This article examines the phenomenon of fossilization in second language (SL) learning and instruction, discussing this process as a form of simplification. Fossilization occurs when particular linguistic forms become permanently established in the interlanguage of SL learners in a form that is deviant from the target language norm and that continues to appear in performance regardless of further exposure to the target language. The article examines specific research on linguistic fossilization, looking at the nature, objects, and manner of fossilization. It also considers the point at which fossilization begins, its persistence, and the types of learners likely to be affected by the phenomenon the most. (MDM)

* Reproductions supplied by EDRS are the best that can be made from the original document. *
FOSSILIZATION AS SIMPLIFICATION?

Larry Selinker
FOSSILIZATION AS SIMPLIFICATION?

Larry Selinker

I was pleased to accept this invitation from Professor Tickoo to try to make some sense of the vast fossilization literature from the point of view of simplification, for the task is long overdue. In truth, this task has proven difficult since, on the one hand, little appears that summarizes the fossilization literature from any point of view, and, on the hand, the fossilization literature is widespread and diffuse. How is one to select from this literature for typicality? What criteria should one use? Most references to fossilization one sees, and there are literally hundreds, are of the type:

"This structure, I (we) conclude, is a good candidate for fossilization, because .....".

That is, one rarely sees an attempt to link the particular conclusion presented with other potential fossilization events which have been discussed in the literature. As Kellerman (1989) points out, even where there are attempts at explanation, "they do not lead to predictions about what linguistic features of the interlanguage are candidates for fossilization". He reminds us of the dire straits of things in this area when he concludes that:

There has been virtually no discussion as to why certain 'accents' may come to typify a whole community of language learners irrespective of differing proficiency levels within that community.

(Kellerman 1989, 88)

It turns out, then, that there are very few general principles which have been proposed to cover fossilization. This essay is an attempt to push the notion fossilization in terms of one such concept: simplification - which was first brought up in Selinker (1972) - knowing full well that any one unitary explanation will fail, but maybe we can clear the collective air a bit.

In the invitation to the volume, Prof Tickoo points out that only one type of simplification is linguistic, and that there are others: pedagogic, psycholinguistic, and perhaps even others. I am sure that what I see most often described in the literature can be termed "linguistic simplification". I see no way to unambiguously define this concept, but what I mean can be gleaned, perhaps, from
an example by Schachter (1988, 1990). She believes that one of the key issues for second language acquisition (SLA) theory is that of "completeness", that non-native grammars, no matters how target like, will be incomplete in interesting ways.

For example, Schachter concludes that cleft structures in English, even when known by non-native speakers (NNSs), are used less frequently by them than by native speakers (NSs), which matches my perception. That is, learners and other interlanguage (IL) speakers overuse structure (1a), even when the discourse calls for structures of type (1b):

1a. I painted the house yesterday.
1b. What I did was paint the house yesterday.
   It was the house that I painted yesterday.

Or, in more academic language, consider (1c) vs. (1d) where it is my experience that concerning various types of cleft sentences, NNSs rarely have a clue.

1c. Corder consistently emphasized that IL is "normally unstable" (e.g. 1981, 16). Now, Klein (1984), in my view, provides the most important conceptual link between fossilization and simplification.

1d. What Corder did consistently was to emphasize that IL is "normally unstable" (e.g. 1981, 16). Now it is Klein (1984) who, in my view, provides the most important conceptual link between fossilization and simplification.

[Sentences in (1d) actually appear in the paper below and the reader might wish to compare the illocutionary force of (1d) to (1c) in each case.]

Schachter’s explanation for the overuse of structures in (1a) and (1c) is interesting here: structures of type (1a) and (1c) are more frequent among NNSs because they are of the more simple canonical word order: subject-verb-object (SVO) than are those of (1b) and (1d). That is structures of type (1b) and (1d), by any measure, are linguistically more complex.

In the document referred to above, Tickoo goes on to point out that in linguistic simplification one is interested in:

... events or developments in which the primary focus of attention has been language - its systems, structures or discourse.
Now, I would contend that, with the help of the dictionary, the concept fossilization is clear (though of course which IL samples are exemplars is up for grabs): the unabridged *Random House Dictionary of the English Language* (1987) has kindly defined fossilization for us under "fossilize":

Ling. (of a linguistic form, feature, rule, etc.) to become permanently established in the interlanguage of a second-language learner in a form that is deviant from the target-language norm and that continues to appear in performance regardless of further exposure to the target language. (p. 755)

There are several important points to be made here: first, it is clear that learner-created permanent IL plateaus, often far from the target language, is the norm in SLA. Second, it appears to be the case, that fossilized ILs exist *no matter what* learners do in terms of further exposure to the TL. Third, given the latter, at any point in time it is nonetheless very difficult, if not impossible, to tell, at a particular point in time, if a learner’s stabilized IL is in fact fossilized. Thus, it is common in SLA discussion to distinguish theoretically "permanent fossilization" from "temporary stabilization" of the IL. Fourth, it is generally agreed (cf. Larsen-Freeman & Long, 1991) that the most reasonable empirical way of studying fossilization is to look longitudinally for what remains in IL speech (or writing) over time. Finally, a solid theoretical explanation of such permanent plateaus is at present lacking. Now, to simplification.

Corder (1981) produced a truism that one cannot simplify what one does not know and argued that SLA is primarily an example of linguistic complexification, that ‘simple codes’ do not necessarily mean simplification has occurred. This is insightful, but as I have argued elsewhere (Selinker, 1984) the argument is too general. I would like to claim that learners, sometimes, and maybe quite consciously, simplify target language (TL) information and that could perhaps lead to fossilization. There are many examples in the literature: as, for example, the French immersion learners who use one form of the verb for the whole paradigm. This also happens in clearly fossilized West African French adult IL so that it may be a general strategy.

It may also be the case that observed "careful" learners "...often make use of only those aspects or features of morphology or syntax which (they are) sure of at the time of writing" (Tickoo, op cit). If so, this may also turn out to be a type of simplification, and that it extends to oral production in some domains, perhaps formal ones. Now it is Klein (1984) who, in my view, provides the most important conceptual link between fossilization and simplification. Klein takes an ecological approach and makes a good case for fossilization as beneficial to learners if the
"freezing" does not take place too early, because (a) fossilized systems are often simple, and (b) therefore, they are more easily learnable systems (interestingly moving the argument from linguistic simplification to psycholinguistic simplification), and (c) one can do many things with these limited means (for further discussion, see Selinker, 1992, 252 and passim.)

In order to study fossilization with this perspective in mind, we must have a vision of various research approaches that could be taken in principle concerning topics such as the objects of possible fossilization, the onset and persistence of stabilized forms, and age variables, all of which are worthy of consideration. Longitudinally-gathered data from second language learners should be carefully examined to see which aspects become stabilized over time. Related issues such as nativization and contextualization of new non-native varieties need to be discussed. Stabilized local varieties, e.g. Navajo-English or Filipino-English, must also be considered. We need an open mind about fossilization, especially the strong claims about "inevitability" and genetic matters.

Selinker & Lamendella (1978, Table 2) provided a sketch of fundamental "Research Problem Areas" which it may be useful to review here. The first area covered is the NATURE of fossilization. By this, it was meant to investigate whether fossilization is a phenomenon peculiar to SLA or a more general cognitive condition, relevant to other types of learning. One unresolved problem in this dimension is whether fossilization is a positive process of halting further IL development in the ecological way suggested above or is the absence or loss of some abilities, which would surely involve cognitive simplification. The second research area involves the SOURCE of fossilization. Is the basic explanatory domain in terms of which fossilization can be described, that of factors external to the individual learners or factors internal to the individual? Another possibility which we should now have tools to discuss, is that of external factors which may be filtered through the current information processing systems of the individual.

A third research area involves the OBJECTS of fossilization. Which aspects of a learner's IL are susceptible to fossilization? Are they single surface items? Particular rules? Subsystems? Might the linguistic objects involve new combinations of grammatical items that are not usually linked in the comparative NL theoretical linguistics that we have now? That is, would we come up with novel linguistic units that would be fossilized in a new comparative IL theoretical linguistics? Is it the case that some linguistic features are more susceptible to premature stabilization than others? In particular, are phonological units especially liable to fossilize? There is an issue of correctness here for it may be reasonable, at times, to view features which are "incorrect" relative to the TL as more susceptible
to fossilization, thereby invoking an early Corder (1967) idea of "errors as a learning strategy". In terms of interaction, can units of communicative competence fossilize independently of the linguistic form of the IL? It seems likely that linguistic forms can fossilize independently of communicative competence and here we may have made some serious progress; see, for example, Selinker and Lakshmanan (In Press) where some suggested fossilizable structures from the literature are presented: e.g. empty categories and tenseless clauses.

A fourth area of potential research involves the MANNER of fossilization and here we can ask if there are particular sequences in which given linguistic features fossilize and if any of these sequences are universal vs. language specific vs. IL specific vs. learner specific? Or, if, in FLL in classrooms, fossilization would occur differentially from SLA in naturalistic settings? Also, in this area of research, we would want to know if fossilization is an abrupt event or a gradual process occurring over a span of weeks, months, or years. We would also want to know here if it is indeed possible (as Selinker and Douglas (1985, 1989) have claimed) that fossilization can occur by discourse domain, where in one domain a structure is fossilized whereas in another, it is still developing.

A fifth area is that of the POINT at which fossilization begins. When, along the learning process, will fossilization "set in" for a given aspect of the learner's IL? Is there any absolute "lower bound" on which fossilization could possibly occur? Is there an absolute "upper bound" by which fossilization necessarily occurs, or does the learner's IL continue to be indefinitely permeable?

A sixth area involves the PERSISTENCE of stabilization. Can it be determined for a given learner whether IL stabilization is merely a temporary plateau or a permanent condition? Is it possible for a person to "de-fossilize" at some point and, if so, under what conditions, internal/external to the learner? Here one would want to ask, if the general conditions of the learner change drastically, does it matter how long the learner had remained fossilized. Would age matter here?

A final area of research involves CANDIDATES for fossilization: which types of learners may be identified in advance as likely candidates for premature fossilization at some great distance from TL norms? Specifically, why do some adults fossilize at a greater distance from TL norms than do others? In terms of child L2 acquisition, is it reasonable to conclude that some children fossilize while others do not?
Now, in order to adequately discuss these possibilities in a large literature, one would need more than the space provided in this essay. The next best thing would be to "walk the reader through" a very few representative works on the topic, hoping that the list is indeed representative. Here we will make a start.

First, an area where several of the above overlap, involves the important debate on TERMINAL 2's AND COMMUNICATIVE VS. GRAMMATICAL TEACHING. Higgs & Clifford (1982), coining the term based on the 1 to 5 scale of the Foreign Service Institute, claimed that grammatical accuracy must be stressed based on their observations of FSI students who had first learned their foreign language "in the streets" and were apparently stuck at a low level of grammatical achievement. In our terms used here, the terminal 2's seemed satisfied with a simpler grammar than that of the TL. There has been a lot of argument about the facts here and its relation to SLA vs. FLL. Terrell (1989) wrote a response since he said that, everywhere he went, people brought up the terminal 2 conclusion of Higgs and Clifford as an argument against the natural approach and the emphasis on comprehensible input. Terrell pointed out the difference in data types between FSI students and those in an undergraduate American University and, importantly, that one could fossilize at the 2, 3, or even 4 levels. Terrell saw the role of grammar as part of complexification, where attention of the learner was focussed on particular elements in the input, and reinforced by the textbook, with the claim being that, with the natural approach, such elements would be acquired more readily.

A second area, which we will too briefly consider, is that of TORONTO FRENCH IMMERSION. Data from immersion programs are interesting because what they involve are children who appear to create ILs under certain sociolinguistic conditions. From the large literature on this subject, Harley & Swain (1984) is a particularly useful paper for in it, they cite detailed empirical work. For example, concerning possible objects of fossilization, they produce an important simplification example of the inappropriate equating by learners of English personal pronoun ('I') with French personal pronoun PLUS auxiliary ('j'ai'). They relate this result to multiple effects, citing language transfer and salience in the input as two effects working in tandem to produce interlingual identifications. They accept the possible reality of fossilization, but interestingly cite a lack of "positive evidence" to date.

The next area involves FOSSILIZATION AND LEARNING STRATEGIES. Sims (1989) is a Ph. D. exam paper and incomplete though this may be as an argument, there is an interesting positing here of a continuum of fossilization, from "soft" to "hard" and that simplifying learning strategies may play a crucial role, in establishing where on this continuum, a learner may end up. Sims
is the first, to my knowledge, to write of this potentially important connection, and, as far as I know, it has not been taken up in the literature.

Mukkatash (1986) considers the possibility that EVEN WITH "SYSTEMATIC ERROR CORRECTION AND EXPLICIT GRAMMATICAL EXPLANATION", fossilization persists. In the context of a teaching situation in Jordan, he looks at Arabic-English IL from an error analysis point of view and considers several possible objects of fossilization. One clear example of simplification in his data is what he calls "BE-deletion" in the context where BE is taught and where errors are systematically corrected. If the rules are known, this must be a case of simplification where "deletion" from what is known is involved. He considers such IL sentences as:

2a. CA predicts errors that X only right or materialize.
2b. The government X trying to make easy life for everyone.

(X is here "used to indicate the position of the deleted elements")

Note these are two different linguistic contexts and it is claimed that we have a case here of contextual conditioning. For the source of the deletion, these sentences are related to "facts of Arabic" where such deletion is the norm. This fits into the MEP perspective mentioned above, where with language transfer being one effect, this will help stabilize the IL and this is discussed in Selinker and Lakshmanan (In Press).

In a different teaching situation, composition teaching in New York City, Yorio (1985) claimed that in the written ILs dealt with, fossilization "has become the most pressing issue that I have to face" (emphasis in the original). He explores the nature of fossilization, looking at composition data of urban U.S. learners to try to gain insights for teaching and comes up with a memorable phrase: "a student should be considered stabilized until proven fossilized". He presents data where a learner's "control of grammar fails him" and, interestingly, much of the data as in the previous example, seems to involve simplified deleted forms, such as:

3a. for short time
    I used make a year book

in Korean-English. This case is more complex because it is clearly variable:

3b. in a row
    I used to have a dog.

20 9
Yorio reviews the fossilization literature and finds it lacking in useful principles, which is one of the main points of this essay.

Now we briefly consider EUROPEAN ENGLISH, which is being more and more described as a dialect(s) in its own right. Could it be that, as large parts of Europe become more and more like one country, that English is taking on the sort of status that it has taken for India or Nigeria, removing the political need to choose which of the local languages should be the "national" language for specific, but important purposes? If this is indeed the case, then we can expect, just as we find Indianized features, we will see Europeanized features in the English(es) involved. From the point of view of nativization (see below), Berns (1988) provides a sociolinguistic description of the status of "a German variety of English", where English may be "more akin to second language for many Germans who use English every day" in interpersonal uses with other Europeans and/or American military forces.

From the point of view of SLA, Sharwood Smith (1989) looks at persistence in Dutch-English of certain verb complementation structures in spite of a large amount of exposure to English. He considers things that regularly "go wrong" with advanced learners of English in the verb complementation area with, for example, the overgeneralization of for/to infinitives. Kellerman (1989) looks at fossilization, also by Dutch speakers of English, of the imperfect conditional despite a "high level of linguistic achievement". Kellerman studies the use of "would" by Dutch speakers, as in:

4. If it would rain, they would cancel the concert in Damrosch Park.

His studies are interesting in this regard, as they clearly involve the MEP. We have reanalyzed his results elsewhere (Selinker and Lakshmanan, In Press) and see his results as occurring because of three effects: NL avoidance, symmetry of structure and what we call "affect", for lack of a better word, i.e. the NNS trying to make the TL "better" or "more precise", an effect we see occurring with advanced IL speakers in most if not all IL situations.

In general, it is important to consider clearly established NON-NATIVE VARIETIES (NNV's) AND THE NATIVIZATION/ FOSSILIZATION DEBATE. In numerous publications, Kachru has taken a strong anti-fossilization position. For example, in a (1988) statement, he produces: "Fallacy IV: The international non-native varieties of English are essentially 'interlanguages' striving to achieve 'native-like' character". It is interesting to quote the relevant paragraph in full:

21
This hypothesis has several limitations as has been shown by Sridhar and Sridhar (1986) and Nelson (1988). Whatever the validity of this hypothesis in second language acquisition in general, its application to the institutionalized varieties of English in the Outer Circle (i.e. of his linguistic-cultural continuum) needs reevaluation. (Kachru, 1988, 4)

We need to be careful here and ask ourselves, first, why colleagues at times appear emotional about this topic. Empirically, are we dealing with different phenomena in the above quote or different labels for the same phenomena? Or, in some way, can both be true at times? We surely want to avoid the pejorative implication that using SLA concepts necessarily implies "deficient versions of some NS (Native Speaker) standard" (as Williams puts it), though we must not avoid the theoretically important "completeness" issue (See below).

The issue of the relationship of these varieties to IL is directly faced by several authors. Davies (1989), for example, uniquely makes a connection between "two major developments in applied linguistics and language-teaching studies in recent years", the institutionalized varieties and interlanguage. He sees both as "necessary approaches" with the former often "simplified in some way" and the latter an account of individual variation. Williams (1987), uses the term "non-native institutionalized varieties of English" (NIVEs) and looks carefully at Singapore-English data and concludes that NIVEs are "an important and growing acquisitional phenomenon", thus linking two important ideas: language use and acquisition, which appear to me to be lacking in the Kachru approach above. Usefully Williams notes that there actually are different uses of the term "nativization", which leads to confusion of the underlying issues involved and interested readers are urged to work these distinctions out for themselves carefully.

Zuengler (1989), using (NNV's), is perhaps the strongest detailing of the issues, looking at data concerning the important socio-psychological concept of "identity". She carefully considers the important concept of the "overall target" of learners and concludes that "many IL speakers do, indeed, have target model alternatives". That is, Zuengler usefully concludes that the learner's final outcome (as fossilized?):

...is not merely a reflection of the input he is most frequently exposed to, but is the result of a selection of model by the learner. (Zuengler, 1989, 82; emphasis in the original)
As I read Zuengler, personal identity relates to final IL outputs as simplified, but again, the interested reader is urged to go to the original sources to work out the complicated suggestions. Important to the above debate, she argues against the assertion made by Kachru and others that "language learning in IL settings and non-native variety settings is fundamentally different", going carefully through the different arguments and rejecting each one.

Not usually involved in this debate are questions of: does FOSSILIZATION occur IN AN "EMERGING ETHNIC SOCIAL DIALECT"? AND WITH PREADOLESCENTS? MacDonald (1938) answers empirically "yes" to both questions in her study of the English of second generation Cubans in Little Havana in Miami.

Perhaps this debate should be related more carefully to the important issue of EVIDENCE FOR SYNTACTIC FOSSILIZATION. Besides the bits and pieces found throughout the SLA literature, we have with Apte's (1988) work an attempt to "trace and describe" grammatical fossilization, one of the first such detailed attempts in the literature. Although he discusses one of the uses of nativization in the literature and his data is Indian-English, he couches his argument primarily in SLA terms. There is some nice material here on learning strategies discovered through secondary data/retrospective interviews.

The previous issue is of course linked to the much discussed relationship of UNIVERSAL GRAMMAR (UG) AND SLA. Schachter (1990) looks at fossilization from the point of view of the "completeness" issue. Can non-native speakers have "complete" grammars? (She answers "no".) Schachter's study involves a careful investigation of one parameter (the UG principle of subjacency) in three groups located in a continuum where NLs differ with regard to the parameter in interesting ways: Korean "showing no evidence of it" and Chinese and Indonesian "showing partial evidence of it". She uses as a comparison group Dutch which "shows the full range of subjacency effects that English does". This constraint involves "movement rules", where (5a) is grammatical while (5b) is not, due to what are known as "islands":

5a. What did Sue destroy?

5b. *What did Sue destroy a book about?
The results show that all groups correctly judged the grammatical sentences which contained islands, but only the Dutch group correctly judged the ungrammatical sentences which contained subjacency violations. Thus, each group used the subjacency information, if available in their NL.

To me, Schachter provides a language transfer simplification conclusion as to why simplified and fossilized incomplete grammars are a necessary outcome of IL learning:

It would appear that the learner has only the input and knowledge of the native language as guides in figuring out the structure of the target language.

As with the MEP above, language transfer is an essential factor. Other variables such as age of first exposure to the target language, number of years of target language study and number of months in the target language country were shown to have no effect.

To conclude, as is often the case in this field, we return for wisdom to the important collection of Corder's papers (Corder, 1981; cf. discussion in Selinker, 1992, especially, chapter 6). What Corder did consistently was to emphasize that IL is "normally unstable" (e.g. 1981, 16), and to provide a conclusion as to why this should be the case. When a learner is not understood, he or she "has a motive to bring his behaviour into line with conventions of some social group, if he is able". This is an important idea and one that is widely accepted. This means that IL learning will cease when learners believe that they are able to get intended messages across with the IL system they have. There is then the interesting idea that learners "may not be able" to match the norms of a target social group which brings up important questions of "inevitability" and "innateness" of fossilization. In the Corder Festschrift referred to above, it was mentioned that "Among fossilized Francophones in Canada, there is no English plural for generations", which brings us back clearly to fossilization as simplification linked with language transfer. But careful discussion of this will have to await another occasion and is taken up in Selinker and Lakshmanan (In Press). There we propose that the multiple effects principle provides a partial explanation to the problem of plateaus in SLA and that the literature is clear in suggesting that, concerning various possible SLA factors, language transfer is a central one. This paper suggests that simplification is as well.
REFERENCES


CORDER, S P (1967). The Significance of Learner’s Errors. IRLAL 4:


Notes

1. During the final preparation of this essay, I was most fortunate to be a Gastprofessor at Universität Kassel and wish to thank Prof Dr Hans Dechert of Kassel and the Fulbright-Kommission in Bonn for making my visit possible. I also wish to thank the Kassel students in my SLA theory seminar for their insightful comments on some of this material.

   While preparing this essay, then, it turns out that I was rethinking SLA concepts while in the process of creating a German IL. I mean that I was seriously rethinking concepts such as those that appear in this essay, as well as other much debated concepts such as "comprehensible input", but this time in terms of variables in my input, such as "fast speech rules" vs. more morphologically-based speech as it was daily varying in the input. It is a sobering experience.

   For what it is worth, in terms of "internal-IL transfer", my diary is full of examples of transfer happening from my written IL to various oral domains, such as some discussed in this essay. I also had the unpleasant realization that many of my SLA and language teaching colleagues are regularly not in the process of learning an L2. It is an open secret that many colleagues in the States have never learned an L2. But, I also realized that most of the rest, even if they have in the past learned an L2, it was years ago, and that they are not presently struggling with all those "learner problems" we write about. I found myself thinking such thoughts as I was reading for this essay all those hyper-neat conclusions in the literature (only a few of which have made their way into this essay), including, of course, much of my own work in the past.
2. A similar point is made by Davies (1984) in the introduction to the Corder Festschrift, when he discusses IL in terms of "partial knowledge".

3. This of course involves "Orwell's Problem", first brought to the attention of the SLA world by Hale (1988). Hale points out that there is in SLA an element of what Chomsky, in political debate, has referred to as Orwell's problem, basically the opposite of the much studied Plato's problem of how we know so much, given the paucity of input. Orwell's problem relates to the question why we know so little despite so much evidence. Hale specifically raises the following in the context of SLA: Why, where it is so, does a marked parameter setting persist in the grammar of a fluent L2 learner despite ample evidence for the unmarked setting in the input. See Selinker and Lakshmanan (In Press), where we propose the "multiple effects principle" (MEP) as a partial answer to Orwell's problem as it applies to the SLA context. The MEP discusses cases where two or more SLA factors work in tandem, and where, it is claimed, there is a greater chance of stabilization of IL forms leading to possible fossilization.

4. From a research methodology point of view, it is important to emphasize that "Research Problem Areas" are not the same animal as "Research Questions". The latter involve precise questions which can, in principle at least, be answered in an empirical study.

5. Technically, 5a is analyzed as:
   \[
   \text{What} \quad \text{did Sue destroy 1]}
   \]
   \[
   S
   \]
   and 5b as:
   \[
   \ast \text{What} \quad \text{did Sue destroy [ a book about 1]}
   \]
   \[
   S \quad \ast \text{NP}
   \]