

Corrective Feedback in L2 Assessment: Negative Evidence and Interactive Practice

Norbert Francis

Northern Arizona University

norbert.francis@nau.edu

Recent research in the area of instructed second language (L2) acquisition has suggested that certain kinds of corrective feedback can be exploited effectively by learners. If this suggestion turns out to be correct, systematic feedback on learner error could be integrated into L2 teaching practice as part of ongoing assessment. More generally, calling attention to errors can be considered as one type of negative evidence: information regarding which aspects of the developing interlanguage system do not form part of the target language (TL) system. One point of view might propose that the usefulness of negative evidence is an empirical question to be settled on a case-by-case basis (which types might be useful to the learner and which types not). This paper will examine the effectiveness of corrective feedback that is provided by machine in an evaluation of one widely utilized computer program for second language learning. One of the central issues in debate regarding corrective feedback (both by hand and by machine) is the reliability of feedback to the learner. If grammatically correct patterns produced by the learner are marked incorrect (for example if the instructor perceives a pragmatic difficulty, or if the computer program cannot discriminate between a syntax error and a semantic problem) the learner receives conflicting and contradictory information about the target language grammar. The paper will conclude with a discussion of ways in which L2 educators can improve the reliability of corrective feedback and how they can use technology more effectively in L2 teaching. Since Computer Assisted Language Learning (CALL) is an invaluable resource today for self-teaching, especially for the student without access to classes, what design features should learners look for to maximize efficiency and effectiveness?

INTRODUCTION

Of the many kinds of assessment, corrective feedback provided to language learners can be considered as a formative type. As learners form hypotheses about the second language grammar, they receive information about which aspects of this developing system (sometimes called an interlanguage system) should be reconstructed or abandoned. To the extent that the relevant research questions can be formulated clearly, they will be among the most important areas of investigation for years to come, as well as ones that will continue to evoke some of the most interesting controversies in the field of language teaching. One reason why they are so important and interesting is because, unlike other questions in second language learning that primarily are either of theoretical or practical relevance, resolving the corrective feedback question is equally relevant to both theory and classroom practice.

The present discussion will touch on three related topics:

- (1) In line with the theme of corrective feedback as a kind of formative assessment, there is the more general concept of negative evidence,
- (2) The research on the effectiveness specifically of corrective feedback in second language learning, and
- (3) The problems of corrective feedback by machine.

NEGATIVE EVIDENCE AND POSITIVE EVIDENCE

Corrective feedback can be considered as belonging to the broader language input category of negative evidence, defined as all type of information available to the language learner about which patterns are not grammatical in the target language. Thus, in addition to direct and indirect feedback from “external” sources, negative evidence might also be generated from self-correction, and metalinguistic reflection, “internal” sources,¹ so to speak. In turn, the distinction between negative evidence and positive evidence in regard to the role that each plays in language learning is of central theoretical importance, being one of the key dimensions when considering the differences between first language (L1) development and second language learning. The complex questions of which acquisition/learning mechanisms are shared between L1 and L2, and in what way grammatical development differs between the two, directly inform the discussion on the role of negative evidence in L1 and L2, as well as the related issue of the role of explicit learning and direct instruction.

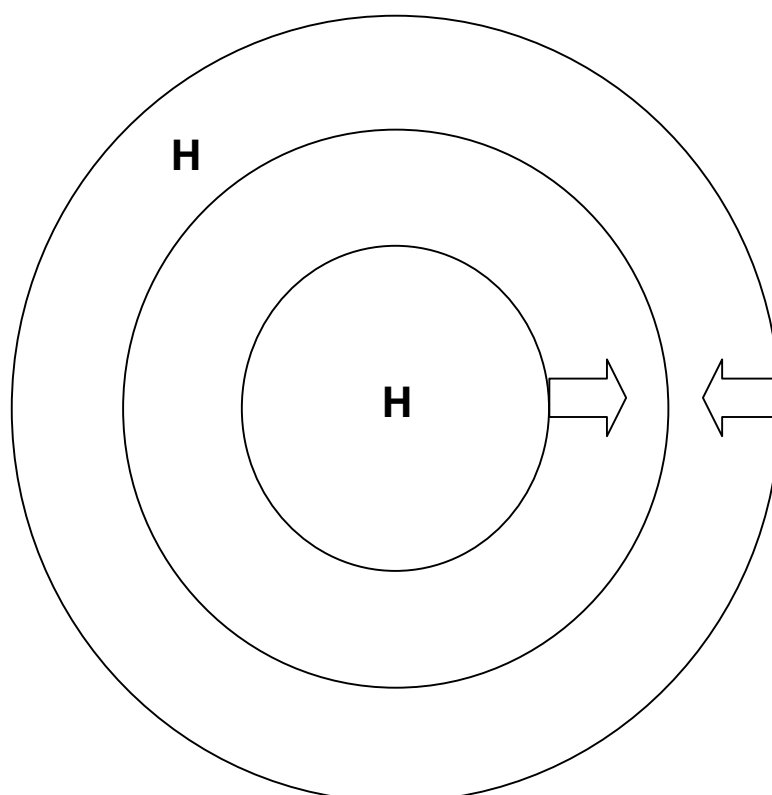
One approach to conceptualizing input and evidence for language development is to consider the learnability circumstances of both the young child and the older L2 learner in relation to what is known as the Subset Problem. In a slight modification of Pinker’s (1990, p. 208) model, Figure 1 tries to represent graphically an important aspect of the acquisition/learning task. A potential virtue of one and the other variant of the model is that they correspond nicely to both L1 and L2 situations. Perhaps simplifying somewhat, we can use the term “hypothesize” to refer to the process of developing a mature or more complete grammatical system, and “learner” to both L1 and L2 subjects. In the inner most of the concentric circles the learner hypothesizes a grammar, incomplete as it always is during the early stages of development. Here the task is to build upon the incomplete grammar, indicated by the expanding arrow. At the same time, the learner forms hypotheses that are “beyond” the target language. In Figure 1, the TL is bounded by the intermediate circle. In other words, the learner hypothesizes a language that is also “too large” (overgeneralizations and other constructions that need to be abandoned somehow), hence the title of the figure: an acquisition problem in which the hypothesized language (H) is, at the same time, a subset and a superset of the TL. Complementary to the task of expanding subset-H, the learner needs to retreat from superset-H, the contracting arrow.

Assuming for a moment the application of general learning strategies for language, expanding the incomplete hypothesis grammar should be more or less straightforward: positive evidence should suffice. The learner expands the interlanguage system through immersion: extensive exposure to correct examples. Some aspects of the grammar might require more and “richer” examples than others. With access to the same general learning strategies, however, retreating from superset-H is not as easy with exposure to positive evidence alone; exceedingly difficult it would be in some domains according to theories based on the Subset Principle. Here, negative evidence would be necessary. What facilitates first language acquisition for young children, who face the problem of a severe shortage of usable negative evidence, is access to a specialized language-specific acquisition mechanism, that complements, boosts, or compensates for the inadequate general learning strategies, again according to some theories (Anderson & Lightfoot, 2002; Maratsos, 1999).

In the case of second language learning, a long-standing debate revolves around the question of whether the above scenario still applies (Ayoun, 1996; Braidı, 1999; Clahsen & Muysken, 1996; Schachter, 1998). If it does not apply, the superfluous negative evidence of first language acquisition is no longer superfluous for learning second languages, because L2 learners no longer have access to the same language acquisition device as children who are acquiring

their L1. In regard to this proposed explanation, there is an interesting potential convergence with opposing theoretical models that, for example, reject the concept of specialized language-specific acquisition mechanisms. If L2 learners must rely on general learning strategies for all

FIGURE 1
The Subset-Superset Problem in language acquisition



domains of the target language, because that's all that they have at their disposal (in L1 as well), positive evidence alone would still be inadequate. And finally, we should keep in mind that the Subset-Superset Problem should not be taken as the only framework to consider in judging the claims of one point of view or another regarding the usefulness of negative evidence in L2 learning.

Figure 2 presents a schema for comparing the different proposals on the purported L1-L2 difference. All theories accept that positive evidence is necessary for both L1 and L2: without a rich immersion in the target language, development is severely impaired in the first case, and cannot advance beyond the most rudimentary knowledge in the second. Cutting across theoretical lines again, most researchers would agree (for different reasons) that, for L1, negative evidence is not forthcoming in sufficient quantity and reliability; and even if it were it is doubtful that young cognitively immature children would be able to exploit it productively. In addition, we might also be able to eliminate, as necessary input factors, other "secondary" type abilities and advanced learning resources such as metalinguistic awareness, direct instruction, and literacy.

FIGURE 2
Positive evidence and negative evidence

**What effect do different "input factors" have on: L1 acquisition, L2 learning?
Which are essential?
Which contribute toward more efficient and effective mastery (e.g. in relation to time & ultimate attainment)?**

<u>L1 acquisition</u>	<u>L2 learning</u>
Positive evidence	Positive evidence
.....	
Negative evidence (in general)	Negative evidence (in general)
Corrective feedback (in particular)	Corrective feedback (in particular)
Metalinguistic awareness	Metalinguistic awareness
Direct instruction	Direct instruction
Literacy	Literacy

But under the second column of Figure 2 (L2 learning) a different picture emerges. Strong naturalistic approaches (Krashen, 1998), which tend to draw a broader equivalence between L1 and L2 acquisition processes, generally favor the view that the same basic acquisition mechanisms are deployed, more or less. Simplifying somewhat, for the development of *linguistic competence*, if positive evidence is sufficient for L1, it should be sufficient for L2. The input factors and learning circumstances below the dotted line may contribute to conscious declarative knowledge, awareness of rules and constructions, and the ability to monitor speech and written expression, but they do not serve as input to implicit language knowledge (competence).

In contrast, approaches to second language learning that (for different reasons) conceive of it as fundamentally, or significantly, different from naturalistic L1 development, will view the “non-primary” explicit learning input factors (below the dotted line in Figure 2) as potentially important. They would contribute significantly to the end state mastery of the second language, not just to a peripheral declarative knowledge, but to competence itself. From this point of view, it is interesting to note that the question of corrective feedback (CF) and negative evidence in L2 learning is part of the more general problem of determining what role “secondary” learning resources play, “secondary” here referring to the broad category of language learning mechanisms that involve deliberate awareness, attention to forms and patterns of language, reflection upon learning processes and outcomes, declarative knowledge and explicit learning of skills (DeKeyser, 1998). These contrast to the “primary” resources that, hypothetically, suffice for the development of core linguistic competence of the mother tongue and the basic language abilities exemplified in face-to-face conversational performance.

CORRECTIVE FEEDBACK IN L2 LEARNING

First of all we should specify what the claim about corrective feedback or negative evidence actually should be in this discussion. It clearly is not enough to investigate whether correction or focus on form “is effective,” “can facilitate,” or “may play a role” in L2 learning, unless the hypothesis to reject is that corrective feedback verifiably *retards* learning, representing an inherent obstacle to learning. Conversely, few researchers would defend the claim that no demonstrable gains in L2 acquisition are possible without explicit learning and correction; that purely naturalistic acquisition is impossible. Rather, the proposal to be tested would have to be specific (and reasonable), something along the lines of: some form of corrective feedback, or systematic negative evidence more generally, results in a measurable difference in overall rate of learning and ultimate attainment. In contrast would be the competing hypothesis, proposing that there will be no significant difference over naturalistic, input-based, instruction of the simple-immersion type (positive evidence only).

This way of framing the problem may help advance the theoretical discussion, but in practical terms, the alternatives that learners face pose new and in some ways more difficult questions. For example, which types of negative evidence should we expect to be most effective (if the general claim turns out to be correct)? On the one hand, self-correction and reflection might steer the learner away from incorrect hypotheses. But at present we know very little about the effect on learning of this kind of negative evidence, or about how teachers might be able to implement it in a systematic way. On the other hand, interlocutor-correction includes: recasts and modeling, request for clarification, repetition of the error, and in the more explicit category, direct correction, metalinguistic explanation, and prompt to elicit the target form. More attention in the research has been devoted to these types of correction.

In regard to the implicit-explicit distinction, studies have attempted to define these categories more precisely, something that has proved to be more difficult than first imagined. For example, even recasts (corrected rephrase of the learner utterance) can be formulated in a didactic manner, and be perceived by the learner as explicit (Ellis & Sheen, 2006). How recasts are interpreted is an issue that remains unresolved, pending the application of new methodological approaches (Carpenter et al., 2006). And then there are the difficult research questions of noticing, how the learner “takes up” instances of negative evidence, and the degree to which they can be processed with the effect of advancing mastery. Again, an important topic for future research is: which aspect of mastery advances, competence or learned meta-knowledge?

A specific question that has arisen in both classroom research and controlled experiments is the relative effectiveness of recasts versus prompts. On the one hand, evidence from learner “uptake” suggests that the more explicit learner-generated repair in response to prompt is more effective than recast, for which the corrective function may often be ambiguous (Lyster, 2004). However, the perception of recasts by learners may depend to a large extent on the degree to which the overall instructional program is oriented toward attending to grammar patterns as opposed to exclusively communicative activities (Sheen, 2004). Other factors, such as learner disposition, aptitude, and L2 proficiency level would be relevant as well.

Now, if we provisionally accept as a working assumption the first hypothesis, that ultimate attainment is favored by some types of corrective feedback, a number of critiques still need to be addressed. At least until findings from comparative studies are more conclusive (Loewen & Erlam, 2006; Shehadeh, 2003), the practical objections to corrective feedback deserve our attention. For example, it has been observed that the feedback provided by language

teachers can be sporadic, infrequent and inconsistent, factors that conceivably might account for conflicting and inconclusive results from one study to the other. While “noisy” and “impoverished” input (of the positive evidence type) probably does not handicap first language acquisition in any significant way, second language learners’ “tolerance” for unreliable feedback, especially of the direct corrective type, may be much lower. This conjecture leads us to a series of speculations about the reliability of corrective feedback in L2 learning and a proposal for future research on this point.

The problem of consistent corrective feedback in fact is shared among all CF categories, implicit and explicit, recasts, prompts and metalinguistic explanation; and it might turn out in the end to be more serious than we expect. In one sense it is less of a theoretical problem than one of implementation: how in the context of the give and take of largely uncontrolled classroom interaction does the teacher deploy corrective feedback routines that are reliable, accurate, and processable by learners? For example, even highly proficient native speakers of the target language routinely fail to discriminate in their CF interventions between errors of morphology and syntax and non-native constructions that involve subtle semantic distinctions and pragmatic infelicities, the latter, despite their subtlety, often presenting themselves as more salient. Complex considerations related to the componential nature of grammar and its interface with other domains of knowledge (Coltheart, 1999; Jackendoff, 2002) are necessary to take into account sometimes in trying to disentangle which linguistic subsystems and non-linguistic knowledge components are contributing to the confusion. Not surprisingly, non-native sounding, grammatically well-formed, learner speech might receive feedback that is either simply incorrect, ambiguous, incomplete or confusing. Again, neither implicit nor explicit corrective feedback is immune to the reliability problem.

CORRECTIVE FEEDBACK IN CALL

Two reports at a recent conference at Tamkang University on L2 writing reminded us that the CF reliability problem is also shared by Computer Assisted Language Learning programs (Chen, 2006; Masumi, 2006). If learners are to receive feedback, it is important that it be consistent, this consistency perhaps being even more important in the CALL environment. With this idea in mind, as a part of our program at Northern Arizona University in second language teaching methods, we have critically examined the corrective feedback features of a widely used CALL software package, the *Rosetta Stone Language Library*, produced by Fairfield Language Technologies. The problems posed by the writing section are illustrative of some of the dilemmas and possibilities in machine corrective feedback.

One might venture to describe the *Rosetta Stone* feedback options as simple and limited, restricted to a highly closed-ended type of writing task, that of dictation. First, the learner listens to an audio version of a short text, accompanied by an image that serves as context support. In a word processing window, the learner transcribes, as in traditional dictation tasks. The corrector consists of a basic spell-checker that matches the learner’s typed version to a single correct model, including punctuation. As illustrated in Figure 3 with hypothetical learner responses, the cursor moves along the text, highlighting each error in succession, prompting self-correction. Following along, the learner analyzes each error in turn and attempts self-correction with the help of a repetition of the audio version. Progressively, texts become longer and more complex, taxing both working memory and grammatical processing ability.

One may be tempted to characterize this feedback procedure as exclusively focused on spelling patterns. While most prompts to self-correct may involve purely orthographic errors,

knowledge of and reflection upon grammatical patterns greatly facilitates self-correction, as the examples in Figure 3 illustrate. That even in the application of the most primitive of text-processing programs imaginable, in this case, learners are prompted to reflect on word and sentence-level grammar, and in the self-correction of punctuation more explicitly so. Be that as it may, admittedly, this feedback feature cannot provide any information to the learner beyond what the exact words of the template provide, much less provide any information to the learner about higher-order discourse patterns.

FIGURE 3
Sample CALL writing activity with corrective feedback

- (1)
Student hears this prompt:
The woman has jumped.
Responds:
The woman as jumped.
The corrective feedback cursor highlights: **as**
- (2)
Student hears this prompt:
The horse is jumping.
Responds:
The horse is jump in.
The corrective feedback cursor highlights: **in**
- (3)
Student hears this prompt:
The child is going to jump.
Responds:
The childs going jump.
The corrective feedback cursor highlights: **childs**, and then **jump**.
- (4)
Student hears this prompt:
They should have studied syntax before the exam.
Responds:
They should of studied syntax before the exam.
The corrective feedback cursor highlights: **of**
-

All of this, however, underscores the problem of feedback quality alluded to earlier. A predicament presents itself in a kind of “trade-off,” analogous to that in language assessment in general between the considerations of validity and reliability, what Davies (1990) referred to in terms of the requirements of “being explicit” and “controlling uncertainty.” In corrective feedback, discrete-point type and closed-ended tasks might be questioned for the restricted domain of language ability for which they provide the learner with practice. The more closed-ended, the further removed they seem to be from meaning, for example (the analogy to validity). Consistency of feedback (the reliability question), however, is generally favored by narrowing the domain of ability. Conversely, the more open-ended the task, the more difficult it becomes to provide learners with frequent and reliable feedback. Interestingly, this relation of “trade-off” seems to apply to both live instructor and machine correction, although it applies, and can be

compensated for, in different ways in each case. So, we might also venture to say that the reliability of feedback in closed-ended prompt and correction type programs, like the one described here, is a feature that CALL in general should value positively, and seek to build upon. In building upon this advantageous feature, the challenge before us would then be to find ways to overcome some the limitations specific to closed-ended tasks.

For us, it is not a question of trying to determine which alternative is better in the abstract. Rather, in designing or selecting CALL activities, the limitations of each option should be taken into account in a deliberate way. Thus, we are in a better position to know what to expect from different learning activities, in particular what kind of feedback is most effective for word and sentence level skills as opposed to the discourse level. To the more integrative abilities might correspond feedback that is more open-ended; not all feedback needs to be corrective.

Generally, as is widely recognized, learning activities of the integrative type, requiring the application of discourse-level abilities tend to be more intrinsically motivating for learners. A factor that educators should keep in mind, it is also one that we can set aside, for now. In fact, the degree to which learners find one or another activity enjoyable or interesting is probably not a good starting point for designing a curriculum and selecting materials. Language learning research also needs to consider what are the most effective and efficient methods for grammar and vocabulary development, for example, taking the factor of motivation as a given. The same applies to learner dispositions toward corrective feedback in particular. Again, language educators cannot be indifferent to how individual students and groups of learners (e.g. who share the same cultural background) perceive teacher correction of different types, for example of varying degrees of directness. But at the same time, it is important to determine the effect on learning of different kinds of negative evidence *in principle*, independently of contextual constraints that vary from one second language learning situation to another. In fact, computer-based corrective feedback is likely to be able to successfully cut across cultural differences regarding the constraints on attending to grammatical form during language use. Speculation aside, and taking care not to throw caution to the wind, we should acknowledge that the pragmatics of human-machine interaction is still not well understood.

The dilemma of “consistency versus open-endedness” brings us to the proposal for research alluded to earlier. If closed-ended, discrete-point CF programs have the potential of providing completely reliable feedback to the learner that is always unambiguous and accurate, this feature presents CALL researchers with an opportunity. Tentatively, or for the time being, we might set aside the issues of “validity,” and “higher-order versus lower-order,” not because they are unimportant, but for purely temporary methodological considerations:

(1) to allow designers to push the limits, so to speak, of the closed-ended type task and determine, in practice, how complex, to what degree higher-order, and how “integrative” closed-ended CALL activities can become, without sacrificing reliability. On-line multiple-choice formats, today, offer learners opportunities that were not available just a few years ago. For example, distracter options can represent degrees of approximation to the correct form, with immediate feedback explanations (and hints, without giving away the answer) on why a response might be partially correct, totally wrong, etc. Without the ability to predict future advances in computing in the short and medium term, this question must be left open. It should be kept in mind that one of the critical features of closed-ended protocols, as in Figure 3, despite their inherent limitation in other respects, is that they are characteristically highly interactive, requiring the learner to be actively engaged, as opposed to observing, as is the case in some open-ended CALL programs.

(2) If reliability of negative evidence, in fact, turns out to be a confounding variable in the CF research, then maintaining it constant should help us direct our attention better to the remaining central issues. The elimination of the CF reliability problem can be thought of as a desirable goal, even if the price to pay for it is high, in part because the problem doesn't apply to all feedback, and for our purposes eliminating all inconsistency and ambiguity is provisional and tentative. At the same time, as a result perhaps, we will be better able to keep an open mind on the still elusive questions, such as the relative merits of recasts and prompts and degree of explicitness in general. And it should go without saying that no part of this proposal should be taken as disfavoring open-ended CALL tasks in general. No counterposition is implied. Rather the research problems associated with closed-ended tasks appear to be somewhat more tractable at this time.

CONCLUSION

In many domains of teaching, CALL will never replace the language teacher, and at the same time the language teacher will never have the opportunity and access to resources for providing learners with the extensive practice that they need, a necessity that can be partially fulfilled by high quality interactive on-line resources.

On the research side, the investigation of the effect of different types of feedback in CALL (corrective and non-corrective) promises to put a new perspective on the continuing discussion on the role of negative evidence in L2 learning. Studies have tentatively suggested, for example, that the electronic environment may even facilitate certain kinds of focus on form reflection among learners. Students have more time to develop and refine their observations, as in the case of requests for clarification and "negotiation moves," and more time to formulate repairs (Morris, 2005). As was mentioned earlier, in the more controlled setting of direct prompts, non-native speakers have the time and the opportunity to reflect and repair (including multiple trial-and-error attempts) that would be difficult to have access to in real-time face-to-face feedback with an actual, live, native-speaker tutor, for example. The learner never has to be concerned with trying the patience of a software program in requesting a repetition, testing a hypothesis once too many times, or persisting in a fossilized error pattern; and the computer will provide as much corrective feedback, for as long as only *one* participant can endure, the one often with more at stake in attaining mastery in the target language.

Exercises that require production can be designed to maximize attention to specific features. If the nature of the task allows for discrete-point feedback that is always reliable, followed by confirming feedback upon supplying the correct response, progressively more difficult grammatical patterns can be presented in a programmed fashion. Some evidence exists that learner uptake is highly favored by online corrective feedback of the elicitation (prompt) type (Heift, 2004). A future challenge for designers will be how to enhance and enrich feedback, as the author of this study points out.

We can now conclude this discussion with a review of the points that led to the proposal for further research. One starting point in even considering the usefulness of corrective feedback was the assumption that second language learners avail themselves, because they have to, of learning strategies that they did not need to, and were not able to, deploy in first language acquisition. Alternatively, if the above assumption can be shown to be unfounded, language learners might still find corrective feedback and other types of negative evidence useful in mastering those aspects of language ability beyond core grammatical knowledge and basic conversational skill (e.g., abilities related to literacy and academic proficiency).

The current research on the effectiveness of corrective feedback in second language learning, focused on implicit and explicit types and other kinds of negative evidence, is of paramount practical and theoretical importance. The question of CF reliability needs to be investigated as well as it addresses an interesting and pertinent critique of corrective feedback by proponents of naturalistic approaches. CALL environments offer a special opportunity to evaluate the effect of controlling this factor. On the practical side, educators need to develop sound criteria for critically evaluating the various on-line programs available to their second language students. These can potentially become the most important source of extra-curricular *interactive* practice, supplementing the comprehensible input that learners are exposed to in the electronic media.

NOTE

1. In a series of recent studies we have examined the tendencies across grade level in bilingual children's ability to deliberately attend to different grammatical and discourse-level features in their writing (Francis & Navarrete Gómez, 2003; Francis, 2005, 2006). In a separate investigation of self-correction in reading (Francis, 2004) the same general research questions came up: (1) how is the development of metalinguistic awareness different from how other kinds of language ability develop, for example the interactive and context-embedded language skills that are acquired spontaneously and uniformly in children's first language? (2) To what extent is the negative evidence that children provide to themselves in literacy tasks involving self-correction a necessary component of school language ability (in this case, academic literacy)? (3) If it is necessary, what kinds of instructional design promote efficient and effective self-correction strategies, as part of the broader development of advanced metalinguistic abilities? Self-correction and reflection could also be considered as an aspect of formative assessment, in this case as an ongoing evaluation on the part of learners of their own learning, providing themselves with useful negative evidence. The assumption here is that without direct instruction, these strategies do not emerge spontaneously and uniformly in all children, not even among all literate children. See Birch (2002) and Grabe & Stoller (2002) for a general review of the problems of second language literacy, and Bitchener et al. (2005), Ferris (2004) and Truscott (1996) for research findings and discussion related to the ongoing debate on corrective feedback in L2 writing in particular.

REFERENCES

- Anderson, S. & Lightfoot, D. (2002). *The language organ: Linguistics as a cognitive physiology*. Cambridge: Cambridge University Press.
- Ayoun, D. (1996). The subset principle in second language acquisition. *Applied Psycholinguistics*, 17, 185-213.
- Birch, B. (2002). *English L2 reading: Getting to the bottom*. Mahwah: Lawrence Erlbaum.
- Bitchener, J. Young, S. & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14, 191-205.
- Braidi, S. (1999). *The acquisition of second language syntax*. London: Arnold.
- Carpenter, H.; Seon Jeon, K.; MacGregor, D. & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, 28, 209-236.

- Chen, H.-J. (2006). Examining the scoring mechanism and feedback quality of MyAccess. Paper presented at the *Tamkang International Conference on Second Language Writing*, Tamkang University, Taipei, December 2, 2006.
- Clahsen, H. & Muysken, P. (1996). How adult second language learning differs from child first language development. *Behavioral and Brain Sciences*, 19, 721-723.
- Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Science*, 3, 115-120.
- Davies, A. (1990). *Principles of language testing*. Oxford: Blackwell.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.
- Ellis, R. & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28, 575-600.
- Ferris, D. R. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime?). *Journal of Second Language Writing*, 13, 49-62.
- Francis, N. (2004). Nonlinear processing as a comprehension strategy: A proposed typology for the study of bilingual children's self-corrections of oral reading miscues. *Language Awareness*, 13, 17-33.
- Francis, N. (2005). Bilingual children's writing: Self-correction and revision of written narratives in Spanish and Nahuatl. *Linguistics and Education*, 16, 74-92.
- Francis, N. (2006). The development of secondary discourse ability and metalinguistic awareness in second language learners. *International Journal of Applied Linguistics*, 16, 37-60.
- Francis, N. & Navarrete Gómez, P. R. (2003). Language interaction in Nahuatl discourse: The influence of Spanish in child and adult narratives. *Language, Culture and Curriculum*, 16, 1-17.
- Grabe, B. & Stoller, F. (2002). *Teaching and researching reading*. New York: Longman.
- Heift, T. (2004). Corrective feedback and learner uptake in CALL. *ReCALL*, 16, 416-431.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Krashen, S. (1998). Comprehensible output? *System*, 26, 175-182.
- Loewen, S. & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, 19, 1-14.
- Lyster, R. (2004). Differential effects of prompts and recasts on form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432.
- Maratsos, M. (1999). Some aspects of innateness and complexity in grammatical acquisition. In M. Barrett (Ed.), *The development of language* (pp. 191-228). East Essex: Psychology Press.
- Masumi, N. (2006). Changes in Japanese EFL writing through a native English reader's feedback. Paper presented at the *Tamkang International Conference on Second Language Writing*, Tamkang University, Taipei, December 1, 2006.
- Morris, F. (2005). Child-to-child interaction and corrective feedback in a computer mediated L2 class. *Language, Learning & Technology*, 9, 29-46.
- Pinker, S. (1990). Language acquisition. In D. Osherson & H. Lasnik (Eds.), *Language: An invitation to cognitive science* (pp. 199-241), Cambridge: MIT Press.

- Schachter, J. (1998). Recent research in language learning studies: Promises and problems. *Language Learning, 48*, 557-589.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research, 8*, 263-300.
- Shehadeh, A. (2003). Learner output, hypothesis testing, and internalizing linguistic knowledge. *System, 31*, 155-171.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*, 327-369.