

A CRITICAL EXAMINATION OF L2 WRITING PROCESS RESEARCH

JULIO ROCA DE LARIOS, LIZ MURPHY, & JAVIER MARÍN

Universidad de Murcia (Spain)

Abstract. The present chapter is intended as a critical analysis of the most relevant recent research into the cognitive processes underlying second language written composition. After an introduction of the research domain, a number of relevant methodological aspects are briefly discussed. These include the data collection procedures used, the assessment of writers' command of the second language, the evaluation of written products, the context of the research, the type and number of participants involved, the type of tasks used, and the way reliability has been reported in the different studies. The substantive part of the research has been analyzed by isolating its main theoretical frames. Each of these frames has allowed us to derive a number of research sub-domains under which the studies have been grouped: the comparison of skilled and unskilled L2 writers, the development of L2 writing skill, the comparison of L1 and L2 writing processes, and the relationship between writing ability and L2 proficiency. A systematic analysis of the findings within each category has led us to identify a number of areas in need of further research: the notion of L2 writing skill, the formulation process, the temporal character of composition, the cognitive mechanisms involved in the transfer of writing abilities across languages, and the situated nature of L2 writing.

Key words: L1 and L2 writing processes, writing skill, transfer of writing skills, skilled vs. unskilled writers; relationship writing skill and L2 proficiency.

1 INTRODUCTION

Second language writing research has shifted its focus of concern among the three basic elements that comprise the writing activity: the socio-cultural context where the writer writes and learns to write, the text the writer produces, and the acts of thinking the writer engages in to produce such text (Cumming, 1998). These three theoretical orientations are usually referred to as writing as socializing, writing as product, and writing as process. The original drive behind the process movement was pedagogical in orientation: it was believed that for teaching procedures to be

effective, they should be based on a theory that accounted for what writers actually do in the process of composing a text. In this regard, the word process was understood in two different ways that correspond to two different trends within the process movement: the expressivist and the cognitivist (Faigley, 1986). For the former, writing was seen as the process that allowed an expression of the self provided the student was encouraged to write freely on those topics that matter to him/her. As such, it was basically a paradigm of L1 writing that had little effect upon L2 writing research (for further information, see Johns, 1990; Schreiner, 1997; Timbur, 1994). For the cognitivists, the word 'process' referred to the mental operations writers use when they are trying to generate, express and refine ideas in order to produce a text: this is the conception of process that we shall use to guide the review of the literature that follows. From this perspective, L2 writing research seems to have evolved into a research domain addressing a homogeneous set of problems which can be seen as the attempt to answer certain questions: (1) Do the reported findings on L1 writing processes apply equally to L2 composing processes? (2) If writing in a non-native language imposes extra constraints on learners, do these constraints create additional competing demands for attention? Do they also influence the type of problems attended to and the strategies used to solve them? (Manchón, 1997). From an applied perspective, answering these questions has been regarded as an essential step for teaching methods to be based on a sound theory that may help practitioners avoid the risks of blindly transposing recommendations for the teaching of L1 writing to the L2 situation (Silva, 1993; Zamel, 1983).

Individual studies within the domain have produced increasing amounts of data in response to those questions. Given their cumulative nature, summarizing them is a necessary precondition for indicating what remains to be done. In this respect, substantial accounts have been undertaken within the L2 writing process literature (Cumming, 1998, in press; Grabe and Kaplan, 1996; Krapels, 1990; Manchón, 1997, in press; Mc Donough, 1999; Silva, 1989, 1993) which, to a greater or a lesser extent, have adopted a critical stance as to the studies' methods, reporting and interpretation practices. The present review is an attempt to follow this critical tradition by approaching the domain covered by the research in a theoretically-motivated manner which allows us to present the research findings, discuss limitations and suggest areas for future research in a systematic fashion.

For this review, all relevant studies that were located, both published and unpublished in the last twenty years, were included for analysis (for a similar approach, see Silva, 1993). The total number of studies analyzed was sixty-five (see appendix). In reviewing these studies, we have identified critical aspects of the research domain by isolating the theoretical frames used and the different research areas derived from their explicit or implicit application. These will be presented, together with an account of the procedures used for their analysis, in section three, after a brief discussion of some methodological aspects of the studies that may affect the comparability of findings (section two). We will proceed in the next four sections with the presentation and discussion of the main findings within each research area, and conclude the study (section eight) with a discussion of the main limitations identified and an examination of some of the main problems we believe L2 writing process research should address in the future.

2 METHODOLOGICAL CONSIDERATIONS ON THE STUDIES REVIEWED

Variations in the methodology, purpose and quality of the studies reviewed impose some constraints on the interpretation of their results. Although a deep methodological analysis of the studies is beyond the scope of this chapter, the following considerations should be borne in mind when reading the sections that follow. First, data have been collected using a variety of techniques ranging, among others, from dual-task procedures, direct observation, analysis of written texts and computerized tracking devices to introspective methods, each of them with its own pros and cons in the potential to access and interpret data. Dual-task procedures deliberately seek to interfere with the train of thought by asking participants to carry out the writing task while simultaneously listening to irrelevant speech or recalling series of random digits (Ransdell & Levy, 1998; Ransdell, Arecco & Levy, 2001). Direct observation, text analysis and analysis via computers, although minimally disruptive of the writing process and adequately informative in some respects, fail to capture some of the covert process of writers at work. They cannot, for instance, provide information about how ideas are generated or what alternatives are evaluated. In contrast, introspection procedures (either on-line or retrospective) are reputed to provide a better picture of what goes on in the mind when tackling a task, although there are drawbacks associated with them, too. For one thing, on-line techniques such as the think-aloud method have raised doubts as to whether they may distort writers' normal composing processes or whether the process of composing aloud may in fact be distinct from that of composing silently (Bosher, 1998; Sasaki, this volume; see also Smagorinsky, 1994, for a review). On the other hand, both immediate retrospective techniques, such as stimulated recall (Bosher, 1998; Sasaki, 2000, this volume), and delayed retrospection techniques, in the form of questionnaires (Sasaki & Hirose, 1996) or interviews (Porte, 1997; Silva, 1992), rather than allowing access to actual mental processes, run the risk of showing 'the way participants situate themselves vis-à-vis a particular question or the person asking it' (Block, 2000: 760). In addition, most studies using on-line or retrospective techniques have not reported the segmentation criteria for the analysis of the resulting protocols, leaving the units of analysis in need of clarification (for exceptions, see Cumming, 1989; Manchón, Roca de Larios & Murphy, 2000; Porte, 1997; Roca de Larios, Marín & Murphy, 2001; Roca de Larios, Murphy & Manchón, 1999; Uzawa, 1996).

Second, although the participants' command of the second language has generally been reported (for exceptions, see Friedlander, 1990; Lay, 1982, 1988; Moragne e Silva, 1988; Ting, 1996), this variable has been assessed in a variety of ways which include institutional status, in-house assessments and standardized tests, each procedure representing different assumptions of what constitutes proficiency (Thomas, 1994). Many studies assumed that being registered for a certain academic course presupposed a specific level of L2 proficiency. The explanation usually given for using this type of institutionally-based assignment is the degree of verbal ability (Arndt, 1987; Fagan & Hayden, 1988; Tarone, Downing, Cohen, Gillette, Murie, & Dailey, 1993; Zamel, 1983) associated in principle with the course in question. This association is usually based on teachers' impressions (Fagan & Hayden, 1988; Miller, 2000; Pennington & So, 1993; Swain & Lapkin, 1995; Uzawa, 1996) or

exam systems (Chenoweth & Hayes, 2001; Henry, 1996; Valdés *et al.*, 1992). The procedure is thus highly economical since it requires no extra work on the part of the researcher or the participants other than registering the position of the latter in each course. Yet, standards in academic institutions are far from homogeneous and even among individuals in the same group there may be large disparities in L2 proficiency level. A second procedure used to assess participants' proficiency involves locally designed tests. This kind of in-house assessment, in contrast to the previous one, allows all participants within the sample to be tested uniformly and guarantees some kind of internal consistency. A typical example of an operational definition runs like this: 'We based level of linguistic competency on a comprehensive four-skills (reading, writing, speaking, and listening) test administered at the end of the students' second year of French studies' (Whalen & Ménard, 1995: 386). Although most studies using this procedure have isolated L2 proficiency as a variable in its own right, on some occasions the assessment of writers' L2 proficiency has been intermingled with that of their L2 writing ability (Hall, 1990), or has been pulled out of the grammar section of an otherwise writing-oriented test (Berman, 1994). In one particular case, the administration of this assessment type was supplemented with participants' self-evaluation of linguistic abilities (Faraco, Barbier & Piolat, this volume). Being locally developed, however, the results of in-house assessment cannot be extended in principle to learners outside the sample. In addition, together with institutional status, this type of assessment tends to establish arbitrary cut-off points between proficiency levels that may affect the comparability of findings. Finally, standardized tests represent multi-trait assessment procedures that are used to identify stages of overall language development, of which writing may be one part. These types of tests are taken to be one of the most objective ways to assess L2 proficiency because their content and validity are subject to on-going scrutiny, which helps to legitimate the generalizability of research results (Thomas, 1994). The tests used in the studies reviewed include the TOEFL (Carson & Kuehn, 1992; Devine, Railey & Boshoff, 1993; Hall, 1991; Kasper, 1997; Qi, 1998), the CELT (Kobayashi & Rinnert, 1992; Sasaki, 2000; Sasaki & Hirose, 1996), the MTELP (Raimés, 1985, 1987; Silva, 1992), the MELAB (Bosher, 1998), the IELTS (Shaw & Ting Kun Liu, 1998), the OPT (Manchón *et al.*, 2000; Roca de Larios *et al.*, 1999; Roca de Larios *et al.*, 2001), the SLEP (Sasaki, this volume), the Cambridge FCE (Victori, 1995), and Carroll's (1980) scale (Smith, 1994). One particular study (Jones & Tetroe, 1987) made use of both the Carroll and the TOEFL tests at different moments in their longitudinal design. In other cases (Ransdell & Levy, 1998; Ransdell *et al.* 2001) the TOEFL was used as the benchmark for participants' self-ratings of proficiency in English, and a further study (Akyel, 1994) made use of the TOEFL interchangeably with a locally designed test considered to be equivalent to the MTELP. Although in many studies informants were partitioned following the specifications of the test used, in some cases students were selected on the basis of further composition scores. One of the main drawbacks of standardized measures is that they are only available for English, leaving researchers with no other option but to rely on more local solutions when they have to assess writers' command of other languages. Another drawback is that they are essentially psychometric in orientation, which prevents them from assessing many aspects of communicative competence.

Third, variability is also observed in the way compositions have been evaluated. A few studies, in spite of claiming that their participants' written products had been analyzed, failed to report the instrument used (Zamel, 1983; Uzawa & Cumming, 1989). In contrast, many studies reported having used the ESL composition profile (Jacobs, Zingraf, Wormuth, Hartfiel & Hughey, 1981) either in its original version or modified to suit their research purposes (Bosher, 1998; Cumming, 1989; Cumming, Rebuffot & Ledwell, 1989; Pennington & So, 1993; Sasaki, 2000, this volume; Sasaki & Hirose, 1996; Uzawa, 1996). The TWE (Carson, Carrell, Silberstein, Kroll, & Kuehn, 1990; Silva, 1992), the SSQS (Ransdell & Levy, 1998; Ransdell *et al.*, 2001) and the HKCEE scoring procedure (Sengupta, 2000) were other standardized tests used less frequently. Other studies fell back on in-house instruments of either a holistic (Akyel, 1994; Friedlander, 1990; Kasper, 1997; Raimes, 1985; Skibniewski, 1988; Victori, 1995) or an analytical nature (Berman, 1994; Hall, 1991; Jones & Tetroe, 1987; Koda, 1993; Qi, 1998; Smith, 1994; Tarone *et al.*, 1993) or a mixture of both (Kobayashi & Rinnert, 1992). Special mention should be made of those studies which, deriving information on composing processes from the analysis of written texts, designed purpose-built text assessment categories with an emphasis on discourse aspects (Carson *et al.*, 1990; Carson & Kuehn, 1992), form and function (Cohen & Brooks-Carson, 2001), and conceptual and linguistic complexity (Yau, 1991), or have developed measures of overall quality, accuracy and fluency comparable to the ACTFL guidelines (Henry, 1996; Valdés *et al.*, 1992; Way, Joiner & Seaman, 2000). The use of such a variety of instruments indicates that there is no single standard of evaluation of writing quality for writing studies. This lack of homogeneity is further complicated if we bear in mind three considerations: (1) some of the rating scales used were originally intended for the placement or admission of learners into academic institutions, not for achievement purposes; (2) the common practice of rating students' compositions holistically has tended to conflate different dimensions of written productions into a single score, thus hiding the very elements that should be highlighted; and (3) the stance of the evaluator has a decisive influence on the score produced (see 8. 1. below). Unfortunately, although some work has been done on the objective indicators that should correspond to levels of L2 writing ability, there is still no consensus on the matter (Cumming & Riazi, 2000).

Fourth, the contexts of most studies fall into two distinct categories: second language and foreign language, although one study (Ransdell *et al.*, 2001) exceptionally included participants from both contexts. The difference is important because the presentation of self as conforming to social values is stronger in a second language than in a foreign language situation (Widdowson, 1983). In addition, each situation offers learners qualitatively different kinds of interaction, input and affective relations, as well as quantitatively different proportions of them, each of which may dramatically affect the acquisition and uses of the target language (Ellis, 1994). Approximately half of the studies reviewed here involve writing in a foreign language context and this tendency has gradually increased in the last ten years. Most studies in this context were devoted to English, but there were also several studies focused on French (Chenoweth & Hayes, 2001; Cohene & Brooks-Carson, 2001; Faraco *et al.*, this volume; Way *et al.*, 2000), Russian (Henry, 1996), German (Chenoweth &

New Directions for Research in L2 Writing

Ransdell, S.; Barbier, M.-L. (Eds.)

2002, X, 277 p., Hardcover

ISBN: 978-1-4020-0538-1