I. INTRODUCTION

Language features high on the list of daily human activities, so popular ideas about it inevitably develop, many of which are little short of “language myths” (Bauer & Trudgill 1998). As Crystal (2011: 3.6) points out, “an important part of the linguist’s job is to eliminate popular misconceptions” or “demythologize” linguistic concepts. The same applies to education, where every politician, parent and child has their own ideas about the ‘best method’ – especially perhaps in language teaching. Here even teachers, despite (or perhaps because of) their training and heard-earned practical experience, may be inclined to favour their intuitions about good practice rather than keep up with research. Teachers and researchers are notoriously bad at talking to each other (M. McCarthy 2008) – and who is to say that the expert researcher is better placed for this than the expert teacher (or vice versa)? Intuition may have been creatively productive in the 1970s, where “the so-called designer methods movement – Suggestopedia, the Silent Way, Community Language Learning, and so on… were ‘data free’, drawing sustenance from rhetoric rather than empirical support” (Nunan 2007: 9). But even today, many research papers in applied linguistics are essentially based on the (teacher-)researcher’s personal experience, subjective intuitions, selective observation and anecdote. Though they may describe innovative tools, techniques and practices and feature the traditional ‘literature review’, the body of many such papers is mainly reflective or descriptive with little if any rigorous data collection. Many current tenets of language teaching theory have come in for criticism for neglecting evidence or even ignoring it – communicative language teaching (Swan 1985; Decoo 2001), constructivism (Kirschner et al. 2006; Matthews 2003; Mayer 2004), task-based learning (Swan 2005), noticing (Truscott 1998; Robinson 1997), induction (Decoo 1996), the lexical approach (Lindstromberg 2003), and the use of authentic documents (Widdowson 2000: Cook 2001; Gilmore 2007), to name but a few.

II. CORPORA IN LANGUAGE LEARNING / TEACHING

This paper looks at the nature of evidence as applied to uses of corpora in language teaching and learning, often referred to as “data-driven learning” or DDL, a term first coined by Tim Johns (1990). In this approach, corpora are used not merely to inform the syllabus, materials or language input (M. McCarthy 2004), but themselves represent a language resource for the learners to explore via dedicated software or materials overtly derived from them (Boulton 2010a). For many, this is strongly rooted in the concordance, but computers can also help to
present the data in other ways that may be beneficial to language learners – frequency lists of word forms or lemmas, clusters and collocates, distribution plots over the corpus, keywords in different subcorpora or for a specialist corpus compared to a reference corpus, and so on. The ability to switch between this kind of quantitative view and the wider context is also increasingly being used to promote a more discourse-oriented approach to corpus use for language learning (e.g. Thorne et al. 2008), and interest is developing in multi-media corpora such as ELISA (http://www.uni-tuebingen.de/elisa/html/elisa_index.html), SACODEYL (http://www.um.es/sacodeyl/) and BACKBONE (http://134.2.2.16:8080/backbone-search/). While DDL may thus be difficult to pin down exactly (Boulton in press), the associated activities draw largely on a constructivist approach to inductive, problem-solving discovery learning from naturalistic pattern-recognition in authentic language data. As such, it might be supposed to promote cognitive and metacognitive skills, autonomy, life-long learning, learner-centeredness and individualisation, and so on; and indeed such claims are often found in the literature.

It is of course one thing to see the potential of corpora and associated software and techniques (or indeed any other approach or technology), quite another to demonstrate their effectiveness, let alone their efficiency. Evidence is essential for any innovation, but for anything related to corpus linguistics – that most empirical of fields in linguistics (Carter 2007) – its lack would be ironical in the extreme (cf. Boulton 2010b). Curiously, then, it has become a commonplace to lament the lack of empirical DDL studies (e.g. Ma 1993: 24; Aston 1998: 14; Gaskell & Cobb 2004: 302; Conrad 2005: 401-402; O’Keeffe et al. 2007: 24; Johansson 2009: 41; Chang 2010: 61). Certainly there seems to be no end to papers extolling the potential advantages (and occasional outright criticism, e.g. Dellar 2002); but what of scientific investigation? Without this, as P. McCarthy (2006: n.p.) puts it, “texts on corpus analysis that ignore quantitative evidence are doomed to endless instances of phrases such as ‘it seems to me’.”

With broad definitions of DDL (i.e. “language users exploiting language corpora”; Boulton in press: n.p.) and of empirical studies (i.e. subjecting “some aspect of DDL to observation or experimentation with some kind of externally validated evaluation other than the researchers’ own intuition”; Boulton 2010b: 130) it has been possible to collect over 90 empirical studies of DDL (see Boulton 2010b [supplement] for a complete list of references, overview and summaries). It is of course not possible to go into much detail here, but one angle is to group them according to their major objectives (most studies have more than one focus):

a) 54 studies of learners’ **attitudes**: on the whole, learners are generally receptive to the approach and perceive corpora favourably; data usually collected via questionnaires or other feedback, qualitative and quantitative;

b) 44 studies of learners’ **behaviour**: on the whole, learners work successfully with corpora, thinking and acting like corpus linguists according to the ‘learner as researcher’ metaphor; mostly a qualitative analysis based on observation and feedback;
c) 32 studies of learning outcomes: on the whole, learning seems to happen and compares favourably with control groups or other approaches – i.e. DDL is both effective and efficient (cf. Boulton 2010b: 139); mainly quantitative data obtained from tests, written samples or other productions;

d) 24 studies of using corpora as a reference resource: on the whole, learners are able to refer to corpora to help with their writing, error-correction or translation for specific purposes (e.g. research papers), though as with dictionaries and other tools, this says nothing about whether learning does or does not happen; data mainly based on written samples or other productions.

Though the overwhelming majority report largely favourable findings, it will be noted that all four categories above are prefaced by ‘on the whole’. This is partly because a very small number of researchers are clearly disappointed with the results overall (e.g. Estling Vannestål & Lindquist 2007). More importantly, and with very few exceptions (e.g. Koosha & Jafarpour 2006), most individual papers are mitigated in their conclusions – DDL does generally do what the researchers thought or hoped it would, but not always as well as they might have expected, or on all the points covered. There are a number of reactions one might have to this, the first being rather subjective: if you begin with a positive attitude to corpora, you are probably more likely to note the positive points; if you are sceptical, you are more likely to pick up on the negative aspects. But this doesn’t really take us much beyond the initial problems inherent in non-empirical evidence.

While it is certainly possible to find evidence to back up any such pre-existing position, it is more fruitful perhaps to focus on the bigger picture – not any one study in isolation, but the weight of evidence as a whole. One approach to this would be to conduct a meta-analysis, combining the results of different studies as is frequently the case in medicine, for example, but relatively rare in fields related to ours (e.g. Masgoret & Gardner 2003 for motivation in language learning; Means et al. 2010 for on-line learning). The reason this is not more widespread is not hard to see, and the empirical studies in DDL are a case in point: they are so diverse in terms of setting, participants, tools, materials, design, data gathering, aims, and so on that they are insufficiently comparable to be combined into a single meta-analysis. However, if such a meta-analysis were possible, it seems very probable that they would, together, produce highly significant results.

III. TOWARDS A COMPLEX SYSTEMS PARADIGM FOR DDL

A third reaction is simply to wonder what is going wrong. The researchers and authors in these studies are overwhelmingly corpus users and enthusiasts: if they can’t get significant results, who can? In fact, the question seems to provide part of the answer: the enthusiasts are those who cannot afford to be overly enthusiastic, and the various studies abound in hedging such as:

> Overall, given that the students were advanced and the items already partially known it is possible to conclude, albeit tentatively, that, given language items at the right level, DDL has an observable (though slight) positive effect on actual use. (Cresswell 2007: 280)
But actually there may be nothing special about DDL in this regard. The vast majority of empirical research in applied linguistics designed to investigate different methods and approaches, tools and techniques, are similarly hedged. And this is where it gets interesting. Larsen-Freeman and Cameron (2008: 243), for example, ask:

Who can say, for example, on the basis of a pre-test / post-test design that a particular experimental treatment works or does not work? If the results are non-significant, the effects of the treatment may not yet be manifest; if the results are significant, they may have followed from an experience prior to the pre-test.

De Bot (2008: 173) puts this into perspective with the example of a study designed to evaluate the impact of L1 use in class:

In a carefully designed quasi-experimental study, no differences among the conditions were found, which was of course disappointing for the researchers, but not really surprising: in an FL classroom many variables play a role in the acquisition of vocabulary, and the use of the L1 is likely to be secondary to many other factors that play a role in the process. In a way, it would have been surprising if such a single factor had explained differences in learning success. This is not a critique of the… study, which is well designed and carried out carefully. The null effect found merely supports the fallacy of focusing on a single explaining factor in a setting in which there are clearly many potentially relevant factors.

These authors are proponents of a new research paradigm in applied linguistics, commonly and interchangeably known as dynamic systems theory (DST) or complex adaptive systems. Though drawing on earlier foundations, the movement began in earnest in the 1990s with work by Larsen-Freeman (1997), attaining widespread interest with the book Complex Systems and Applied Linguistics (Larsen-Freeman & Cameron 2008) and a number of special issues of reputable journals: Applied Linguistics (Ellis & Larsen-Freeman, 2006); The Modern Language Journal (de Bot, 2008); Language Learning (Ellis & Larsen-Freeman, 2009).

In this paradigm, language (and, a fortiori, language learning) is “dynamic, complex, nonlinear, chaotic (at times), unpredictable, sensitive to initial conditions, open, self-organising, feedback sensitive, adaptive, characterised by strange attractors, which are fractal in shape” (Ellis & Larsen-Freeman 2006: 576). DST recognises the “unknowableness and interconnectedness of systems” (Larsen-Freeman & Cameron 2008: 232), including language whose emergent properties cannot be reduced to a simple series of rules (Beckner et al. 2009: 2). The same is true of language learning, hence the inherent difficulty of isolating variables or establishing direct cause-effect relationships. Language acquisition is essentially “exemplar based” (Ellis & Larsen-Freeman 2006: 565), a bottom-up process deriving from formal or informal exposure to the language where “every instance of language use changes an idiolect’s internal organization” (Beckner et al. 2009: 16). While this highlights the difficulties facing the researcher, it does tie in well with DDL which entails massive exposure to the language via the corpus, including through organized extracts in the form of concordances (cf. Gaskell & Cobb 2004: 304).

The new research framework goes a long way towards explaining the problems inherent in traditional research, but exactly what contribution it can make to overcoming those problems is admittedly “challenging” (Larsen Freeman 2006: 594). The large-scale mathematical models proposed by Beckner et al. (2009: 12) are likely to be nigh on impossible to implement in many cases, not least because researchers in applied linguistics simply do
not have the necessary skills and resources to implement them. Of more immediate relevance is the insight that qualitative and quantitative approaches need to be combined to obtain a more complete picture, though precisely how has yet to be determined. Many of the early empirical DDL studies tended to be small-scale and qualitative in nature (Chambers 2007): on the one hand, they provide useful insights to the processes / performance and not just the product / outcomes (Moreno Jaén 2010: 243) and allow a certain depth of understanding of various factors involved; on the other, they tend towards “excessive complexity that might discourage use through its impenetrability and unwieldiness” (Leakey 2011: 251) and which often makes it difficult to extrapolate to other cases. Quantitative research, by contrast, may be more generalizable as it irons out some individual differences; but that is also its disadvantage as it can result in “over-simplicity, [making] it a blunt and meaningless instrument” (Leakey 2011: 251). Variation between individuals is crucial and not just “noise” (Ellis & Larsen-Freeman 2006: 564), and over time, “it is difficult to imagine a dynamic systems study that does not have a prominent longitudinal aspect” (Dörnyei 2009: 242).

Partisans of a DST approach applied to language learning recognise these difficulties, but remain optimistic. For Dörnyei (2009: 245), “the absence of ready-made research models and templates is not an indication of the inadequacy of a dynamic approach but only of the transitional problems that are bound to accompany a major paradigm shift.” We should perhaps expect fewer studies that are exclusively quantitative in nature in their attempt to isolate variables and establish direct causal relationships, but that does not mean that a quantitative approach should be abandoned altogether. According to de Bot (2008: 169), “some of the proponents of DST are overstating their claims, and there is not yet enough substantial evidence to abandon traditional cognitive science in favour of a DST based approach.” Though he admits that it is an impossible endeavour to control all potential variables, it should be able to reduce them to something more manageable and relevant (p. 175). In the same vein:

Arguments for complexity are not arguments against simplicity. Some things can be understood by the analytic and reductionist program and where that program works it has done great service in elucidating causality. The problem is that it works where it works and it does not work everywhere. (Byrne 2002: 101-102, cited in Dörnyei 2009: 241)

Research to date is thus not without interest and the results should not be ignored: they provide precious insights to many aspects of language learning. Quantitative research is currently seen as more prestigious, at least insofar as it dominates empirical research in many prestigious journals as highlighted in two recent analyses (Benson et al. 2009; Richards 2009) – just as it does in the study of learning outcomes from DDL presented above. But as Byrne points out in the preceding quotation, the methodology is limited and constraining. Quantitative research by definition focuses on the quantifiable, with a corresponding emphasis on short-term outcomes (cf. Leakey 2011); in the case of DDL, it is notable that much of the research so far focuses on targets that are easy to measure in a highly controlled experimental environment – short-term learning outcomes in vocabulary and lexico-grammar, as well as error-correction and Likert-scale questionnaires of learner attitudes, etc. At the same time, there is a notable dearth of studies looking at the major advantages that are generally attributed to DDL as seen earlier – the long-term effects on learner autonomy, responsibility, life-long learning, constructivism,
cognitive and metacognitive development, language awareness, skills and communicative ability, the capacity to work with authentic data, and so on.

IV. CONCLUSION AND IMPLICATIONS

A complex dynamic systems perspective does not in itself imply that empirical research should be abandoned, though it may need tackling in new ways (Dörnyei 2011). As the DST paradigm evolves, we can expect new types of research, combining qualitative and quantitative into an integrated whole rather than being treated separately, which may in turn lead to new research questions and open up new areas to explore. It remains to be seen exactly how: to date, only Thorne et al. (2008) have worked in this direction in DDL. In the meantime, the body of empirical DDL research suggests there is good reason to believe that, thoughtfully applied with sensitivity to the local context, corpora can be useful and relevant to many learner populations and individuals with different needs and for different purposes. Current research provides a solid background to inform ideas of what is likely to work, but each individual study is inevitably specific in that it derives from a particular context, while the overview is inevitably generic and may not apply in every case. Contrary to popular belief, there can be no ‘best method’ (Prabhu 1990), and one size never fits all: no innovation will suit all learners perfectly for all language points on all occasions in every context, so it is a local decision as to whether a new technique is ‘sufficient’ to warrant further exploration for given purposes with given learners in given conditions. DDL researchers certainly make no claim it is a panacea (cf. Boulton 2009), or that it should be used to the exclusion of other techniques. Against this background, teachers / researchers who are interested in seeing what DDL (or indeed any other innovation) can bring to their own learners should not ignore the evidence available, but ultimately have but one option: to try it out.

References


http://linguistlist.org/issues/17/17-3033.html


