types involved less such interruption pressure, and therefore facilitated production of longer sentences. For number of self-corrections the problem-solving task again stood out. There was only one (1) self-correction per minute (as compared to 4.8 and 5.8 self-corrections per minute in the role-play and authentic tasks). This led Tong-Fredericks to speculate that "because speakers are so engrossed in working towards a solution [on the problem-solving task] [...] everything is subordinated to the goal, i.e. finding the solution" (1984, p. 138-139). As for speed of speaking, the problem-solving task

showed slightly fewer words per minute (111 wpm as compared to 121 and 127 wpm in the other two task types). However, a follow-up qualitative analysis showed that the problem-solving tasks had high incidence of repetition and echoing, which served to fulfil various functions such as hesitating, agreeing, acknowledging and confirming. Tong-Fredericks suggested that this was a sign that the problem-solving task involved the learners in more hypothesis testing behaviour. Finally, it should be clear that in all the interpretations of results, which Tong-Fredericks provides, he makes a link between measures of the quantity of language use and the processing conditions associated with task types.

A later study, by Newton and Kennedy (1996), illustrates a somewhat different way to link language use and cognitive processes. These researchers hypothesized a difference in the types and frequency of prepositions and conjunctions in the learners' language use on *split* and *shared* information tasks (a distinction which is similar to Doughty and Pica's (1986) *optional* versus *required exchange* tasks - cf. section 2.1.2). A further distinction was made between *spatial* versus *decision-making* tasks, and the resulting four conditions were: 1) split/decision-making, 2) split/spatial description, 3) shared/decision-making, and 4) shared/spatial description. Based on a concordance analysis of the data, and after making manual adjustments for what grammatical functions the concordance analysis yielded, the researchers made the following conclusions:

- prepositions made up a significantly ($p < 0.05$) larger proportion of the words in both of the spatial description tasks, as compared to both of the decision-making tasks.
- prepositions made up a significantly ($p < 0.05$) larger proportion of the words in the split/spatial description tasks, as compared to the shared/spatial description task.
- conjunctions made up a significantly ($p < 0.05$) larger proportion of the words in both of the shared tasks, as compared to both of the split tasks.

Newton and Kennedy argue that spatial description tasks contain more use of prepositions because of the need to mark locative relationships, and that there were more prepositions in the split/spatial, as compared to shared/spatial, tasks because of the information exchange that was involved. Finally, Newton and Kennedy argue that the shared information tasks contain more conjunctions (and especially subordinating conjunctions) because of the higher frequency of reasoning, persuading and arguing, thereby creating a greater need to mark interpropositional relationships. Hence, the link between language use and cognitive processes, which Newton and Kennedy make is somewhat more intricate than that made by Tong-Fredericks (1984; cf. analysis in the above paragraph). That is, the link they make is actually a link between the grammatical features used by learners, and the processing conditions associated with the relevant task types.

A recent group of studies have made an even more sophisticated link between language use and the cognitive processes associated with different task types. This research has explored the effect of task types on the fluency, accuracy and grammatical complexity of learners' language use (cf. Skehan, 1996; 1998). A study by Robinson (1995) is a good illustration of the underlying rationale for this focus (using a task done by an individual learner, however). Robinson compared a 'here-and-now' narrative task, where the learner could look at a set of pictures while doing the narrating, and a 'there-and-then' narrative task, in which the learner had to give the set of pictures back to the researcher before beginning to narrate. Robinson hypothesized that the availability of the picture prompts in the 'here-and-now' condition would lighten the cognitive processing load of the task, and that this, in turn, would lead to more fluent language use. By contrast, he hypothesised that the 'there-and-then' task would constitute a higher cognitive processing load, and, surprisingly maybe, that the added effort in the 'there-and-then' task would lead to more accuracy and grammatical complexity in the learner's language use. Robinson's results confirmed his hypotheses.

Three recent studies on learner interaction have used such a focus on the fluency, accuracy and complexity of learner's language use in tasks. The first of these, a study by Foster and Skehan (1996), largely confirms Robinson's hypotheses about how processing conditions of different task types can affect the fluency, accuracy and complexity of learners' language use. Foster and Skehan designed three different tasks, each of which were considered to constitute a slightly higher processing load (in terms of familiarity with the topic, and clarity of the task's structure). Just as had been hypothesized, the learners' language use was the most fluent on the task with the most familiar topic, and the clearest structure (narrating a fictional personal experience). Furthermore, the learner language was of a significantly higher complexity in the two tasks with the higher processing load (a picture narrative and a decision-making task). However, the measure for accuracy showed a less predictable pattern.

Another study by the same researchers (Skehan & Foster, 1997) used three tasks that were considered to be fairly similar to those in the earlier study. However, the resulting fluency, accuracy and complexity of the learners' language use were somewhat confounding. Whereas some of the results were consistent with Robinson's original hypothesis, other results showed an entirely reverse relationship for the different task types (for a discussion of this apparent dissociation cf. Skehan, 1998, chapter 5). Hence, the only thing that can be said for certain from this research is that there is a 'trade-off' between fluency, accuracy and complexity in learners' language use on tasks.

A study by Bygate (1999a) illustrates the difficulty of finding the exact nature of this 'trade-off', between fluency, accuracy and complexity, which a particular task promotes. Bygate approached this problem by taking a detailed look at the grammatical complexity of

learners' language use on two task types, a narrative task and an argumentative task. His analysis proceeded in steps, beginning with the following sets of findings (Bygate defined a T-unit as "a grammatically defined structure, consisting of an independent finite clause plus any finite or non-finite clauses depending on it" (1999a, p. 196)):

1. The narrative tasks elicited significantly more words per T-unit* than the argumentation tasks (all p values < 0.01).
2. Subordination per T-unit was not significantly different between the two task types.

Bygate considered this first set of findings contradictory. He reasoned that since the length of a T-unit is only a rough measure of grammatical complexity, and because the number of subordinate clauses in a T-unit may be seen to contribute more directly to this complexity, no firm conclusions could be made about differences in grammatical complexity of pupils' language use on the two task types. Even so, the question of why the T-units were longer in the narrative tasks was, from a processing perspective, still unanswered. Additional analysis on a representative sub-sample of the data showed that:

- 3) The narrative task data contained a significantly ($p < 0.001$) greater frequency of verb arguments per finite verb.

From this, Bygate reasoned that some of the extra words per T-unit in the narrative task could be going towards producing more such verb arguments around finite verbs. This finding, therefore, seemed to indicate that the narrative tasks facilitated language use of greater complexity, after all. However, yet another step of analysis showed that the argumentation data contained a greater total, as well as a greater range, of verb groups. In addition, the argumentation data also seemed to contain more formulaic language use. Since formulaic language use is processed in chunks it adds little to the measure of grammatical complexity. In sum, Bygate found it difficult to determine which task type was associated with greater grammatical complexity. The only clear conclusion that could be drawn, it seemed, was that the two task types produced different types of language practice for the students.

2.2.3 Critique and Implications for the Research Aims

The research on learner interaction within the cognitive perspective shows that different task types are associated with different cognitive processing. In other words, different task types may, in fact, constitute different processing conditions, in turn affecting learners' language use. However, just as for the main findings from the research within the negotiation of meaning perspective, the cognitive perspective also relies on statistical analyses of frequency

data. Hence, the findings have no direct implications for the present aims, and it is again necessary to look at the limitations of the research to find more indirect implications.

The intricate nature of some of the analyses within the cognitive perspective points to a possible limitation of the research reported in this section. That is, although the research does show a relationship between the processing conditions of task types, and measures of language use, at increasing levels of detail it becomes a challenge to keep track of all the possible variables. The study by Bygate (1999a), reported on above, goes a long way toward rising to this challenge, but at the same time illustrates this possible limitation. This has an interesting implication for research on learner interaction general, as well as the research aims for the present research. That is, at a certain level of detail the cognitive perspective is both clear and convincing. In the same way, there may be a level of detail at which an account of the dynamics of learner interaction is clear and convincing. Similarly, as the research reported in this section seems to suggest, there might be levels of detail at which any account of the dynamics of learner interaction is less clear, and less convincing.

Another potential limitation, or strength, depending on one's perspective, of the cognitive perspective is the explicit task-based nature of the research. A task-based focus has the strength that it makes a close link between theory and pedagogy (Bygate, et al., 2001). At the same time, this very feature has been criticised by some authors as being untenable. For example, Coughlan and Duff remark that underlying task-based research "is the belief that these tasks, and their resulting behavior, are scientifically controllable and measurable" (1994, p. 173). However, Coughlan and Duff make their critique from the perspective of individual cases of learners' activity, where they observed a great deal of unpredictability. Whereas the critique may be valid when one is interested in such individual cases of learner interaction, the critique somewhat misses the point when considering that task-based research, such as that reported on in this section, samples from a population of learners, and generalises to a population of learners. Nevertheless, the wide appeal that Coughlan and Duff's critique seems to have in the literature brings added urgency to the question of what unpredictable situational dynamics are involved in learner interaction. That is, it validates the research aim of visualising the dynamics of learner interaction.

2.3 Sociocultural Perspective

This section begins with a sub-section providing a brief background to the sociocultural perspective. The next sub-section reviews some representative research on learner interaction within this perspective. A final sub-section provides a critique of the sociocultural perspective, and an assessment of any implications for the aims of the present research.

2.3.1 Background to the Sociocultural Perspective

In language education research more generally, the sociocultural perspective draws on two major sources. The first influence is the concepts for exploring human cognitive development suggested by Vygotsky (cf. 1978; 1986). This includes 1) the conceptualisation of language mediating thinking and activity, 2) the notion of knowledge structures, including language, starting life on an inter-psychological plane, and 3) that only if ‘presented’ within an individual learner’s Zone of Proximal Development (ZPD) do knowledge structures, and language, become part of the learner’s repertoire of individual cognitive resources. Another influence is the six features of scaffolding suggested by Wood, Bruner and Ross (1976). These features describe different ways in which potentially deliberate interaction can facilitate learning. Finally, activity theory is also sometimes cited as an influence (e.g., Lantolf, 2000). However, this influence is subtler, and has not resulted in many clear applications (but cf. McCafferty, Roebuck & Wayland, 2001).

In research on interaction between language learners the sociocultural perspective only emerged in the mid to late 1990s. However, as was indicated in the critique of the negotiation of meaning perspective in sub-section, a strand of research prompted by Swain’s (1985, 1995) output hypothesis has many commonalities with the more directly sociocultural research. In fact, the more recent contributions of Swain and colleagues explicitly draw on sociocultural theory (cf. Swain & Lapkin, 1998). Hence, the review of sociocultural research on learner interaction will begin with a sub-section describing this strand of research.

The remaining research, which is reported under the heading of the sociocultural perspective, has mirrored developments in the field of language education more closely. That is, the research has focused on learners’ talk as mediating interaction, the identification of features of scaffolding in learner interaction, and the relationship between learners’ interaction and their ZPD. This remaining research is reported in a second sub-section.

2.3.2 Research on Functions of Output in Learner Interaction

The research discussed in this sub-section is directed at the identification of the three functions of output in learners’ talk-in-interaction. These three functions, as suggested by Swain (1995) include a) learners’ noticing gaps in their language resources, b) learners’ testing hypotheses about how to use the language, and c) learners’ talking about language (as in metatalk).

The analysis of the functions of output relies on the identification of so-called *language-related episodes* (LREs) (cf. Kowal & Swain, 1994; Swain & Lapkin, 1995). The identification of these episodes was, in part, influenced by the similar *critical episodes* used by Samuda and Rounds (1993) to make sense of talk between adult ESL learners in a spot the

difference task. Samuda and Rounds defined a critical episode as “the segments of the task in which a difference is being discussed” (1993, p. 126). Furthermore,

the beginning of each critical episode was defined as the point at which a participant nominated an item that was a difference; the end was defined as the point at which they moved on to another part of the picture. (1993, p. 127)

Samuda and Rounds also distinguished between ‘successful’ critical episodes, where the participants agreed that they had found a difference, and ‘unsuccessful’ ones, where the learners failed to realize they had a difference. This distinction was motivated by Hawkins (1985) concern about whether negotiation of meaning constituted an ‘appropriate response’ to non-comprehension (cf. discussion in section 2.1.3). The rationale was that if a transcript was segmented into episodes corresponding to the learners’ discussion of picture differences, the pictures in the task materials could provide a reference point for whether particular episodes of interaction indeed resulted in learners’ comprehension.

The use of LREs in the research on functions of output in learner interaction abandoned this close link between the structure of a task, and learners’ interaction. Instead, LREs are defined as episodes of interaction in which some aspect of the target language is the focus of learners’ talk. Hence, this research proceeds by first identifying LREs, and then conducts in-depth analyses of excerpts of transcription representing the LREs for evidence of any of the three functions of output.

Three studies, which are representative of the research on functions of output in learner interaction, are reported below. The studies were all set in grade 8 French immersion classrooms, and were part of a larger project, the purpose of which was to find out what types of cooperative tasks were more related to the three functions of output (Kowal & Swain, 1997). The latter aspect of this project correctly implies that the research was focused on tasks, or task-based.

A first of these studies, by Kowal and Swain (1994), used a dictogloss task. The researchers describe this task as follows:

a short, dense text is read to the learners at normal speed; while it is being read, students jot down familiar words and phrases; the learners work together in small groups to reconstruct the text from their shared resources; and the various versions are then analysed and compared in a whole class setting. (1994, p. 28)

Kowal and Swain’s analysis of the learner interaction began by separating out LREs from the transcribed data. This process led the researcher to distinguish between three types of LREs: 1) meaning-related episodes, 2) grammatical episodes, and 3) orthographic episodes. Out of a

total of 224 LREs, there were 69 meaning-related episodes, 93 grammatical episodes, and 62 orthographic episodes.

In-depth analyses of the LREs showed that the learners were indeed noticing gaps in their language resources. The analyses also showed that when gaps were noticed, the learners would search for solutions, and that this search for solutions involved learners' talking about language, as well as learners' testing hypotheses about how to use the language. Nevertheless, Kowal and Swain also observed that learners sometimes lacked the necessary metalanguage to verbalize their thinking. They also observed that there was considerable variation between individual learners, and groups of learners, in the extent of noticing, hypothesis testing and metatalk. Furthermore, in heterogeneous groups (in terms of language level) the lower level learners were often given less opportunity to test their hypotheses, or to engage in metatalk. Note that since the dictogloss task did not require transfer of information, this finding is not necessarily contradictory to Varonis and Gass's (1985) finding that heterogeneous pairs negotiate more (cf. section 2.1.2).

A second study, by Kowal and Swain (1997), included a second task condition. This was a cloze exercise focusing on the correct formation of verb tenses. Furthermore, in this task the teacher tried to avoid heterogeneous groups. This latter strategy did seem to allow the weaker learners to play a more active role. However, the learner interaction in the cloze exercise differed from the dictogloss task in a different respect. Kowal and Swain comment that in the first task (the dictogloss) "it was the students who identified and set the agenda of grammatical items to be discussed. In the second [the cloze exercise], it was the teacher and the focus was much narrower." (1997, p. 305). Finally, the researchers did not comment on the possibility that the narrow focus of the cloze exercise might have made it easier for the weaker students to contribute.

A third study, again as part of the same overall project, constitutes a considerable narrowing of any remaining gap with the sociocultural research to be reported upon in the next sub-section. That is, in this study, by Swain and Lapkin (1998), the research focus was on how the talk of *one* pair of language learners served to 'mediate' learning through noticing, hypothesis testing and metatalk. The analysis made a strong case for the fact that the learners' talk indeed served to mediate learning through the functions of output. However, the pair of learners whose interaction was analysed was the pair whose written product was assessed as the best-written product of an entire class (a total of 12 pairs). Hence, the findings of this last study may only apply to 'good' learners.

2.3.3 Research on Mediation, Scaffolding and the ZPD in Learner Interaction

The research discussed in this sub-section picks up where the study by Swain and Lapkin (1998), reported at the end of the previous sub-section, left off. That is, a common method used in sociocultural research is to explore the mediational role language use plays in cognition and learning (cf. Ahmed, 1994; Vygotsky, 1997).

Just like the research reported upon in the previous sub-section, sociocultural analysis of learner interaction also relies on the selection of episodes of interaction. However, the sociocultural research differs somewhat in that it lets the categories of analysis direct which episodes of the transcribed data are examined in-depth.

One instance of research that explores the mediational role language plays is a study by Brooks and Donato (1994). Their study, with high school learners of Spanish as participants, involved learners doing a two-way information exchange task. Their in-depth analysis of the learner interaction in this task suggested the presence of three types of so-called *semiotic mediation*, including *speaking as object regulation*, *speaking as shared orientation* and *speaking as goal formation*.

Brooks and Donato's study does not provide results and findings in the traditional sense. However, the researchers argue that learners' activity during the task was defined as much by their speaking as object regulation as it was by any *a priori* stated task requirements or instructions. In addition, learners achieved the necessary level of intersubjectivity through speaking as shared orientation. Finally, speaking as goal formation ensured that the learners worked in a purposeful manner towards completing the task. Overall, this painted a picture of learners "attempt[ing] to control the problem-solving task actively by constructing it verbally and orienting themselves to the language and task demands as they understand them" (Brooks & Donato, 1994, p. 271). This is what Vygotsky (1986) refers to as *regulation*, and the researchers argue that this regulation is not only visible in learners' talk, but also inherently cognitive.

DiCamilla and Anton (1997) suggest *repetition* as another type of semiotic mediation, or regulation, in learner interaction. In their study of interaction between adult language learners, they found that repetition was a pervasive feature in the learners' talk. Moreover, based on their in-depth analyses of the transcribed interaction, they argue that repetition can help scaffold the learners' performance. More specifically, DiCamilla and Anton suggest that repetition "distributes the scaffolded help throughout the activity, and thereby holds the scaffold in place, as it were, creating a cognitive space in which to work (e.g., think, hypothesize, evaluate), and from which to build (i.e., generate more language)" (1997, p. 627).

The description which DiCamilla and Anton provide is similar to observations made in a study by Donato (1994), that talk between learners can externalise cognitive processing, act to simplify the task, and in turn lead a group to accurate use of language features which any *one* of the students could not have managed on their own. However, Donato develops this into a fully-fledged application of the six features of scaffolding (cf. Wood et al., 1976). In his in-depth analysis of the transcribed talk of three university students' planning of French sentences (which they were to use in an upcoming task), he found the six features of scaffolding could be related to specific situations, as follows:

1. Recruiting interest in the task: This arose from learners' explicit requests for assistance, which would prompt other learners to think of ways to overcome the stated problem.
2. Simplifying the task: This feature seemed to arise in two ways: 1) it arose in a situation where the learners had to construct a very complex verb form, which a learner at this level would have difficulty processing by herself. The group-talk, however, externalised the processing, and their joint attention acted as a 'collective scaffold' facilitating the eventual correct use of the verb form; 2) on an individual level it arose through 'private speech'.
3. Maintaining pursuit of the goal: This feature also arose in two different ways: 1) it arose from the learners own questioning of each other's suggestions. This questioning would generate a continual 'stream' of goals and sub-goals for the group to pursue; 2) it arose when one learner's cues or hints (e.g. providing part of a word or sentence) enabled another learner to maintain her belief in ultimately being able to find/construct the relevant word or sentence.
4. Marking critical features and discrepancies between what has been produced and the ideal solution: This feature arose in a situation where each learner appeared to control a specific aspect of the needed language form, and where the learners, in a 'collaborative spirit', pieced together the full form.
5. Controlling frustration during problem solving: This feature arose from individual learners' confidence in being able to draw on the collective resources of the group.
6. Demonstrating idealized version: This feature occurred when one speaker knew the correct language form, allowing him/her to act as a 'tutor'.

A final approach associated with the sociocultural perspective is to assess the learner interaction in terms of how it relates to the learners' Zone of proximal Development (ZPD). The ZPD can be conceptualised as the 'space' between what a learner can do with the help of an expert, and what the learner can do by herself (cf. Vygotsky, 1978). The difficulty in applying this concept to learner interaction is that neither, or none, of the learners engaged in the learner interaction are likely to be experts (or otherwise they would not be learners).

However, in-depth analysis of extracts of transcribed data can reveal episodes of learner interaction where a first learner knows something that a second learner does not, making the first learner a 'de facto' expert, however briefly it may last.

A study by Ohta (1995), which analysed the transcribed role-play interaction between two adult learners of Japanese, illustrates such an analysis of the ZPD. In particular, Ohta's analysis showed how the expert-novice relationship would alternate between two learners. The relationship was in part determined by speaker roles and initiative, but also by which of the learners acted as the expert in the traditional sense, i.e., which learner could demonstrate relevant knowledge. Another observation Ohta made was that the higher proficiency learner seemed to maintain her interest in the task by frequently bringing extra elements and language into the role-play, thereby continually redefining the task goals. One could say that the higher proficiency learner moved the task into her own higher ZPD, all the while scaffolding her interlocutor's experience so that she would not leave him behind. This behaviour seemed to be highly deliberate in Ohta's data, as well as involving a high level of empathy for the other learner's experience of the task.

2.3.4 Critique and Implications for the Research Aims

The sociocultural research on learner interaction, including the research on the functions of output, represents a major shift from the negotiation of meaning and cognitive perspectives. The shift is away from frequency data and statistical analyses, to in-depth analyses of learners talk-in-interaction in episodes of transcribed data. Furthermore, in the studies which draw more directly on sociocultural theory there is a potentially important contribution to the present research aim. This contribution is related to Brooks and Donato's (1994) claim that learners' talk acts to mediate learner interaction. That is, the researchers claimed that learners' activity during a task was defined as much by their talk-in-interaction as it was by any *a priori* stated task requirements or instructions. This is a dynamical account because it recognises that learner interaction unfolds in the time-dimension, and suggests that learners' talk is central in this contingent and temporal process.

Another implication for the present research aim is related to the rich insights that the in-depth analyses of episodes of learner interaction provide. These in-depth analyses of learners' talk-in-interaction provide crucial understanding of the local dynamics that occur on the timescale of episodes. However, the focus on in-depth analyses of talk-in-interaction is not exclusive to sociocultural research. An early application of the negotiation of meaning perspective (cf. Varonis & Gass, 1985) also used in-depth analysis of episodes of transcribed data. However, in retrospect this served merely illustrative purposes, and the negotiation of meaning perspective quickly abandoned in-depth analyses in favour of frequency counts and

statistical analyses. The implication, then, is that in-depth analysis of extracts of transcribed data may help to validate whatever method is developed for representing and analysing the dynamics of learner interaction.

There is also a limitation associated with an exclusive focus on single episodes of learner interaction, whatever the focus of the in-depth analysis. The limitation relates to the ambiguity there may be in how the single episodes relate to the learner interaction as a whole. Long points to this limitation when he argues that sociocultural research is characterised by the lack of “descriptive statistics pertaining to the normalcy and variation of isolated cited examples and excerpts, and the consequent unknown typicality or status of those examples” (1997, p. 320). Needless to say, this is only a limitation if the analyses of episodes of learner interaction are claimed to be representative of something more than the episode itself. Such claims, however, are desirable, because they widen the appeal of whatever research findings one reports.

A related limitation is how the episodes are selected for in-depth analysis in the first place. In the research on the functions of output in learner interaction the selection has relied on the identification of language-related episodes, or LREs. Whereas this criterion improves the chances that the analyst will actually encounter the functions of output she is interested in, it also disconnects the selection from the activity that the learners are doing. It is interesting to note that the critical episodes used by Samuda and Rounds (1993; cf. discussion in sub-section 2.3.2), were identified as a feature of the task itself (each episode represented the learners’ negotiation of one difference). In relation to the task the learners were doing, this allowed Samuda and Rounds to make clearer claims about the typicality and status of the episodes they analysed. The more strictly sociocultural studies, by contrast, freely select episodes that contain those uses of learners’ talk that they are interested in. While this ensures that the selected episodes of interaction will contain whatever features the analyst is interested in, it also maximally exposes the research to Long’s critique about typicality and status of data, and severely limits the ability of these researchers to generalise beyond the episodes they have selected for analysis.

The implication of the above set of limitations, having to do with the focus and selection of episodes of data for in-depth analysis, is at present somewhat unclear. However, if the method for accounting for the dynamics of learner interaction incorporates episodes in some way, any findings will be subject to the same limitations as those outlined in the above. That is, unless the typicality and status of episodes can be revealed by the method.

A final limitation of the sociocultural research has to do with how the pairs, or groups, of learners, whose interaction is analysed, are selected. For example, the study by Swain and Lapkin (1998), reported at the very end of sub-section 2.3.2, selected the single pair whose written product was assessed to be the best one in an entire class. Other studies within the

sociocultural perspective are less transparent about how the participants in their research were selected. The concept of *representative sampling* is associated with more positivist research programs. Even so, the lack of any principled approach to selecting participants raises a question about the validity of both findings and insights. In the case that the eventual application of a dynamical account of learner interaction is limited to only a few number of participants, the implication should be that the selection of these participants should follow some principled procedure.

2.4 Implications for the Research

The review of established perspectives in research on learner interaction has yielded a number of implications for the present research aims.

The most important implication that has emerged is that existing findings lend some support for the development of a dynamical account of learner interaction. This is most clearly illustrated by the critiques of the negotiation of meaning perspective. That is, research within the negotiation of meaning perspective has tended to focus on a limited population of learners, to rely on less than full data sets, and to let research agendas control the settings and task conditions closely. The research by Foster (1993, 1998) and Jacob et al. (1996) indicates that these three ways of ‘controlling’ data collection and analysis may hide the unpredictable situational dynamics that are involved in interaction in actual classrooms. A method for analysing the dynamics of learner interaction could contribute insights about the extent to which this is the case. Furthermore, the critique of the task-based research more generally, by Coughlan and Duff (1994), and the unpredictable situational dynamics which sociocultural research takes as underlying, further validates the present aim of developing a method for analysing the dynamics of learner interaction.

Another implication of the observations made by Foster (1993, 1998) and Jacob et al. (1996), about the problems associated with controlling data collection too tightly, is that care should be taken how learner interaction data is collected for the present study. That is, the dynamical nature of learner interaction may be best expressed in data collected from a relatively unperturbed classroom setting. However, in doing so, the collection of data should follow clear procedures for selecting participants. This is to avoid the critique levelled at the sociocultural perspective for either selecting ‘good’ learners only, or otherwise being ambiguous about how ‘representative’ the participants are.

A further implication for the research aims relates to the critique of the cognitive research on learner interaction. That is, there was a level of detail where the effect of processing conditions on language use was both clear and convincing, and other levels where such clarity seemed to be lacking (cf. discussion in sub-section 2.2.3). As this may be a potential

limitation of any method of research, this should be a discussion point when outlining the potential contributions of visualisation as a method for researching learner interaction.

The potential value of making a close link to what the learners are actually doing, in meaningful terms, in the activity they are engaged in, is also an implication for the present research. The value of doing so is suggested by the ambiguous findings on the effect of different task types on the negotiation of meaning between learners. That is, it was suggested that these ambiguous findings might be related to how the negotiation of meaning research invariably defines task types in abstract terms (cf. discussion in sub-section 2.1.3). This may also relate to the lack of any principled way of establishing the typicality and status of episodes of interaction in sociocultural research. That is, it may be that by making a closer link between the analysis of episodes of interaction and what the learners are doing in an activity, such as was done by Samuda and Rounds' (1993) early use of critical episodes (cf. discussion in sub-section 2.3.2), would be a better approach to revealing the typicality and status of episodes of interaction. This, combined with a more concrete description of the activity the learners are doing, might make the effect of the task type involved more transparent.

Finally, the sociocultural research on learner interaction represents two potentially important contributions to the present research aims. The first contribution is how sociocultural research focuses on learners' talk as mediating, or regulating, learner interaction. The sociocultural research only uses the regulative function of learners' talk to explore the timescale of single episodes of interaction. Nevertheless, this suggests that a focus on the time-ordered organisation of learners' talk, or regulative activity, could be part of a dynamical account of learner interaction across the duration of classroom activities. This possibility will be pursued in more detail in the next chapter outlining a dynamical perspective on learner interaction. The sociocultural research also suggests the value of in-depth analysis of single episodes of learner talk-in-interaction. This can be combined with the observations made by Hawkins (1985) and Aston (1986) that no firm conclusions can be made about whether negotiation of meaning leads to comprehension without some other form of data or analysis to confirm this. Similarly, it might be difficult to make firm conclusions based on a visual method of analysis. Hence, the implication is that whatever may be the outcomes of visualisation, these outcomes may need to be validated by some form of in-depth analysis of the learners' actual talk during the classroom activity in question.