Studies in Second Language Acquisition **41** (2019), 503–527
doi:10.1017/S0272263119000275

Research Article

LANGUAGE REFLECTION FOSTERED BY INDIVIDUAL L2 WRITING TASKS

DEVELOPING A THEORETICALLY MOTIVATED AND EMPIRICALLY BASED CODING SYSTEM

Sonia López-Serrano*

University of Murcia/Pompeu Fabra University

Julio Roca de Larios

University of Murcia

Rosa M. Manchón

University of Murcia

Abstract

There has been a growing interest in the study of writing from the perspective of its potential contribution to language development. However, scant attention has been paid to key methodological considerations regarding the analysis of the connection between L2 writing processes, reflection on language while writing, and language learning. In an attempt to advance in this domain, and informed by models of L2 writing, and cognitive L2 writing research framed in the problem-solving paradigm, this study provides a comprehensive description of the language reflection individual writers engage in when solving the linguistic problems they face while completing writing tasks in their L2. The think-aloud protocols generated by 21 EFL learners while writing an individual argumentative essay were analyzed on the basis of a reconceptualization of language-related episodes as problem-solving strategy clusters. The result is a comprehensive, theoretically motivated, and empirically based coding system that is offered as a basis for future research in the domain. We discuss the methodological implications of our analytic approach and advance some theoretical implications for future debates on the language learning potential of individual writing tasks.

This study was financed by the Spanish Ministerio de Economía y Competitividad (research grant FFI2016-79763) and by Fundación Séneca, Murcia, Spain (research grants 19463/PI/14 and 20832/PI/18). We would like to thank the volume editors and the anonymous reviewers for their insightful suggestions for improvement of this article.

*Correspondence concerning this article should be addressed to Sonia López-Serrano, Pompeu Fabra University, Department of Translation and Language Sciences, Roc Boronat 138, 08018 Barcelona. E-mail: sonia.lopez@upf.edu

Copyright © Cambridge University Press 2019

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

Different positions originating within the second language (L2) writing and second language acquisition (SLA) fields have recently claimed that writing can provide learners with special opportunities to restructure, extend, and refine their L2 knowledge (Harklau, 2002; Ortega, 2011; Williams, 2012). Earlier, Cumming (1990) had posited that the natural disjuncture between the written product and the mental processes required for its generation and revision might help learners to focus on form-meaning relationships. The availability of time in the writing mode and the permanence of written texts have also been considered to facilitate leaners' use of explicit knowledge, attention to form as well as meaning, cognitive comparisons, and noticing the gap between their interlanguage and the L2 (Manchón & Williams, 2016).

These predictions have received empirical support in research on collaborative writing. Thus, it has been found that in collaborative writing conditions L2 users tend to pool their resources, reflect on and deliberate over alternatives, and give each other immediate feedback. As a result of these operations, learners may deepen their awareness of the relationship between meaning, form, and function and stretch their current linguistic resources toward their potential developmental level (Storch, 2013, 2016, for recent reviews of this body of work).

However, the controlled nature of the tasks used in research on collaborative writing and the oral character of peer interaction during collaboration may result in forms of linguistic and strategic processing that may not be directly extrapolated to the individual writing of complex texts (Manchón, 2011). Yet, individual writing may provide distinct favorable conditions for activating writers' own problem-solving processes, which are deemed to lead to L2 learning (Cumming, 1990). Surprisingly, the language learning potential of language reflection in individual writing has remained practically unexplored, with the notable exceptions of the pioneering attempts by Cumming (1990) and Swain and Lapkin (1995). With the present study we intend to start filling this gap through the provision of a theoretically motivated and empirically based coding system that attempts to capture in a comprehensive way the language processing activity learners engage in when solving the linguistic problems they face during individual writing tasks. A valid coding system is considered a necessary preliminary step for future explorations of the language learning opportunities afforded by individual writing tasks.

Although the ultimate aim of our work is methodological in nature, the study is also intended to contribute to our understanding of the nature of the language reflection associated with individual writing and to its language learning potential. Given the exploratory nature of our study, our reflections in this domain are presented as working hypotheses to be put to the empirical test in future studies.

THEORETICAL FRAMEWORK

Theoretical and empirical interest in the "writing-to-learn-language" dimension of L2 writing (Manchón, 2011) has grown exponentially in the last 10 years (as reviewed in Manchón & Vasylets, 2019; Manchón & Williams, 2016). A central concern in the domain is to ascertain what is unique about writing that can lead to advancing language competences. The theoretical prediction is that the pace and permanence of writing facilitates L2 writers' control of their attentional resources and prioritization of linguistic concerns. These processes are also deemed to be prompted by the problem-solving

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

nature of the act of composing: it is posited (see Byrnes & Manchón, 2014; Manchón & Roca de Larios, 2007 for further elaborations) that the deeper language processing associated with the problem-solving activity inherent to most forms of writing will facilitate engagement in crucial language learning processes, such as noticing, metalinguistic reflection, and analysis of explicit knowledge. The permanent nature of written texts makes them potentially useful for testing one's own knowledge of the L2 given that L2 users can engage in the process of cognitive comparison. This is so because "the cognitive window is open somewhat wider and learners have a richer opportunity to test their hypotheses when they write than when they speak" (Williams, 2012, p. 328). Finally, the visibility of the written text is also purported to aid L2 users' attempts to produce "pushed output" (Swain, 1995, 2005) as such visibility may prompt the L2 writer's noticing of either mismatches between intended meaning and its linguistic realization, or the need to improve such linguistic output (Manchón & Williams, 2016).

Cumming (1990) and Swain and Lapkin (1995) were pioneering attempts to look into writing and language learning. Drawing on Schmidt's (1990) noticing hypothesis and Swain's (1985) output hypothesis, these studies assumed that the "problems that arise while producing the second language can trigger cognitive processes that are involved in second language learning" (Swain & Lapkin, 1995, p. 371). Through qualitative analyses of what Swain and Lapkin (1995) labeled as *language-related episodes* (LREs), it was concluded that L2 writers frequently face problems when trying to express their thoughts in writing, primarily as a consequence of having to pay attention concurrently to both the ideas to be expressed and the linguistic representations of these ideas. As a result, L2 writers activate search processes (accessing and retrieving lexical and morphosyntactic items from memory and constructing connections between them) that enable them to generate alternative linguistic options and to assess the forms generated. These search and assessment cycles may also result in the modification and consolidation of L2 knowledge or even in the internalization of new knowledge (ibid.).

Surprisingly, despite the obvious relevance of these insights for the relationship between cognitive activity while writing and L2 learning, subsequent research on L2 writing processes paid little attention to such relationship and to the affordances of individual writing processes from the perspective of language learning. As a result, although the completion of writing tasks has been viewed as necessarily involving different degrees of problem solving, very few empirical attempts have resulted in explicit coding proposals that can account for the combined activation of strategies when writers try to solve each individual problem addressed (for exceptions, see Roca de Larios, Murphy, & Manchón, 1999, and elsewhere). Similarly, Cumming's (1990) original distinction between "more" extensive and "less" extensive searches, and Swain and Lapkin's (1995) distinction between "complex" and "simple" thinking seemed to indicate that individual writers' attention to language does involve different levels of engagement and processing. Yet, these levels have not been further elaborated, either theoretically or empirically. In parallel, the few theoretical models of cognitive processing in individual L2 composition (e.g., Chenoweth & Hayes, 2001; Grabe & Kaplan, 1996; Wang & Wen, 2002; Zimmermann, 2000) have paid scant attention to the lexical and grammatical processes involved in the conversion of thoughts into linguistic form, and even less attention has been devoted to whether such processes are conducive to language learning.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

THE PRESENT STUDY

It follows from the arguments presented so far that it is highly relevant to explore further individual writing in connection with its language learning potential. However, to conduct empirical studies that explore such relationship, we first need a comprehensive coding system that captures the language reflection fostered by individual writing. Hence, the setting up of such coding system may be regarded as a theoretically and empirically relevant endeavor. In our view, this task ought to be accomplished in three ways. First, the analyses should be situated in the inner workings of the text-generation processes responsible for translating ideas into linear strings of language. This is so because, as mentioned previously, text-generation processes are purported to be crucially linked to the language learning that may ultimately occur (Kormos, 2012; Manchón, 2014; Roca de Larios, 2013). Second, progress can be made in the domain if the translation of ideas into linear form is approached through a conceptualization of strategies as mental actions that come in clusters (against prevailing atomized views of strategies) and that give an added value to learners' linguistic knowledge when they are flexibly and adequately orchestrated in the course of the activity (Macaro, 2014). Finally, this conceptualization of strategies, we suggest, ought to guide the analysis of strategy use in relation to the language problems addressed by learners. This approach should facilitate ending up with a coherent, process-oriented description of the levels of processing involved (Leow, 2015).

In addition, the previously mentioned analytic dimensions ought to be framed in a relevant model of written production. Kellogg's (1996) model of writing[1] was selected for this purpose. Originally propounded as a framework to account for L1 writing processes, the model has nevertheless been used as a suitable theoretical framework for the analysis of L2 writing processes on the basis of its descriptions of the linguistic encoding processes that are supposed to generate considerable cognitive demands for L2 writers (e.g., Révész, Kourtali, & Mazgutova, 2017). Kellogg emphasizes the dynamic character of written text production by assuming that three main subprocesses—formulation, execution, and monitoring—interact with one another in a nonlinear, recursive fashion.[2]

Formulation involves, first, setting up goals as well as generating and organizing ideas so as to elaborate a coherent plan for the content of the text and, second, translating these ideas into linguistic form through an "amalgam of linguistic processes" (Kellogg, 1996, p. 61) that include lexical access and retrieval, syntactic framing, expression of cohesion, and conversion of phonemes into graphemes. Translation, as mentioned earlier, is inherently related to the natural disjuncture between written products and the mental processes needed for their generation (Cumming, 1990). In this respect, Kellogg argues that writers may mentally try out sentences or parts of sentences before writing them down. Depending of its degree of completeness, this tentative process would correspond either to the notion of "inner speech" (Vygotsky, 1962), when the output of translation is reduced to word meanings and their phonological representations with a very sketchy syntactic structure, or to the idea of "pretext" (Witte, 1987), when the output is more developed at both lexical and syntactic levels. The output of translation is programmed for use in the execution phase through the appropriate activation of the motor movements required to produce a handwritten or typed piece of text.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

Finally, monitoring allows writers to ascertain whether the output of previous processes in the form of ideas conforms to their intended meaning. When mismatches are identified, feedback is sent to the process(es) in question, that is, planning (in the case of text-organization problems) or translation (in the case of spelling, lexical, or syntactic problems). This means that, as is the case with translation, monitoring loops can occur both before and after the text is transcribed. As we will explain in the method section that follows, our approach to the analysis of writing processes and their potential for language learning takes into consideration the three writing subprocesses suggested in this model. Following Izumi (2003), we have assumed that these subprocesses may provide learners with successive representations of their output broken down into its semantic, syntactic, morphological, and phonological dimensions by means of feedback loops. As mentioned previously, the slow pace of writing and the permanent record left (Williams, 2012) are supposed to make it easier for learners to use these loops to reflect on their output representations, compare them with their knowledge sources, and notice, if appropriate, holes or gaps in their interlanguages.

In what follows we describe how we proceeded while setting up our coding system for language reflection. The raw data, which were taken from a previous research program (Manchón, Roca de Larios, & Murphy, 2009, for an overview) and reanalyzed according to our current research purposes, were the think-aloud protocols generated by 21 EFL learners while completing an argumentative writing assignment under time-constrained conditions.

METHOD

PARTICIPANTS

The participants were 21 Spanish learners of English as a foreign language (EFL), aged 16–24 with a pre-intermediate to advanced level of L2 proficiency, as measured by the Oxford placement test (OPT) (Allan, 1995). The pre-intermediate group consisted of seven high school students aged 16–17. Their OPT scores ranged from 100 to 108 (B1 in the Common European Framework of Reference for Languages [CEFR]). The intermediate group was composed of seven university students, aged 19–20, who were in the third year of a degree in education and whose scores were between 140 and 157 (B2 in the CEFR). Finally, the advanced group consisted of seven participants aged 23–24. They had an OPT score of between 174 and 190 (C1 in the CEFR) and they had just graduated from a five-year degree in English studies. This division into proficiency groups allowed us to obtain a more comprehensive sample of the different dimensions of analysis of LREs in individual writing, hence facilitating generalizability of findings across proficiency levels.

TASK AND DATA-COLLECTION PROCEDURES

The participants were asked to complete an argumentative composition individually under think-aloud conditions based on the following prompt (provided in English):

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

Success in education is influenced more by the student's home life and training as a child than by the quality and effectiveness of the educational programme. Do you agree or disagree? (Raimes, 1987, p. 465).

Although the think-aloud technique is not without limitations (see Bowles, 2010), it is widely accepted that verbal protocols provide a useful and valid window into processes and processing during task completion, in general, and writing tasks, in particular (see Leow, 2015; Manchón, Murphy, & Roca de Larios, 2005).

Due to practicalities, the participants were audio recorded in groups in the language lab. Two people were present during the recording sessions: the English teacher of that group (or ex-teacher in the case of the most proficient group) and a member of the research team. This procedure was a safeguard against differences between the recording sessions and in principle established a less threatening environment. To familiarize participants with the think-aloud technique, they were asked to verbalize everything that came to their minds while writing a mock composition during a practice run that lasted about 20 minutes. Deliberately, no modeling was provided so as to prevent them from restricting their verbalizations to the type of examples they had seen modeled (Smith, 1994). They were told that, although they had to write their texts in English, they were free to use the language of their choice while thinking aloud. Once they were accustomed to thinking aloud, they were given the writing prompt and were told they had one hour to handwrite their texts. They were audio recorded in individual booths, and no dictionaries or requests for help were allowed because the main aim of the study was to explore the ways in which the students tried to solve the linguistic problems they faced in the course of writing.

DATA CODING

The 21 think-aloud protocols that were generated were transcribed using ordinary spelling conventions. They were then subjected to a qualitative analysis through an iterative process described in the following text (see Figure 1). The point of departure for this analysis was Swain and Lapkin's (1995) original definition of LRE in individual writing:

any segment of the protocol in which a learner either (i) spoke about a language problem he/she encountered while writing and solved it either correctly or incorrectly [or left it unresolved] or (ii) simply solved it without having explicitly identified it as a problem (p. 378).

Identification of LREs and Their Linguistic Focus

Two coders (first and second authors) proceeded to identify the LREs in six of the think-aloud protocols, two per proficiency group. The first step was to identify the sections in the protocols in which writers were focused on formulating or revising text (as opposed to processes such as reading the prompt, organizing the text, or generating ideas). Then, our participants' verbalizations of linguistic concerns were isolated. To that end, we looked for examples of metalanguage, defined as "all language that is used to talk about language" (Fortune, 2005, p. 26), paying particular attention to problem indicators, which varied in terms of their degree of explicitness. For the most part, these indicators

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

included explicit comments, such as overt questions (How do you say…?), remarks revealing lack of knowledge of the L2 (I don't know how to say…), or questionings and evaluations (I'm not sure if this is right). Problem indications were at times more implicit, such as verbalizations of inner speech (Let's see…) or suprasegmental features (e.g., intonation). In addition to problem indicators, LREs could be preceded by goal-setting comments, such as "Let's try to find a better word."

For each problem indicator or goal-setting comment, the linguistic unit that constituted its area of concern (from spelling to suprasentential levels) was identified. Following previous research (e.g., Fortune & Thorp, 2001; Swain & Lapkin, 1995), it was decided that each independent LRE (a) would focus on only one specific language concern and (b) would end when the writer provided a written solution to the problem at hand (or left

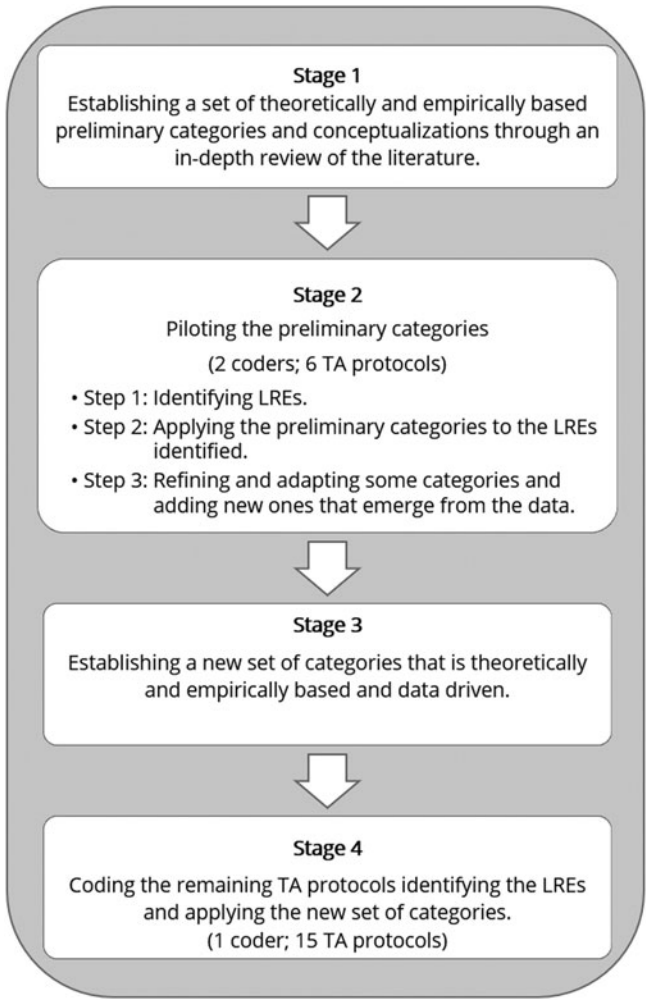


FIGURE 1. Stages in the data coding process.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

it unresolved) and moved on with text composition. If attention was temporarily directed to a different area of linguistic concern while engaged in a particular linguistic problem, it was counted as a different LRE, which was also isolated and analyzed as an *embedded* LRE.

Following these criteria, the two coders individually identified the LREs in the six selected protocols (322 and 310 LREs, respectively). An agreement estimate was conducted that indicated 92.7% of intercoder agreement. Discrepancies were discussed and collaboratively resolved. Once all the episodes in the six protocols were isolated, they were subjected to subsequent coding. The initial categories and conceptualizations for this coding are explained in the following text and graphically presented in Figure 2.

Coding Procedures: Resolution and Nature of LREs

The review of earlier empirical research evidenced that, in addition to their linguistic focus, the analysis of LREs should include two dimensions hardly considered in existing work investigating individual L2 writing processes: (a) their resolution and (b) the nature

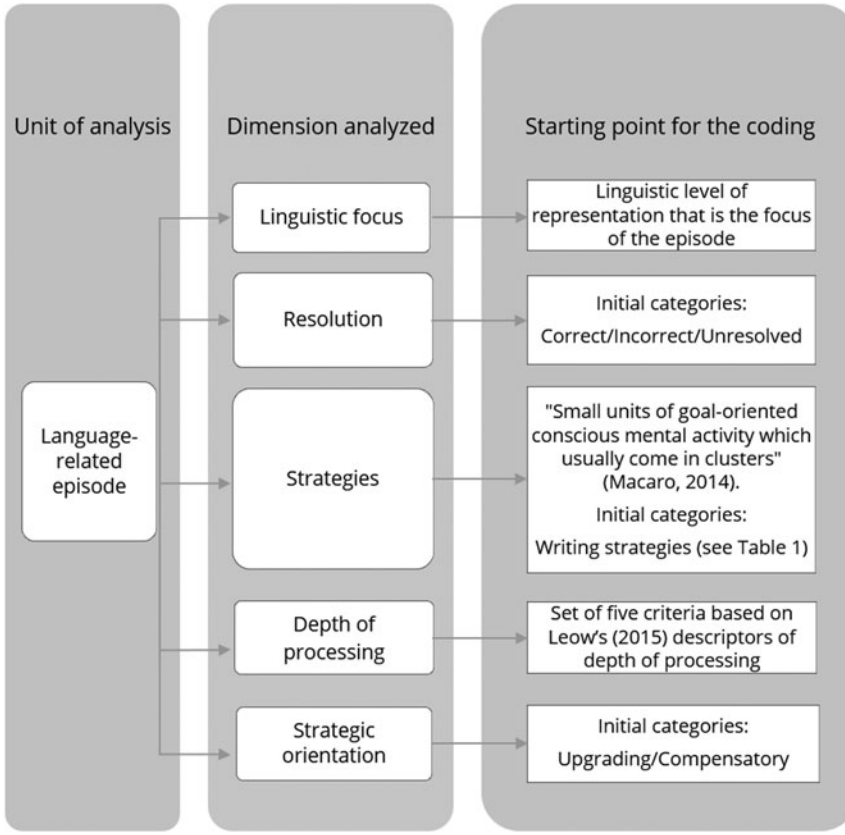


FIGURE 2. Initial categories and conceptualizations used in the coding process.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

of linguistic reflection. For the resolution of LREs, previous analyses of collaborative LREs (e.g., Fernández-Dobao, 2012) served as the basis to establish our initial categories *correct*, *incorrect*, or *unresolved*, which we then tested against our data. The final decision on correctness was based on the characteristics of the final text written down, that is, without considering whether other forms verbalized within the LRE would have been correct or acceptable. For the exploration of the nature of linguistic reflection involved in the different episodes, we assumed that three aspects were to be considered: (a) the writing strategies used, together with the possibility of considering the combination of strategies in each LRE as a problem-solving unit; (b) the orientation of the strategies used in each unit; and (c) the depth of processing demonstrated in each LRE.

Strategies

Our analysis of strategies relied on the categories propounded in previous research to account for L2 formulation processes in individual writing, as shown in Table 1.

With these sets of categories in mind, each episode identified in the six protocols was analyzed as follows (see exemplification in Table 2): first, we coded each language form in focus within the episode, taking into account whether it was a "text" (form written) or a "pretext" (form verbalized but not written) (Witte, 1987) and whether it had been produced in the learner's mother tongue or in the L2. Second, we identified (a) every strategy applied, (b) the knowledge sources and reasons verbalized during the decision-making process (see Step 6 in Table 2), (c) the technical metalanguage involved in these decisions (if any), and (d) the changes introduced in the written text. By way of illustration and for clarification purposes, the example shown in Table 2 has been broken down into these different steps.

TABLE 1. Writing strategies identified in previous research used as starting point in the coding

Generation and selection of **lexical alternatives** (rehearsal, lexical searches, rhetorical refining).	Cumming (1990); Murphy & Roca de Larios (2010); Raimes (1987); Sasaki (2000, 2002, 2004); Swain & Lapkin (1995).
Producing and comparing **cross-linguistic equivalents**, synonyms and translations.	Cumming (1990); Macaro (2007); Murphy & Roca de Larios (2010); Swain & Lapkin (1995); Wang (2003).
Generation and selection of **syntactic alternatives** (rehearsal, restructuring, and recombining of constituents).	Macaro (2007); Raimes (1987); Roca de Larios et al. (1999).
Use of explicit knowledge (reasoning, knowledge of linguistic rules).	
Use of implicit knowledge (assessing linguistic choices through intuition or appeal).	Cumming (1990); Swain & Lapkin (1995).
Evaluation and revision (assessing in relation to a criterion, standard or rule, evaluating the formulation cluster).	Cumming (1989); Macaro (2007).

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

TABLE 2. Example of an LRE broken down into its component steps.

LRE	Steps	Coding
as I **started** ...	1	L2T1: Generation of forms
no, started not,	2	L2T1: Evaluation: Negative
[CROSSES OUT started]	3	L2T1: Repair: Deletion
as I learnt?	4	L2PT2: Generation of forms
Learnt or began to learn?	5	L2PT2/L2PT3: Considering alternative forms
No, I don't like that one	6	L2PT3: Evaluation: Negative - Appeal
as yes *I I* learnt	7	L2PT2: Evaluation: Positive
(3) How can I say it here?	8	Pause + Problem indicator
This is a very relevant aspect in my case, *as I*, *as I* (6)	9	Rereading + Pause.
was yes	10	L2T4: Generation of forms: Evaluation: Positive
(3) have been	11	Pause + L2PT5: Generation of forms
was no	12	L2T4: Evaluation: Negative
[CROSSES OUT was]	13	L2T4: Repair: Deletion
have been	14	L2PT5 > L2T5
as I have been, I have been	15	L2T5: Sounding out
an **eager learner since my early childhood**	16	L2T5: Message elaboration
yes! Good!	17	L2T5: Evaluation: Positive

Notes: The following conventions were used: all words in normal font are TA talk by the writer; segments in **bold** indicate text written down; *italics* indicate re-readings and repetitions of that text; numbers in parentheses indicate pauses of 3 seconds or more; pauses of less than 3 seconds are shown by dots; statements inside square brackets record changes in the text; pretexts were coded as PT and texts as T (these codes were preceded by the labels L1 or L2, and followed by a figure that indicated their order in relation to the total forms generated in the episode).

This analysis of the individual steps within each LRE showed that, as anticipated by Macaro (2014), strategies rarely appeared in isolation. Quite the contrary, they frequently worked in clusters activated to solve the language problem at hand. It was thus decided that each cluster of strategies could be accommodated within the notion of problem space. Following Newell (1980), an LRE was considered a language-related problem space and, accordingly, it was conceived of as (a) a gap (of various types and a variety of levels) between an initial state (the way the writers represented the linguistic problem to themselves in the first place) and a goal or final state (the possible solution to the problem), and (b) a sequence of strategies or steps leading from the initial state to the goal state.

Depth of Processing

Accepting that deeper engagement with language prompts more learning opportunities (e.g., Leow, 2015), we decided to analyze whether the LREs identified in our data showed different levels of processing. For this purpose, we turned to Leow's (2015) recent proposal of depth of processing, conceptualized as "the relative amount of cognitive effort, level of analysis, elaboration of intake together with the usage of prior knowledge, hypothesis testing and rule formation employed in decoding and encoding

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

some grammatical or lexical item in the input" (ibid., p. 204). Accordingly, every episode was analyzed in terms of the following five indicators:

1. Length of the LRE and pausing behavior.
2. Number of alternatives generated and assessed (both in the L1 and the L2).
3. Analysis and manipulation of different levels of linguistic representation (from analyses of discrete elements such as morphemes to syntactic and semantic analyses).
4. Amount and variety of strategies deployed to solve the LRE.
5. Use of metalanguage revealing: (a) connecting items to rules or to prior knowledge of the L2, (b) analyzing form-meaning relationships (L1-L2 or L2-L2), and (c) formulating hypotheses about the L2.

The length of an LRE was taken as an initial clue of the participant's potential engagement in solving the problem at hand: a short LRE would usually involve fast access to a solution or the abandonment of the search process. Pausing behavior was carefully interpreted in the context of each LRE; long (10 seconds or longer) or repeated pauses were considered an indication of possible covert linguistic reflection only when most of the other parameters of engagement were found in the episode. The number of alternative forms formulated and assessed (both in the L1 and the L2) were interpreted following Kellogg's (1996) model of writing mentioned in the preceding text. Accordingly, we assumed that the higher the number of forms generated, the more the participants accessed their long-term memory and the higher the number of loops through the monitoring system. The syntactic and semantic analyses in which manipulations of different structures were closely related to the achievement of the participants' intended message were considered instances of deeper processing, being informed here by theoretical approaches such as the involvement load hypothesis (Laufer and Hulstjin, 2001) and Craik and Lockart's (1972) levels of processing model. In turn, the use of different strategies to solve the problem at hand was interpreted in relation to Craik's (2002) concept of elaboration, that refers to the ways in which language processing can be enriched through the use of "within-item elaboration" strategies (reflecting on different characteristics of the item, e.g., physical, semantic) and "between-item elaboration" strategies (e.g., accessing previously stored knowledge and establishing semantic relations with other items, appealing to episodic memory). Finally, instances of metalanguage use were regarded as instances of explicit L2 knowledge, a type of knowledge that has been connected by some scholars with higher levels of awareness and with the creation and strengthening of form-meaning relationships (Kormos, 2006, 2011).

Strategic Orientation

Finally, our conceptualization of LREs as problem spaces in which strategies can work in clusters allowed us to add a new dimension to the analysis, namely the exploration of whether LREs evidenced a particular *strategic orientation*. The orientation of an LRE would be defined as the purpose for which writers engaged in strategies from the initial to the final state of the problem space (e.g., Murphy & Roca de Larios, 2010; Roca de Larios, et al., 1999). Previous studies had identified two main orientations of strategy

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

clusters: (a) compensatory (making up for a lack of lexical knowledge or difficulty in accessing the necessary items), and (b) upgrading (refining and improving the already available L2 structures). Hence, we speculated that the semantic, stylistic, and linguistic concerns involved in the upgrading orientation of strategy clusters, as compared to their compensatory orientation, might trigger the expression of more complex or precise ideas. In this respect, we accepted the theoretical claim that these ideas are supposed to activate lemmas associated to more complex or precise syntactic encoding procedures and that, as a result, they may be taken to foster vocabulary, morphological, and syntactic development (Kormos, 2011). Consequently, we observed the use of strategies from the initial to the final state of each problem space to ascertain whether they were used for compensatory or for upgrading purposes.

The coding of these different dimensions was conducted by the two coders on the six protocols mentioned previously under the principles of collaborative coding (Smagorinsky, 2008). This involved the use of both inductive and deductive approaches that allowed us to apply, refine, or adapt some of the categories established in previous research as well as to create new ones with the aim of capturing the nature of individually produced LREs. As explained, discrepancies between the two coders were resolved through discussion until complete agreement was achieved. The remaining 15 protocols were analyzed by the first author, who identified and coded all the LREs occurring in them. The 21 protocols contained 1,044 LREs, with an average of 50 episodes per participant. The categories that emerged from this theoretically motivated and empirically based coding process are presented in the text that follows.

RESULTS

In what follows we present the dimensions that make up our multidimensional coding system of LREs, namely (a) linguistic focus, (b) resolution, (c) strategies, (d) orientation, and (e) depth of processing. The list of subcategories can be found in Figure 3. Three LREs (one from each proficiency level) are provided in Figure 4 and commented to illustrate how the categories are applied and work in combination. Readers are referred to the Online Supplementary Materials (OSM) for a detailed exemplification of all the subcategories (Tables 3–8) and a checklist to aid in the analysis of LREs (Table 9).

LINGUISTIC FOCUS

The data showed that the LREs in our data involved a multiplicity of levels of linguistic representation, ranging from spelling to discourse concerns (see Table 3 in the OSM). Regarding morphology, our participants produced LREs dealing with both bound morphology (particularly focusing on verb forms and tenses and word number) and free morphemes (which included concerns about prepositions, determiners, and conjunctions). Their most frequent concerns were at word level, with a large number of the episodes focusing on lexical issues. This category included all the LREs that dealt with the four types of free, open-class morphemes (verbs, nouns, adjectives, and adverbs) (Examples 1 and 2). In addition, student-writers' LREs frequently dealt with small groups of words constituting phrases and phrase fragments. Probably due to the open nature of the task, our participants also paid attention to larger structures such as clauses and sentences (Example 3). Dealing with these structures usually involved the

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

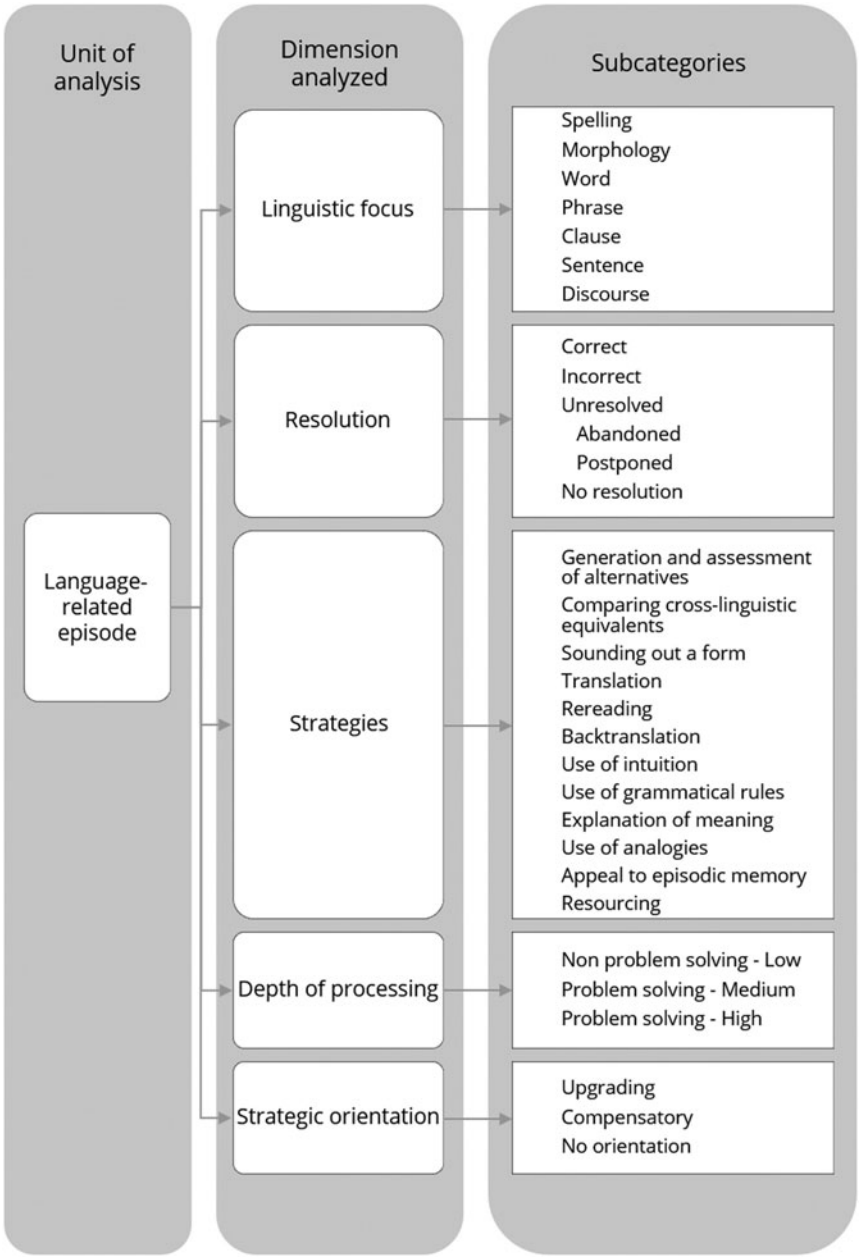


FIGURE 3. Dimensions and subcategories in the coding system.

occurrence of *embedded* LREs that focused on specific words or phrases that were part of the clauses (see Table 4 in the OSM). The analysis of the protocols also indicated that our participants were aware of the importance of discourse concerns, such as maintaining coherence and cohesion or avoiding unnecessary repetitions throughout the text.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

RESOLUTION

Three main categories in the resolution of LREs were identified: correct, incorrect, and unresolved. The majority of the episodes in our data had a *correct* outcome, despite the fact that our participants had no external resources available (Examples 1 and 3). *Incorrect* episodes were those in which the participants produced solutions that differed from the L2 norms (Example 2). LREs were left *unresolved* for two main reasons: linguistic and nonlinguistic. In the first case, the participants felt that they were not able to find a solution to their problem, whereas in the second case they stopped their problem-solving behavior and moved on due to external reasons such as task management concerns (e.g., awareness of time constraints). *Unresolved* LREs included *abandoned* and *postponed* LREs. While the former was subsequently forgotten, postponed LREs were frequently revisited and writers engaged in problem-solving several times until they found what they thought was a correct answer. Therefore, postponed LREs usually generated *recurrent* LREs that sustained learners' attention to form during the whole task (Fortune & Thorp, 2001).

Finally, it is important to note that not every LRE necessarily had to finish with a clear outcome; some LREs were reflections on form that did not involve the provision of a solution (e.g., evaluations of the already written text) (see Table 5 in the OSM).

STRATEGIES

The strategies employed by our participants to solve their LREs are listed here (examples of each strategy are provided in Table 6 in the OSM):

- Generation and assessment of forms (through the L2, the L1, or both)
- Comparing cross-linguistic equivalents
- Translation
- Sounding out a form or trying it out in context
- Rereading
- Backtranslation
- Use of intuition
- Use of grammatical rules or technical metalanguage
- Explanation of meaning of a word or expression
- Use of analogies
- Appeal to episodic memory (i.e., recollection of specific learning experiences)
- Resourcing (their own notes or the prompt)

As mentioned in "Method," these strategies were usually applied in combination to solve the LREs (see examples in Figure 4).

DEPTH OF PROCESSING

Our participants engaged in different levels of linguistic reflection while writing their texts and their LREs evidenced three types of behaviors: (a) non–problem solving (NPS), (b) problem solving with a medium depth of processing (PS-M), and (c) problem solving with a high depth of processing (PS-H). NPS episodes corresponded to instances of momentary attention to language that prompted an immediate application of knowledge

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

in an attempt to generate the necessary L2 form or to evaluate the form produced. In these episodes, the first alternative formulated was considered faulty by the monitor because it was an L1 item, an incorrect L2 item, or two competing forms were activated and one needed to be discarded. A solution (that may be overtly or covertly monitored) was retrieved and no further analysis was conducted. Therefore, these episodes were considered low in terms of their level of processing.

In turn, problem-solving LREs occurred when a linguistic concern was identified and, to find a solution, the student-writer needed to consciously activate a number of strategies to solve the problem at hand. Two main levels were identified: medium and high depth of processing. PS-M episodes were characterized by the deployment of a low number of strategies and the generation and evaluation of a small number of alternatives (usually a maximum of three) (Examples 1 and 2). PS-H episodes involved spending a considerable amount of time trying to solve the LRE, producing and evaluating a relatively high number of forms, and deploying a series of strategies in a flexible way. In turn, they could also contain manipulations of long syntactic structures that prompted syntactic and semantic analyses of the smaller components of such structures. The reasoning employed to evaluate the alternatives available (if any) usually contained metalanguage that showed reflection on cross-linguistic equivalents, formulation of hypotheses, and application of L2 rules to achieve complex rhetorical purposes (Example 3) (see Table 7 in the OSM).

ORIENTATION

As evidenced in previous research on L2 writing processes, our participants engaged in problem-solving processes for both compensatory and upgrading purposes. However, there were episodes in which no clear orientation could be identified, especially when the texts were evaluated or translated from their L1.

Our participants tried to compensate for their lack of linguistic L2 resources (Examples 1 and 2) by orienting their strategy clusters toward a variety of goals which, in consonance with the terminology used by Dörnyei and Kormos (1998), included any of these trajectories in each LRE:

- Using a superordinate or a general term as a substitute for a target item (approximation)
- Using an L1 or L3 term and adjusting it to L2 pronunciation or morphology (foreignizing)
- Creating a nonexisting L2 word by applying L2 formation rules (coinage)
- Inserting full L1 words within their L2 texts (code switching)
- Translating literally a word, a phrase, or a clause from the L1 into the L2 (literal translation)
- Exemplifying or illustrating the attributes or properties of the target item (circumlocution)
- Reducing certain grammatical specifications of the lemma to make it more in accordance with the writer's linguistic repertoire (grammatical reduction)
- Transferring to the lemma certain grammar specifications and argument structures from the L1 (grammatical substitution)
- Deleting parts of the message to avoid problematic linguistic items or topics (message reduction)
- Replacing parts of the message with new ones (message replacement)
- Giving up the preverbal message as a whole (message abandonment)

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

EXAMPLE 1

mucha disciplina mucha ¿mucha? ¿mucha? ¿mucha? more (...) más eso es más more es más too much ¿no? ¿sería? demasiado, eso es, demasiado pongo demasiado *yo he tenido* **too much discipline** (P-INT7-E23)

Focus	Resolution	Strategies	Depth of processing	Orientation
Word	Correct	Generation of forms (L1) (x2) Generation of forms (L2) (x2) Comparing cross-linguistic equivalents Backtranslation	Problem solving: Medium	Compensatory

EXAMPLE 2

I have a lot of **friends who are** (...) las víctimas venga creo recordar que víctimas era eso de casualties del accidente pero aquí no sé yo si eso se podrá decir (...) sí, eso es solamente para accidentes, entonces (5) [CROSSES OUT are] *quienes who* que han sufrido (...) *have a lot of friends who* **have suffered** suffered *have suffered* **negative effect from the educational system** (INT7–E17)

Focus	Resolution	Strategies	Depth of processing	Orientation
Word > Clause	Incorrect	Generation of forms (L1) Generation of forms (L2) Explanation of meaning Repair Backtranslation Generation of forms (L1) Rereading Translation Generation of forms (L2) Circumlocution	Problem solving: Medium	Compensatory

EXAMPLE 3

aprovechándose de lo que tiene bueno y descartando lo que tenga de malo (4) *but of course you can try to adapt yourself to it* **by taking advantage** taking advantage of (6) of (3) de cuantas cosas buenas pueda ofrecer *taking advantage* **of whatever** whatever (3) o sea *puedes intentar adaptarte a él aprovechándose de cuantos* aspectos positivos pueda ofrecer of *whatever* favourable vamos a cambiar un poco favorab **favourable aspects it may offer** it may offer **to you** *que pueda ofrecerte* and y lo contrario **and** avoiding evitar **avoiding those which** que no puedan serte beneficiosos *which those which* **can be** be **of no benefit** benefit **to you** (ADV5-E58)

Focus	Resolution	Strategies	Depth of processing	Orientation
Clause	Correct	Generation of forms (L1) Rereading Translation Generation of forms (L1) Translation Backtranslation Message elaboration Translation Backtranslation Generation of forms (L1+L2) Translation Message elaboration	Problem solving: High	Upgrading

FIGURE 4. Examples of language-related episodes.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

The data also indicated that LREs were oriented toward upgrading the texts produced in terms of ideational and discourse-related concerns (Example 3). The former took different forms, ranging from changes to the informational focus of a sentence (reconceptualization) and turning potentially ambiguous or incomplete initial statements into clearer or more informative ones (elaboration), to episodes in which participants simply made sure that the message conveyed was true (ensuring truth value). Upgrading orientation of LREs was also visible at a discourse level. Writers were found to (a) control the coherence-cohesion of the discourse by handling the connections that link propositions or clauses, or changing the functional role of clauses/sentences; (b) adjust the message to the register requirements of the written language, avoid repetitions, and upgrade initial formulations in translation; and (c) manipulate the structure of the information to be conveyed by means of postponement of information and manipulation of topic-comment relations (see Table 8 in the OSM for further examples).

GLOBAL ANALYSIS OF THREE LRES

As noted previously, one of the novel approaches in our analysis was the reconceptualization of LREs as problem spaces in which strategies were put to work in combination to solve the linguistic concern at hand. In what follows, we present three examples analyzed as global units taking into account all the dimensions included in our coding system.

Example 1 illustrates a compensatory lexical search conducted by a pre-intermediate participant. In this search, she generates a short series of L1 forms that she briefly compares with their L2 equivalents. In particular, the writer formulates the L1 phrase "mucha disciplina" (a lot of discipline) but immediately realizes that she cannot retrieve the L2 equivalent for the Spanish "mucha." Probably due to similarities in the initial sound "m," the participant has access to the word "more" that she immediately discards as it is the equivalent for the L1 "más." The next form retrieved is "too much" (probably triggered as well by similarities between "much" and "mucha"). An explicit evaluation, however, helps her realize that the real meaning of "too much" is not "mucha" but "demasiado." Even though her search is motivated by the learner's lack of access to the correct L2 form, she considers that "too much" may also be an adequate alternative to express her original intended meaning, therefore generating a correct solution. The LRE is considered to be medium in terms of depth of processing due to the limited number of alternatives generated and feedback loops involved, as well as the simplicity of the processes of cross-linguistic comparison.

In turn, Example 2 shows how an intermediate-level participant searches for the L2 equivalent of the Spanish word "víctima" in the particular context of education. Interestingly, the episode shows that the compensatory search is not motivated by a lack of knowledge. Quite the contrary, the writer has already acquired the specific form-meaning mapping of "casualties" as "víctimas" in the particular context of accidents, but the possibility of extending this mapping to the context of education is not at all clear for her. She is aware of her limitations in the expression of "víctimas del sistema educativo," which she doesn't assimilate to the type of "víctimas" (casualties) she already knows. Thus, she has been able to notice "a gap in her own ability" (Izumi, 2003). As a result, she opts for expressing the idea at hand through restructuring her text and using

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

a circumlocution that includes alternative, although not fully correct, lexical items that share semantic features with the target word ("suffered negative effect from the educational system").

Finally, Example 3 illustrates a problem-solving-high upgrading LRE in which an advanced-level participant strategically decomposes and translates a clause originally formulated in the L1. The episode involves the fragmentation of the initial Spanish message "aprovechándose de lo que tiene de bueno y descartando lo que tenga de malo" into smaller and more manageable units or chunks, which are translated and adapted into the L2 until reaching the final text "by taking advantage of whatever favourable aspects it may offer to you and avoiding those which can be of no benefit to you." The three initial chunks ("aprovechándose," "de lo que tiene de bueno," "y descartando lo que tenga de malo") are frequently elaborated to produce more refined and formal versions of the original message. For instance, in the second chunk, the verb "tiene" (have) is turned into "pueda ofrecer" (it may offer) while "bueno" (good) becomes "favourable aspects." In turn, in the third chunk ("descartando lo que tenga de malo") "lo que tenga de malo" is further reformulated into "those (factors) which no puedan serte beneficiosos" using a combination of both the L1 and the L2 to elaborate the message. In this case "no serte beneficiosos," an expression with still some strong overtones from the spoken register, is rendered into the more conventional "of no benefit to you." We can thus see that the use of the strategic use of the L1 in this particular episode, as opposed to the previous examples, allows the writer to upgrade or refine parts of the initial message representation to satisfy the pragmatic, stylistic, and register concerns that she has set for herself in consonance with the demands of the task. The writer has been able to orchestrate the use of strategies and the manipulation of linguistic elements in a more flexible manner (Macaro, 2014) and, as a result, "exert more control over language choices" (Roca de Larios et al., 1999, p. 17).

DISCUSSION

With the aim of advancing methodological approaches to the study of language reflection fostered by individual writing tasks, the present study attempted to contribute a theoretically motivated and empirically based coding system for LREs to be used as a basis for future research in the domain. To that end, we analyzed think-aloud protocols through the lens of relevant theoretical perspectives and by means of coding procedures that combined insights from previous research on both individual and collaborative writing. This coding system is thus conceived as a preliminary and, in our view, necessary step to explore, first, the effect of different variables such as L2 proficiency or task characteristics on the occurrence and nature of individual LREs, and, second, the various ways in which LREs may potentially contribute to learners' L2 development. As is the case with the categorizations employed in the analysis of LREs produced in collaboration (see Storch, 2013 for a review), our coding system can be used to conduct quantitative analyses of the data (e.g., calculating the occurrence of specific categories across proficiency levels or tasks) and/or to elaborate qualitative accounts that describe particular phenomena within LREs (e.g., descriptions of successful/unsuccessful orchestrations of strategies).

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

In what follows, we interpret our findings and their significance with respect to the ultimate methodological aim guiding the study. Additionally, from a more theoretical angle, we discuss how our analytic proposal may aid in the interpretation of the language learning potential of individual LREs.

In relation to the *unit of analysis*, the coding procedures employed, informed by the tenets of the problem-solving literature and by recent theoretical explorations of the notion of strategy, enabled us to reconceptualize LREs as language-related problem spaces. Our contention would be that one advantage of this reconceptualization is that LREs are preserved intact, as global entities integrated by strategies and seen as macro-units of analysis under which the different discreet units of analysis so far used in the literature on writing processes and collaborative writing could be subsumed (see Murphy & Roca de Larios, 2010). A second advantage is that LREs, as shaped by strategy clusters, can now be seen as having a particular orientation (compensatory or upgrading), a distinction that (although not reported in previous research on LREs) has important consequences for the exploration of the language learning potential of LREs (see following text). Finally, it must be noted that this characterization of LREs as problem spaces fits in nicely with the subprocesses suggested by the writing model discussed previously (Kellogg, 1996). More precisely:

1. Writers generate an intended message (already constructed or in the process of construction), but when they intend to translate it into linguistic form, a number of different problems may arise: the lexical items, syntactic frames, or graphic procedures needed for that translation process to occur may not be easily accessed and retrieved.
2. This leads writers to engage in a series of steps or strategies: the intended message is cut up into chunks that can be matched with and distributed over the lexical items present in the lexicon. This is done by means of search operations that may engage all subprocesses of writing production and provide the learner with successive semantic, syntactic, morphological, or graphic representations of her output, before or after execution, by means of monitor or feedback loops.
3. When a final state (in the form of a solution, approximation or abandonment of the problem) is reached, the different cycles of formulation, execution, and monitoring come to an end.

In the course of this search process, and by means of these feedback loops, writers may reflect on these representations, compare them with their explicit and implicit knowledge sources and notice, if appropriate, holes in their interlanguages. These noticing processes have been taken to provide impetus for language development (Swain, 1995 and elsewhere) given their likely impact on noticing input and shaping intake (Izumi & Hanaoka, 2012).

As for the *linguistic focus* of LREs, a crucial question is how the complexity of meaning and the lack of transparency of form-meaning mappings of L2 target structures may affect writers' attentional demands. Complexity of meaning appears, for example, when the semantic system of the L2 is different from that of the L1, as is usually the case with verbal aspect, or when equivalent notions are not expressed in a similar way across languages (Jiang, 2000). Lack of transparency of form-function mappings is usually due to factors such as (a) redundancy, that is, the form is not semantically necessary because its meaning is also expressed by other elements in the sentence; (b) optionality, that is, the

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

alternating presence or absence of the form concerned makes the relationship between form and meaning difficult to establish; and (c) opacity, that is, different forms stand for the same meaning and different meanings stand for the same form (DeKeyser, 2005). Learners with limited L2 knowledge, may experience cognitive overload when paying attention to this kind of structures (Skehan, 1998) and will probably rely on the use of semantic information rather than morpho-syntactic cues, with the consequence that their attention to form will be limited (Izumi, 2003). It may be hypothesized that, as L2 proficiency increases, complex structures progressively become more salient and, as a result, they are more likely to prompt noticing, stimulate processing of form, and enhance awareness of linguistic gaps (Schmidt, 1990).

In terms of *strategies*, we also assume that, when manipulating the different components of the writing task, (a) all writers adopt some level of strategic behavior, which should be contemplated in terms of clusters of strategies combined and evaluated against their own personal goals, and that (b) opportunities for language learning arise from the interaction of these strategies with the ideas to be expressed and the linguistic knowledge that writers bring to the task (Macaro, 2006, 2014). However, as a function of the flexibility in the use the strategy clusters activated, a great deal of individual variation should be expected in the added value that strategic behavior brings to linguistic knowledge (Macaro, 2014). Our study represents a step forward in this direction with our operationalization of this flexibility in terms of the variety of strategies used within LREs. We may, therefore, hypothesize that the episodes with a greater language learning potential will be those in which the range of strategies used is wider. As the variety of strategies used in each LRE increases, writers are expected to have more opportunities to use their prior knowledge and engage in the analysis and evaluation of tentative formulations through hypotheses testing and metalinguistic reflection, as well as to increase the cognitive effort invested in solving the problem at hand (see Leow, 2015). Yet, this is a hypothesis that needs to be empirically tested.

As for *depth of processing*, we should recall that the learning potential of writing may fully occur only when the psycholinguistic processes of lexico-grammatical encoding and monitoring are coordinated, which means that the three main phases of writing production suggested by Kellogg (formulation, execution, and monitoring) must be jointly activated for such potential to be deployed. If they are dissociated, as is the case with drilled-like or mechanical writing tasks, natural production mechanisms are likely to be disrupted and their effects will probably not materialize (Izumi, 2003). However, the different feedback cycles activated in the process of composition may provide writers with a representation of their output that is sent back to the corresponding processes if mismatches are identified, that is, the process of translation, in the case of lexico-grammatical or morphological problems, or the process of planning, if the generation of ideas is at stake. We may thus conclude that the more the lexico-grammatical encoding phases and monitoring cycles are involved in LREs, the more they will be serving as "internal priming devices" to enhance the grammatical consciousness of writers (ibid.).

In turn, we would also hypothesize that non-problem-solving and problem-solving episodes present different potentials for language learning. In our data, NPS-LREs, which involved the immediate application of knowledge after the identification of a language concern by the monitor, involved very little processing time. Therefore, even though output modification occurs, the episode is unlikely to leave traces in the writer's

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

memory due to the low engagement required in its resolution. Although the learning potential of these episodes in terms of expanding L2 knowledge may be low, they may still contribute to L2 development by fostering knowledge automatization through frequent practice. These episodes appear to be similar to some of the examples described as "simple inspection" by Swain and Lapkin (1995) and basically involve intuitive ways of dealing with form-meaning mappings, which may nevertheless contribute to the consolidation of existing knowledge.

By contrast, problem-solving LREs, and particularly those that evidence high depth of processing, can be equated to those episodes labeled "extensive searches" (Cumming, 1990) or "complex thinking" (Swain & Lapkin, 1995). These kind of LREs involved instances of the three main cognitive activities identified by Cumming (1990) in his analysis, namely (a) conscious reflection aimed at searching out L2 lexical, phrasal, and syntactic alternatives; (b) cross-linguistic equivalents comparisons for a variety of purposes and at various levels of sophistication; and (c) reasoning about linguistic choices. Concerning cross-linguistic comparisons, it must be noted that, according to Cumming (1990), they are not to be understood only as mere translation attempts, but rather as processes that involve "a mental equation of semantic, lexical, syntactic and pragmatic categories across languages" (p. 494). In turn, this may lead to the integration of L1 and L2 knowledge both at the semantic level and at the level of lexemes (Jiang, 2000). In addition, both searching for alternative forms and comparing cross-linguistic equivalents involve the testing of hypotheses against the feedback provided by the learners' internalized knowledge of the L2 by means of feedback cycles (Izumi, 2003). The semantic and/or morphosyntactic reprocessing of output involved in these testing procedures are taken to represent instances in which learners experiment with new forms and expand and exploit their resources in creative ways (Swain, 1995).

In terms of the *orientation* of strategic behavior, the semantic, stylistic, and pragmatic concerns involved in the upgrading orientation of strategy clusters may, as compared to the compensatory orientation, trigger more complex concepts and extend the L2 system further. Complex concepts, in turn, may activate lexical items associated with syntactic building procedures that require complex encoding procedures, hence resulting in vocabulary, morphological, and syntactic development (Kormos, 2011). These new structures may be consolidated through repeated practice because each time they are used while performing a task, procedural knowledge of the corresponding encoding procedure is activated and automatization is aided (ibid.). We may thus infer that, in terms of the affordances for language development, strategic behavior oriented toward upgrading concerns may be superior, in principle, to compensatory behavior, another hypothesis that ought to be put to the empirical test.

Bearing these theoretical and empirical considerations in mind, we interpret our data as suggesting that LREs in self-sustained writing tasks will hypothetically involve more language learning potential if (a) the complexity of the linguistic unit attended to is in consonance with the proficiency level of the writer; (b) the use of writing strategies is flexible and upgrading in orientation; and (c) the production of the structure concerned involves high depth of processing and repeated engagement in the subprocesses of writing, with special emphasis on the stages of translating and monitoring, which allow learners to search the available knowledge (explicit and implicit) they may possess. However, these are still hypotheses that need to be empirically tested in future research.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Despite the potential methodological and theoretical contribution of our study, a number of limitations should be acknowledged. First of all, due to practicalities, the writing task employed was a one-shot, laboratory-type task. Future studies should include examples of various types of tasks and genres to capture the inherent complexity and diversity of writing. In addition, even though the analysis of our data focused solely on LREs, we acknowledge the importance of exploring the meaning making and rhetorical nature of argumentative writing more fully.

Another limitation of the study is its reliance on participants' online verbalizations, texts, and notes as the only data source. In this respect, the combined use of data from multiple sources and modalities (i.e., stimulated recall interviews, eye-tracking, keystroke logging, digital screen capture) has been advocated as an alternative to strengthen and complement the analyses and interpretation of writing processes (Gánem-Gutiérrez & Gilmore, 2018; Révész & Michel, 2019). Future research should explore whether the application of these new methodologies to the units of analysis suggested in the present study provide a more accurate understanding of the composition process, especially at the moment-by-moment levels.

Concerning our participants, we included different educational and age groups in the study to make sure that as many phenomena as possible could be identified in the LREs (not just those problems/strategies restricted to a particular level). However, this decision meant that the participants' writing ability was not controlled for. As a result, the data obtained might in total or in part reflect native language writing ability, overestimating L2 writing performance, a limitation that future studies ought to overcome.

Furthermore, we mentioned above that, for the LREs to afford opportunities for language learning, the complexity of the forms addressed by L2 writers should be in consonance with their L2 proficiency level. As a function of this level, L2 learners will probably differ in the attention paid to complex structures and in their efforts to refine the alternatives produced with the aid of their repertoires of L2 lexical items. In addition, as L2 proficiency level "logically leads to the role of prior knowledge in depth of processing" (Leow, 2015, p. 221), they will probably handle the translation and the monitoring processes involved with different degrees of depth and efficiency. Further research needs to confirm these predictions with participants representing a wider range of L2 proficiency and through a variety of tasks as suggested in the preceding text.

Finally, further research should also contemplate the role played by individual differences in the different stages of the written language production process and the potential for language learning associated to them. As suggested by Kormos (2012), individual, cognitive, and motivational factors may have a bearing on the generation, formulation, and monitoring of ideas and their linguistic expression and, as a result, on how learners exploit the language learning potential of writing by means of noticing, knowledge internalization and consolidation processes, hypothesis testing, and problem solving.

Despite these limitations, it is hoped that the coding system offered in our study represents a worthy methodological contribution to the analysis of language reflection fostered by individual writing, and that our research provides the impetus for the exploration of the previously mentioned future lines of inquiry into writing processes.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0272263119000275

NOTES

[1]Although, as suggested by one of the reviewers, the variety and interaction of components included in its control, process, and resource levels makes Hayes's 2012 model more elaborate and cognizant of recent advances in research and theory than Kellogg's (1996), the way the latter describes the different linguistic encoding and monitoring processes involved in the generation of texts seemed to us more appropriate for the theoretical perspectives we pursued.

[2]As Rijlaarsdam and van den Bergh (2006, p. 51) acknowledge in their considerations on building empirically based writing models, data obtained through think-aloud protocols tend to offer a linear view of the writing process. In this sense, all research on the cognitive processes of composing is limited to a certain extent, as it is not able to capture the multiplicity of parallel thinking processes that take place during writing.

REFERENCES

Allan, D. (1995). *Oxford placement test*. Oxford, UK: Oxford University Press.

Bowles, M. (2010). *The think-aloud controversy in second-language research.* London, UK: Routledge.

Byrnes, H., & Manchón, R. M. (2014). Task-based language learning: Insights from and for L2 writing: An introduction. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 1–23). Amsterdam, The Netherlands: John Benjamins.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, *18*, 80–98.

Craik, F. I. M. (2002). Levels of processing: Past, present, and… future? *Memory*, *10*, 305–318.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, *11*, 671–684.

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, *7*, 482–511.

DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, *51*, 1–25.

Dörnyei, Z., & Kormos, J. (1998). Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in Second Language Acquisition*, *20*, 349–385.

Fernández-Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair and individual work. *Journal of Second Language Writing*, *21*, 40–58.

Fortune, A. (2005). Learners' use of metalanguage in collaborative form-focused L2 output tasks. *Language Awareness*, *14*, 21–38.

Fortune, A., & Thorp, D. (2001). Knotted and entangled: New light on the identification, classification and value of language related episodes in collaborative output tasks. *Language Awareness*, *10*, 143–160.

Gánem-Gutiérrez, G. A., & Gilmore, A. (2018). Tracking the real-time evolution of a writing event: Second language writers at different proficiency levels. *Language Learning*, *68*, 469–506.

Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistics perspective*. Harlow, UK: Pearson Education.

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, *11*, 329–350.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication 29*, 369–388.

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, *24*, 168–196.

Izumi, S., & Hanaoka, O. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, *21*, 332–347.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*, 47–77.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. E. Ransdell (Eds.), *The science of writing* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum Associates.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Erlbaum.

Kormos, J. (2011). Speech production and the Cognition Hypothesis. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 39–60). Amsterdam, The Netherlands: John Benjamins.

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, *21*, 390–403.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*, 1–26.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centred approach*. London, UK: Routledge.

Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *Modern Language Journal*, *90*, 320–337.

Macaro, E. (2007). Language learner strategies: Adhering to a theoretical framework. *The Language Learning Journal*, *35*, 239–243.

Macaro, E. (2014). Reframing task performance: The relationship between tasks, strategic behaviour, and linguistic knowledge in writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 53–77). Amsterdam, The Netherlands: John Benjamins.

Manchón, R. M. (2011). Writing to learn the language: Issues in theory and research. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 61–82). Amsterdam, The Netherlands: John Benjamins.

Manchón, R. M. (2014). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 27–52). Amsterdam, The Netherlands: John Benjamins.

Manchón, R. M., & Roca de Larios, J. (2007). Writing-to-learn in instructed language learning contexts. In E. Alcón Soler & M. P. Safont Jordá (Eds.), *Intercultural language use and language learning* (pp. 101–121). Dordrecht, The Netherlands: Springer.

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. W. Schwieter, & A. Benati (Eds.), *The Cambridge handbook of language learning*. Cambridge, UK: Cambridge University Press.

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 567–586). Berlin, Germany, and Boston, MA: De Gruyter Mouton.

Manchón, R. M., Murphy, L., & Roca de Larios, J. (2005). Using concurrent protocols to explore L2 writing processes: Methodological issues in the collection and analysis of data. In P. Matsuda & T. Silva (Eds.), *Second language writing research: Perspectives on the process of knowledge construction* (pp. 191–206). New York, NY: Routledge.

Manchón, R. M., Roca de Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing: Implications for theory. In R. M. Manchón (Ed.), *Foreign language writing: Learning, teaching and research* (pp. 102–124). Clevedon, UK: Multilingual Matters.

Murphy, L., & Roca de Larios, J. (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL writers. *Journal of Second Language Writing*, *19*, 61–81.

Newell, A. (1980). Reasoning, problem solving and decision processes: The problem space as a fundamental category. In R. S. Nickerson (Ed.), *Attention and performance* (pp. 693–719). Hillsdale, NJ: Erlbaum.

Ortega, L. (2011). Reflections on the learning-to-write and writing-to-learn dimensions of second language writing. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 237–250). Amsterdam, The Netherlands: John Benjamins.

Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning*, *37*, 439–468.

Révész, A., & Michel, M. (2019). Introduction. *Studies in Second Language Acquisition*, *41*, 491–501.

Révész, A., Kourtali, N. E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviours and linguistic complexity. *Language Learning*, *67*, 208–241.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press

Rijlaarsdam, G., & van den Bergh, H. (2006). Writing process theory: A functional dynamic approach. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 41–53). New York, NY: Guilford Press.

Roca de Larios, J. (2013). Second language writing as a psycholinguistic locus for L2 production and learning. *Journal of Second Language Writing*, *4*, 444–445.

Roca de Larios, J., Murphy, L., & Manchón, R. M. (1999). The use of restructuring strategies in EFL writing: A study of Spanish learners of English as a foreign language. *Journal of Second Language Writing*, *8*, 13–44.

Sasaki, M (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, *9*, 259–291.

Sasaki, M. (2002). Building an empirically-based model of EFL learners' writing processes. In S. Ransdell & M. L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 49–80). Amsterdam: Kluwer Academic.

Sasaki, M (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning*, *54*, 525–582.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 206–226.

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

Smagorinsky, P. (2008). The method section as conceptual epicentre in constructing social science research reports. *Written Communication*, *25*, 389–411.

Smith, V. (1994). *Thinking in a foreign language: An investigation into essay writing and translation by L2 learners*. Tubingen, Germany: Verlag.

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol, UK: Multilingual Matters.

Storch, N. (2016). Collaborative writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 387–406). Berlin, Germany: De Gruyter Mouton.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–153). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Applied linguistics. Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford, UK: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, NJ: Lawrence Erlbaum.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step toward second language learning. *Applied Linguistics*, *16,* 371–391.

Vygotsky, L. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Wang, L. (2003). Switching to first language among writers with differing second-language proficiency. *Journal of Second Language Writing*, *12*, 347–375.

Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing*, *11*, 225–246.

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, *21*, 321–331.

Witte, S. P. (1987). Pre-text and composing. *College Composition and Communication*, *38,* 397–425.

Zimmermann, R. (2000). L2 writing subprocesses: A model of formulating and empirical findings. *Learning and Instruction*, *10*, 73–99.

https://doi.org/10.1017/S0272263119000275 Published online by Cambridge University Press