The present study addresses the reactivity of two types of verbal protocols in SLA research. It expands on the work of Leow and Morgan-Short (2004), who found nonmetalinguistic verbalization during a second-language reading task to be nonreactive for beginning learners’ text comprehension, intake, and production of the targeted morphological form. The present study investigated the reactivity of both metalinguistic and nonmetalinguistic protocols, using a syntactic structure and advanced language learners of Spanish. Results indicated that neither type of verbalization significantly affected text comprehension or written production of old or new exemplars of the targeted structure when compared to a control group, although metalinguistic verbalization appeared to cause a significant decrease in text comprehension over nonmetalinguistic verbalization. Furthermore, both types of verbalization significantly increased the amount of time on task.

Since the 1980s, there has been a marked increase in the use of verbal reports to study cognitive processes, so much so that “both concurrent and retrospective verbal reports are now generally recognized as major sources of data on
subjects' cognitive processes in specific tasks" (Ericsson & Simon, 1993, p. xi). To gather information about subjects' cognitive processing, thought processes, and strategies, the collection of verbal reports has become standard in many fields, ranging from accounting (Anderson, 1985) and marketing (Biehal & Chakravarti, 1989) to software engineering (Hughes & Parkes, 2003) and language research (Cohen, 1986; Leow, 1997; Rosa & O'Neill, 1999). Such verbal accounts, it is argued, are valuable because of the difficulty inherent in capturing the essence of mental processes. Notwithstanding the apparent value of verbal reports, the validity of such reports as a measurement of cognitive processes has been debated in non-SLA literature for decades, especially in relation to the issue of reactivity (i.e., the supposition that verbalization might alter the cognitive process). One recent study (Leow & Morgan-Short, 2004) empirically addressed the issue of reactivity in SLA research methodology and found nonreactivity for concurrent nonmetalinguistic verbalization at the beginning level of foreign or second language (L2) experience. Given that the issue of reactivity is strongly related to the level of internal validity of studies premised on the role of attention in SLA, the present study expands on Leow and Morgan-Short's study by empirically investigating the role that type of verbal report (nonmetalinguistic vs. metalinguistic) elicited during an L2 reading task plays in the subsequent comprehension and written production of learners with an advanced stage of L2 experience.

Prior to the publication of Ericsson and Simon's (1984, 1993) seminal review of verbalization, many investigators tended to treat alike all studies instructing subjects to verbalize their thoughts. However, as Ericsson and Simon pointed out, not all verbalizations are alike; they differ according to the conditions under which they are obtained. In their typology, Ericsson and Simon categorized verbal reports as either concurrent or retrospective based on the temporal frame in which the reports are collected. Concurrent protocols are those collected as subjects verbalize while performing the task in question, whereas retrospective protocols are collected when subjects verbalize some time after performing the task. Ericsson and Simon advised that retrospective protocols be used with caution because it is impossible to “rule out the possibility that the information [subjects] retrieve at the time of the verbal report is different from the information they retrieved while actually performing the experimental task” (1993, p. xii). They advised that concurrent protocols be collected whenever possible “to avoid this problem of accessing information at two different times—first during the actual cognitive processing and then at the time of report” (1993, p. xiii).

In addition to categorizing verbal reports in terms of temporal space, Ericsson and Simon (1993) also distinguished between other types of protocols. After all, “being instructed to explain verbally each step of a solution is quite different from focusing wholly on solving the problem efficiently while verbalizing one's thoughts concurrently” (p. xvii). Ericsson and Simon distinguished between protocols that require subjects to verbalize their thoughts per se and those that require subjects to verbalize specific information, such as reasons
and justifications. For the purposes of the present study, verbalizations of thoughts per se will be referred to as nonmetalinguistic and those requiring verbalization of additional specific information will be referred to as metalinguistic.1

Two major issues surrounding the validity of verbal reports surface from this typology. One is veridicality—namely, whether the information in verbal reports accurately represents the thought process it is designed to capture. Some evidence for nonveridicality has been found for retrospective protocols, which can “yield substantial forgetting or fabrication in all tasks” (Russo, Johnson, & Stephens, 1989). Concurrent protocols, on the other hand, are not subject to the same critique, as they are collected during the task. The second issue is reactivity—that is, whether the act of thinking aloud alters the end state of the cognitive process (accuracy of task performance). Ericsson and Simon’s (1993) model predicted that verbalizations that require subjects to verbalize additional specific information (metalinguistic verbalization) could alter subjects’ thought processes and, therefore, potentially have reactive effects on task performance. However, they cautioned that if such justifications are “generated as part of the normal process of solution,” then such verbalization should not have an effect on task performance (p. xxxiii). Therefore, implicit in this statement is the supposition that the requirements of a task might have a crucial impact on the reactivity (or lack thereof) of metalinguistic verbalizations.

VERBAL REPORTS IN SLA RESEARCH

Introspective methods, including verbal reports, have been used extensively as a data elicitation technique in SLA. Since SLA was first studied systematically in the early 1970s, there has been some debate over their use: Researchers like Selinker (1972) proposed that only learners’ production data should be used in formulating theories and conducting research about SLA, whereas others, such as Corder (1973), argued that production data from language learners provide only a small piece of the language learning puzzle. As Corder pointed out, in order to understand how language learning works, it might also be necessary to determine what learners think about their own production. Corder believed that this type of information could only be gathered through introspective methods.

Although the debate over the use of verbal reports continues today, many researchers in the field continue to use verbal reports—both concurrent and retrospective—in an effort to gain some information about learners’ cognitive processes as they interact with the L2. As Gass and Mackey (2000) pointed out, “understanding the source of second language production is problematic because often there are multiple explanations for production phenomena that can only be assessed by exploring the process phenomena” (p. 26).

Verbal reports have been used as a data elicitation technique in SLA attentional research (studies premised on the role of attention in SLA) for less than a decade; the first published mention of them is by Alanen (1995), who combined concurrent verbal reports with two other offline data collection measurements to address the role of noticing in her study. Recently, several empirical studies premised on the role of attention and awareness, or both, in L2 development have begun to employ concurrent (online think-alouds in Leow, 1997, 1998, 2000, 2001a, 2001b; Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999) or retrospective verbal reports (stimulated recalls in Mackey et al., 2000) to gather data on learners’ cognitive processes while they interacted with the L2 data. These data were used to establish the role of attention and awareness in L2 processing during exposure to the L2 before investigating its effects on L2 development. However, the use of verbal reports in L2 acquisition has been critiqued by a number of sources on the grounds that the requirement to think aloud might be reactive (Ellis, 2001; Jourdenais, 2001); that is, verbalization might alter the cognitive processes employed by learners while performing an L2 task. Consequently, the validity of verbal reports in SLA research methodology has been questioned. However, these claims appear to be based largely on anecdotal evidence, as there has been only one empirical SLA study (Leow & Morgan-Short, 2004) to date that investigated the reactivity of think-aloud protocols.

Leow and Morgan-Short (2004) investigated the effects of nonmetalinguistic verbal reports during the reading process on 77 beginning Spanish learners’ text comprehension of a Spanish article (384 words) and intake and written production of a morphological structure (the impersonal imperative in Spanish). Participants in the control condition read and completed the tasks silently, whereas participants in the experimental condition read and completed the tasks while thinking aloud. Results found nonreactivity for concurrent nonmetalinguistic verbalization, as the two groups did not differ significantly on either text comprehension (measured by a multiple-choice recognition task) or postassessment task performance (on a controlled written production task). Although Leow and Morgan-Short (2004) were the first to empirically address the methodological issue of reactivity in SLA research methodology, whether
type of verbal report (metalinguistic vs. nonmetalinguistic) plays a role in reactivity has not been empirically investigated in the field of SLA, and further exploration of this issue is clearly warranted.

**PREVIOUS NON-SLA RESEARCH**

Research on the effects of verbal reports on problem-solving and decision-making tasks has been conducted in the field of cognitive psychology since at least the 1950s. A comprehensive review of the non-SLA research into the reactivity of verbal reports is outside the scope of this article, but a synthesis of dozens of studies in cognitive psychology that have investigated the reactivity of nonmetalinguistic verbal reports indicates that such verbalizations do not seem to influence cognitive processes when compared to silent control groups (Ericsson, 2002). This finding of nonreactivity suggests that nonmetalinguistic verbalization might be a valid method of capturing internal thought processes. However, the findings indicate that, overall, verbalization tends to be reactive for latency (solution time) because the additional time needed for verbalization increases the overall solution time.

With regard to metalinguistic verbalization, an early review of studies (Ericsson & Simon, 1993) found mixed results with regard to reactivity. The majority of the studies involving metalinguistic verbal reports examined the effects of metalinguistic think-alouds compared to a silent control group and found metalinguistic think-alouds to be reactive for accuracy in most cases (Berry, 1983; Berry & Broadbent, 1984; Bower & King, 1967; Davis, Carey, Foxman, & Tarr, 1968; McGeorge & Burton, 1989; Short et al., 1991; Stanley, Mathews, Buss, & Kotler-Cope, 1989; Wilder & Harvey, 1971; Wilson & Schooler, 1991). For many of the studies, time (latency) was not reported, but Ericsson and Simon’s prediction that metalinguistic verbalization requires extra processing time (as compared to other types of verbalization or to a control) was borne out in almost all of the studies that reported latency (Ahlum-Heath & di Vesta, 1986; Allison, 1990; Carpenter, Just, & Schell, 1990; Fidler, 1983; Gagné & Smith, 1962; K. Robinson, 2001).

Several more relevant studies were published before and after Ericsson and Simon’s (1993) book; the results of these studies (see Table 1) will be briefly summarized in the following section.

Studies investigating the reactive effects of metalinguistic versus nonmetalinguistic protocols have generally compared the following conditions: control (silent), nonmetalinguistic verbalization, and concurrent metalinguistic verbalization. Participants were usually undergraduate students, and experimental materials employed in the research designs included problem-solving tasks (e.g., the Tower of Hanoi, Katona Card problem, computerized forest-fire fighting simulator, etc.) and Vygotsky tests (e.g., Ladd and PAMs tests). The findings of these studies are inconclusive. Three of the six studies (Berardi-Coletta, Buyer, Dominowski, & Rellinger, 1995; Brunk, Collister, Swift, & Stay-
Table 1. Verbal reports and reactivity: Overview of empirical research studies comparing metalinguistic and nonmetalinguistic verbal reports

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Task(s)</th>
<th>Groups</th>
<th>Design</th>
<th>Reactivity</th>
<th>Nonreactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berardi-Coletta et al. (1995)</td>
<td>Experiment 1</td>
<td>109 undergraduate students</td>
<td>Tower of Hanoi</td>
<td>Between-subjects</td>
<td>Accuracy, time</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 TA [+ meta], 2 TA [− meta], 1 control (silent)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>64 undergraduate students</td>
<td>TA [+ meta], TA [− meta], TA + meta, TA [− meta], TA (silent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40 undergraduate students</td>
<td>TA [+ meta], TA (silent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15 undergraduate students</td>
<td>TA [+ meta], TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>64 undergraduate students</td>
<td>TA [+ meta], TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>147 undergraduate students</td>
<td>Control [+ meta], TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>60 undergraduate students</td>
<td>Control (silent), TA + meta, TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24 undergraduate students</td>
<td>Control (silent), TA [− meta], TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>343 high school students</td>
<td>Control (silent), TA + meta, TA [− meta], TA + meta, TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>60 undergraduate students</td>
<td>Control (silent), TA + meta, TA [− meta], TA + meta, TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24 undergraduate students</td>
<td>Control (silent), TA [− meta], TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>343 high school students</td>
<td>Control (silent), TA [− meta], TA + meta, TA [− meta]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. TA = think-aloud; [+ meta] = metalinguistic; [− meta] = nonmetalinguistic.

Time not reported.

The metalinguistic verbalization group performed significantly worse than the control group; the nonmetalinguistic verbalization group performed between the two extremes, only slightly worse in terms of accuracy than the control.
ton, 1958; Dickson, McLennan, & Omodei, 2000) found that verbalization affected task performance, whereas two of the six (Brehmer, 1974; Norris, 1990) found that verbalization had no effect on task performance. Only one study (Russo et al., 1989) found mixed results, which indicated that verbalization affected performance for some tasks but not for others. However, all of the studies reviewed (with the exception of Russo et al., 1989) employed nonverbal and problem-solving tasks, or both, limiting their relevance to SLA.

In fact, of all the studies on reactivity found in the literature, only six used verbal tasks (Lass, Klettke, Lüer, & Ruhlender, 1991; Mathews et al., 1989; Rhenius & Deffner, 1990; Russo et al., 1989; Short et al., 1991; Stratman & Hamp-Lyons, 1994). For a brief overview of the methodologies of these studies, see Table 2. Of the six studies, only one (Short et al., 1991) found verbalization (in this case, metalinguistic) to be reactive for accuracy. However, these studies showed great heterogeneity in terms of task type, with experimental tasks ranging from anagrams and verbal analogies to sentence assembly, text revision, and artificial grammar activities. Furthermore, they included participants from diverse pools, ranging from elementary school children to college-age students. As such, conclusive results cannot be drawn from this small sample of studies with regard to reactivity of verbal reports on L1 verbal tasks, although the results in general support the finding of nonreactivity exhibited by nonverbal tasks.

Given the paucity of research involving L1 verbal tasks, it is not surprising that there has been even less research on reactivity of verbalization during L2 verbal tasks. To date, only one recent study (Leow & Morgan-Short, 2004) of those cited previously has examined the reactivity of nonmetalinguistic verbal reports during the processing of L2 tasks. However, the effects of metalinguistic verbalization on L2 tasks and the differential effects of type of verbalization (both nonmetalinguistic and metalinguistic) on more advanced language learners have yet to be examined.

The issue of time (latency) also appears to be inconclusive. In the studies that reported time (Lass et al., 1991; Rhenius & Deffner, 1990), verbalization is usually shown to take longer than a silent group. However, the majority of studies did not report times on task.

Given the inconclusive findings and heterogeneity of task types in reactivity studies in the field of cognitive psychology and the lack of a base of research with L2 tasks, more studies are clearly needed to address the validity of verbal reports in SLA.

THE PRESENT STUDY

The present study, motivated by the previous research in cognitive psychology and the need to continue addressing the internal validity of research designs in SLA studies, expanded on Leow and Morgan-Short’s (2004) study by empirically investigating the role that type of verbal report (nonmetalinguistic...
Table 2. Verbal reports and reactivity: Overview of empirical research studies employing verbal tasks

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Task(s)</th>
<th>Groups</th>
<th>Design</th>
<th>Reactivity</th>
<th>Nonreactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lass et al. (1991)</td>
<td>70 undergraduate students</td>
<td>Anagrams</td>
<td>TA [- meta], control (silent)</td>
<td>Between-subjects</td>
<td>Latency</td>
<td>Accuracy</td>
</tr>
<tr>
<td>Mathews et al. (1989, Exp. 2)*</td>
<td>168 undergraduate students</td>
<td>Artificial grammar activities</td>
<td>Time 1 &amp; time 2: silent; all verbalized</td>
<td>Between-subjects</td>
<td>Accuracy</td>
<td></td>
</tr>
<tr>
<td>Rhenius &amp; Deffner (1990)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td>24 students</td>
<td>Sentence assembly</td>
<td>TA [- meta], control (silent)</td>
<td>Between-subjects</td>
<td>Latency(^b)</td>
<td>Accuracy</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>48 students</td>
<td>Word and geometrical puzzles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 3</td>
<td>21 students</td>
<td>Raven's matrices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russo et al. (1989)</td>
<td>24 undergraduate students</td>
<td>Mathematical task</td>
<td>Control (silent), concurrent TA</td>
<td>Between-subjects</td>
<td>Accuracy for TA</td>
<td>Accuracy for TA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(mental addition), verbal task (anagrams),</td>
<td></td>
<td></td>
<td>[± meta] on gambles and addition; time for concurrent group</td>
<td>[± meta] on anagrams and Raven's matrices</td>
</tr>
<tr>
<td></td>
<td></td>
<td>numerical task (gambles), pictorial task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Raven's matrices)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short et al. (1991)*</td>
<td>94 bright, average, learning disabled, and</td>
<td>Spatial and verbal analogies</td>
<td>Time 1: silent;</td>
<td>Within-subjects</td>
<td>Accuracy for all students on both types of analogy (most profound effects for bright and average students)</td>
<td>No statistical analysis(^b)</td>
</tr>
<tr>
<td></td>
<td>developmentally handicapped fifth-graders</td>
<td></td>
<td>time 2: TA [+ meta]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stratman &amp; Hamp-Lyons (1994)</td>
<td>12 undergraduate and graduate students</td>
<td>Revision of text</td>
<td>TA [- meta], silent</td>
<td>Within-and between-subjects</td>
<td>No statistical analysis</td>
<td></td>
</tr>
</tbody>
</table>

Note: TA = think-aloud; [+ meta] = metalinguistic; [- meta] = nonmetalinguistic.

\(^a\)Time not reported.

\(^b\) Ability to detect and fix organizational errors in the text decreased in the verbalization condition. Detection of pronoun errors and occurrence of word-level errors increased.
guistic vs. metalinguistic) elicited during an L2 reading task plays in the subsequent text comprehension and written production of learners drawn from an advanced stage of L2 experience. Given that previous research has found that reading strategies vary with proficiency level (O’Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985), advanced L2 learners were chosen for this study to explore whether the previous finding of nonreactivity for beginning learners in Leow and Morgan-Short’s study would be replicated. Additionally, the present study investigated the effects of verbal reporting on system learning (the ability to apply recently learned grammatical information to new exemplars). Finally, the study sought to address the issue of time (latency); that is, whether the amount of time it took to complete the task was significantly affected by type of verbalization. Specifically, the study was guided by the following research questions:

Does the type of verbalization (nonmetalinguistic vs. metalinguistic vs. control)

1. have an effect on advanced L2 readers’ comprehension of an L2 text?
2. have an effect on advanced L2 readers’ ability to produce old exemplars of the targeted L2 structure?
3. have an effect on advanced L2 readers’ ability to produce new exemplars of the targeted L2 structure?
4. significantly affect the amount of time it takes to complete the task?

Participants

The original pool of participants consisted of 70 college-level students enrolled in a fifth-semester Spanish course who received extra credit in exchange for participation. At the time of the experiment, the targeted syntactic structure (the Spanish pluperfect subjunctive) had not been formally presented in class to the participants. However, participants had almost certainly encountered the form at some point when reading authentic texts as part of the communicative language program. To ensure that only participants with minimal knowledge of the targeted structure were included in the final sample, only those participants who scored 2 or less out of 10 on the pretest and who indicated on a debriefing questionnaire at the end of the study that they did not initially understand the rules for using the structure were included. Participants were also eliminated for (a) not completing all parts of the experiment and (b) not providing audible think-aloud protocols. Of the original pool of 70 participants, 25 were eliminated on the basis of these criteria, yielding a final sample of 45 participants. These participants were randomly assigned to either a control group or to one of the verbalization groups (metalinguistic or nonmetalinguistic). There were 17 participants in the control group, 18 in the nonmetalinguistic group, and 10 in the metalinguistic group.
Targeted Structure

The targeted linguistic structure was the pluperfect subjunctive in Spanish. This structure was chosen because it poses difficulty even for advanced learners of Spanish and had not been formally presented to the participants at the time of the study. Additionally, the experimental text provided numerous opportunities for contextual guessing of the structure’s meaning and use. To understand the basic message of the text, participants almost certainly had to realize that the pluperfect subjunctive forms in the text conveyed contrary-to-fact statements in the past tense (i.e., Si George Walker Bush hubiera sido un argentino descendiente de algún inmigrante italiano, quizá hoy sería un tal Jorge Ángelo Fossati. “If George Walker Bush had been an Argentine descendant of some Italian immigrant, perhaps today he would be named Jorge Ángelo Fossati.”).

Materials

The text used in the present study was an 861-word article from the Argentine magazine El Correo. The article, titled Si Bush hubiera sido porteño (“If Bush had been Argentine”) contained 28 instances of the pluperfect subjunctive, making it ideal for this study (see the Appendix).

A computer audio recording program (Audio Hijack version 1.2.4) was used to record participants’ verbalizations as they read the text and completed the tasks. Participants wore headsets during the experiment and spoke their thoughts aloud into the microphone. The software recorded their voices as 64-bitrate MPEG Audio Layer 3 (MP3) files, which could be easily stored and quickly retrieved for subsequent analysis.

Assessment Tasks

To measure participants’ comprehension of the text, a 10-item multiple-choice comprehension task was designed to elicit pieces of information provided through the pluperfect subjunctive in the text. The multiple-choice questions were provided in the participants’ L1 (English) to ensure that participants were being tested on their comprehension of the Spanish text alone and not on their comprehension of the questions themselves.

To measure participants’ written production of the targeted structure, two fill-in-the-blank written production tasks were prepared. The first task was designed to measure participants’ ability to produce the targeted structure in familiar contexts. This task consisted of 20 sentences, all taken directly from the Spanish text. In each sentence, a verb was deleted and its English equivalent was provided in parentheses, as in (1).
Es probable que sus abuelos (would have arrived) ______ al país en la oleada de 1880.

"It is likely that his grandparents ______ to the country during the wave of 1880."

Participants were instructed to write the appropriate verb form to complete each sentence. Of the 20 sentences, 10 required the targeted structure and 10 were distracters that required other verb forms.

The second task was designed to measure participants’ ability to produce the targeted structure in new contexts. This task consisted of 20 sentences not found in the text. A contextualizing sentence was provided in Spanish, followed by a sentence with blanks in both the “if-clause” and the “result clause,” as in (2).

Juan trabajó veinte horas la semana pasada y ganó $600. Juan (GANAR) ______ más dinero la semana pasada si (TRABAJAR) ______ cuarenta horas.

"Juan worked twenty hours last week and earned $600. Juan (EARN) ______ more money last week if he (WORK) ______ forty hours."

The infinitive of the appropriate verb for each blank was provided in parentheses. Of the 20 sentences, 10 required the use of the targeted structure and 10 required the use of other verb forms.

**Testing Procedure**

The pretest, which consisted of the first written production task, was administered 2 weeks before participants were exposed to the experimental input. During that 2-week period, participants were not formally exposed to the targeted structure.

On the day of the experimental exposure, participants were randomly assigned to either the control group or to one of the verbalization groups. Participants in both verbalization groups reported to the language laboratory, whereas participants in the control group remained in their classrooms. All participants were randomly assigned a packet containing (a) an agreement form, (b) instructions, (c) the experimental text, (d) the comprehension task, (e) the production task testing usage in familiar contexts, (f) the production task testing usage in unfamiliar contexts, and, finally, (g) the debriefing questionnaire. Materials were identical for all groups, except for the instructions relating to verbalization. Participants in the control group were simply instructed, "You will read a text from an Argentine magazine and answer some questions about it. Please complete the questions in order.” Participants in the two verbalization conditions were given the same instructions as the control group, plus explicit instructions on how to think aloud. In the nonmetalinguistic condition, participants were instructed as follows:
As you read the article and answer the questions, *please think your thoughts aloud*. That is, say whatever passes through your mind while you read and answer the questions. You may speak in either English or Spanish.

Participants in the metalinguistic condition were instructed as follows:

While you read the text that follows, we would like you to verbalize every thought and every detail of your thought process, including what information you are looking at, what thoughts you are having about any piece of information, how you evaluate different pieces of information, and why. And as you answer the questions that follow the text, please do the same, providing your reasoning or justification for each response.

Please complete the questions in order. As you read the article and answer the questions, *please think your thoughts aloud*. That is, say whatever passes through your mind while you read. As you answer the questions, *give a justification out loud for your answers and explain your reasoning*. You may speak in either English or Spanish.

As they read the text and completed the tasks, participants in the two verbalization groups were reminded to continue thinking aloud. The control group followed the same procedure as both of the verbalization groups, with the exception of not having to provide think-aloud protocols. Participants were encouraged to speak in either L1 (English) or L2 (Spanish) as they wished so that their think-alouds would not be constrained by their language proficiency.

**Scoring Procedure**

For the comprehension task, one point was awarded for each correct answer and zero points for each incorrect answer, for a total possible score of 10. For the first production task, only those 10 sentences that necessitated the targeted structure were counted; one point was awarded for each correct answer and zero points for each incorrect answer, for a total possible score of 10. For the second production task, again only the 10 sentences that necessitated the targeted structure were counted. Participants received one point if they correctly produced the targeted verb form (*hubiera* + past participle) in the appropriate clause and zero points if they did not produce the targeted structure in the appropriate clause. Participants were not penalized for making spelling errors or for overgeneralizing irregular past participles; that is, *hubiera* *volvido* (for *hubiera vuelto*) would be scored correct.

**Protocol Coding Procedure**

To ensure that participants in each verbalization group accurately represented their respective group in regard to the production of metalinguistic or nonmetalinguistic protocols, one-third of the protocols were randomly selected and coded [± metalinguistic] by two raters. Given that interrater reliability
was 100% for these protocols, the remaining protocols were subsequently coded by one rater.

RESULTS

Prior to conducting statistical analyses with participants’ test scores, reliability analyses were performed on the comprehension task and on the two production tasks. In all cases, the reliability coefficient was computed using Cronbach’s alpha. Reliability on the comprehension task was found to be fairly low (~.45), whereas reliability on the two production tasks was high (~.88 for the production of old exemplars and .87 for the production of new exemplars).

Research Question 1

To measure the effects of type of verbalization on comprehension, raw comprehension scores (see Table 3) were submitted to a one-way analysis of variance (ANOVA) (see Table 4). Results showed a significant main effect for group, \( F(2, 45) = 3.86, p < .05 \). A post hoc Ryan test was performed to identify which contrasts contributed to the main effect. Results revealed that the significant main effect was the result of a difference, significant at the .05 level, between the nonmetalinguistic and metalinguistic groups, with the nonmetalinguistic

Table 3. Comprehension: Descriptive statistics

<table>
<thead>
<tr>
<th>Condition</th>
<th>Comprehension M</th>
<th>Comprehension SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>5.59</td>
<td>1.938</td>
</tr>
<tr>
<td>[- meta]</td>
<td>6.67</td>
<td>1.879</td>
</tr>
<tr>
<td>[+ meta]</td>
<td>4.80</td>
<td>1.135</td>
</tr>
</tbody>
</table>

Table 4. Comprehension: One-way ANOVA for type of verbalization

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sums of squares</th>
<th>Mean squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of verbalization</td>
<td>2</td>
<td>24.1935</td>
<td>12.0967</td>
<td>3.857*</td>
</tr>
<tr>
<td>Within-groups</td>
<td>42</td>
<td>131.7176</td>
<td>3.1361</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*\( p < .05 \).
group performing significantly better on the comprehension task than the metalinguistic group. In other words, when compared, participants who thought aloud nonmetalinguistically while reading an L2 text appeared to have comprehended significantly more of the text content than participants who thought aloud metalinguistically. However, no significant difference was found between the control group and either experimental group on the comprehension task \((p = .079, p = .270)\) for the nonmetalinguistic and metalinguistic groups, respectively. As such, the presence of either type of verbal reports did not appear to have a detrimental effect on participants’ comprehension of the L2 text when compared to a silent control group.

**Research Question 2**

To measure the effects of type of verbalization on participants’ ability to produce old exemplars of the targeted structure, a one-way ANOVA was first conducted on raw pretest production scores to determine whether the three experimental groups were comparable at the outset of the study. The results revealed no significant differences between the groups, \(F(2, 42) < 1\). Consequently, it could be assumed that any gains in the production scores of old exemplars from the pretest to the posttest were due to the type of verbalization rather than to preexisting differences between experimental groups.

The pretest and immediate posttest raw scores for each group were subsequently submitted to a \(3 \times 2\) repeated-measures ANOVA (for descriptive statistics, see Table 5). Type of verbalization (metalinguistic vs. nonmetalinguistic vs. control) was entered as the between-subject factor, and time (pretest vs. posttest) was entered as the within-subject factor. The ANOVA (see Table 6) only showed a significant main effect for time, \(F(2, 42) = 56.72, p < .05\), revealing neither a significant main effect for type of verbalization, \(F(2, 42) < 1\), nor a significant interaction between time and type of verbalization, \(F(2, 42) < 1\). In other words, although all three groups significantly improved from the pretest to immediate posttest, type of verbalization did not appear to have an effect on participants’ ability to produce old exemplars of the targeted structure.

**Research Question 3**

To measure the effects of verbalization on participants’ ability to produce new exemplars of the targeted structure, raw scores on the new exemplars production task were submitted to a one-way ANOVA (for descriptive statistics, see Table 7). Results showed no significant main effect for type of verbalization, \(F(2, 45) < 1\) (see Table 8). In other words, as in research question 2, type of verbalization did not appear to have an effect on participants’ ability to produce new exemplars of the targeted structure.
Research Question 4

Finally, to determine whether type of verbalization had a significant effect on the amount of time it took participants to complete the reading task, times spent on task were submitted to a one-way ANOVA (for descriptive statistics, see Table 9). Results showed a significant main effect for type of verbalization, $F(2, 45) = 8.74, p < .05$ (see Table 10). Post hoc Scheffé tests indicated that this main effect was due to significant differences between the control and both experimental groups. Effect sizes calculated for the control/nonmetalinguistic comparison and control/metalinguistic comparison were large (0.98 and 1.65, respectively). In other words, both experimental groups (metalinguistic and nonmetalinguistic) spent significantly more time processing the L2 text when compared to the control group. The amount of time spent in the two verbalization groups did not differ significantly ($p = .405$).

DISCUSSION

Research Question 1

The first research question, which asked whether type of verbalization (nonmetalinguistic vs. metalinguistic vs. control) significantly affected readers’ text

Table 5. Production of old exemplars: Descriptive statistics

<table>
<thead>
<tr>
<th>Condition</th>
<th>Pretest M</th>
<th>Pretest SD</th>
<th>Posttest M</th>
<th>Posttest SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>0.35</td>
<td>0.702</td>
<td>4.35</td>
<td>2.597</td>
</tr>
<tr>
<td>[- meta]</td>
<td>0.44</td>
<td>0.784</td>
<td>5.17</td>
<td>3.666</td>
</tr>
<tr>
<td>[+ meta]</td>
<td>0.40</td>
<td>0.699</td>
<td>4.70</td>
<td>4.448</td>
</tr>
</tbody>
</table>

Table 6. Production of old exemplars: ANOVA for type of verbalization and time

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sums of squares</th>
<th>Mean squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of verbalization</td>
<td>2</td>
<td>3.602</td>
<td>1.801</td>
<td>0.310</td>
</tr>
<tr>
<td>Time</td>
<td>1</td>
<td>395.510</td>
<td>395.510</td>
<td>56.722*</td>
</tr>
<tr>
<td>Type of verbalization × time</td>
<td>3</td>
<td>2.300</td>
<td>1.150</td>
<td>0.165</td>
</tr>
</tbody>
</table>

*p < .05.
comprehension, was answered affirmatively. There was a statistically significant difference between the comprehension scores of the two experimental groups, with the nonmetalinguistic group performing significantly better than the metalinguistic group. Asking even advanced L2 readers to verbally provide additional information that might be extraneous to the reading process appears to have a reactive effect on their comprehension. Indeed, this conclusion is supported by qualitative data from the think-aloud protocols of some participants in the metalinguistic group. A few participants in this group (who scored low on comprehension and postassessment tasks) commented that it was difficult to follow the meaning of the text while verbalizing their thoughts and justifications. One participant was particularly explicit, saying, “OK, I’m not taking any of this in basically . . . I don’t know what any of this is going on, because I’m really not paying attention. This is distracting to have to talk while I’m doing this.”

No statistically significant difference was found between the control and either experimental group, supporting Leow and Morgan-Short’s (2004) findings at a lower language experience level that nonmetalinguistic verbalization was not reactive for text comprehension when compared to a control group. However, given that the comprehension task was found to have relatively low reliability, these findings need to be viewed with some caution.

Table 7. Production of new exemplars: Descriptive statistics

<table>
<thead>
<tr>
<th>Condition</th>
<th>New exemplars</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Control</td>
<td>3.06</td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>[- meta]</td>
<td>3.61</td>
<td>3.22</td>
<td></td>
</tr>
<tr>
<td>[+ meta]</td>
<td>2.40</td>
<td>3.56</td>
<td></td>
</tr>
</tbody>
</table>

Table 8. Production of new exemplars: One-way ANOVA for type of verbalization

<table>
<thead>
<tr>
<th>Source</th>
<th>$df$</th>
<th>Sums of squares</th>
<th>Mean squares</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of verbalization</td>
<td>2</td>
<td>9.5810</td>
<td>4.791</td>
<td>0.477</td>
</tr>
<tr>
<td>Within-groups</td>
<td>42</td>
<td>421.6190</td>
<td>10.039</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>421.6190</td>
<td>10.039</td>
<td></td>
</tr>
</tbody>
</table>
Research Questions 2 and 3

The second and third research questions, which asked whether type of verbalization significantly affected the production of old and new exemplars, respectively, were answered negatively. No significant difference was found among any of the groups for production of either old or new exemplars of the targeted structure. These results support and build on the findings of Leow and Morgan-Short (2004), who similarly found no significant difference in performance between nonmetalinguistic and control groups on intake or production tasks for a morphological structure. The present study, using a syntactic structure rather than a morphological one, also found both metalinguistic and nonmetalinguistic verbalization to be nonreactive for production, in apparent contradiction to Ericsson and Simon’s (1993) model, which predicted that verbalizing additional specific information such as justifications would affect task accuracy. It appears that in this study, the requirement to verbalize justifications did not have a significant impact on participants’ ability to produce either old or new exemplars of the targeted structure.

In addition to contradicting Ericsson and Simon’s (1993) model, this finding of nonreactivity for the metalinguistic group seems counterintuitive. After all, it stands to reason that if students—especially at this level of language

<table>
<thead>
<tr>
<th>Time on task (latency): Descriptive statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
</tr>
<tr>
<td>Control</td>
</tr>
<tr>
<td>[+ meta]</td>
</tr>
</tbody>
</table>

Table 9.

<table>
<thead>
<tr>
<th>Source</th>
<th>$df$</th>
<th>Sums of squares</th>
<th>Mean squares</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of verbalization</td>
<td>2</td>
<td>477.318</td>
<td>238.659</td>
<td>8.735*</td>
</tr>
<tr>
<td>Within-groups</td>
<td>42</td>
<td>1147.482</td>
<td>27.321</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.
experience—are required to provide justifications for their thought processes as they read a text and are subsequently asked to provide reasons for each of their answers, they will be pushed to think more analytically and possibly even to come up with a rule for the use of the targeted structure. However, in this case, although participants were asked to provide such justifications, their performance on production tasks was neither improved nor hindered.

Data from the think-aloud protocols were helpful in providing further insight into the similarities between the two verbalization groups. Participants in each group were sorted according to their performance on postassessment tasks to allow high and low scorers in each group to be compared. Interestingly, the protocols that the high and low scorers in each of the groups produced while reading the text were remarkably similar. High scorers from both the nonmetalinguistic and metalinguistic groups shared a common trait: They all reported awareness of the targeted structure at least at the level of meta-awareness as they read the text. Meta-awareness is an intermediate stage between awareness at the level of noticing and awareness at the level of understanding (Leow, 1997). According to Leow, a protocol is coded [+ meta-aware] when participants demonstrate a report of a cognitive change related to the targeted structure but do not verbalize the underlying grammatical rule. Compare, for example, the following protocols in (3) and (4).

(3) Metalinguistic high scorer
[Reading the text aloud, and correctly translating into English, Es inevitable preguntarnos también qué hubiera pasado si Stalin... ] “It’s inevitable to ask ourself what would have happened if Stalin... So hubiera is that would have stuff.”

(4) Nonmetalinguistic high scorer
[Reading the text aloud, Si George Walker Bush hubiera sido un argentino descendiente de algún inmigrante italiano... ] “Wait, so what is this saying? If George Walker Bush... hubiera sido un argentino... what is hubiera sido? Oh, so maybe it’s if George Bush had been an Argentine... because, yeah, maybe this is about what if he was Argentine and not American.”

Similarly, the low scorers in both groups did not appear to report awareness even at the level of noticing, as they did not comment on the targeted form as they read the text. This similarity between performers in both verbalization groups is reminiscent of P. Robinson’s (1995, 1997a, 1997b, 2001) suggestion that it might not be only the learning condition that affects allocation of attentional resources, level of awareness, and the extent of long-term learning but, rather, a combination of variables, including learners’ differing cognitive abilities. It also supports the findings of other studies conducted within an attentional framework that find a facilitative role for awareness in L2 learning (Leow, 1997, 2000, 2001a; Rosa & Leow, 2004a; Rosa & O’Neill, 1999).

The protocols of the high and low scorers in both groups on the postassessment tasks revealed another interesting trend. Although only the high scorers
in each group reported awareness of the targeted structure as they read the
text, as a group the participants in the metalinguistic condition (both high
and low scorers) generally showed some awareness of the function of the
unknown structure as they verbalized their justifications for their answers to
the production tasks; that is, the requirement to verbalize justifications for all
answers on the production tasks pushed the participants to have a certain
level of awareness. Compare, for example, the following protocols from par-
ticipants in the metalinguistic group as they completed the production tasks,
in (5) and (6).

(5) Metalinguistic low scorer
[Reads the sentence, “Quizás (he would have returned) _____ todos los días a su
casa a almorzar” and says:] “Quizás he would have returned, well, first of all, I
think quizás is [requires] subjunctive, but I’m not gonna mess with that because I
don’t really know. Would be, umm . . . I think the best thing to do is the imperfect,
umm although it sort of looks conditional because of the would, but it is in the
past for sure, so would have, wait, no, I think the best is . . . So here’s how we’re
gonna do it, just because I can do what I can . . . I think it’s habría ummm . . . volvido.”

(6) Metalinguistic high scorer
“So you use it [the targeted structure] in the past tense when there is doubt or
uncertainty, so it’s like, uh, maybe a past conditional.”

In the process of completing the production tasks and having to verbalize
the reasons for filling in each blank, both of these participants reported aware-
ness of the function of the targeted structure. However, the low scorer did
not report awareness of the targeted structure as he read the text, so he was
unable to use the targeted form from the text to fill in his gap of knowledge to
complete the production tasks correctly. He knew that the forms he was writ-
ing were incorrect, as evidenced by his statement, “So here’s how we’re gonna
do it, just because I can do what I can.” However, given that he did not dem-
onstrate a relatively high level of awareness while processing the targeted
structure, this might have contributed to his inability to complete the produc-
tion task with the correct forms. This explanation is plausible because the
high scorer did report awareness of the structure while reading the text and
was able to incorporate that knowledge with the justification to complete the
production tasks correctly.

Research Question 4

The fourth research question, which asked whether type of verbalization (non-
etalinguistic vs. metalinguistic vs. control) significantly affected the amount
of time on task, was answered affirmatively. Both verbalization groups spent
significantly more time on task than the silent control group. Moreover, the
metalinguistic group required more time on task than the nonmetalinguistic
group, although, in this study, the difference between the two was not statistically significant. This is not surprising in light of the predictions of Ericsson and Simon (1993) and the findings of reactivity studies in cognitive psychology (e.g., Berardi-Coletta et al., 1995; Brunk et al., 1958).

LIMITATIONS AND FUTURE RESEARCH

One limitation of the present study is the relatively small sample sizes of participants and the uneven distribution of participants in the metalinguistic group when compared to the other two groups. Additionally, the reliability on the comprehension task was found to be fairly low and caution is needed when interpreting the results of this task.

The finding that metalinguistic protocols are reactive for text comprehension when compared to nonmetalinguistic protocols should be examined in future studies, not just extrapolated to all situations. The text used in this study was quite long (over 800 words) and also included a difficult syntactic structure to which the participants had not been formally exposed. These factors could have had an impact on the results of the study, and future studies should therefore address factors such as text length and type, text difficulty, and linguistic structure. Although Leow and Morgan-Short (2004) and the present study have shed some light on the validity of verbal reports during L2 reading, future studies are warranted to investigate the reactivity of such protocols on different types of tasks, particularly because verbal reports have already been used in SLA research to measure learners’ attention and awareness during L2 problem-solving tasks (Leow, 1997, 1998, 2000, 2001a; Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999). Such investigations into reactivity might be even more crucially needed in light of the fact that several studies in cognitive psychology have found differential effects for reactivity among a common set of participants according to the task being performed (Rhenius & Defner, 1990; Russo et al., 1989). This has led some researchers to suggest that “the causes of reactivity are not general but due jointly to the demands of the task and to verbalization” (Russo et al., p. 763).

In line with the idea that the causes of reactivity are not general is the notion that individual differences might affect reactivity. This is a promising area for future research to investigate, and moderator variables such as working memory, cognitive speed, aptitude complexes, and learning style might be fruitful sources of investigation in this methodological strand of investigation. As one anonymous SSLA reviewer pointed out, if learners’ patterns of cognitive abilities or aptitudes differ, with some patterns facilitating and others inhibiting learning and performance under certain conditions, or on certain tasks (P. Robinson, 2001), it stands to reason that abilities and aptitudes might also affect verbalization under different conditions and on different tasks. Empirical data are clearly warranted to support or refute this premise.
CONCLUSION

The present study was designed to address the issue of reactivity of nonmetalinguistic and metalinguistic protocols in SLA research methodology and its related issue of time (latency). It expands on the work of Leow and Morgan-Short (2004), who found that nonmetalinguistic verbalization during an L2 reading task did not have any significant effects on first-semester Spanish learners’ text comprehension, intake, or production of the targeted morphological form. Using a syntactic structure rather than a morphological one and using advanced (fifth-semester) Spanish learners, the present study investigated the reactivity of both nonmetalinguistic and metalinguistic protocols. Results indicated that, compared to a control group, neither metalinguistic nor nonmetalinguistic verbalization significantly affected either text comprehension or written production of old or new exemplars of the targeted form. However, metalinguistic verbalization did appear to cause a significant decrease in text comprehension when compared to nonmetalinguistic verbalization. Finally, both metalinguistic and nonmetalinguistic verbalization had a significant effect on the amount of time spent processing the L2 text when compared to a control group.

Although verbal reports have been used to measure attention and awareness in a number of SLA studies, the validity of these reports is just beginning to be examined. If they are proven to be valid and to accurately represent learners’ mental processes, such protocols will provide insights into the language learning process, information that is unavailable in a classical pretest-posttest research design. However, given the inconclusive findings even in cognitive psychology, where studies on reactivity have been conducted for over 50 years, one should be careful in interpreting the findings of studies that employed verbal reports. This is especially the case given that the reactivity of verbal reports might vary according to a number of factors, including task type, text length and type, text difficulty, linguistic structure, and individual differences. There is clearly a need for future research in this strand of research methodology in SLA.

(Received 29 November 2004)

NOTES

1. Given that participants in the metalinguistic group were requested to comment aloud on “every thought and every detail of [their] thought process” in addition to providing a justification for their reasoning, the term metalinguistic could also be viewed in the present study to include metacognitive processes.

2. As suggested by one anonymous SSLA reviewer, the more sensitive Ryan test was used instead of the Scheffé test for the comprehension test because of its low alpha (.45) and its less than 20 within-cell sample size.
REFERENCES


APPENDIX

Experimental text

Si Bush hubiera sido porteño

¿Qué hubiera pasado en USA si George Bush hubiera nacido en La Boca? ¿Y si José Stalin hubiera obedecido a su madre, quien lo obligó a ser seminarista a los 15 años cuando su padre murió? Ni hablar si Adolfo Hitler hubiera sido aceptado en la Academia de Bellas Artes de Viena. Las vidas de estas personalidades no se puede cambiar, así como el fracaso de Vincent Van Gogh como pastor protestante y el sueño frustrado de Leonardo Da Vinci de ser un excelente Chef. Pero el futuro de la humanidad aún es incierto y está en nuestras manos.

Si George Walker Bush hubiera sido un argentino descendiente de algún inmigrante italiano, quizás hoy sería un tal “Jorge Ángelo Fossati” o el “Sr. Jorge Luigi Marincioni.” . . . es decir un tipo cualquiera sin trascendencia alguna (salvo para su familia y la gente del barrio).

Es probable que sus abuelos hubieran llegado al país en la oleada de 1880, siendo unos entre los miles de trabajadores expulsados por la primera Gran Depresión del sistema capitalista, a pesar de que desde 1850 se expandía a un ritmo galopante. Las riendas del liberalismo económico estaban en las manos de las potencias como Gran Bretaña, que iban conquistando a todo vapor nuevos mercados de ultramar, alimentando su ambición con el prestigio que representaban las colonias, a las que luego de varios años tuvieron que renunciar (aunque hay que admitir que a algunas naciones imperialistas todavía les cuesta aceptarlo).

Imagínense que Bush hubiera sido fruto del amor entre algún grotesco italiano y una robusta gallega. Si en vez de crecer en USA, bajo un techo republicano y piso de algodón, hubiera sido criado en el barrio portuario de La Boca, con su aroma particular de siempre, tangos de arrabal y en un escenario rodeado de humeantes chimeneas.

Si George Bush hubiera jugado a la pelota en la vereda, tomado mate en vez de milk shakes y hubiera visto ATC en vez de CNN, ¿Alguien conocería hoy a Bin Laden? Quizás “la nona” lo hubiera mandado al pequeño “Jorgito” a la escuela de la esquina con guardapolvo inmaculado y su padre le hubiera comprado con mucho esfuerzo un Manual Santillana para que estudiara las guerras mundiales ilustradas y las atrocidades de la historia universal condensadas en tan sólo unas hojas con mapas de colores.

Quizás, hubiera vuelto todos los días a su casa a almorzar, fascinado con los torpedos que lanzaban los submarinos alemanes a los buques norteamericanos en el Atlántico Norte durante 1917 o hasta le hubiera gustado jugar con soldaditos de plomo no de carne y hueso a la hora de la siesta. Quizás se hubiera conmovido con la historia de las 1000 grullas, que forma parte de la tradición oriental y consiste en hacer plegados artesanales para salvar las vidas de las víctimas de la radiación que liberaron las bombas nucleares en Hiroshima y Nagasaki (1945).

Tal vez, con el tiempo, el adolescente “Jorge” se hubiera estremecido al enterarse de la agonía de los 80.000 inocentes que fallecieron en los hospitales colmados de Tokio.
y las camadas de jóvenes emprendedores que perdieron sus vidas en la encarnizada lucha de Vietnam. Quizás no hubiera entendido la impaciencia de Truman por terminar la Segunda Guerra Mundial y mucho menos lo acordado por las potencias en Postdam, Alemania.

Es inevitable preguntarnos también qué hubiera pasado si Stalin le hubiera hecho caso a su madre, como la mayoría de los hijos solían hacer en ese entonces. Sin embargo, el desafiante “José”, fue un georgiano huérfano quien se convirtió en el “salvador de los soviéticos” y sucesor de Lenin en 1923 (pese a que éste especificó en su testamento, que por prudencia, no quería entregarle el poder). ¿Qué sería de Rusia si en el convento en el que estudiaba para ser cura, no hubieran descubierto que participaba en movimientos antizaristas en 1904? Es probable que Stalin terminara dando misa en una capilla perdida de algún pueblo polvoriento de Georgia (aunque su temperamento y su poca fe en Dios, no lo hubieran permitido).

También sería interesante saber qué hubiera pasado si a Adolfo Hitler lo hubieran aceptado en la Academia de Bellas Artes de Viena en Austria, su tierra natal. Quizás se hubiera evitado que volcara su frustración en el campo militar y esa mezcla incandescente de resentimiento, paranoia y delirios de grandeza no hubiera entrado en erupción, tras el ridículo Tratado de Versalles que firmaron las potencias vencedoras en 1919, aplastando a los alemanes e instalando un aire de incertidumbre, y temor que acechó la paz del mundo. (Se podría decir que hoy se respira un aire parecido, gracias a la urgencia de Colin Powell y Bush).

Si el joven George Bush hubiera vivido más cerca de lo que Nixon y Kissinger hicieron el martes 11 de septiembre de 1973 en la Casa de Gobierno chilena en la que estaba Salvador Allende, quizás hubiera entendido algunas de las razones por las que el indelible 11 de septiembre de 2001, se pagó con sangre inocente. Fue atroz, no hay duda, el terrorismo no debería existir. Pero existe y quizás deberíamos preguntarnos ¿por qué?

Source: El Correo
February 17, 2003, Buenos Aires