MARK GARNER
JOHN MCKENNY
Northumbria University
Reino Unido
mark.garner@unn.ac.uk
john.mckenny@unn.ac.uk

Getting real: can an ecological theory of language help in the development of materials for language learning?

Introduction

"There was a time when members of the linguistics corps regarded a corpus as a corpse". Sture Allén addressed these words to a gathering of corpus linguists at the 82nd Nobel Symposium in 1991. This clever juxtaposition of three cognates from the Latin root *corpus* contains a reference to earlier reservations about corpus linguistics held by transformational -generative / grammarians.

But it also reminds us of a paradox which still remains with us. If we take naturally occurring language from its milieu and computerise it, are we not in some way killing the thing that we love? This is an extreme version of the Observer's paradox. Does a node word or phrase in a concordance line with a span of four, five or ten words retain enough of its original co-text to make it fully intelligible to the peruser of the concordance page or screen? Many corpus linguists (and we count ourselves among them) would say that in most central and typical cases, a concordance with a span of four words is enough to reveal the most important collocates of the node word. They might add that the charge of decontextualization is overstated because current text retrieval software can supply the full co-text from which each concordance line is taken. Indeed, in seconds, a great deal of sociolinguistic contextual information can be called up, via a header or from a separate file, for all texts in the corpus: e.g. the gender, age-group, social class, dialect and educational level of each speaker or writer is registered together with the size of the audience or readership.

The sensations of the writer or speaker is one facet of communication which is not usually recorded in corpora. Another dimension not usually reported is how the utterances or written texts are received and the effect they produced on the listener or reader.

If we look too much at lexis we need to continually stand back to avoid losing sight of discourse***.*** . We can see the trees or the wood but not both. When we affirm that the whole of discourse is greater than the parts, we hope we don't appear unduly sceptical about the empirical enterprise because we are fully committed to it. When we do corpus linguistics, we compare wordlists and make concordances. When we extract words, N-grams and concordance lines from their co-text, the co-text and context are temporarily forgotten. This dissection of language need not necessarily transform my corpus into a corpse. We have worked mainly with written texts for the last six years and am less familiar with spoken corpora and the large body of research on them Dare we make a virtue out of ignorance and suggest that we can thus look more naively and pre-theoretically at the spoken data? Were we to say that we could be more corpus-driven (Tognini-Bellini 2003), we could expect the retort that there are no such 'innocent' looks.

Our data comprise twelve first encounters at La Trobe University in Australia. These were volunteer undergraduates, previously unacquainted with one another, whose initial conversations, in pairs, were recorded and transcribed. (We are aware that the communicative situation was only partly naturalistic. Nonetheless, in the post-conversation interviews, all participants indicated that they had established some form of rapport with their conversational partners, and the evidence of the tapes is that, after some initial stiltedness, almost all conversations appeared to proceed fluently and naturally.

The data were collected as part of a pilot study to test whether analysis of real-life communication could lead to the production of ecologically valid teaching materials. The data were interpretatively annotated for any form of patterning of interactive behaviour, by two researchers working independently. The aim was an ecological description of initial conversations. The original study did not include computer analysis, but we have recently decided to apply both approaches to the analysis in the hope of discovering synergies or complementarity of views.

Might the idea of context be enriched or enhanced so that the corpus linguist could approach the corpus she wishes to interrogate with a set of macro-questions derived from an all-encompassing theory or cluster of such theories? Stubbs (2001:242-244) attempts to outline just such a blend of the theories of Popper, Giddens and Searle. We would like to suggest the marriage of corpus linguistics and an ecological theory of language***.*** One of the hoped-for offspring is more ecologically valid language teaching materials.

Section 1

Corpus-based approaches to linguistics were faulted because they modelled the wrong aspect of language—performance. According to Chomsky (1965), for example, the proper object of study for linguists was the knowledge or competence underlying such performance data. Partington (1998:145) suggests that a corpus is neither performance nor competence but supplants the differentiation between the two concepts. He reminds us

that the concept of competence, the ideal speaker's knowledge of the language, raises certain philosophical problems because claims about 'ideal knowledge' are not falsifiable by any evidence. But the individuality of performance needs to be transcended.

If the speaker's competence could be made conscious then it begins to resemble the type of knowledge that we propose instilling in language students, a kind of holistic, strategic pre-lingual stance to the communicative act. But what corpus analysis uncovers is not immediately relevant to the language teacher. From corpus to classroom is a long road.

Sinclair and Renouf underline the efficacy of the computer as a scientific instrument for linguistic study, but one which works quite differently from the minds of the students in language courses:

Retrieval systems, unlike human beings, miss nothing if properly instructed-no usage can be overlooked because it is too ordinary or too familiar ... The human being, contrary to popular belief, is not well organized for isolating consciously what is central and typical in the language; anything unusual is sharply perceived, but the humdrum everyday events are appreciated subliminally (Sinclair and Renouf 1988: 151)

This quotation carries an implicit warning about the mind's limited capacity to learn from patterns. It would appear that decontextualized chunks of language cannot simply be reinserted into the warp and woof of language. If it is easier for us to detect differences and appreciate regularities contrastively, this suggests (pace Michael Lewis, 1993) that showing regularities to language learners might not engage their attention as effectively as the usual bifurcatory either/or rules of pedagogic grammars. A favourite manoeuvre by lecturers in ELT methodology is to act out a scenario which represents typical practice in an unreconstructed ELT classroom. Maley and Duff amusingly describe the teacher who asks the question, "Is this a pen?" and asks the reader to imagine what would happen if you approached a London docker with a pen and posed the same question. Widdowson (1972:16) lampoons the teaching of the relatively infrequent present progressive form of verbs ("I am walking towards the door, I am opening the door..."). He suggests that only specialized activities would use such a grammatical construction (e.g. a bomb disposal expert or a TV chef). According to Mindt (1995) only 5% of verb occurrences have progressive aspect.

The implication of this kind of critique of textbook language is that the language presented to learners should resemble real-life language as much as possible. With the advent of computer corpora our capacity to capture real language has been hugely increased but even some wielders of corpora have questioned the wisdom of such direct transplantation. Carter (2003) talks of a classroom reality which might require cleaning up and simplifying real-life language to improve language students' learning.

Could a computer analysis of the smallish La Trobe corpus contribute to the aim of developing ecologically valid language teaching materials for adult beginners? This is what I set about investigating using Wordsmith Tools Mark III. Having a corpus which is of a readable length means that hunches are easier to follow up. The question still remains of how second language pedagogy can benefit from the findings of corpus or any other analysis of the target language. The beauty of corpus linguistics is that we feel that we can just go forth boldly and sort out the theory later.

Barbieri (2003) contains an analysis of what she calls the "new" quotatives, expressions such as BE+LIKE, BE+ALL and GO which introduce direct speech and seem to be supplanting the more traditional SAY in the speech of younger speakers of English. In my analysis of the La Trobe corpus I found that approximately 30% of the occurrences of WENT were quotatives as were c. 20% of the uses of GOING. Barbieri provides a sociolinguistic/dialectological explanation of this aspect of the data: the language is undergoing diachronic change which is picked up in the smaller La Trobe corpus and in the much larger MICASE corpus. I checked the usage of the speakers aged from 0-24 in the BNC and found only a sprinkling of these quotatives. Most of the recordings for the spoken part of the BNC are from at least a decade ago.

N	Concordance of WENT in La Trobe corpus
1	it's excellent and so we went and had a look like but
2	oh so we just like er went down and just thought oh yeah
3	yeah but um yeah we went out last night to celebrate
4	hands up and sort of they went jeez so there's so
5	yeah it was we actually went to the exhibition first
6	ah I know exactly well um I went there and then because
7	ages to go but um it just went really quick like um
8	yeah no I'm I just went to uni up here yep
9	done but I don't know how I went because I mean I just could
10	as just too easy S; went with each person? oh
11	good I look - I was er I went down ah to the - um
12	Caesar \ ahm and we went to Melbourne and they had
13	I saw the play it just sort of went woof it just came to life
14	when I was in High School I went to Melbourne and we actually
15	he'd eaten all this stuff and he went awr crikey good
16	is morning at ten o'clock and went should I have something
17	oh my God oh and I went oh shit I loved it

18 operation and felt sick oh and he went like he went I can't eat
19 sick oh and he went like he went I can't eat and he …
20 ugh 99 I came back again and went part time finished off
21 mm um when I actually went to see him I don't like
22 cars and things and he went no car haha like
23 the course at BRIT um I went for a year without a computer
24 I skipped the first year and went straight in to the second year
25 that's not a bad life though I went over there myself […] on the
26 Indian pacific um um like I went to Melbourne and Sydney
27 it's beautiful out there we went down er to Cooperpede
28 gone down by the time we went through yeah you
29 the last huge floods that went through yep and we
30 week and sort of hang out I went over there when I was in ye
31 else for two years? ah I went part time oh ok yep

I began to suspect that the role these strings play in the communication is a form of self-affirmation. It is in this sense that Wray **(**2001) sought to explain the use of formulaic sequences. Quotatives might be a way to affirm or even advertise the self. The speaker suggests that this is the world from their point of view, this is what they saw and felt. Quotatives might then be considered a subset of self-revelatory devices where the listener is made privy to the thoughts and feelings of the speaker. This may be an aspect of a wider cultural phenomenon: there is, for example, a new culture of management (derived ultimately from therapy), in which laying bare one's feelings and avoiding conflict provide the platform for interpersonal and organizational communication. In the future we would like to explore these reported thoughts and feelings using the taxonomy suggested by McIntyre, Bellard-Thomson, Heywood, McEnery, Semino and Short (2003).

Another feature of the language of the La Trobe corpus that became apparent from doing an N-gram analysis was the function of the most common trigram *I don't know.* Only a minority (23.6%) of occurrences express ignorance. The majority of instances are followed by an affirmation of some kind and could be viewed as a hedging or framing device whereby the speaker, as it were, renounces dogmatism before affirming something. Their role within the communication appears to be that they leave open the possibility of discovering or negotiating common ground.

N	Concordance of **I DON'T KNOW** in LaTrobe Corpus
1	Australia and sort of you know I don't know a field trip to some
2	what's in Chicago that you'd I don't know just to see all those places
3	so did Mark yeah I know I don't know why
4	t it gets commercial then yeah I don't know how * might even be
5	really small and everything but I don't know I think you'd put up with
6	plenty of space yeah well we - I don't know I'd be a little bit worried
7	did you I got it all done but I don't know* how I went because I me
8	he's over in the Business building and I don't know* if she does that was
9	oh you might very well might oh I don't know maybe see I've always
10	you see the Tempest when you're I don't know yeah you
11	… I think you're whole concept of I don't know of life itself does
12	blah I was just so mad at her I don't know you do alter you
13	gives you a different mm: oh I don't know I felt sorry for the poor
14	totally employed down in the city I don't know but a lot of people I know
15	be about fifty per cent in arts I say I don't know about the other courses
16	ah yeah you see that book so I don't know maybe if you see the
17	is fantastic so whether that helped I don't know but then I noticed I
18	o'clock or a bit earlier even ah I don't know it's very dark these
26	don't know about living by myself I don't know yeah how good
27	feel about living in Melbourne oh I don't know I've um um I trai
28	it's not really a country town it's ah I don't know yeah you could say th
29	you call Albury a country town I don't know ah it's not really a
30	thought about post grad but um I don't know but you can do an
31	that and I just thought like aah I don't know whether you could write
32	them if they don't ring you yeah I don't know I think I'll er I'll hang o
33	you know what room is she in I don't know blue or I think it'
34	and I didn't get that much money I don't know* how people can raise

'I don't know' occurs within several distinct phraseologies which serve to disambiguate the various senses of the prefab. The colligations with *why, how* and *if* are the only expressions of ignorance. The colligations with *whether* (line 31) *or I* (lines 5,6,13,19,27,32, express uncertainty: 'I'm not sure...'

We found the LaTrobe speakers pulling off virtuoso performances: using limited resources they were inviting and giving empathy, making their own and their listeners' thoughts and feelings the underlying motif, although not the ostensible subject matter, of their interactions. We have set ourselves the question; to what extent is it possible and advisable to attempt to teach adult EFL beginners the same strategies? If we can find a way to transmute corpora to teaching materials this would pave the way for the development of more culturally appropriate locally usable materials.

Section 2

As mentioned in the introduction, the data were collected as part of a pilot study, funded by La Trobe University in Australia, using the perspective of language ecology (Garner, 2004). The aim was to discover communicative patterns underlying certain types of interaction. Apart from its potential contribution to social psychology and linguistics, the project aimed to contribute to the development of appropriate LT materials. The ability to create and use certain linguistic forms is only a part of the ecology of communicative competence. A dialogue must communicate to the satisfaction of the parties involved, which is inseparable from the interpersonal relationship and the cultural definition of the situation that is appropriate to the given exchange. We are interested in the relevance to language teaching of a fuller understanding of what constitutes culturally and situationally appropriate behaviour, along with the language in which it is manifested.

A major weakness of most so-called "communicative" approaches to language teaching is that they are based on an inadequate conception of communication. Linguistic elements remain their overwhelming preoccupation; they rarely consider the ecology of communication, even though it is at least as important as the linguistic elements, which interact with the environment as the participants jointly construct the meanings. The task of developing situationally and interpersonally relevant communication skills is left to the learner, at the point at which he or she engages in real-life interactions with native speakers.

For example, most language courses include a unit on "getting to know you", in which various language structures and vocabulary are presented explicitly, and various interactive strategies are presented implicitly. If students are taught how to introduce themselves by name and to ask for the interlocutor's name, along with the language patterns they are being taught that these patterns of communicative behaviour are appropriate within an initial conversation. But what are the grounds for assuming that they are? *When* do native speakers engage in exchanging names? Or, to start from the other end, how do speakers who are getting acquainted interact? What do they talk about, why, and how? Do they, for instance, introduce themselves by name, give information about their nationality, work, place of living, and so on, and ask for the same information from the interlocutor?

From an ecological perspective, the unit of analysis is not the language, but the situated interaction. The language is, to borrow a concept from complexity theory, an emergent property. The analysis itself is not primarily a search for theoretical knowledge, but a form of praxis or "practical theory", whose outcomes are focused on how to interact rather than on what to say. It is "concerned with the way embodied persons in a real world act together to create patterns of practice that constitute their forms of life" (Cronen (1995, p. 231)

It is not our purpose in this paper to present the current findings of the analysis; rather it is to illustrate the approach using one or two of the results to date. First, the most common communicative function of the interactions is what might be styled as "establishing common ground". The bare minimum for keeping a conversation going is, of course, something to talk about, and one feels that this is the primary reason for teaching language learners how to ask for and give information. Initiating conversations, however, are about much more than that. Participants want not just to talk about anything but to find experiences, knowledge, beliefs, values, and so on, that they share. For example:

OHT 1 here

Where information is given or asked for, it is not an end in itself, but part of the search for commonality.

OHTs 2 & 3 here

The capacity to introduce oneself by name does not seem to be particularly important. In only two of the interactions did a participant begin by introducing him- or herself by name; in one of those two, the other did not respond by giving her name. In fact, if one reflects on one's own experience, there are many situations of initiating contact—striking up a conversation on an aeroplane, for example—in which it seems to be quite inappropriate to exchange names. In others, names may be exchanged near or at the end of the interaction, only after common ground has been established.

Common ground is established in a wide variety of ways, including, for example:

- discovering experiences that both participants in some way can claim to have had, even if the similarities between them are apparently far-fetched;
- discovering people—individuals or types of individuals—that both participants can claim to know or know of;
- discovering which places both participants know
- creating 'stepping stones' i.e., the tenuous second-hand links in next paragraph

In a number of instances in these data, the commonality was second-hand, as it were: one participant knew someone whose experience was similar to that of the other participant. This was still regarded as an appropriate basis for sharing views.

As John indicated earlier, this raises the fascinating question of how much it is possible and desirable to use these insights as the basis for materials and methods. In what sense can practical theory inform language teaching practice and materials? Any given communicative function is performed by a wide variety of linguistic forms, and it would be unrealistic to expect that all or even a significant number of these forms could be taught to beginners. An alternative approach, however, is worth investigating: can learners be taught to enter into conversations with the intent of achieving a particular function? This would imply that the learner would learn to recognize, using the target culture's norms, certain types of interaction, and would develop an understanding of the processes of interaction appropriate to them.

It seems to us that ecological analysis of situated communication and corpus analysis of situated speech have the potential to inform and support each other, to span the gap between corpus and classroom, and to provide some answers to the challenging question raised by Kohonen, *et al*.:

> How could foreign language teaching be designed so that it promotes the development of the learner's holistic personal and intercultural competence? (p. 2)

References

1. Allén, S. 1991 In Svartvik, J. 1992 (ed.) *Directions in corpus linguistics: proceedings of Nobel symposium No. 82, Stockholm*, *4-8 August 1991* Berlin:Mouton de Gruyter.
2. Barbieri, F. 2003 'The 'new' quotatives in American English: A cross-register comparison.' Paper delivered at Corpus Linguistics 2003 Conference, Lancaster University*. University Centre for Computer Corpus Research on Language Technical Papers Volume 16-Special Edition*
3. Carter, R. 2003 In B. Seidlehofer *Controversies in Applied Linguistics.* Oxford: Oxford University Press.
4. Chomsky, N.1957 *Syntactic Structures*. Cambridge, Mass: MIT.
5. Chomsky, N. 1965 *Aspects of the theory of syntax*. Cambridge: Mass.: MIT.
6. Cronen, Vernon E. "Practical theory and the tasks ahead", in Leeds-Hurwitz (Ed.) (1995) pp. 217-242.
7. Duncan, S., & Fiske, DW 1977 *Face-to-face Interaction: Research, Methods, and Theory* Hillsdale, NJ: Erlbaum
8. Garner, Mark 2004 *Language: An Ecological View*. Oxford: Peter Lang.
9. Goffman, E. 1981 *Forms of Talk* Philadelphia: University of Pennsylvania Press.
10. Kohonen, Viljo, Jaatinen, Rita, Kaikkonen, Pauli, & Lehtovaara, Jorma (ed.) (2001) *Experiential Learning in Foreign Language Education* Harlow: Pearson.
11. Leeds-Hurwitz, Wendy "Introducing social approaches" in Leeds-Hurwitz, Wendy (Ed.) pp. 3-20.
12. Leeds-Hurwitz, Wendy (Ed.) (1995) *Social Approaches to Communication* New York: Guilford.
13. Lewis, M. 1993 *The Lexical Approach.* Hove: Language Teaching Publications.
14. Maley, A. and Duff, A. 1997 Drama techniques in language learning ***Cambridge***: Cambridge University Press.
15. McIntyre, Bellard-Thomson, Heywood, McEnery, Semino and Short 2003 'The construction of a corpus to investigate the presentation of speech, thought aand writing in written and spoken British English'. Paper delivered at Corpus Linguistics 2003 Conference, Lancaster University. *University Centre for Computer Corpus Research on Language Technical Papers Volume 16-Special Edition*.
16. Mindt, D. 1996 An empirical grammar of the English verb system. Berlin: Cornelsen Verlag.
17. Partington, A. 1998 *Patterns and meanings.* Amsterdam: Benjamins.
18. Shotter J. 1984 *Social Accountability and Selfhood* Oxford: Blackwell.
19. Sinclair, J.M. and Renouf, A1988 'A lexical syllabus for language teaching'. In R. Carter and M. McCarthy (eds.) *Vocabulary and Language Teaching.* London: Longman.
20. Stubbs M 2002 *Words and Phrases: Corpus Studies of Lexical Semantics.* Oxford, Blackwell.
21. Tognini-Bellini 2001 *Corpus linguistics at work.* Amsterdam: Benjamins.
22. Widdowson, H. G. 1972 'The teaching of English as communication. *English Language Teaching. 27*, 15-19.
23. Wray, A., 2001 *Formulaic Language and the Lexicon.* Cambridge: Cambridge University Press.