

ISSN 0910-500X

# 英文學思潮

THOUGHT CURRENTS IN ENGLISH LITERATURE

VOLUME LXXVI

2003

THE ENGLISH LITERARY SOCIETY  
OF  
AOYAMA GAKUIN UNIVERSITY

青山学院大学英文学会

## Online Interlanguage Analysis: 'Absence is Presence'

Charles Browne, Larry Selinker, Carol Kinahan

In this article, we present a new tool in language learning: distance online interlanguage analysis (DOILA), and look at a structure much talked about in the second language acquisition (SLA) literature: sentences with items 'missing' in them from the target language point of view, the so-called "empty category" /e/ elements which are common, noticeable, persistent and thus probably fossilizable, in interlanguage sentences<sup>1</sup>.

Importantly, from the point of view of second language teaching, such syntactic elements can be "missing" from student interlanguage even in situations where there is a great deal of positive evidence for an item in the target language teaching and natural input, and even where the student 'knows' in a declarative sense that there should be an item there. This particularly seems to be the case in spontaneous speech, which is notoriously difficult to teach. One example might be: "Thank you very much, I appreciate /e/." Here

- 
- 1) It is important to have some understanding of the linguistics of what can happen to an interlanguage structure of this complexity. It is our view that sometimes different linguistic theories, whether functional or structural, handle different linguistic phenomena better than others and that none is perfect for our purposes. In this case, /e/ sentences with deletions, sometimes symbolized by [ec] and sometimes [t] or a trace which can leave an 'unpronounced copy', are handled adequately only in the universal grammar generative literature. For this reason, it has been argued elsewhere (Selinker, 1996) that teachers must thus 'pay heed' to 'Chomsky linguistics'. The analytical problem is that the interlanguage /e/ is NOT always 'the same' as the various forms in which it occurs in the universal grammar literature, but in fact overlaps. This makes sense since that literature is based primarily on native speaker 'intuitions' and interlanguage, by definition, cannot be. This is important for the 'syntactically active' vs. 'syntactically inert' distinction made in Section III. (See also the extended discussion in Kaplan & Selinker (1997) and the references there.)

there is a "missing" /it/ *from the target language point of view*, i.e. what is 'absent' from one point of view (productive) is 'present' from another (perceptual), a constant occurrence in interlanguage discourse, i.e. the native speaker listener 'fills in' the missing 'it'. Can 'new media' help with this sort of persistent interlanguage problem? This is in general what we explore in this paper.

Section II, discusses briefly the internet 'tool' we use, which is not necessarily restricted to this or any particular interlanguage structure. But, here for purposes of illustration we restrict ourselves to this one interlanguage structure. And in Section III, since apparently the intricacies of this structure are not well-known in the world of 'computer-mediated communication' (CMC) and 'computer assisted language learning' (CALL), we present in some detail its range of potential occurrence discussing the linguistics of the /e/ structure. We avoid for this paper the difficult worry as to whatever may be the target language grammar, *from an interlanguage point of view*, whether interlanguage /e/ is indeed one structure or not — a question which must eventually be sorted in SLA; here we present an introduction to the problem.

First, given some confusion in the language teaching literature, we have to deal briefly with what is and is not 'second language acquisition' and how it has been related to CMC and CALL in the widest sense. This is not as obvious as it first appears, since one sees in the language teaching literature some very glib statements as to what 'mainstream SLA' is and how it is or is not relevant to what teachers are doing. With experience in the field of SLA that spans several decades, the authors do not recognize these characterizations as valid. This is important since looking at one interlanguage structure, as we do here, is rare or even new in CALL and CMC, which was suprising<sup>2</sup>.

---

2) We checked with two experts in the field and received the following email responses:

What is second language acquisition? For us, second language acquisition is an academic discipline which is defined as:

... the study of how second languages are learned. It is the study of how learners create a new language system with only limited exposure to a second language. It is the study of what is learned of a second language and what is not learned. (Gass and Selinker, 2001).

A quick scan of the Table of Contents of introductory SLA books reveals that SLA is a field that explains the learning and processes of SLA from many points of view and is dependent on a variety of source fields from cognitive to social. SLA is as applicable as practitioners wish it to be and to us it is very applicable to what one can do with CALL and CMC. When we speak of this wide view of SLA to teaching colleagues, they are often surprised at the eclectic nature of SLA and thus we have to initially emphasize this point.

One of the most lucid and useful attempts at reviewing the applications of computers to SLA, under the term CASLA ('computer applications in second language acquisition') can be found in Chapelle (2001) though SLA itself is never directly defined. This is important since a prominent reviewer (Bird, 2002) makes a claim about what Chapelle's view of the 'aim' of SLA is:

... Chapelle questions the extent to which computer technology has benefited SLA, a field where the aim, she argues, is centered on improving learners' L2 communicative ability (p. 41). (Bird, 2002, 17).

In fact, Chapelle agrees with us that this is a misreading of her work since:

---

I think you are right that the CMC studies tend to look at negotiation of meaning rather than at development of a syntactic feature. (Carol Chapelle, personal email communication).

AND:

No, I'm not aware of any studies that have investigated the learning of any one particular structure, such as the one you mention. (Mark Warschauer, personal email communication).



The point of p. 41 is to say that areas like computational linguistics have different objectives than the objectives that I'm interested in in the book. So, even though these areas have something to offer, doing computational linguistics is a different activity than doing CASLA. (personal communication).

Importantly, she also argues that she '... would certainly agree that the whole of SLA is not about improving instruction ... In the book, I was trying to focus on evaluation issues so my inclusion SLA perspectives in very narrow.' (personal communication)

Perhaps this is why Chapelle makes a distinction between CASLA and SLA. Our point is that many researchers in SLA have gone out of their way to deny that SLA is primarily a practical field, i.e. SLA is NOT a field where the aim is promoting 'development of communicative L2 ability'. In fact, there was once even a (now) infamous disagreement about the goal of the prestigious annual SLRF (*Second Language Research Forum*) conference and whether any practical papers should be let in at all. Thus, importantly, although social variables are central to SOME forms of SLA, 'negotiation of meaning' studies are far from the whole of SLA. To reiterate, the goal of SLA is to understand how second languages are learned (and not learned) and practical applications are in the eyes of the beholder. This view of SLA in fact is the main rationale for *Second Language Research*, a top journal in the field that is often overlooked by practitioners. This is particularly unfortunate since not only are certain very important interlanguage structures covered primarily there (fn.1) but also one sees special issues of the Journal that concentrate on topics important to practitioners such as: 'Explanations and scientific method in second language research' (Selinker and Lakshmanan, 2001).

As Chapelle points out (fn. 2) 'CMC studies tend to look at negotiation of meaning' and this view of SLA runs throughout the large volume edited by Egbert and Hanson-Smith (1999) with the

topics covered being such matters as interaction, authentic audience, authentic task, opportunities for exposure and production, feedback, cognition, learning styles and motivation, atmosphere and control, all necessary background to second language acquisition, where the goal is to promote SLA (see also the series of studies in Warschauer and Kern, 2000). However, our concern is one of balance: interlanguage analysis is a large part of current second language acquisition practice and this is not reflected in most of the CALL and CMC literature. Specifically, how can one 'promote' SLA if learner language is left out of second Language acquisition? Thus, to coin a slogan:

if SLA is only about linguistic structure, then it is not about second language *acquisition*, and if SLA is not at all about linguistic structure, then it is not about second *language acquisition*.

These are not trivial issues since they lead directly to age-old arguments as to how we know what we know about language and about language learning, and thus how we decide what to teach, whether or not language teaching is a principled discipline, whether there should be a focus on (grammatical) form, etc. These arguments go back to the pre-structural days of 'grammar-translation' teaching and part of the purpose of this paper is to bring these arguments up to date in the context of the role of information technology with a balanced emphasis on interlanguage structure. Although there is not sufficient space to provide the details of that debate here, suffice it to say that we take the position that 'focus-on-form' and organization of teaching by grammatical form, is what many if not most teachers are currently doing in their classrooms<sup>3)</sup>

3) This point was most forcefully made in a talk by Michael Swan (February, 2002 at the Colloquium on 'Explicit Knowledge in Second Language Acquisition', St Mary's College, University of Surrey), where he reported on his researching what a large number of teachers around the world actually do: 'Most of the world's teachers ... seem to follow pre-planned formal syllabuses' and thus do grammar presentations.

and that 'noticing' or 'noticing the gap' between interlanguage structure and target language structure is what many researchers believe (e.g. Schmidt, 1993 and the references there) is necessary for second language acquisition to occur, a point of view which we will come back to.

## II. The Tool: Doila

We have worked on a 'tool' in language learning which we call 'distance online interlanguage analysis' (DOILA). It is our hope to be able to use this tool to have regular virtual meetings of teachers around the world communicating with each other and analyzing 'locally important and difficult interlanguage sentences' of their students. It is also our hope to be able to create large databases or databanks of such locally important interlanguage sentences. Finally, we also hope to be able to expand the notion 'classroom' in a new media age, to help create new forms of virtual classrooms, integrated with traditional classrooms perhaps, where each learner gains an understanding him/her-self from a 'Good and Bad Language Learner' point of view, where learners can become consciously aware of their interlanguageS and how these intersect with their learning strategies. They could thus access these large databases and participate in interlanguage analysis of key locally important interlanguage structures, which they themselves may produce. The syllabuses for GBLL courses given at Michigan, Carnegie Mellon and Birkbeck College, University of London can be accessed at <http://www.bbk.ac.uk/lc/al/larry/>.

The DOILA tool uses an internet relay chat system that produces automatic transcripts which we use for further analytical and pedagogical purposes. (So far, we have had about 10 sessions producing well over 100 pages of transcripts). In order to conduct online interlanguage analyses at a distance, which in the first case, was with participants in Galicia working in real time with partici-

pants in London, we decided to use a freeware synchronous chat program as this would allow a far more natural flow of discussion than would the asynchronous posting of messages on bulletin boards. An additional merit of this approach is that it would also provide transcripts that were used the next day in class.

However, we soon found the term 'synchronous' to be something of a misnomer. For example, if one types in a message and sends it, it may not necessarily appear immediately on the screen and, to our surprise, different screens may sometimes show a different order of responses. This created some difficulties at the beginning, as participants reading the screen could not always tell to whom a particular participant was replying, or to what message. For example, a participant might type 'I agree', and no one would know with what or with whom they were agreeing, resulting in a chorus of people typing 'what do you agree with?'. Also, participants sometimes for technical reasons, had to change terminals, and others found them using not their own nickname but one previously assigned to a particular terminal, at times creating confusion. Also, we did not always have one terminal for each participant and had to make do with sometimes two or three to a terminal with only one nickname. Some participants would chitchat during the discussion, which distracted everyone else. In addition, early discussions tended to follow a very free format, which meant that they sometimes lacked direction, focus and structure.

We have gained much from this experience of trying to make the asynchronous synchronous. We have discovered an anomaly between the computationally synchronous/asynchronous distinction and a linguistic one. The computational distinction, roughly the same as between online and offline chat, in this case is misleading linguistically. If we accept the OED definition of 'synchronous' as 'existing or happening at the same time' (p. 2223, vol. II, 'Shorter Oxford English Dictionary'), it is clear from our transcripts that

linguistically, the computational distinction does not work. Synchronicity is an important coherence principle in ordinary conversation which allows, for example, the phrase 'I agree.' or 'That statement is right.' to be understood in its full referentiality. It is clear to us that in internet chat 'conversation', especially among multiple parties, with the current delayed screen appearances of inputted text online, that the interlanguage analysts, in this case, must be referentially less opaque and even over-precise, saying something like: 'I agree with Jim that the transfer is from Galician and not from Spanish.' There are other linguistic manifestations of this important anomaly in our database that we intend to discuss in a later paper.

In order to help overcome these initial difficulties, after much discussion both on- and offline, we created a discourse analysis model of how to conduct interlanguage analysis on the Web. This model first defines some of the types of talk which occur during online sessions. We isolated four types of talk: 'machine talk', or automatic messages flowing from the program, such as 'Carol has joined the group interlanguage.'; 'traffic talk', talk designed to run the sessions such as: 'Now everyone provide a gloss for this sentence.'; 'social talk', the distracting chitchat we tried to have people avoid such as: 'How's the weather in London?'; and 'content talk', which involves talk about the work, in this case interlanguage analysis. The model includes in 'content talk' about interlanguage such background categories as context, personal background information of the interlanguage speaker, data collection mode, topic of the interlanguage sentence under investigation, interlanguage sentence under study, native speaker gloss, target language gloss, and analysis (ie structure and source) of the interlanguage sentence.

The model has indeed proved useful for pedagogical purposes and is set out and explained on an instructions sheet (which can be found under the filename: 'Online Sessions: Talk and Protocols', at [http:// www.bbk.ac.uk/lc/al/larry](http://www.bbk.ac.uk/lc/al/larry)) disseminated to all participants

of our online analyses, and includes some conventions to help disambiguate the shortcomings of the chat program and ensure a more orderly and comprehensible discussion.

Thus, with the DOILA tool, students and teachers from all over the world have the opportunity to discuss how students create interlanguage and the learning problems related to that. (Cf. Selinker & Kinahan, 2001 for more discussion of this tool where the focus is on language transfer.)

### III. The Structure: 'Empty categories' /e/

It is important that the reader understand the linguistic phenomenon that we are trying to deal with using this internet tool, since its common occurrence in interlanguage is matched by its lack of discussion in CALL and CMC. This structure, as pointed out in the Introduction, is a classic case of 'absence as presence' where the absence of something can be very meaningful. If the learner says:

Thank you very much, I appreciate.

it is fairly obvious that they were trying to say the equivalent native speaker sentence:

Thank you very much, I appreciate it.

But the absence of 'it' can be very meaningful in some cases. Notice the following, taken from the Italian-English diary database referred to below. An Italian speaker who knows English very well omits objects in the discourse domains of most of her spontaneous speech and email, but does not do this in more formal, especially written domains. She said the following to one of us:

How do you come across?

The recipient of this sentence was surprised, thinking of course that she meant: 'How are you perceived?'; 'How do you appear to others?'; 'How do you present yourself?'; and thus answered some-



thing like:

I don't know, you'll have to ask others.

She was confused and frustrated, repeating her interlanguage sentence several times, at last expanding the context to something like:

You know the article by George Lakoff on 'Metaphors of Terror'.  
How do you come across?

The 'penny dropped' as they say in the UK. She had meant something like:

How do you come across it? OR: How do you come across that?

i.e. she was asking: Where do you find the Lakoff article (on the internet) or, structurally:

How do you come across /e/?

The original interlanguage sentence is an example of 'interlanguage ambiguity', the sentence without the 'it' having a fundamentally different meaning (in the native Englishes at least) to the one she intended. We have tested this interlanguage sentence with a number of native speakers and the first reaction is the one stated above: 'How are you perceived?' etc. Because it is grammatically correct, native speakers don't look for another meaning and misunderstand the interlanguage intention. Only after much negotiation of meaning does the intention become clearer. Importantly, the native speakers we tested this on required a lot of context to understand the interlanguage intention.

As will become clear, with interlanguage /e/ we know we are dealing with several different *target language structures*, but in interlanguage terms, it is an open and empirical question as to whether we are dealing with one or several *interlanguage structures*. For the purposes of this initial CALL and SLA discussion, we will treat such absent elements as one interlanguage structure and will

look at it as such from a pedagogical intervention point of view. To understand this structure (or structures) some detail is necessary.

What we do know is that restructuring this missing /e/ structure to the correct 'full' target language structure in interlanguage is difficult and may be impossible for some non-native speakers. There is an input problem here: one reason this phenomenon may be so intractable is that non-native speakers have to face target-language positive evidence in the input, where NATIVE SPEAKERS ALSO DELETE STRUCTURES. In fact, there are jokes using this feature that only native speakers seem to get, at least on first hearing. One example would be:

Context: Conference goer about to leave a conference early yells in farewell to a colleague:

I'm off.

Response: You've been off for years.

with the missing locative phrase ('I'm off [from here]') allowing the joke to go through. Though this type of construction is often used by native speakers, the authors have found that such jokes are difficult to understand for ESL/EFL students and are usually only understood by native speakers without lots of additional explanation. On a practical level, non-native speakers must be at a disadvantage if native speakers do indeed talk this way often, as we think they do. From a social point of view, in terms of interlanguage pragmatics, this type of banter facilitates group cohesion and getting along with people, certainly in the U.S. and the U.K., at least.

Thus, importantly, some sort of native speaker deletion is in the input but it often occurs with different native speaker constraints than is found in interlanguage. One native speaker example in English is sometimes called 'ellipsis', where 'subject omission' is especially prevalent in native speaker finite clauses but not in subordinate clauses, which makes it very different from subject



omission in Italian and Spanish. For example, we get the paradigm:

English: 'I have bought a book', but not 'have bought a book.'

whereas:

Italian: '*Io ho comprato un libro*'. AND '*Ho comprato un libro*.'

In sentences like '*Ho comprato un libro*,' the Italian 'missing' subject is implied. Note that in native English, however, the subject in co-ordinate structures can be deleted, as in:

John arrived and /e/ opened the door.

This is a finite clause phenomenon in native English and not a subordinate clause phenomenon (unlike null-subject in Italian, say) so in the following:

'I need a friend (whom) I can rely on,' but not

'I need a friend (whom) /e/ can rely on.'

In this sentence, for native speakers, the object 'who(m)' is deletable but the subject 'I' in the subordinate clause is not, so in this grammatical context, subject /e/ is not possible. To show the discourse domains nature of native speaker deletion, in native English subjects can be omitted in certain domains of discourse, such as informal talk, notes, postcards, and most famously in diary writing, e.g. the so-called 'Bridget-Jones' diary phenomenon:

Cannot believe have not realized this before.

As pointed out in fn. 1, the linguistics of empty categories has been well studied in the universal grammar generative literature going back to the 1970's in English (e.g. Thrasher, 1977), in a comparative Romance framework (e.g. Rizzi, 1987) and in the first language acquisition generative literature (e.g. Rizzi, 1994). For those teachers wishing to begin dealing with the grammar of this important structure, we recommend as more accessible the detailed descriptive work of Haegeman and her students (Haegeman, 1997,

1999; Haegeman and Ihsane, 1999); in fact, the example sentences in this paragraph are all taken from one of her recent talks (Haegeman, 2002).

One thing native speakers never do is particle deletion when the semantic unit is: verb + (one or two) particle(s), with the total unit becoming an idiom. Take the phrasal unit '*put up*' meaning 'to host' as in '*put up the guests*' vs. the unit '*put up with*' meaning 'to tolerate', as in '*put up with the guests*'. Native speakers will NEVER delete the 'up' or the 'up with' or the 'with' in these units respectively. So, what happens in interlanguage with these empty categories /e/? There is interlanguage deletion with the idiomatic meaning being transferred by interlanguage speakers from the verb + particle structure to the bare verb, i.e. without the particle. However, this phenomenon seems very structured, occurring from the right with interlanguage speakers never deleting a medial particle when there are two particles to the semantic unit. Take a mythical interlanguage speaker who is a particle deleter. With these structures, s/he could produce the interlanguage sentence:

I put the guests.

meaning either target language interpretation. Suppose this speaker produces, as happened with a Bulgarian-English speaker (cf. Selinker, Kuteva and Lakshmanan, 1996).

I cannot put up the guests.

Are we dealing with 'I cannot host those guests.' or with 'I cannot tolerate those guests.?' Note that with the latter meaning, pragmatically, the speaker may also be implying 'I cannot host those guests.' This last implication could be seriously misinterpreted by native speakers. This sort of deletion again is very structured in interlanguage, with the rightmost particle being deleted and thus creating the interlanguage ambiguity here, but the medial one cannot be deleted. That is, interlanguage /e/ sentences of the type:

I cannot put with the guests.

are not attested to in our databases. Also, interestingly, interlanguage /e/ grammar is structured in another way here: with target language non-separable phrasal verbs, such as 'run into' meaning 'meet by chance', the particle is never deleted in interlanguage English. That is, structures of the type:

I ran John.

with the intention, 'I met John by chance' are also unattested to in our databases.

So, what else happens in interlanguage with empty categories /e/? More equivalent native speaker structures are missing than native speakers produce, e.g. prepositions and particles. Sometimes, a number of things are missing in interlanguage and it seems to make no difference, e.g. after the World Trade Center disaster, the NYTimes ends a story about Afghanistan and the USA with the poignant words:

He stopped me by my elbow, making sure his eyes had mine. When he spoke again it was in the slow diction of a man on an excursion into an unfamiliar language, but who wanted to be heard. He nodded, deep enough to be a bow, before raising both hands to eye level and letting them flutter to his waist.

The meaning was obvious to the NY Times reporter, even high in mountains in a distant corner of the earth: the local person's gestures meant 'Towers falling down'. Then he was reported to say:

"New York," the Afghan porter said. "Very sorry us."

with the copula being deleted.

But what do we know empirically? The first detailed and serious attempt in SLA to deal with missing elements in interlanguage is that of Klein (1993) who studied what she calls 'null prep'. Klein looked at such common interlanguage sentences such as:

'Which house did she live last year?' and

'That's the house (that, which) she lived last year.'

where, in accordance with normal generative use, she claims that there exists in these sentences null prep, i.e. omission of a preposition ('in', in this case) in a question or relative clause. These claims are explored in great detail in a review article (Kaplan and Selinker, 1997) which had the puzzled title: 'Empty, null, deleted, missing, omitted, absent... items in interlanguage'. This indicated how confused we were about the theoretical status of such elements for as is well-known in interlanguage studies, how can a learner 'delete' what they don't know? Klein, taking this and other discussions on board, recently updated her account in a well-argued technical paper (Klein, 2001). It would take us far afield to discuss the intricacies of these analyses, but importantly, Klein's discussions of null prep concentrate on only one segment of what we consider a much larger problem with /e/ phenomena and this is what we want to introduce in this section.

For the reader to begin get some idea of the range of the phenomena, take, for example, some of the sentences which were discussed in Kaplan and Selinker (1997): The sentence:

I picked the children from the kindergarten.

was said by a Spanish speaker who knows English very well, where the particle *up* is obviously absent. Further testing showed that he knew that the semantic unit was *pickup* and not *pick*. Had the particle *up* been deleted? Is it absent or missing in the interlanguage? And how do we decide? (see Selinker & Kuteva (1993) and Selinker, Kuteva & Lakshmanan (1996) for more discussion of sentences like this in Bulgarian-English.). We leave for the reader the interesting puzzle as to whether the structure of this sentence is

I picked /e/ the children from the kindergarten.

or:

I picked the children /e/ from the kindergarten.

Next note the sentence:

I want to point that...

which was said by a native speaker of French who was a PhD student in the United States, where the word *out* appears to be missing from the point of view of target-language English at least and, again, where when pressed he claimed he 'knew' that the unit was 'point out'.

Another type of /e/ appeared in a front page article in the NY Times right after the Gulf war:

We will build back.

According to the article, this sentence was being regularly said by Kuwaitis at the time, after the world drove out the Iraqis. From our point of view, one wants to know if in this sentence, after *build* or even after *back*, there is an empty category with a missing object, that is: exactly what do they want to build back? On the other side of the political divide we see the same structure with the native language Hebrew, as reported in the NY Times when describing: 'Opponents of a Golan Heights Deal Rally in Tel Aviv':

Mr. Lemann's son, Daniel, 13, added, "If Clinton wants us to give the Golan to the Syrians, I think he should give back to the Indians in America first."

where the object of 'give back' is missing and has to be inferred.

Particle deletion is endemic in English intergrammars, even for those who know English very well and have been using it for decades. Prime Minister Sharon of Israel has said, more than once in the politically important discourse domain of 'international news speak':

I call the Palestinians today to stop the violence. That is what I call the Palestinians.

Well, we can guess what he 'calls' the Palestinians, but we think he must mean the structure 'call /e/' in the first sentence, i.e. in target language grammar, to 'call on' the Palestinians to do what he wants; in the second he has deleted more, in other words: 'That is what I call on the Palestinians to do'.

Chinese is known in the generative literature as being a language with 'bare' grammar, (Chomsky, 1995) vs. English with a more 'full' grammar (obligatory subject, object...), with other languages at various places along this continuum. One of the authors witnessed a Chinese colleague (who lectures and publishes in perfect academic English) on a visit to China producing in his spontaneous speech and, especially, in the domain of interpreting, interlanguage English sentences such as:

Why has been such change?

for: 'Why has /e/ been such change (in China)?', i.e. in target language terms: 'Why has there been such change?' Another example was:

Before came.

for: 'Before /e/ came', i.e. in target language terms: 'Before he came'. Such sentences, he agreed, he would NEVER utter in the English-speaking country where he lives; they definitely appear domain bound and, in terms of the 'Multiple Effects Principle' (Selinker and Lakshmanan, 1992) which links transfer and fossilization with a pedagogical corollary, such sentences may well show that when back in China the speaker tends to 'backslide' to an earlier interlanguage stage (for example, see Corder, 1981) in certain discourse domains related to spontaneous speech. Such 'syntactically active' /e/ structures as null subjects and objects (as opposed to 'syntactically inert' structures such as null-prep above) are discussed in a generative universal grammar approach to these problems for L2 Chinese learners by Yuan (1997).



There is an important example from the universal grammar generative SLA literature (Hale, 1988), the Italian-English interlanguage sentence:

This allows to conclude that . . . <sup>4)</sup>

where the target language needs a 'small pro' that can be either definite:

This allows me to conclude that . . .

or indefinite:

- 4) This sentence from the literature is one of the most interesting and is discussed in references listed in this paragraph in connection with the important problem for SLA: 'Orwell's Problem', basically the opposite of the well-known and much studied 'Plato's Problem' in child-language acquisition, namely: How is it possible that we know so much despite so little evidence available to us? Orwell's Problem, on the other hand, relates to the question: How is it possible that we know so little despite so much evidence available to us? Sentence (4a) was said by a native speaker of a Romance language, and is reported in Hale (1988). Hale's interesting example concerns a well-known Italian theoretical linguist who must know a great deal about the structures involved, apparently having written on this problem, but still produces sentences such as the one in the text:

This allows /c/ to conclude that LF movement obeys subadjacency after all. (as written by Hale, 1988: 32)

What does the /c/ equal here exactly? Is it some sort of hypothesized empty category? Of what theoretical status in the interlanguage? This sort of sentence of course is produced for the intended target language sentence: "This allows us/one to conclude.". Hale concludes that this speaker despite all that he knows "persists" in:

... the use of "small pro" object (with arbitrary reference) ... Such usage often persists despite the speaker's conscious recognition of the position of English in relation to the parameter recently discussed in detail by Rizzi (1987). (Hale, 1988: 32)

Hale then makes a most important fossilization point, integrating it with types of competence, which in turn becomes a central cognitive point about types of knowledge:

Examples of this sort suggest that certain L1 parameter settings may be extremely difficult to eradicate from an acquired L2, at least at the level of integrated linguistic competence (as opposed to conscious intellectual understanding of surface grammatical facts) ... (ibid)

Hale concludes that this may be particularly so in the case of parameters whose effects are diffuse within the grammatical system as a whole. (Cf. Selinker & Lakshmanan (1992, 200-201) where Hale's discussion is integrated into the multiple effects principle, the principle which appears to link language transfer and fossilization in interlanguage development.)

This allows one to conclude that...

To show how intractable this /e/ structure can be, this Italian-English interlanguage sentence, 'This allows /e/ to conclude that...' as is pointed out in detail in fn. 4 and the references there, was said by an Italian linguist who has written one of the definitive generative accounts of empty categories in theoretical grammar. As in many, if not most of these, language transfer is a factor and the 'Multiple Effects Principle' (MEP), the principle that unites transfer and fossilization, seems to be at work and ordinary teaching may be ineffective. (cf Selinker and Lakshmanan, 1992 for a discussion of this sentence in terms of the MEP and its attendant pedagogical corollary).

To sum up, this phenomenon of interlanguage /e/ is a discourse-domains bound phenomenon (e.g. Selinker and Douglas, 1989, Douglas, 2000 and the references there). What often seems to happen is that in the formal context of academic writing or a formal test-taking domain, where the learner has time to reflect on the interlanguage/target language relationship, these /e/ forces might be kept to a minimum. For some speakers who can handle such formal domains, we often see the occurrence of /e/ most frequently in their spontaneous speech and in its correlate, email writing. There is an important methodological point here: an ordinary empirical design of research might very well *under-represent* the presence of /e/ in important domains<sup>5)</sup> and only a diary ethnographic approach is reasonable, with such case studies of individual learners being preferable, given the huge amount of interlanguage variation. An idea of the potential range of interlanguage /e/ phenomena<sup>6)</sup>, can be found at <http://www.bbk.ac.uk/lle/al/>

5) Of course, it may *over-represent* the presence of full-/e/ in other domains of the interlanguage, the test maker/scorer nightmare; cf. Douglas (2000) for exquisite detail.

6) For reference to a data base which samples the range of interlanguage-English /e/ phenomena from a multitude of native languages, see Appendix II.



larry/, where there appears some of the results of a diary study of the spontaneous speech of several very advanced Italian speakers of English, each of whom knows so much English that their academic writing mostly appears indistinguishable from that of a native speaker.

For example, the analysis goes like this:

Context: A student is very angry at something the administration has done.

Interlanguage sentence: We won't take it to lie down . . . They keep on to move the goal posts on us.

Probable target language gloss: We won't take it lying down . . . They keep on moving the goal posts on us.

Probable analysis: We won't take it /e/ lie down . . . They keep on /e/ move the goal posts on us. This is a case where native speaker deletion occurs but with gerund not infinitive. Also, it is an interesting case where substituting an infinitive for a gerund makes a difference and apparently weakening or even losing the intended idiomatic meaning<sup>7</sup>.

From the segments in this section and from our interlanguage /e/ Italian-English database (<http://www.bbk.ac.uk/llc/al/larry/>), we can see that interlanguage /e/ is common and persistent and what is missing is not always clear to the native speaker, even when there is time for analysis. We hinted that there may be only one interlanguage process here for /e/ but several target language structures involved, which are not isomorphic to the interlanguage. In terms of *target language* grammatical classification, are all word classes equally susceptible? We don't think so. What target language word classes have we seen which are most susceptible to the /e/ process?

7) These sentences are presented as part of Problem 2.2 'Metaphors and Aphorisms' in Gass, Sorace and Selinker (1999a). See also the discussion on this problem in the accompanying 'Teachers' Manual' (ibid, 1999b).

Pronouns? — yes; Prepositions? — yes, especially stranded ones; Particles? — yes (but not 'inseparable ones' or 'medial' ones); Nouns/ Noun Phrases? — yes; Auxiliaries? — yes.

HOWEVER: Verbs? — no (except for copulas); Adjectives? — no<sup>8)</sup>.

Adverbs? — unclear at this point.

But we must insist that teachers not assume that interlanguage speakers are necessarily operating here with these native speaker categories. To imply, for example, that interlanguage speakers KNOW in all cases the difference between particles and prepositions can lead to teaching strategies that may confuse the student more. Specifically, we should not assume that interlanguage speakers always know the difference between particles and prepositions which native speakers know instinctively at an early age, as the difference between 'He [ran into] John.' vs. 'He [ran] into a wall.' or the infamous one: 'The prince [turned into] a frog.' vs. 'The prince [turned] into an alley'. The best approach to take for the moment we think is to regard interlanguage /e/ as 'a cover term' for a number of phenomena which may vary by discourse domain and by individual in various discourse domains but to treat it, in the absence of clear interlanguage SLA evidence as one structure. The key SLA research question is: 'How do you study what is not there?' OR, as Kaplan (pc) puts it: 'How do we figure out if something isn't there?'. But we leave this for other work and concentrate here on our attempts to use our DOILA internet tool to research interlanguage /e/ sentences across the internet. We have done this in two types of settings: a European setting, i.e. the University of Santiago di Compostela, where Spanish and Galician are the native languages in NW Spain, and then in a more distant setting: the Shanghai International Studies University where various

8) We owe some of this analysis to earlier email conversations with Tami Kaplan.

dialects of Chinese are the native languages and interdialects<sup>9)</sup>. Both these settings interacted with London and some individuals in other places, such as Japan, Lebanon and the U.S.

#### IV. The Database: Doila Transcripts

The full transcripts (over 100 pages) of the distance online interlanguage analysis sessions referred to in this section are described and can be accessed on the web page found at <http://www.bbk.ac.uk/llc/al/larry>. In this section, we present some sample analyses which show how we were able to have students and their teachers in several countries analyze locally important interlanguage sentences, sentences that were impervious to pedagogical treatment, where the /e/ structure concerned was brought to the active 'attention' either immediately online or, more usually, using the record of the transcripts described above the next day in the classes. We now discuss several episodes from the transcripts below<sup>10)</sup>:

##### EPISODE 1: 'disguise as a chicken'

This DOILA session which began as a social conversation, turned into a very serious and detailed analysis by advanced learners of an authentic interlanguage sentence actually produced online and is a good example of what can be achieved using this technology. The initial material comes from a session designed to test the link from Spain to London. In response to some very light-hearted

9) We wish to thank our Galician colleagues Ignacio M Palacios Martinez, Maria Rosa Alonso Alonso and Ramiro Cebreiros Alvarez and our Shanghai colleagues Professor Dai Weidong, President of Shanghai International Studies University, Zhou Ailing, and Zhang Xuemei for an exciting time in organizing and following through with these sessions. Of course, as usual we thank our web guru, Jim Tyson.

10) These transcripts were discussed in a seminar given by the second author at The College Of New Jersey (Mallorca campus) in the Summer of 2001. The following students were kind enough to put up with detailed analyses of these /e/ sentences in almost 40 degree temperatures and we are grateful for their insights: Jennifer Baxavanis, Todd Bombard, Jean Colrick, Nancy Harris, Trudi Hodges, Harry Houllis, Alla Karetny, Sue Ko, Akino Sakaoka, Ana Sanchez, Renee Semler.

banter, Rosa in Spain (<aldebar> in the transcript), a linguistically-sophisticated native speaker of Spanish and near-native speaker of Galician, produces the following interlanguage /e/ sentence:

<aldebar> That's funny because it's Carnival here and I'm going to disguise as a chicken; Immediately afterwards, we get Jim as <barukh> in London 'saying':

\*barukh falls off his chair laughing

and before anyone else has time to react, Rosa jokes:

<aldebar> An empty category about empty feathers

and then initiates a serious online discussion about the empty category she has 'produced':

<aldebar> Ok. Let's analyse this: Ramiro says it's performance-related. I agree, I know it but I didnt use it in this context

She then adds:

<aldebar> Rosa: \* I know myself should be placed in that slot but I didn't use the pronoun. Why not? There are no external factors, such as anguish, yet it is missing.

This is an example of one of the situations that learners can find themselves in, which was mentioned both in the Introduction and in Section III. Here, an empty category occurs in spontaneous online chat performance even where the learner 'knows' in a declarative sense that there should be a syntactic element there. As Rosa's comments above make clear, such occurrences can be puzzling, even frustrating, for a learner, especially when 'no external factors' such as anxiety, can account for them. Her immediate interest in analysing the /e/ is thus likely to come from two sources: the natural curiosity of the sophisticated language student coupled with the learner's need to know why the /e/ occurred in order to prevent it from happening again.

There is a potential meaning problem in this situation. The /e/ after the verb 'disguise' in:

That's funny because it's Carnival here and I'm going to disguise

/e/ as a chicken needs in native English some sort of full element depending on who is going to be disguised. If it's the person speaking, one needs a reflexive 'myself' ('disguise myself as a chicken'); if it is someone else there is a choice of noun phrase or pronoun if the referent is clear ('disguise President Bush as a chicken' / 'disguise him as a chicken'). It is not so severe as the latter, since the moderator (an English native speaker, in Spain for these sessions) then spells out the target language structure corresponding to the interlanguage /e/ under discussion:

<angel> rosa means in the above sent. that the reflexive proN is missing, right?

This forum seems appropriate then for turning learner attention toward the potential language problem which in ordinary conversation might have just slipped by and the speaker may not know they have produced an /e/ structure. When the context is too rich, native English speakers may be unlikely to give interlanguage speakers the opportunity to notice their production of this type of empty category.

Jim (a linguist and native speaker of English and interlanguage speaker of Spanish) then suggests a useful way to further the analysis:

<barukh> aldebar: ... the question is whether in one of her NLs, the verb (cognate?) disguise has how many obligatory arguments?

<barukh> if the direct object argument is not obligatory in some NL, then this could be transfer

In response, Rosa immediately asserts that her Spanish native language is the crucial factor behind the empty category in question:

<aldebar> It's transfer. In Spanish reflexive pronoun is part of the verb because it is a reflexive verb: vestirse

However, as the synchronicity of chat programs can lead to the unusual situation (vs. face to face conversations) of several people



in the same virtual room 'talking' online at the same time but with the possibility of being 'heard' by everyone else sooner or later, one of Rosa's co-participants (Ramon), momentarily oblivious to her initial comment, explores the idea that her Galician interlanguage might also be a contributing factor:

<angel]> ramon says you need the reflexive in sp and g

The discussion continues to explore with where this /e/ comes from, this being important to the learners apparently as a way of 'noticing', with the preliminary conclusion confirming the Galician "vestir" (to dress) as a possible source of transfer.

<barukh> you can't say yo voy a vertir como galina?

<barukh> oops sorry, vestir

<aldebar> And in Galician we use a verbal periphrasis together with the pronoun: vou disfrazarme

<barukh> ith my athpirathion

<aldebar> ramiro: No, you have to say, voy a vestirme de gallina gallina

<aldebar> ramiro: its also a reflexive verb.

<barukh> ok so it doesn't look like morphosyntactic transfer...

<aldebar> ramiro: however, in Galician "vestir" is not a reflexive verb

<barukh> it could be a processing / grammar interface effect: the online processing is different for this structure in english and spanish because in spanish the DO pronoun is a suffix on the verb

<barukh> a clitic aha!

<barukh> so this could be transfer from Galician?

<aldebar> ramiro: according to Rosa, both possibilities, that is, vestir as a reflexive verb and as a non reflexive verb are right

Note that in the above segment, Rosa can clearly be seen to be moving on from her first hypothesis (transfer from Spanish) to beginning to consider her Galician as a possible culprit when she is reported as stating that in Galician, "vestir" can be either reflexive

or non-reflexive, thus flatly contradicting Ramiro's earlier assertion that 'in Galician "vestir" is not a reflexive verb'. The moderator then asks (where 'g' is Galician and 'sp' is Spanish):

<angel> q: is g. diff. from sp re: reflexive?

Further discussion finds Rosa confirming:

<aldebar> Absolutely! Some Spanish reflexive verbs are not reflexive in Galician barukh, you were right some castillian verbs require reflexive pronouns where galician doesn't

A little later, Jim's point that

<barukh> of course there is vestir in castilian without reflexive as in "I dressed my child in red" or whatever the translation equivalent is in spanish

is answered by Rosa:

<aldebar> But "I dressed my child in red" is a different interpretation, does not refer to the subject interpretation, does not refer to the subject

leading her to say:

<aldebar> But you can say in Galician, eu visto de galiña (I dress as a chicken) where the pronoun is not required

and add, a little later:

<aldebar> But in Spanish you have to use the pronoun in the "equivalent" sentence in Spanish: Me visto de gallina

The above points made by Rosa herself, showing online progression in learning, in her dissection with her colleagues of her own Spanish native language and Galician interlanguage structures, document her gradual progression away from her first rather knee-jerk hypothesis, and culminate in her embracing a very different one, although there is one reservation. Rosa is now <jajajk> in the transcript:

<jajajk> Rosa: O.K. Jim, the ec is a case of transfer from my Galician IL where the pronoun is optional. But I rarely speak Galician, so how is it that there is transfer from my GIL to my

spoken English?

This is a very good question and one which has only relatively recently begun to be discussed by SLA researchers, usually in the context of an area called 'interlanguage transfer', which Jim begins to answer as follows:

<barukh> can you say in g eu me visto de galinha? it's a complex lexical choice problem that causes this error! the g translation equivalent of disguise [*disfrazarme*] shares a semantic and phonetically similar prefix with english so the english primes the galician lexical item

<barukh> the castilian lexical item is chosen online despite the g lexical item's priming

Rosa's next crucial question:

<jajak> So, why is my NL blocked in favour of my GIL?

relates to class discussion of the importance of interlanguage transfer where it seems that the native language is blocked and is less dominant than one presumes. That is, interlanguages, as well as native languages, can be a source of other interlanguage structures. Rosa's question is followed in the transcript by Jim's expansion of his argument, maybe not seeing Rosa's question:

<barukh> but because disfazar primes vestir in g and spanish primes vestire in g also ... the morphosyntax of the g lexical item influences the IL output

<barukh> i just explained it!

Thus, Jim feels he has worked out where the /e/ comes from and then 'thinks aloud' about the interlanguage transfer issue raised above by Rosa:

<barukh> it's about the relations between cognates in your NL and IL mental lexicons

<barukh> the spanish and the galician are alike enough that they strongly prime each other and either can be available for IL transfer on-line



<barukh> the actual transfer will probably reflect a markedness hierarchy

<barukh> no because \*both\* structures are available in g, and so the overall structure is the shared one

<barukh> you could represent this as a weighted network of connections

<angel]> what do you mean by overall structure?

<barukh> a probabilistic IL grammar linked to performance factors  
Trying to tie up the loose ends, Rosa makes the following point:

<jajjk> Rosa: Galician primes the verbal periphrasis, while in Spanish the lexical item primes, however the periphrasis was not transferred.

which is indeed an interesting learner hypothesis coming from the learner herself. Jim continues to amplify his analysis, and then answers Rosa's point:

<barukh> g allows a periphrastic structure and shares a straightforward structure with s and e

<barukh> so the g periphrastic is only weakly primed compared with the straightforward SAAD sentence

Rosa's final question is:

<jajjk> RosArosa: yes, BUT Galician and Spanish only share the lexical item, not the verbal structure. Yes it's possible that the core element is not transferred, why is it that I transfer what is not marked, taking into account that core elements prime in transfer.

And Jim's answer, and final conclusion, is as follows:

<barukh> the (y) do share the lexical structure but it is optional in g and obligatory in s.

<barukh> my argument is that the primings are such that the g lexical item v is primed and affects the IL syntax, the g verb desfazar is primed and primes vestir in g

<barukh> the spanish translation equivalent of disguise primes vestir in g... all of this leading to the selection of one lexical item

and one syntactic structure but they don't match directly either s or g grammar

Rosa offers no argument to Jim's final conclusion. Throughout this transcript, we witness Rosa's continuous questioning of her own native language and interlanguage knowledge, prompting a joint exploration of the possible reasons behind her production of a particular /e/ interlanguage structure and showing online progression in learning. One advantage of DOILA then is that it allows the learner to notice and analyze in a sophisticated manner their interlanguage sentences and to realize that in the terms of Semler (see fn. 10 above) 'that there may not be one static answer'. DOILA has allowed Rosa to do just that. Semler adds: 'From a teacher's point of view, witnessing a learner's own inquiry into his / her learning, abilities, and capabilities, as Rosa continuously did in this transcript, is remarkable'. (Cf. Selinker and Kinahan, 2001, for further discussion of this example.)

#### EPISODE 2: 'plug this'

The interlanguage sentence: 'We need to plug this.' was said by the technician in the first (1998) sessions. The technician was a PhD student in English linguistics who knew English very well. The moderator, native speaker Larry running the session in Spain asks online a native speaker linguist, Jim, and Italian-English speaker PhD student, Rita, the latter two both in London:

<larry> so, jim, rita, how would you analyze this?

Jim, who is <barukh> in the transcripts says:

<barukh> missing particle: it must be a particle because it is a mobile in the vp, so to speak, in english: you can say we need to plug in the light or we need to plug the light in

Then, more analysis follows, with barukh noting that (like many of the interlanguage /e/ sentences above), the interlanguage sentence as produced could have a meaning in target language English and thus is ambiguous in interlanguage English:

<barukh> but... that's a bit hasty. I can use "plug" without a following particle when im talking about plugging a hole or something

Then comes a request for more context, which is the key apparently for native speakers to get the interlanguage intention, once they realize that the interlanguage sentence does not have the meaning they would normally ascribe to it. Rita then guesses:

<ritam> we need more context. i could be saying we need to plug this in the back of the tv and barukh states:

<barukh> ritam: you're right of course — the use of fragmentary data is always full of dangerous temptations to speculation!

The moderator then gives the instruction:

<larry> from s de c: type in more context for rita

and gets the following from the interlanguage speakers in Spain:

<Ros> We have to plag this into the socket

<Dolo> you just have bought a new computer and have fixed all the different parts, but it won't work, so somebody finds out what happens: We need to plug this...

<judith> this piece of sentence was produced when trying to connect to your terminal last monday.

<malu> there is a clear ec after THIS.context: we need to plug this so that we can connect it to the computer

<raquelT> the context was "to connect" refering to plug and socket

<Ros> ... and the other part of this thing into the video.

<isabel> this sentence was said in the computer room in Santiago a few days ago. Something was not working because it was not properly connected.

<angeles> somebody it is trying to make something work and this person finds out the problem: we need to plug this

<AnaM> the speaker wanted to conect something in the computer

<javi> We need to plug this to see it in the overhead projector.

These interlanguage comments are important since in every case, these interlanguage speakers CONFIRM the interlanguage sentence by re-producing it<sup>11</sup>). With such context, barukh and ritam catch on to the intended interlanguage meaning and the possible interlanguage ambiguity and provide the definitive analysis for the non-native speakers:

<barukh> on the particle analysis, it would be "we need to plug this /e/"

<barukh> if it were a missing prep, we would expect the speaker to have produced a bare NP argument: we need to plug this lead

...

<ritam> i agree with jim the /e/ in after 'this'

<barukh> javi's example still has a missing element when compared to the tl

<barukh> since nothing follows in the utterance, it seems most likely that it is after all a particle

The penny then drops for one of the interlanguage speakers who realizes that with the interlanguage ambiguity, but not in the speaker's

11) The analysis that the interlanguage sentence has /e/ is also confirmed when the interlanguage speakers are asked by the moderator in Spain to give the native language gloss:

<larry> next from s de c type 2. POSSIBLE NL GLOSS:

<angeles> tenemos que enchufar esto

<judith> tenemos que enchufarlo, conectarlo.

<Dolo> hay que enchufarlo

<malu> necesitamos enchufar esto

<AnaM> tenemos que enchufar esto

<Ros> Tenemos que enchufarlo (aquí)

<raquelT> necesitamos enchufarlo

<isabel> necesitamos conectar / enchufar esto

<javi> necesitamos enchufarlo

The analysts in London get the confirmation as well:

<barukh> ok, so the lexical item selected in spanish is not phrasal (not surprisingly!) and takes a bare NP direct object.

<barukh> ritam: yeah, erm, particles don't form a constituent with arguments in VP and form a semantic unit with the verb showing in a contrastive way the English particle learning problem.

native language:

<malu> you really have several different contexts now and for another:

<judith> jim, you mean this is the case of a missing particle, i see your point. yesterday we had some confusion regarding particles and prepositions.

Then the moderator in class the next day, as language teacher, is able to show the students how THEY were able to notice the /e/ structure through their own work using the DOILA tool.

EPISODE 3: 'Yes I'd like.'; 'If you'd like.' 'If you want.'

The interlanguage sentence: 'Yes, I'd like.' was said by another of the local participants who knew English very well. This was said as the set-up to the online session by the moderator:

<larry> ns says to local person: do you want to go to the cinema?

<larry> local person says: yes, i'd like.

All the local English teachers agreed that this was a locally possible English fragment and most of the students, realizing we would be looking for /e/, analyzed the sentence as:  
Yes, I'd like /e/.

But one speaker in Spain, Ros, says on the transcript:

<Ros> My English is not good enough. I would say: Yes, I'd like IT.

This showed us that she was aware that there was /e/ and tried to fill it, though in an idiosyncratic way. Later on in the discussion, it becomes clear that Ros has not only 'noticed the gap' with respect to this structure between her interlanguage, 'Yes, I'd like it' and the target language, but she has also, of her own volition, started using the TL structure itself, and this is after the analysis, where Jim, the native speaker had said:

<barukh> this is a difficult one!

<barukh> yes, i'd like /e/

<barukh> where /e/ is a proS ... which strangely could also be



realised by a bare infinitival "to".

We then find Ros starting to use the target language structure itself:  
 <Ros> Me gustaría ir al cine; I' say this, if I could not go there,  
 but I'd like to.

Soon after Rosa's TL gloss ('Yes, I'd like IT'), <barukh> works out  
 a second TL gloss:

<barukh> or ...

<barukh> yeah, exactly it could be not it but THAT

<barukh> yes, i'd like that

In the ensuing analysis of the interlanguage sentence 'Yes, I'd  
 like /e/', one participant, <isabel> offers the following comment,  
 which shows her integrating her meta-linguistic knowledge with her  
 interlanguage English knowledge:

<isabel> it is a case of an empty category or gapping both native  
 and non native speakers may produce this sentence. The missing  
 particle is to.

She appears to be saying here that native speakers of English  
 would also produce the IL sentence 'Yes, I'd like /e/', which is  
 very unlikely, but there is a situation where native speakers could  
 produce the clause:

... d like /e/.

in reply to the question above:

Do you want to go to the cinema? If you'd like.

This target language fact was brought out by a student later on  
 in the online session. Apparently, the discussion of the empty  
 category in the interlanguage sentence 'Yes I'd like /e/' prompts  
 <javi>, one of the Spanish speakers, to introduce a related ques-  
 tion:

<javi> There might be an *ec* after like. When you answer to a yes/  
 no question, you need a proform like *to* but in *if* clauses it is not so  
 clear if the *to* is needed in the target language.

It turns out that this is perfectly correct, that is you could in

informal conversation get the following target-like response that would work:

Do you want to go to the cinema?

If you'd like /e/.

which interestingly could vary with:

If you'd like to.

This new angle is taken up by <malu>, another Spanish speaker, after the above discussion has concluded, and shows that for these interlanguage speakers of English, possible interlanguage empty categories are a real problem as is variable verb choice, here 'like' vs. 'want'. Prompted by all of this, she cleverly then asks:

<malu> what about IF YOU WANT (TO)?

and then after some more discussion she asks the crucial pedagogical question:

<malu> ok i would like you to explain to me why you can answer this but not I WOULD LIKE

Then we see:

<ritam> larry would you say IF YOU WANT without the TO?

<ritam> about in the previous context of the cinema?

Not realising that the question originally came from <malu>, <larry> answers and gets a useful exchange with Rita, the Italian-English speaker in London.

<larry> rita this is a real interesting q:

<larry> yes, under some conditions i could say IF YOU WANT

<larry> so what we have i think is a general tendency to gap in this way which sometimes leads to nn /e/

<larry> and sometimes leads to what equals ns gapping.

<larry> so the same process is ok sometimes for the tl speaker and sometimes leads to a clear noticeable deviance, in the 1: many interlg sense, is this clear?

<ritam> so we could hypothesise that the nns has different inputs from the ns and sometimes he gets it right and sometimes he

doesn't.

So, one again sees confusion with /e/ structures for non-native speakers because of variable native speaker input, with native speakers sometimes producing a syntactically active /e/ and at other times producing an overt full pro-form.

Interwoven in the above comments from <larry> was the following concurrent discussion showing very active learner interaction with each other and the structure.

<Dolo> regarding Malu's If you want ... Can it be a direct translation from Galician?

<malu> dolo, i think it ['If you want'] can also be a direct translation from spanish

<raquelT> it's not necessarily a translation from spanish . a ns [of English] would say that

<Dolo> is there a contextual difference between If you want (TO) or would a ns vary at will, without taking into account who you talk to, the degree of familiarity, etc?

<isabel> i think IF YOU WANT woul d be correct in Can we go to the cinema? yes if you want

<malu> who is your question for,dolo?

<Dolo> my question is for an English ns

<malu> ok

<raquelT> the degree of familiarity doesn't vary

The session ends with the third author of the paper in London confirming part of what the Spanish speakers want to know:

<carolk> I would say 'if you want' in an informal situation, I think.

The learners are right: there are indeed contextual/pragmatic differences between the several native speaker variants which further complicates the issue and was discussed in class later when the students had the printed transcripts. Finally, in terms of input to the learner, note that native speaker deletion in this context is the norm, whether 'Yes, I'd like to (go to the cinema)', 'If you want (to



go to the cinema)', 'If you want to (go to the cinema)', with the full sentence being overly formal, which is where learning problems may begin.

EPISODE 4: 'keep up'.

Here the moderator in Spain starts the online interaction with the context:

<larry> so, the context: we were walking on a mt top where there was a castle and she (Rosa) was talking about the problems of the city who now owns the castle

<larry> rita, carol, do you get the context?

<ritam> yup

<larry> carol are you there?

<carolk> yes, and Jim's on this terminal too

<carolk> We get the context

Then he produces the interlanguage sentence to be discussed:

<larry> rosa says: they can't keep it.

And requests:

<larry> what do you understand by that?

The following exchange happens with the analysts in London, beginning with Rita and Carol, and the students in Spain, beginning with Malu.

<ritam> anab is joining us and we'll be sharing this pc. over to her now

<carolk> The city can't keep the castle

<malu> possible meanings intended: castles can fall down if you have no money to care FOR them; if you don't keep up your castles you will definitely lose them and they will just be remains

<larry> carol, come up with a paraphrase

<carolk> What did Rosa actually intend to say?

<larry> rosa intended to say that the city needs money to maintain it.

<larry> carol is that what you understood?

<carolk> I wasn't sure, I thought she meant they were going to lose the castle for some reason, have to sell it

<ritam> could it be that if they don't have money someone else will take over them and they can't keep it?

The interlanguage ambiguity then is recognized with no way for the participants to disambiguate the possible /e/ sentence with the structure being either: 'They can't keep it.' with no missing element or 'They can't keep it /e/' with a missing one. Then the moderator asks for native language glosses and interestingly they vary, showing possible learner confusion:

<larry> everybody: 3. POSSIBLE NL GLOSS:

<AnaM> no pueden mantenerlo

<judith> no pueden mantenerlo.

<larry> rosa: no pueden mantenerlo

<malu> no pueden mantenerlos

<Dolo> no tienen dinero suficiente para mantenerlo en pie

<javi> no pueden mantenerlo

<angeles> no pueden costearlo

<raquelT> no podian mantenerlo, no tenian dinero suficiente para sostenerlo

<Ros> No pueden mantenerlo (includes perhaps the consequence: they'll have to sell it)

<isabel> no pueden permitirse sufragar los gastos o no pueden mantenerlo

This possible confusion is later discussed in class. In the transcript, the moderator then goes back to asking for analysis:

<larry> ok, so what would you say carol for rosa's meaning?

<carolk> The 'up' particle together with 'keep' as in 'keep up' the castle appears to be missing

<carolk> Rosa's meaning: They cannot maintain the castle

<ritam> carol, no if she means no pueden mantenerlo because the verb has a different meaning from to keep up

<larry> ok so carol thank you, what i have been telling them has been backed up (maintained!!!) by you.

<carolk> Jim is saying: the pattern that appears to be emerging is that the category that consistently appears to be missing is the particle

The non-native speakers were clearly having a hard time with this one.

Again, this information from the transcripts seemed to the students the next day to be invaluable in raising their consciousness of what is going on here in this interlanguage and how native speakers might have problems with interpreting interlanguage intention.

### V. Conclusion

Briefly, we have tried to demonstrate the potential effectiveness of a new internet tool DOILA (distance online interlanguage analysis) working with a persistent interlanguage structure, which we have symbolized as /e/, where grammatical elements are missing in the interlanguage from the target language point of view and can at times cause important communication problems. We think this structure is 'fossilizable' by discourse domain and can be brought to conscious attention with the DOILA tool both online and using the transcripts in a conventional classroom.

We have felt the need to illustrate this /e/ structure in some detail since it is apparently not known in the world of CMC and CALL. Thus, we have presented a range of this persistent and resistant interlanguage structure, discussing the linguistics of the /e/ structure but avoiding for this paper, the difficult worry as to whether, from an interlanguage point of view, interlanguage /e/ is really one structure or not. That is a question which must eventually be empirically sorted out in SLA theory and here we present an introduction to the problem. We think we know how to do this, but here we leave the detail for another occasion.

On a general level, we believe that the solution to this problem lies in not only helping the learner to 'notice' the gap between missing elements in their interlanguage and the equivalent 'full' target language structure but also to CONFRONT this gap as a serious learning problem. After several years experience, we feel that the DOILA internet tool can help this process. We cannot underestimate the importance of having students do this to lessen miscommunication. What we think we have noticed is that such miscommunication happens, not only when there is a MISMATCH between co-participants in activating discourse domains (for example, see Douglass & Selinker, 1985, Selinker and Douglas, 1989, Douglass, 2000). What is new here is that such miscommunication appears to happen when there is a mismatch in another way. That is, when co-participants (not limited to two, necessarily) mismatch 'syntactically active' and 'syntactically inert' /e/ structures within discourse domains. When co-participants match either syntactically active structures or syntactically inert structures, we hypothesize there will be NO miscommunication from this source. But when there is a mismatch, e.g. with one co-participant producing a syntactically active /e/ structure and another hearing a syntactically inert structure, we hypothesize MIScommunication. (cf. the 'How do you come across?' vs. 'How do you come across /e/?' example or the 'We need to plug this.' vs. 'We need to plug this /e/' examples above). That is, 'absence' for one may be either absence for another or presence for that other and this plays an important role in getting one's message across in a context.

We are convinced then that this computer application provides a useful tool for second language learners using it to understand their interlanguage development and non-development as it relates to many recalcitrant structures at crucial points in their learning, as well as being an aid to researchers understanding interlanguage learning in CALL and CMC conditions.

### Acknowledgements

We want to thank Carol Chapelle, Dan Douglas, Mike Selinker and Mark Warschauer who were kind enough to answer our various queries when we weren't sure about something important here. We appreciate the broad perspective of each.

### References

- Bird, S. (2002). Review of Chapelle (2001). *Language Learning & Technology*. 6.1. 17–20.
- Chapelle, C. (2001). *Computer applications in second language acquisition*. Cambridge: Cambridge University Press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge: CUP.
- Douglas, D. and L. Selinker. (1985) Principles for language tests within the 'discourse domains' theory of interlanguage: research, test construction and interpretation. *Language Testing*. 2: 205–226.
- Egbert, J. and E. Hanson-Smith (Eds.) (1999). *CALL environments: research practice and critical issues*. Washington, D.C.: TESOL.
- Gass, S and L. Selinker (2001) *Second language acquisition: an introductory course*. Second Edition. Mahwah, NJ: Erlbaum.
- Gass, S, A. Sorace and L. Selinker (1999a). *Second Language Learning: Data Analysis*. Second Edition. Mahwah, NJ: Erlbaum.
- Gass, S., A. Sorace and L. Selinker (1999b). *Teachers' Manual to Gass, Sorace and Selinker* (1999).
- Haegeman, L. (1997). Register variation, truncation and subject omission in English and French. *English language and linguistics*. 1: 233–270
- Haegeman, L. (1999). Adult null subjects in non pro-drop languages. In: M. Friedemann and L. Rizzi (eds). *The Acquisition of syntax*. London: Longman.
- Haegeman, L. and T. Ihsane, (1999). Subject ellipsis in embedded clauses in English. *Journal of English language and linguistics*. 3. 117–145.
- Hale, K. (1988). *Linguistic Theory: Generative Grammar*. In Flynn, S. and W. O'Neil, (eds.) *Linguistic Theory in Second Language Acquisition*. Kluwer. 26–33.
- Klein, E. (1993). *Toward second language acquisition: a study of null-prep*. Dordrecht: Kluwer.
- Klein, E. (2001). (Mis) construing null prepositions in L2 intergrammars: a commentary and proposal. *SLR* 17.1.
- Kaplan, T. and L. Selinker. (1997). Empty, null, deleted, missing, omitted, absent... items in interlanguage. Review article. *Second Language Research*. 13: 2, 170–186, focussing on Klein (1993).
- Rizzi, L. (1987). Null object in Italian and the Theory of pro. *Linguistic Inquiry* 17.501–57.
- Rizzi, L. (1994). Early null subjects and root null subjects. In: T. Hoekstra and B. Schwartz (eds). *Language acquisition studies in generative grammar*. Amsterdam: Benjamins.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual review of applied linguistics*. 13: 206–226.
- Selinker, L. (1996). Teachers must pay heed to Chomsky. *English Language Gazette*. July, 1996.
- Selinker, L. and D. Douglas, (1989). Research methodology in contextually-based second language research. *Second Language Research*. 5.1: 1–34.
- Selinker, L. and C. Kinahan. (2000). *Online interlanguage analysis: pilgrimage from Santiago*.



- In: I. Plag & K. P. Schneider (Eds). *Language use, acquisition and language history: (Mostly) empirical studies in honour of Ruediger Zimmerman*. Trier. 83–98.
- Selinker, L. and T. Kuteva (1993). Metaphor as a consciousness raising strategy. In Mair, C., M. Markus and J. Wallmannsberger (eds). *Neue Ansätze in der Kontrastiven Linguistik (New Departures in Contrastive Linguistics): Proceedings of the 1991 Conference*. Universitat Innsbruck.
- Selinker, L., T. Kuteva and U. Lakshmanan (1996). The 'pickup' phenomenon: on researching a tough interlanguage area. In: Sajavaara, K. and C. Fairweather (eds). *Approaches to Second Language Acquisition*. University of Jyväskylä. 43–54.
- Selinker, L. and U. Lakshmanan. (1992). Language transfer and fossilization: the "multiple effects principle". In: S. Gass & L. Selinker (eds). *Language transfer in language learning*. (Revised edition). John Benjamins. 197–216.
- Selinker, L. and U. Lakshmanan. (2001). Explanations and scientific method in second language research. *Special issue of Second Language Research*. 17.4.
- Thrasher, R. (1977). *One way to say more by saying less: a study of so-called subjectless sentences*. Kwansei Gakuin University Monograph Series. Vol. 11. Tokyo: Eihosha Ltd.
- Warschauer, M. and R. Kern. (2000). *Network-based language teaching: concepts and practice*. Cambridge: Cambridge University Press.
- Yuan, B. (1997). Asymmetry of null subject and null object in Chinese learners' interlanguage of English. *Studies in Second Language Acquisition*, 19: 467–497.
- Zhang, Xuemei. (2000). *New advances in on-line interlanguage analysis in China*. manuscript. Shanghai International Studies University.