Problems in Output and the Cognitive Processes They Generate: A Step Towards Second Language Learning

MERRILL SWAIN and SHARON LAPKIN
The Ontario Institute for Studies in Education

This paper argues, and provides data to support the argument, that in producing an L2, learners will on occasion become aware of (i.e. notice) a linguistic problem. Noticing a problem can ‘push’ learners to modify their output. In doing so, learners may sometimes be forced into a more syntactic processing mode than might occur in comprehension. Thus, output sets ‘noticing’ in train, triggering mental processes that lead to modified output. What goes on between the original output and its reprocessed form, it is suggested, is part of the process of second language learning.

INTRODUCTION
In a recent article on second language speech production research, Crookes (1991: 117) states that ‘... the role of output (i.e. production or use) in the development of SL [second language] proficiency has largely been ignored or denied (e.g. Krashen 1989).’ The purposes of this paper are to argue that there are roles for output in second language learning, and to present some relevant data regarding one of those roles. Specifically, in this paper it is argued that problems that arise while producing the second language (L2) can trigger cognitive processes that are involved in second language learning.

The paper is organized as follows. First, we provide some background to our interest in the roles of output in second language learning. Second, a brief discussion of possible functions of output in second language learning is provided. In doing so, we are not minimizing the role of comprehension, or input, in SLA theory; rather, we wish to make the case that sometimes, under some conditions, output facilitates second language learning in ways that are different from, or enhance, those of input. Finally, we describe data we have gathered from young adolescent learners of French thinking aloud while writing in their L2, and our analyses, results, and interpretations of these data.

BACKGROUND
In Canada, and now in a number of locations in the United States, there are programs in which majority language children receive all or part of their education in a second language right from the beginning of school. These programs, referred to as ‘immersion’ programs, provide their students with a rich source of comprehensible input. Through this educational program, the teaching of a second language is integrated with content teaching. The goals of immersion...
education include both academic achievement in the content areas studied, and a high level of proficiency in the second language.

In general, research has shown that students who enter an immersion program in kindergarten score equivalently to native speakers of the target language by the end of elementary school on global tests of listening and reading comprehension. However, in their speaking and writing, they are clearly identifiable as non-native speakers and writers (see, e.g. Swain 1984; Genesee 1987). Although there is continued progress toward target language performance as the immersion students progress through the grades, their inter-language remains sufficiently ‘off-target’ as to be a cause for concern. For a theory of second language acquisition that claims that the only source of second language acquisition is comprehensible input, these results from immersion programs (and other input-based programs, see, for example, Lightbown and Halter 1993) provide strong counter evidence.

Immersion students become quite fluent in the target language and are generally able to communicate what they need and want to say in the academic context of their classroom. In their classroom, their use of the target language is to a large extent with their bilingual teachers who are able to understand the immersion students’ interlanguage, and with their peers, who speak roughly similar interlanguages. Somewhere along the way to target-language proficiency, when the ability to understand and be understood is achieved, the students’ second language development appears to slow down, and the push to develop their French beyond its current state diminishes.1 Our fundamental interest now is in how to help these immersion learners move beyond this stage in their second language development, to a more accurate, target-like French.

A consideration of the role of student output in this process was stimulated by a classroom-based study we conducted. In that study, we tape-recorded a day in the life of ten grade 6 immersion classes, and we discovered that there was surprisingly little French spoken by the students that was of any significant length. In fact, we found that only 15 per cent of the ‘public’ talk of the students was greater than a clause in length. We also found that feedback from the teachers on errors the students made, although not infrequent, seemed haphazard and random, based more on an ‘irritation factor’ than any rational pedagogical approach (Swain 1988; Allen et al. 1990).

FUNCTIONS OF OUTPUT
The relative lack of sustained talk in French in these immersion classes was an unexpected finding and led to a consideration of how output might be important in L2 learning beyond that of simply enhancing fluency. The proposal developed (Swain 1985, 1993) and briefly explicated below, goes as follows: in producing the L2, a learner will on occasion become aware of (i.e. notice) a linguistic problem (brought to his/her attention either by external feedback (e.g. clarification requests) or internal feedback). Noticing a problem ‘pushes’ the learner to modify his/her output. In doing so, the learner may sometimes be forced into a more syntactic processing mode than might occur in com-
prehension. Thus, output may set 'noticing' in train, triggering mental processes that lead to modified output. It is the purpose of this paper to determine if learners' own output does, on occasion, serve as an attention-getting device, and if it does, does it sometimes serve to stimulate the learners to engage in linguistic analysis?

But let us first examine the proposal in a little more detail. Schmidt and Frota (1986: 311) offer a 'notice the gap principle' which states, 'a second language learner will begin to acquire the targetlike form if and only if it is present in comprehended input and "noticed" in the normal sense of the word, that is consciously'. Our hypothesis is that output is one of the triggers for noticing. That is to say, in producing the target language, learners may encounter a problem leading them to recognize what they do not know, or know only partially. In other words, the activity of producing the target language may prompt second language learners to consciously recognize some of their linguistic problems; it may bring to their attention something they need to discover about their L2.

The 'output hypothesis' (stripped to its bare bones) is that even without implicit or explicit feedback provided from an interlocuter about the learners' output, learners may still, on occasion, notice a gap in their own knowledge when they encounter a problem in trying to produce the L2. The communication strategy literature (e.g. Tarone 1977; Faerch and Kasper 1983; Bialystok 1990; Kellerman 1991) provides evidence that learners do notice problems as they speak, and do try to do something about them. But what do they do when they notice a problem? Do they focus on morphology and syntax? In this paper we examine these processes as they are revealed through think-aloud protocols produced by L2 students while writing.

Other research (e.g. Pica, Holliday, Lewis, and Morgenthaler 1989; Iwashita 1993) has described what learners do linguistically when 'pushed' to modify their output. These studies indicate that during the process of negotiating meaning, learners will indeed modify their output in response to such conversational moves as clarification requests or confirmation checks. For example, Pica et al. (1989) found that in response to clarification and confirmation requests, over one-third of the learners' responses were modified either semantically or morphosyntactically. Although no one has yet shown directly that these modified, or reprocessed, responses are maintained in the learner's interlanguage, the assumption is that this process of modification contributes to second language acquisition.

In fact, a study by Nobuyoshi and Ellis (1993) is suggestive. Their study indicates, albeit with a very small sample size of only 3 experimental learners and 3 comparison learners, that 'pushing' these learners to improve the accuracy of their production results not only in immediate improved performance but also in gains in accuracy over time (Nobuyoshi and Ellis 1993: 208). In their conclusions, they suggest that it is useful to distinguish two meanings of acquisition: '1) acquisition as the internalization of new forms, and 2) acquisition as the increase in control over forms that have already been
internalized.' (ibid.: 210). Nobuyoshi and Ellis suggest that pushing learners to modify their output results in acquisition of the second type, that is, acquisition as an increase in the ability to deploy existing grammatical knowledge more accurately.

Pica et al. (1989), however, suggest that in modifying their output, learners may also be engaged in acquisition as the internalization of new forms. In modifying their output, they argue, learners 'test hypotheses about the second language, experiment with new structures and forms, and expand and exploit their interlanguage resources in creative ways' (Pica et al. 1989: 64). In other words, one might consider that the modified, or reprocessed, output represents the leading edge of the learner's interlanguage.

This is precisely the point made by Tarone (1993) who relates variation theory to second language acquisition through communication strategies. She argues that 'a learner's use of communication strategies can function to stretch an IL system beyond its current limits, resulting in free variation as the learner tests new hypotheses in the search for an appropriate word or structure' (Tarone 1993: 18). She bases her argument on data presented by Liu (1991) who studied a young Chinese child, Bob, over two years as he learned English. In his interactions with the researcher, relative to his interactions with peers or teachers, Bob's linguistic structures were the most complex and represented the widest variety of language functions. Constructions that Bob used in interaction with the researcher suggested that in these contexts Bob stretched his interlanguage system: he took 'risks with the interlanguage system in attempting to communicate' (Liu 1991: 16). Structures that appeared first in interaction with the researcher appeared later in other settings, providing strong evidence that, for this child at least, the researcher–Bob context is the one in which development occurs.

Tarone speculates as to what makes this context so positive for interlanguage development. After arguing that the provision of finely tuned input seems unable to explain the findings, she proposes that another possible explanation was Bob's attempt to produce comprehensible output . . . In this view, it is in those interactional contexts where the learner needs to produce output that the current interlanguage system cannot handle that the learner pushes the limits of the interlanguage system to MAKE it handle that output, thus keeping that system permeable and open to change. (Tarone 1993: 17)

Thus, research is beginning to accumulate evidence supporting the theoretical claim that 'pushing' learners beyond their current performance level can lead to enhanced performance, a step which may represent the internalization of new linguistic knowledge, or the consolidation of existing knowledge. In this paper, we wish to examine what mental processes learners engage in as they move from the original output to their modified output; as they move from encountering a linguistic problem in L2 production to developing a solution.

It has been demonstrated that the processes involved in producing language can be quite different than those involved in comprehending language. Van Dijk
and Kintsch (1983) have shown with native speakers that comprehension will sometimes rely on comprehension strategies rather than on a closed, logical system of rules required to produce a grammatical utterance. For example, in comprehending an utterance, a native speaker may make guesses about the probable structure of what they are hearing (or reading) based on syntactic or semantic clues. One strategy, for example, is 'try and make sense of a sentence using content words alone'.

Krashen (1982: 66) has pointed out that 'In many cases, we do not utilize syntax in understanding—we often get the message with a combination of vocabulary, or lexical information plus extra-linguistic information.' Similarly, Gary and Gary (1981: 3) state, in speaking of the comprehension-based approach to language teaching, that

Speech requires linguistically more complex tasks than comprehension. Comprehension—at least all but the most advanced levels—allows many linguistic signals to be ignored: redundant grammatical and semantic functions such as concord, definite/indefinite distinctions, singular/plural distinctions, etc., can very often be ignored without seriously distorting the message being comprehended.

In concluding their research report about a comprehension-based ESL program in New Brunswick, Lightbown and Halter (1993: 23) state:

In a sense, it is hardly surprising that students left on their own to acquire language purely from exposure to comprehensible input seem to need help with certain aspects of the language...[for example,] focused instruction and corrective feedback can help to fill these gaps and enhance their performance... The findings of this study also lend support to the view expressed by Swain (1985, 1988), Sharwood Smith (1986) and others that the kind of processing which is necessary for comprehension is different from the kind of processing which is required for production and, ultimately, for acquisition. As Cook (1991) has expressed it, the ability to decode language, that is, the ability to understand the meaning conveyed by a particular sentence, is not the same as code breaking, that is, discovering the linguistic systems which carry that meaning.

On the basis of theoretically and empirically based arguments such as these, Swain (1985, 1993) suggested that perhaps one function of output in second language learning might be to force the learner to move from the semantic processing prevalent in comprehension to the syntactic processing needed for production. It might be that producing language forces learners to recognize what they do not know or know only partially. This may trigger an analysis of incoming data, that is, a syntactic analysis of input, or it may trigger an analysis of existing internal linguistic resources, in order to fill the knowledge gap (see Figure 1, on page 388). The latter is essentially what Cumming (1990) found when he examined a subset of decision-making processes adult second language learners made while composing. He argued that the subset he identified might have potential for second language learning. We will return to Cumming's finding shortly, as similar issues are addressed in the study presented below.
THE STUDY

This study is a partial replication of Cumming's (1990) study. However, his learners were intermediate and advanced adult learners attending university, whereas our learners were young adolescent students in a French immersion program. Both groups of learners were being instructed through the medium of their second language. Cumming was interested in the composing processes of his learners and while classifying his data noticed reports of thinking processes which he inferred, from previous research and theory, might relate to processes of second language acquisition. We, however, began with a specific interest in whether our young learners would identify problems as a result of trying to produce the target language and what they might report doing to overcome them. Could what they reported be related to processes of second language learning? Would we see evidence of grammatical analysis as they struggled to produce the target language?

Thus, our research questions in this study were as follows:

1. As young adolescent learners produce their L2, do they ever become aware of the gaps in their linguistic knowledge? In other words, can producing the L2 lead to a conscious awareness by a learner of what he/she does not know or knows only partially?
2. If young adolescent learners do become aware of linguistic gaps in their knowledge as a result of attempting to produce the L2, what do they do? Specifically, are cognitive processes ever triggered which theoretical and empirical accounts suggest are involved in second language acquisition?
3. Do these learners ever engage, in trying to solve their linguistic problems, in grammatical/syntactic analyses?

METHODOLOGY

1. Subjects

The subjects in this study were 18 students from a grade 8 early French immersion class of 21 students. The 18 students were those whose parents had signed permission forms allowing their children to participate in the study. Students in this class were from a lower-middle to middle class socio-economic background (the school is located in the inner city of a large metropolitan area in Ontario, Canada). The students, whose average age was 13, had a wide range of academic abilities.

The students also had had generally similar educational backgrounds in that they had all participated in an early French immersion program since kindergarten. Immersion teachers are encouraged to use approaches that are similar to those being used by their English language counterparts. In Ontario, where this study was conducted, these approaches reflect a provincially recommended learner-centered philosophy of education and a whole-language approach to language teaching which emphasizes process writing. In the early
grades (up to grade 3), the students were instructed entirely in French. In later
grades, instruction in English was introduced for as much as 50 per cent of the
day. In grade 8, the students were taking history, geography, mathematics, and
French language arts in French.

In spite of the overall experiential nature of L2 learning in the immersion
classroom, formal grammar instruction is often an important feature of the
French immersion classroom. It is thus highly likely that as these students
entered their grade 8 year, they would have been exposed to an eclectic second
language teaching approach consisting of learner-centered activities fortified
with a regular dose of traditional grammar activities (Kowal and Swain 1994).

Of the 18 students from whom we collected data, 9 were selected for data
analysis. The procedures for selecting these 9 students were as follows. At the
end of the year, their French language arts teacher ranked her students
according to their overall proficiency in French. The top two students and the
bottom two students were chosen for inclusion in our sample. These students
are referred to as the 'most proficient' and 'least proficient' students respec-
tively. Then 5 additional students were randomly selected from among the
middle group. The final sample consisted of 6 girls and 3 boys.

2. Procedure
The writing task was developed in consultation with the students' homeroom
teacher. It was decided that students should be given a theme that they had
covered in class. Their familiarity with the topic, it was hoped, would allow the
students to focus attention on their writing and the need to think aloud. An
environmental theme was therefore chosen.

Each student met individually with the researcher in a small, quiet room.
Procedures suggested by Ericsson and Simon (1993) for eliciting think-alouds
were followed. Students first received a brief introduction in English outlining
the task which they would be expected to do. They were told that the researcher
was interested in knowing what they were thinking as they were writing. In order
to accomplish this, they would be asked to write a short article in French of one
or two paragraphs, on an assigned topic. As they performed the writing task,
they were to think aloud. To illustrate, the researcher thought aloud while
solving a multiplication problem. The students were then given a similar
problem to solve.

The students were then given the specific writing task: 'You are a reporter for
your local newspaper. You have been asked to write an article about an environ-
mental problem facing your community. In your article, describe the problem
and offer one or two possible solutions to the problem.'

The students were told that they must write in French, but that it would be fine
for them to use either French or English as they thought aloud. They were asked
not to erase any mistakes that they might make, but simply to cross them out and
continue. They were told that if they stopped talking for very long, the
researcher would remind them to think aloud.

Whenever students made a change without commenting on it, they were
promoted to verbalize what they were thinking. The probe used was ‘what are you thinking?’ Finally, students were advised that they could not have access to a dictionary or any other aid, and that the researcher would not be able to help either. These last conditions were imposed because we were interested in seeing what students would do without feedback; whether they would try to work out solutions on their own. They did, but they occasionally found the situation frustrating as indicated by this comment made by S8: ‘I want to say the major cause, Uhm. OK. La cause, la cause, uhm. I would use a dictionary if I had one right about now!’

Upon completion of a draft, students were asked to edit their work using a red pen to record any of their changes, and to think aloud while they were editing.

A tape-recorder was placed on the table where the student was working. The researcher remained in the room, sitting close to the student so she could see when the student crossed out something. Each session lasted approximately one hour.

LANGUAGE-RELATED EPISODES

1. Definition
Language-related episodes were identified in the think-aloud protocols of our sample of students. We defined a language-related episode as any segment of the protocol in which a learner either spoke about a language problem he/she encountered while writing and solved it either correctly (see example 1) or incorrectly (see example 2); or simply solved it (again, either correctly or incorrectly) without having explicitly identified it as a problem (see example 3).

Example 1 (translations in italics)
S17: [S17 has written an article about how phosphates released into lakes and oceans cause plants in them to grow quickly to such an enormous size that they will kill all the fish. She struggles in the following think-aloud episode with how to say ‘kill all the fish’] ‘et mort (and dies). I don’t know. I don’t know because mour . . . mourir les poissons (to die the fish), it’s like mourir is something that you do. It’s not something that someone does to you. So it’s more like they’re being murdered and not dying. So, uhm, et tue toutes les poissons (and kills all the fish), or something like that.’

In example 1, the student has produced ‘mort’ (die), and is not pleased with this lexical choice. Her explanation for the change she makes is in fact a rather sophisticated one, relating to the need to use a transitive verb ‘tuer’ (to kill) rather than an intransitive one ‘mourir’ (to die). Recognizing her difficulty, she searches her own linguistic knowledge for information which might help to solve her problem. Of course, we cannot tell whether the analysis she conducts reflects some generalized knowledge she has of transitivity/intransitivity that she is possibly applying for the first time to the difference in meaning between ‘mourir’ and ‘tuer’, or whether she is struggling consciously for the first time with the concept as she senses the difference in meaning between the two verbs. If the former, then what she would seem to be doing is consolidating her own linguistic knowledge in applying it to a known, or possibly new, context. If the latter, then
what she would seem to be doing is working out, on-line, a sophisticated linguistic rule based on a difference she senses in the meaning of the two verbs.

Example 2

S3: [S3 has just written: ‘Il y a trop d’utilisation des chemicaux toxiques qui détruisent l’ozone’. (There’s too much use of toxic chemicals which destroy the ozone layer.) In his think-aloud, we hear him trying to produce a noun form of the verb he has just used:] ‘La dé...truc...tion. Et la détruction. No, that’s not a word. Démolition, démolisson, démolition, démolition, détruction, détrusion, détrusion, la détrusion des arbres au forêt de pluie (the destruction of trees in the rain forest).’

It is difficult to know in example 2 the extent to which the student’s attempts to produce ‘la détruction’ are English- or French-based. Both languages are probably influencing his choice of words (e.g. demolish/démolir). It is clear from his pronunciation, however, that the endings he tries out are French, being either the noun suffix ‘on’ or ‘tion’ (Grevisse 1980). His final solution, ‘la détrusion’ is wrong, but he has made use of his knowledge of French by using the stem of the verb he has just produced and adding a French-sounding suffix. This example is revealing, because the incorrect solution allows us to conclude that new knowledge has been created through a search of his own existing knowledge. His search began with his own output which he heard as incorrect.

Example 3

S3: [S3 is writing about how products such as aerosols destroy the environment. He then says aloud the following, while writing:] ‘C’est pour ça qu’on doit arrêter . . . qu’on doit les arrêter. (That’s why one must stop . . . one must stop them.)’

In example 3, the sentence would have been better rendered as ‘On doit arrêter de les utiliser’. However, the student appears to have recognized the need for ‘arrêter’ to have a complement, and inserts one to refer to ‘the products’. The initial omission, or incorrect placement, of the direct object pronoun is a relatively common error made by immersion students (Selinker, Swain, and Dumas 1975), and this student’s self-correction to include it could be considered as a step in his linguistic development, if only as a consolidation of procedural knowledge.

Returning now to our definition of language-related episodes, as we have said, each one is related to a problem the student had with the production of the target language. In almost all of our examples, the resolution of the problem involved drawing simultaneously on both gist (meaning) and language use (Scardamalia and Paris 1985). This was also the case in the episodes reported by Cumming (1990) in discussing ones that he thought might reflect cognitive processes underlying second language learning among adult learners. However, in a few instances in our data, only language use was focused on, and these tended to be in reference to the written form of the target language (see example 4). We have included these as language-related episodes because we consider them to be examples of the consolidation of linguistic knowledge as the learner applies existing knowledge to old contexts, or as would seem likely in example 4, to new contexts.
Example 4
S16: [S16 writes 'Pour solvè'] 'OK, I shouldn't have an accent there. I should have an “r” because it's l’infinitif.'

These language-related episodes are quite different from another type of episode, which we have not analyzed, but which appeared in our students’ think-aloud protocols. These episodes, dealt largely with the generation of ideas, usually in preparation for composing (see example 5):

Example 5
S9: ‘Ok, she [the reporter] can uhm, write about, like cutting down trees and how it's really a big factor and then she can write, uhm she could tell them how they could like use other products. Like if you want a table made out of wood it would make more sense, ok, well, if you wanted, like, a table out of wood, uhm you could uhm like take a tree that’s fallen down say in a rain storm or a wind, uhm, a hurricane or something like that, and since it’s already dead that would make more sense in taking it than chopping down new healthy grown trees that still have a life to live.'

Each language-related episode dealt only with one linguistic item. Sometimes episodes overlapped, and sometimes one was embedded in another. In example 6, there are three episodes. One episode continues throughout: what tense should be used for 'envoyer', and how is it formed. Embedded in that larger episode are two others: one where the student changes 'de chemiques' to 'des chemiques'; and one in which 'qui' is inserted as subject of 'envoyer':

Example 6:
S12: [S12 has just written: 'Cette problème est causé par de chemiques envoyé' (This problem is caused by chemicals sent)]. 'Uhm, is it passé composé or what? Cette problème est causé par des ... just had de, so you change it 'cause it's des chemiques. Envoyé ... sent ... well it might be passé composé because xxx. It's in the past. So, I ... ‘cause it says that are sent. So I could change this. Cross out envoyé and write uhm des chemiques qui sont envoyé. If I did it that way [the way she originally had it] then I would have to put it in either uhm imparfait or, or in infinitive or a verb tense, but I think it would be better in passé composé, and I have to put uhm who are. I can't just put ... I have to put which are, or whatever, so I have to add the qui, and then because it's passé composé I add the uhm sont before the verb. Wait, would it be sont? Uhm, Ya.'

2. Reliability
Four researchers independently identified the language-related episodes in one think-aloud protocol and subsequently discussed their decisions. Where there was disagreement initially, it was possible to reach a consensus among the researchers as to what precisely a language-related episode was and how to identify it as detailed in the preceding section. Subsequently, another protocol was examined by the same four researchers, and, although there were minor discrepancies about when an episode actually began or ended, there was full agreement on what constituted an episode.

The next step was to categorize the language related episodes.
3. **Classification**

We proceeded to classify the data—the language-related episodes—in the following way. The same four researchers who identified the language-related episodes in the two protocols mentioned above, independently categorized each episode from one protocol. No categories were pre-established; rather they were entirely data-dependent.

Our intent was to categorize each language-related episode according to how the learners solved, in the way that they did, the linguistic difficulties that they identified as they produced the target language. In other words, *we wanted to try to arrive at the mental processes we thought were reflected in the changes the students made to their output.* In our first pass through the data, we tried to stay as close to the data as possible. A second-level analysis presented later relates these categories to cognitive processes that have been identified in the SLA literature as potentially involved in second language learning.

The descriptive categories were discussed at a meeting in which consensus emerged among the four researchers as to the categories and labels for them. These seven descriptive categories are listed below along with examples:

1. Sounds right/doesn’t sound right
   a) lexical
      
      Example 7
      S12: ‘No, I’m not going to write cette *(this)* because cette ozone layer doesn’t go very well. Uh, I know, I could write cette problème.’
   b) grammatical
      
      Example 8
      S8: ‘I was gonna write les droits uhm d’animaux, but it doesn’t sound right so I said les droits des animaux *(animal rights).*’

2. Makes more sense/doesn’t make sense
   
   Example 9
   S15: ‘Puis il reconstruit *(Then it rebuilds)*. Puis il les reconstruit *(Then it rebuilds them)*. I added that because it makes more sense to me. Uh, because it rebuilds them.’

3. Applied a grammatical rule
   
   Example 10
   S12: ‘Porter du *(To wear some)*. No not du because it’s des. It’s plural. Des vêtements *(some clothes).*’

4. Lexical search
   a) via English
      
      Example 11
      S11: ‘utiliser la ... jar. I’m trying to translate a word in my head and I’m trying to find the word for jar. Jar. Jar. Jar. Uh, contenu.’
PROBLEMS IN OUTPUT

Example 12
S9: 'Peut-être ils sont trop, (Maybe they are too) uhm, lazy, lazy, lazy, lazy, lazy, lazy, uhm, trop. I'm thinking of lazy. Ils . . . no. I don't know how to write lazy and I'll never be able to figure that out so I have to change the structure of the sentence so I can write something else instead of lazy. Peut-être ils n'ont pas assez d'énergie pour marcher. (Maybe they haven't enough energy to walk.)'

b) via French
Example 13
S8: ' . . . laissent échapper des chemiques qui sont, qui, no, qui créent (let chemicals escape which are, which, no, which create). I was going to say ces usines laissent échapper des chimiques qui sont harmful to the environment but I decided to put ces usines laissent échapper des chimiques qui, uhm, créent de la pluie acide (these factories let chemicals escape which create acid rain).'

c) via both
Example 14
S9: 'J'ai assis sur un banc, (I sat down on a bench) une . . . She can sit like on a bench to start writing her article about cutting down the trees, but I don't know how to say banc, so I'll have to write bench. Can I write bench? Or do I have to write it in French? Un banc, chaise, I'll just write chaise (chair).'

5. Translation (phrase or greater)
Example 15
S8: 'The problem of acid rain is generally becoming worse. Le problème de la pluie acide uhm dans notre communauté devient de plus en plus, plus sévère. (The problem of acid rain in our community is becoming more and more serious.)'

6. Stylistic
Example 16
S12: 'A cause de ça, ceci. (Because of this.) Because you're not supposed to use ça when you're writing.'

7. Spelling
Example 17
S12: 'I'm thinking about how to spell faible (weak). It might be the way you spell it in English but I don't know.'

4. Reliability
The four researchers independently classified the language-related episodes in the second protocol and attained a consensus on how to classify each episode. Following that, one researcher classified the remaining data (7 protocols). The episodes of one of those protocols, selected at random, were independently classified by a second researcher. Agreement was 82 per cent.

5. Comparison of our categories with Cumming's categories
As mentioned above, Cumming (1990) categorized the decision-making processes used by adult second language learners while composing that he thought might have potential for second language learning. His categories were:
(1) searching out and assessing improved phrasing; (2) comparing cross-linguistic equivalents; and (3) reasoning about linguistic choices.

According to Cumming, the most conspicuous cognitive activity in the first category 'involved searching for the “right” word to express an idea' (Cumming 1990: 491). This category overlaps with our lexical search category (see examples, 11, 12, and 13). Our lexical search category, however, incorporates examples of Cumming's second category, comparing cross-linguistic choices (see example 14, as well as example 15 under the category of translation). As Cumming argues, 'these efforts to equate expressions in their two languages [can be interpreted] as productive strategies wherein learners try to gain greater control over the knowledge they have previously acquired about the foreign language, integrating it with their knowledge of their mother tongue ...' (ibid.: 496).

Cumming's third category, reasoning about linguistic choices, is reflected in our categories of applying a linguistic rule (see examples 1, 4, 6, and 10). It probably also overlaps with what we have labelled stylistic (see example 15).

What is not reflected in Cumming's categories (but appear in some of his data: personal communication 1994) are our first two categories of (1) sounds right/doesn't sound right and (2) makes more sense/doesn't make sense. Reasoning in these ways seems important to our subjects. There are at least two reasons why our learners may rely more on sound and ‘sense’ than Cumming’s learners. First, our subjects are young adolescent learners, not adult learners, and may thus in general be less analytic about language. Second, our subjects are early immersion students whose initial experience with their second language was at an early age, and was highly experiential and communicative. As with one's native language, there are many times one is aware that something is non-native-like, but cannot say why. It would be interesting to know if late immersion students, who begin their intensive exposure to their second language at a later age and are immediately taught about French, would be as reliant on sound and ‘sense’ for their decisions about linguistic correctness.

FINDINGS
The first result is the finding that young adolescent second language learners do indeed become aware of gaps in their linguistic knowledge as they produce their L2. Furthermore, when they encounter difficulties in producing the target language, they do engage in thought processes of a sort which may play a role in second language learning. They do so even when external feedback is unavailable to them. It will take further research to trace the effect of these cognitive processes on learning. However, the thought processes identified represent processes similar to those other theorists and researchers (e.g. Selinker 1972; Cohen and Robbins 1976; Corder 1981; Kellerman and Sharwood Smith 1986; McLaughlin 1987; Larsen-Freeman and Long 1991) have hypothesized to be involved in second language learning: extending first language knowledge (particularly meanings) to second language contexts; extending second language knowledge to new target-language contexts; and formulating
and testing hypotheses about linguistic forms and functions. In doing so, these learners sometimes engage in grammatical analysis which, though not essential to comprehension, is essential to accurate production.

It needs to be pointed out that, for these learners at least, the substance of their thoughts was sometimes faulty, leading to incorrect hypotheses and inappropriate generalizations, suggesting that relevant feedback could play a crucial role in advancing their second language learning.

Other results relate to the quantification of these adolescent learners' thoughts, and as with their identification, are offered as a description of what occurred. As noted in the discussion of methodology, students first wrote a draft of their article and then were given an opportunity to make corrections (editing phase). During the original drafting phase, 113 language-related episodes were noted. Seventy-seven were noted during the editing phase. Combined, they provide 190 occasions in which students consciously recognized a linguistic problem as a result of producing, or trying to produce, the target language. Our claim is that on each occasion, students engaged in mental processing that may have generated linguistic knowledge that is new for the learner, or consolidated existing knowledge. The descriptive categories begin to identify the types of cognitive processing that occurred.

The data are summarized in Tables 1 and 2. Table 1 shows the percentage of language related episodes that fall into each category for the drafting phase and the editing phase. Table 2 shows the same information, but limited to the two highest-proficiency students and the two lowest-proficiency students. Both tables also indicate the total number of language-related episodes that occurred for each phase of the writing task. Table 1 shows that lexical searches account

<table>
<thead>
<tr>
<th>Table 1: Percentage of language-related episodes during draft and edit phases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive categories</td>
</tr>
<tr>
<td>Sounds right/doesn't sound right</td>
</tr>
<tr>
<td>—lexical</td>
</tr>
<tr>
<td>—grammatical</td>
</tr>
<tr>
<td>Makes more sense/doesn't make sense</td>
</tr>
<tr>
<td>Applied a grammatical rule</td>
</tr>
<tr>
<td>Lexical search</td>
</tr>
<tr>
<td>—via English</td>
</tr>
<tr>
<td>—via French</td>
</tr>
<tr>
<td>—via both</td>
</tr>
<tr>
<td>Translation</td>
</tr>
<tr>
<td>Stylistic</td>
</tr>
<tr>
<td>Spelling</td>
</tr>
<tr>
<td>Total (per cent)</td>
</tr>
<tr>
<td>Total N</td>
</tr>
</tbody>
</table>

Copyright (c) 2004 ProQuest Information and Learning Company
Copyright (c) Oxford University Press
Table 2: Percentage of language-related episodes during draft and edit phases for the two most-proficient and the two least-proficient students

<table>
<thead>
<tr>
<th>Descriptive categories</th>
<th>Draft phase</th>
<th></th>
<th>Edit phase</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'Most-proficient'</td>
<td>'Least-proficient'</td>
<td>'Most-proficient'</td>
<td>'Least-proficient'</td>
</tr>
<tr>
<td>Sounds right/doesn't sound right</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>—lexical</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>—grammatical</td>
<td>9</td>
<td>10</td>
<td>24</td>
<td>18</td>
</tr>
<tr>
<td>Makes more sense/doesn't make sense</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Applied a grammatical rule</td>
<td>19</td>
<td>20</td>
<td>48</td>
<td>15</td>
</tr>
<tr>
<td>Lexical search</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>—via English</td>
<td>30</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>—via French</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>—via both</td>
<td>9</td>
<td>15</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Translation</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stylistic</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Spelling</td>
<td>13</td>
<td>10</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Total (per cent)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total N</td>
<td>47</td>
<td>20</td>
<td>21</td>
<td>20</td>
</tr>
</tbody>
</table>

for 50 per cent of the language-related episodes during the drafting phase, but only for 6 per cent of them during the editing phase. During the editing phase, however, more attention is focused on whether it makes sense, how it sounds, and whether it is grammatically accurate. Indeed, these categories account for 75 per cent of the language-related episodes during the editing phase.

The data from only four students are shown in Table 2. We therefore consider Table 2 only as a source of hypotheses rather than as data from which any conclusions might be drawn. What is interesting is that during the drafting phase, the most-proficient students produced in absolute frequency more than twice the number of language-related think-aloud episodes as the least-proficient students ($N = 47$ vs. $N = 20$). Furthermore, in the editing phase, the most-proficient students appear to pay considerably more attention to grammar than the least-proficient students, both in terms of what sounds right (24 per cent vs. 18 per cent of language-related episodes) and in their application of grammatical rules (48 per cent vs. 15 per cent). We also think that it is worthy of note that the difference is much greater in the category of application of grammatical rules than in that of what sounds right. We wonder if this is evidence of greater analytic skills among the higher-proficiency students? Another way of looking at these particular data is that the higher-proficiency students are twice as likely to rely on applying a grammatical rule (48 per cent) than on what sounds right (24 per cent); whereas the lower-proficiency students
are about equally as likely to rely on either (18 per cent vs. 15 per cent) to solve their linguistic difficulty.

These findings are consistent with other research which suggests that conscious knowledge of rules is involved with greater L2 accuracy (Hulstijn and Hulstijn 1984). Hawkins and Towell (1992: 109) suggest that a plausible explanation is that ‘conscious knowledge of rules facilitates in some fashion the development of subconscious knowledge’, although they do not rule out another possible explanation that ‘more advanced learners ... are able to convert subconscious acquired knowledge into verbalisable statements, and less advanced learners are not’. In either case, the association between conscious knowledge of rules and greater L2 accuracy is worthy of note, and one that we will continue to explore in our future research.

In an attempt to see the data-driven set of categories derived from the language-related episodes in a more abstract way, we have extracted three general categories of cognitive processes implicated in second language learning: generating alternatives, assessing those alternatives, and applying the resulting knowledge. These general categories are shown along the left-hand side of Table 3. The columns in Table 3 represent aspects of language that might be focused on: spelling, morphology, syntax, vocabulary, discourse, register, genre, etc. We have filled in Table 3 with the number of each example provided in the body of this paper so that readers can examine for themselves directly the relationship between the examples and the categories of mental processing. That we have been able to do so suggests that the ways in which students solved problems as they identified them while producing the target language may have potential for language learning. Gaps in Table 3 probably reflect foremost the nature of the task that students were asked to do, and secondly, the areas that are not of major attentional focus for students with the sort of second language learning background they have, at this age of early adolescence, and at their level of second language proficiency. Instruction might be able to focus their attention elsewhere.

Finally, in Figure 1, we provide a sketch of second language learning from an output perspective. It is important to reiterate that output is not the only source of second language learning. Rather, under some conditions—that is, when, as a result of producing the target language, learners ‘notice’ a problem, they conduct an analysis leading to modified output. That is, noticing may occur because of either internal or external feedback which may prompt, for example, the generation of alternatives and assessment of them through simple inspection through to complex thinking. When learners cannot work out a solution, they may turn to input, this time with more focused attention, searching for relevant input. Or, they may work out a solution, resulting in new, reprocessed output. What goes on between the first output and the second, we are suggesting, is part of the process of second language learning.

In summary, this study has suggested that the communicative need engendered by the task did force the learners into thinking about the form of their linguistic output. In other words, it moved learners from semantic to
Table 3: Cognitive processes implicated in second language learning and triggered by output

<table>
<thead>
<tr>
<th>Cognitive processes</th>
<th>Aspects of language focused on</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spelling</td>
</tr>
<tr>
<td>Generate alternatives (search own knowledge; generate hypotheses)</td>
<td>17*</td>
</tr>
<tr>
<td>Assess alternatives (puzzle over; mull over; test hypotheses)</td>
<td>17</td>
</tr>
<tr>
<td>Apply:</td>
<td></td>
</tr>
<tr>
<td>- existing knowledge to known contexts</td>
<td>8?, 10</td>
</tr>
<tr>
<td>(increase in control; consolidation)</td>
<td></td>
</tr>
<tr>
<td>- existing knowledge to new contexts</td>
<td>4, 8?</td>
</tr>
<tr>
<td>(increase in control; consolidation)</td>
<td></td>
</tr>
<tr>
<td>- new knowledge (∼internalization of new heuristics/rules/forms/knowledge)</td>
<td>2</td>
</tr>
</tbody>
</table>

*The numbers represent the number of the example in the text of this article.
grammatical processing. The data have allowed us to derive a set of descriptive categories for the language-related think-aloud episodes. Reflection on the categories led to our identification of two different dimensions: cognitive processes involved in second language learning, and aspects of language focused on. Furthermore, the fact that in the domain of grammatical analysis, there appeared to be important differences between higher- and lower-proficiency learners may have implications for instruction in the second language classroom.

(Revised version received January 1995)

ACKNOWLEDGEMENTS
Drafts of this paper have been presented in seminars at Cornell University and the University of Pittsburgh, and at the 1994 AAAL conference in Baltimore. We would like to thank Elizabeth Langs who helped us collect the data, and who, along with Sylvia Spielman, helped us identify and classify the language related episodes. Also, we would like to thank our friends and colleagues who read and commented on earlier versions of this paper: Andrew Cohen, Alister Cumming, Birgit Harley, Christina Bratt Paulston, and Elaine Tarone. We are particularly grateful to Kees de Bot who wrote a detailed response to our interpretation of Example 1 and to Helen Moore who wrote a critical and insightful term paper about an earlier draft of our paper (Moore 1994). We wish to express our thanks, too, to the three anonymous reviewers of this paper. Each reviewed our paper from a different theoretical stance: cognitive, sociocultural, and neuro-biological. Each found something of value in our paper; each found things to question and criticize. Although we recognize the utility of a theoretical account which provides more of a role for the social context, or a neurological basis for SLA, the present paper remains embedded in cognitive terms for reasons of comparability with previous work, and because we consider it a worthwhile intellectual and empirical exercise to examine if there might be identifiable functions of output in second language learning that can be isolated and understood in cognitive terms. In this paper, our major goal is to determine if, as a result of trying to produce the L2, internal thought processes are triggered which are implicated in second language learning. From a sociocultural perspective, the think-aloud protocols we have obtained are an attempt to capture the learner’s internal dialogue without interruption, without disturbing what she/he knows (or does not know). Last, but
by no means least, we would like to thank the Social Sciences and Humanities Research Council of Canada for funding this and other research on 'The Output Hypothesis' (Grant # 410-93-0050).

NOTES
1 In a recent paper, Tarone and Swain (in press) suggest a possible motivation for the slow-down in the students' second language development, namely the increased social need to talk to each other in a vernacular. They argue that in the immersion classroom, a diglossic situation develops in which the vernacular is the student's L1, English. The percentage of English language use thus increases over time in peer–peer interactions, possibly causing a loss of interest and motivation to develop the French L2 superordinate beyond its current state.

2 The student meant *et meurt*, the correct third person singular form of the present tense of the verb *mourir*.

3 *Contenu* technically means ‘contents’; but the student probably means *contenant* or ‘container’.

4 The student may have been thinking of the word ‘feeble’.

REFERENCES


Gary, J. D. and N. Gary. 1981. 'Talking may be dangerous to your linguistic health: The case for a much greater emphasis on listening comprehension in foreign language instruction.' *IRAL* 19: 1–14.


