

Comprehensible Output

Stephen Krashen

System 26: 175-182, 1998

The comprehensible output (CO) hypothesis states that we acquire language when we attempt to transmit a message but fail and have to try again. Eventually, we arrive at the correct form of our utterance, our conversational partner finally understands, and we acquire the new form we have produced.

The originator of the comprehensible output hypothesis, Merrill Swain (Swain, 1985), does not claim that CO is responsible for all or even most of our language competence. Rather, the claim is that "sometimes, under some conditions, output facilitates second language learning in ways that are different from, or enhance, those of input" (Swain and Lapkin, 1995, p. 371). A look at the data, however, shows that even this weak claim is hard to support.

THE SCARCITY ARGUMENT

A problem all output hypotheses have is that output is surprisingly rare (Krashen, 1994). In the case of CO, the problem is especially severe.

A recent confirmation of the scarcity of output is Ellis, Tanaka, and Yamazaki (1994), who examined vocabulary acquisition under three conditions, tasks in which EFL students heard (1) "premodified" input (input recorded from a task performed with a native speaker and non-native speaker who could request clarification), (2) interactionally modified input (the non-native students could interact with the native speaker), or (3) unmodified input (input recorded from a native speaker doing the task with another native speaker). Of interest to us here is the finding that "of the 42 learners in the IM (interactionally modified) group, only seven engaged in meaning negotiation. The others simply listened" (p. 211).

Even when acquirers do talk, they do not often make the kind of adjustments the CO hypothesis claims are useful in acquiring new forms.

Pica (1988) concluded that instances of comprehensible output were "relatively infrequent" (p. 45). In her study of ten one-hour interactions between low level ESL acquirers and native speakers (teachers), only 87 potential instances of comprehensible output were found, that is, interactions in which the native speaker

requested "confirmation, clarification, or repetition of the NNS utterance" (p. 93). These 87 interactions contained only 44 cases in which the non-native speaker modified his or her output (about four per hour), and of these 44, only 13 modifications involved grammatical form, about one per hour.

In Pica, Holliday, Lewis, and Morgenhaller (1989), intermediate ESL acquirers interacted with native speakers. Because the situation in Pica (1988) did not promote negotiation (an interview), some of the conversations in this study were in situations designed to require negotiation and comprehensible output. Of 1952 native speaker utterances, 327 were "signals indicating clarification or confirmation of what the NNS had said" (p. 74). In reaction to these 327 utterances, the non-native speakers produced 116 responses containing "modified output." In other words, they produced comprehensible output in response to about 6% of the native speaker's utterances (116/1952).¹

Interactions were also contrived to promote negotiation in Van den Branden (1997). Eleven and 12 year old speakers and acquirers of Dutch interacted with peers or with a teacher in an activity in which speakers had to describe a drawing to a partner. In peer-peer dyads, 51 instances of negotiation of meaning were recorded, and of these, the speaker modified their output 20 times. In peer-teacher dyads, there were 49 instances of negotiation of meaning and 20 instances of alteration of output. In both cases, this amounted to about one every five minutes. We do not know if the alteration improved the grammatical accuracy of the output; we are only told that "these modifications often involved, or included, formal modifications ..." (p. 616). Even if every case resulted in improvement, however, this data confirms that even in contrived situations, comprehensible output is infrequent.

Lyster and Ranta (1997) recorded 18.3 hours of French immersion language arts and subject matter lessons involving fourth and fifth graders. The lessons contained a total of 3268 student turns, of which about 1/3 (1104, or 34%) had at least one error. While teachers provided some kind of feedback to 62% of these errors, only 73 were in the form of a clarification request, "a feedback type that can refer to problems in either comprehensibility or accuracy, or both" (p. 47). Of these 73, 20 were followed by student repair, or correction. This amounts to about one per hour, a result very close to that reported by Pica (1988) for conversations. In this case, however, only the one student producing the repair had the benefit of comprehensible output.

The situation in writing is similar. Cumming (1990) examined the think-aloud protocols of second language writers, hypothesizing that instances in which writers appeared to be attending to both form and meaning at the same time are potential instances for language acquisition, according to the comprehensible output hypothesis. Only 30% of the verbal reports made by the writers in his sample were

of this kind (p. 490). In addition, the nature of the episodes makes it unlikely that they play a major role in language acquisition: Most of the episodes were writers' searching for the right word, or searching for first language equivalents. The latter is the familiar strategy of falling back on the first language when competence is lacking in the second language (Newmark, 1966).

In Swain and Lapkin (1995), grade 8 early French immersion subjects were asked to write a short essay in French (one to two paragraphs) and then edit it, and to "think aloud" as they made decisions. For the draft and editing stage combined, there were "190 occasions in which students consciously recognized a linguistic problem as a result of producing, or trying to produce, the target language" (p. 384). This amounts to an average of 10.6 per student. If students wrote a short essay everyday (which they do not), this would mean about ten chances to improve through writing per day - not very much. As was the case with Cummings' study, many of the decisions were lexical (looking for the right word), not grammatical (50% in the first draft). In addition, Swain and Lapkin note that there was no evidence that any of the episodes they described led to improvement.

ACQUISITION WITHOUT OUTPUT

There are numerous studies that confirm that we can develop extremely high levels of language and literacy competence without any language production at all (Krashen, 1994). Laboratory studies show that subjects typically acquire small but significant amounts of new vocabulary knowledge from a single exposure to an unfamiliar word in a comprehensible text (Nagy, Herman, and Anderson, 1985), enough to account for expected vocabulary growth, and similar results have been reported for second language development (Pitts, White, and Krashen, 1989; Day, Omura, and Hiramatsu, 1991; Dupuy and Krashen, 1993). It has been argued that a similar effect exists for spelling (Krashen, 1989). In addition, there are case histories of those who have developed very high levels of competence from input alone (Richard Boydell suffered from cerebral palsy and acquired language through listening and reading alone, see Krashen, 1985; Malcolm X and Richard Wright credit their literacy development to wide reading, discussed in Krashen, 1993).

Ellis (1995) is an additional analysis of Ellis et. al. (1994), discussed above, and provides another instance of acquisition without output. The "premodified" group, a group that did no speaking at all, made modest but clear gains in vocabulary, gaining, in fact, more words per minute than the group that interacted with the native speaker.²

DOES CO LEAD TO LANGUAGE ACQUISITION?

Nobuyoshi and Ellis (1993) claim to have provided data showing that comprehensible output results in actual improvement. In their study, six adult EFL students in Japan of "fairly low-level proficiency" but who were "capable of using at least some past tense verb forms correctly" (p. 206) were asked to participate in a jigsaw task with their teacher in which they described actions in pictures that, they were told, occurred the previous weekend or previous day. During the first session of the study, the three experimental subjects received requests for clarification if the verb was not in the past tense or if the past tense was incorrectly formed. During the second session, one week later, they received only general requests for clarification (when the teacher did not understand). The three comparison subjects received only general requests for clarification each time.

Nobuyoshi and Ellis report that comparison subjects did not improve their past tense accuracy. Two experimental subjects (E1 and E2) were able to improve their performance in response to requests for clarification at the first session, but the third experimental subject (E3) did not. Nobuyoshi and Ellis claim that E1 and E2 sustained their gains to time 2, with E1 increasing accuracy from an original level of 31% to 89% and E2 increasing from an original 45% to 62%. Nobuyoshi and Ellis conclude that their study "provides some support for the claim that 'pushing' learners to improve the accuracy of their production results not only in immediate improved performance but also in gains in accuracy over time" (p. 208).

As Nobuyoshi and Ellis point out, however, their conclusions are based on a very small sample size. In addition, they are based on a very low number of obligatory occasions. E1, who showed the clearest gains, went from 4 correct out of 13 at time 1 to 8 correct out of 9 at time 2. E2 went from 9 correct out of 20 at time 1 to 16 correct out of 26 at time 2. E1's gains are statistically significant (Fisher Exact Test, 2 tail, $p = .0115$) but E2's gains are not ($\chi^2 = .69$). Thus, for one subjects there was no evidence of the value of comprehensible output (E3), and for another, gains were not statistically significant. Data supporting a central hypothesis should be made of sterner stuff.

Note also that all three subjects had studied the past tense rule, and had been clearly focused on it in session 1. It is reasonable to expect that when subjects are focused on form, then put back in the same environment, they will be focused on form again, especially if the conversational partner is their teacher. The significant effect on E1, in other words, may have been a performance effect - E1 was simply more inclined to try to use a consciously learned rule for the past tense and was a more

successful Monitor user than E2 or E3.

Van den Brandon's subjects (Van den Brandon, 1997, discussed earlier) who participated in sessions that encouraged negotiation of meaning increased their output relative to a control group that did not engage in interaction, but were not superior in grammatical accuracy. Each subject, however, only had seven to nine minutes of interaction.

Tarone and Liu (1995) suggest that CO may have played a role in the second language development of Liu's subject "Bob." Bob was recorded interacting with peers, with teachers, and with an "adult-friend" (Liu). Tarone and Liu note that language use was much more complex in the latter interactions, and, in general, "new structures appear first in interactions between Bob and the researcher, spread to the interactions with his peers, and appear last in his interaction with his teacher" (p. 119). They note that it is likely that Liu provided Bob with more complex input, but also suggest that Bob's attempts to produce CO in interacting with Liu played a role. While interacting with Liu, Bob used English in a much wider range of speech acts than in the other situations, and this may have pushed Bob to "go beyond the limits of his interlanguage competence" in production" (p. 121). Tarone and Liu show that the CO hypothesis, as well as the Input Hypothesis, is consistent with what is known about Bob's development. As they note (p. 123), data is lacking on the frequency of CO, which prevents us from determining whether CO resulted in language development and whether Bob produced significant quantities of CO.

The DISCOMFORT OF CO

The CO hypothesis predicts that we acquire language when there is a communicative breakdown and we are "pushed to use alternative means to get across .. the message ... precisely, coherently, and appropriately" (Swain, 1985, pp. 248-249). In addition to the research that shows that CO is an unlikely candidate, there is additional evidence that "pushing" students to speak is unpleasant for them. When asked what aspects of foreign language classes are the most anxiety-provoking, students put "talking" at the top of the list (Young, 1990). Laughrin-Sacco (1992) reported that for students in beginning French classes, "for nearly every student ... speaking was the highest anxiety-causing activity" (p. 314).

Ten "anxious" foreign language students interviewed by Price (1991) stated that their greatest source of anxiety "was having to speak the target language in front of their peers" (p. 313). Of great interest here is the finding that another source of stress "was the frustration of not being able to communicate effectively" (p. 105).

These results suggest that it is "pushed output," having to utilize structures they have not yet acquired, under demanding conditions, that students find uncomfortable. Methods based on comprehensible output put students in this situation constantly.

CO AND THE INTERACTION HYPOTHESIS

The CO hypothesis is linked to what is sometimes called the "interaction hypothesis," the hypothesis that we acquire language from interacting with others. As stated in this way, the interaction hypothesis is vague - Is interaction necessary or just helpful? Is it the only way to acquire language or one way to acquire language? Also, what occurs during interaction that causes language acquisition?

I have argued that a part of interaction that does not contribute to language acquisition is the output produced by the language acquirer. In addition, there is evidence that a strong version of the interaction hypothesis, one that asserts that interaction is necessary for language acquisition, is not correct. Such a hypothesis denies that acquisition can occur from reading and listening. In addition to the massive data showing that reading can promote language development, the results of Ellis et. al. (1994), discussed above, confirm that acquisition is possible without actually participating in the interaction. A weaker version of the interaction hypothesis is that interaction can be a good source of comprehensible input (Krashen, 1982).

THE NEED HYPOTHESIS

The CO hypothesis is closely related to the "need hypothesis." I have never seen the need hypothesis discussed explicitly in print, but it is widely assumed to be true. The need hypothesis says that we acquire language only when we "need" to communicate, when we need to make ourselves understood. If this hypothesis is correct, language acquirers must be forced to speak the second language. The need hypothesis thus implies that "submersion" is a good thing, in that it forces students to try to communicate.

The need hypothesis is not correct. An excellent counter-argument was presented by Garrison Kieler on the Prairie Home Companion, in a segment called "The

Minnesota Language School." The Minnesota Language School operates on the assumption that we acquire language when we need to use it. Their method is to take someone who speaks no German at all, fly them up in a helicopter, and threaten to push them out if they don't start speaking German immediately. If the need hypothesis were correct, this would work.

According to the input hypothesis, need can be helpful when it puts the acquirer in a position to get comprehensible input. All the need in the world, however, will not result in language acquisition if there is no comprehensible input. In addition, interesting and comprehensible input will result in language acquisition whether need is present or not.

SUMMARY AND CONCLUSION

The comprehensible output hypothesis has numerous difficulties.

- Output and especially comprehensible output is too scarce to make a real contribution to linguistic competence.
- High levels of linguistic competence are possible without output.
- There is no direct evidence that comprehensible output leads to language acquisition.

In addition, there is some evidence that suggests that students do not enjoy being "pushed" to speak.

The original impetus for the CO hypothesis was the observation that students in French immersion, despite years of input, were not as good as observers felt they should be in grammatical aspects of their second language (Swain, 1985). Input, it was suggested, was therefore not enough. It can be argued, however, that we haven't yet given comprehensible input a real chance. We have yet to see how students will do if their classes are filled with comprehensible input, if they have access to a great deal of very interesting reading and listening materials (films, tapes), and if the acquisition situation is genuinely free of anxiety. (There is evidence that children in French immersion do very little pleasure reading in their second language; Romney, Romney and Menzies, 1995).

Given the consistent evidence for comprehensible input (Krashen, 1994) and failure of other means of developing language competence, providing more comprehensible input seems to be a more reasonable strategy than increasing output.

NOTES

1. Swain (1995) notes that in this study "in response to clarification and confirmation requests, over one-third of the learners' utterances were modified either semantically or morphosyntactically" (p. 131). This is correct. My concern here, however, is how frequent CO is in general. CO in response to requests for clarification was frequent in this study, but not overall.
2. As noted earlier, only seven of the 42 subjects in the interaction group actually spoke: Ellis et. al. found, however, that these seven "did not enjoy a clear advantage in either comprehension or vocabulary over those who just listened" (p. 212).

REFERENCES

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

Day, R., Omura, C., and Hiramatsu, M. (1991) Incidental vocabulary learning and reading. *Reading in a Foreign Language* 7, 541-551.

Dupuy, B. and Krashen, S. (1993) Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning* 4, 55-63.

Ellis, R. (1995) Modified oral input and the acquisition of word meanings. *Applied Linguistics* 16, 409-441.

Ellis, R., Tanaka, Y. and Yamazaki, A. (1994) Classroom interaction, comprehension, and L2 vocabulary acquisition. *Language Learning* 44, 449-91.

Krashen, S. (1982) *Principles and Practice in Second Language Acquisition*. Englewood Cliffs, NJ: Prentice Hall.

Krashen, S. (1985) *The Input Hypothesis: Issues and Implications*. Beverly Hills, CA: Laredo Publishing Company.

Krashen, S. (1989) We acquire vocabulary and spelling by reading: Additional

evidence for the input hypothesis. *Modern Language Journal* **73**, 440-464.

Krashen, S. (1993) *The Power of Reading*. Englewood, CO: Libraries Unlimited.

Krashen, S. (1994) The input hypothesis and its rivals. In Ellis, N. (Ed.) *Implicit and Explicit Learning of Languages*, pp. 45-77. London: Academic Press.

Loughrin-Sacco, S. (1992) More than meets the eye: An ethnography of an elementary French class. *Canadian Modern Language Review* **49**, 80-101.

Lyster, R. and Ranta, L. (1997) Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition* **19**, 37-66.

Nagy, W., Herman, P., and Anderson, R. (1985) Learning words from context. *Reading Research Quarterly* **20**, 233-253.

Newmark, L. (1966) How not to interfere with language learning. *International Journal of American Linguistics* **40**, 77-83.

Nobuyoshi, J. and Ellis, R. (1993) Focused communication tasks and second language acquisition. *ELT Journal* **47**, 203-210.

Pica, T. (1988) Interactive adjustments as an outcome of NS-NNS negotiated interaction. *Language Learning* **38**, 45-73.

Pica, T., Holliday, L., Lewis, N. and Morgenthaler, L. (1989) Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition* **11**, 63-90.

Pitts, M., White, H., and Krashen, S. (1989) Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language* **5**, 271-275.

Price, M. (1991) The subjective experience of foreign language anxiety: Interviews with highly anxious students. In Horwitz, E. and Young, D. (Eds.), *Language Anxiety*, pp. 101-108. Englewood Cliffs, NJ: Prentice Hall.

Romney, J.C., Romney, D., and Menzies, H. (1995) Reading for pleasure in

French: A study of the reading habits and interests of French immersion children. *Canadian Modern Language Review* **51**, 474-511.

Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass, S. and Madden, C. (Eds.), *Input in Second Language Acquisition*, pp. 235-256. New York: Newbury House.

Swain, M. (1995) Three functions of output in second language learning. In Cook, G. and Seidelhofer, B. (Eds.) *Principle and Practice in Applied Linguistics: Studies in Honor of H.G. Widdowson*, pp. 125-144. Oxford: Oxford University Press.

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

Tarone, E. and Liu, G-Q. (1995) Situational context, variation, and second language acquisition theory. In Cook, G. and B. Seidelhofer, B. (Eds.) *Principle and Practice in Applied Linguistics: Studies in Honor of H.G. Widdowson*, pp. 107-124. Oxford: Oxford University Press.

Van den Branden, K. (1997) Effects of negotiation on language learners' output. *Language Learning* **47**: 589-636.

Young, D. (1990) An investigation of students' perspectives on anxiety and speaking. *Foreign Language Annals* **23**, 539-553.