

Editorial - A research note on corpora and discourse: Points to ponder in research design

Mautner, Gerlinde

Published in:
Journal of Corpora and Discourse Studies

DOI:
[10.18573/jcads.32](https://doi.org/10.18573/jcads.32)

Published: 01/01/2019

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):
Mautner, G. (2019). Editorial - A research note on corpora and discourse: Points to ponder in research design. *Journal of Corpora and Discourse Studies*, 2, 1 - 13. <https://doi.org/10.18573/jcads.32>

GERLINDE MAUTNER

WIRTSCHAFTSUNIVERSITÄT WIEN

(VIENNA UNIVERSITY OF
ECONOMICS AND BUSINESS)

CITATION

Mautner, G. (2019). A research note on corpora and discourse: Points to ponder in research design. *Journal of Corpora and Discourse Studies*, 2:1–13

KEYWORDS

discourse; corpora; research methodology;
project design

DOI

10.18573/jcads.32

ISSUE DOI

10.18573/jcads.v2

A RESEARCH NOTE ON CORPORA AND DISCOURSE: POINTS TO PONDER IN RESEARCH DESIGN

ABSTRACT

Corpus-assisted discourse studies (CADS) has developed into a mature and popular approach to the study of text and talk. Yet researchers still need to face certain key methodological challenges every time they design a research project. These include: finding the right fit between the research question, data, method and theory; preventing analytical tools from dictating the research process; maintaining the distinction between quantitative and qualitative tools; ensuring that empirical claims are commensurate with the representativeness of the data; and avoiding the misinterpretation of findings. Although none of these challenges can be overcome entirely, significant benefits can be reaped from addressing them critically and self-reflectively.

CONTACT

Gerlinde Mautner, Institute for English Business Communication, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria.
gerlinde.mautner@wu.ac.at

ORCID

0000-0002-5529-3184

LICENSE

© The authors. Available under the terms of the CC-BY 4.0 license

A research note on corpora and discourse: Points to ponder in research design

Gerlinde Mautner

Wirtschaftsuniversität Wien

1. Corpora and discourse studies: The state of play

Like many such events, the launch of the *Journal of Corpora and Discourse Studies* one year ago was not just a fresh departure but also a sign of maturity and consolidation. It seems fair to say that twenty-odd years after it first appeared on the methodological fringe, corpus-assisted discourse studies, the field served by JCaDS, has gravitated towards the mainstream. Conversations preceding the journal's launch suggested a considerable variety of views on both mission and method. Yet most researchers would probably be happy to go along with Partington (2010, p. 88), who describes corpus-assisted discourse studies as 'the investigation and comparison of features of particular discourse types, integrating into the analysis, where appropriate, techniques and tools developed within corpus linguistics'. And, for a particular kind of project — typically located at the language and society interface and using large data sets — it is now arguably the 'go-to' approach. (At least among researchers with a background in linguistics: in large areas of the social sciences, computer assistance remains confined to content analysis and the use of software such as nVivo [QSR International, 2018]). What better moment, then, than this first anniversary to assess the state of play in corpus-assisted discourse studies?

Given what has been achieved, perhaps that should now be capitalized as Corpus-Assisted Discourse Studies, or acronymized as CADS. From other fields (such as CDA, Critical Discourse Analysis, and ELF, English as a Lingua Franca), we know that the moment when an 'approach' — usually code for a combination of theory, method, and mindset — morphs into a handy acronym also heralds the birth of a shared scholarly identity. For some, such an identity is simply a vehicle for teaming up with like-minded peers in order to muster increased momentum for research and to share ideas, data and tools. For others, sadly, a strong academic identity is twinned with an almost religious belief that their chosen method is the road to salvation for everyone. That said, as far as I can see (admittedly from a very biased vantage point), most CADS-oriented researchers have so far managed to avoid the worst excesses of academic tribalism. And long may it remain so.

However, such developments, important though it is to keep an eye on them, are not the focus of this paper. Instead, I want to take a step back from the fray of doing and assessing CADS research in order to address a few basic issues involved in research design. I make no claims of novelty here; in principle, all the points I discuss will sound familiar. In practice, however, basic questions of method are sometimes allowed to slip into the background — perhaps because they are considered *too* basic, and the answers taken for

granted. So, before identifying what I consider to be five key challenges for researchers at the project design stage, I will examine why, in broad terms, fixing on a methodology can be so problematic for researchers of discourse — however they may define it.

2. Getting started on ‘doing discourse’

Reflections on method are a cornerstone of research. They may not excite people’s passions in the same way that its content does, but what such reflections lack in immediate appeal they amply make up for in relevance. Of course, as an applied linguist (and perhaps particularly if you are of a critical persuasion), you initially embark on research projects because you care about what language does in (and to) society, not because you are intrigued by the methodological challenges involved. Yet these challenges have a troublesome habit of getting in the way and, if left unattended, can seriously interfere with the content of one’s research as well.

Decisions on method are thus invariably milestones on the research journey — whichever paradigm one works in, and no matter whether one is an experienced or novice researcher. Perhaps surprisingly, junior colleagues tend to be more outspoken in raising foundational questions about method. In doing so, they often flag up complex issues that their senior colleagues are also struggling with (but may be too self-confident to admit it). To me, it was something of an eye-opener, at a recent symposium, to hear a PhD student ask an established, widely published speaker who had been less than explicit about his discourse-analytic method, ‘but *how* did you *actually* do it?’ The answer the speaker gave was rather fuzzy, to say the least. And he seemed visibly surprised that the question even arose. It would appear that routine and a sense of security can be barriers to reflexivity and insight, whereas naiveté and insecurity can be powerful drivers of critical inquiry.

Although, as a field, CADS can reasonably be considered both successful and mature, it is still perfectly natural for CADS researchers to struggle with the methodology, and in particular with epistemological questions (though in the spirit of this paper, *engage with* rather than *struggle* might be the more appropriate verb). For example, what exactly is the nature of corpus-based evidence, and how does it differ from other text-based evidence? How can one identify meaningful links between features of texts and the context in which they were produced and received? Faced with the myriad linguistic phenomena that one might investigate, how does one pick the ones that are most interesting and relevant, but without prematurely dismissing features as irrelevant? And the simplest question of all is anything but: *Where does one start?* Indeed, the title I originally had in mind for the present paper was: ‘Here’s my corpus. And now what?’

To that question there is quite a range of possible answers. For, as we know, studying discourse is common practice now: not just in applied linguistics, but also in many social sciences influenced by the so-called ‘linguistic turn’, such as sociology, psychology and

management. In many different disciplines, then, we find researchers who ‘do discourse’. What exactly they mean by this can differ significantly, depending on which particular school of thought they are affiliated with. On the other hand, they generally share three convictions: that discourse and society are mutually constitutive; that meaning is ‘a product of social practices’ (Angermuller, Maingueneau and Wodak, 2014, p. 3); and that social phenomena can be usefully studied through the textual fall-out they generate. However, from these shared foundational assumptions, diverse traditions have branched out, and ‘doing discourse’ now means rather different things to different people. Gee and Handford’s 2012 handbook, for example, includes an impressive twelve different varieties of discourse analysis. Putnam and Fairhurst (2001), too, acknowledge the extent of diversity. The very fact that such handbooks and review articles exist points to the establishment of discourse studies as a field, however fuzzy its edges may be. Yet the heterogeneity can be rather bewildering.

As a result, newcomers to the field — and not only they — often find it hard to answer a number of fundamental questions. How can they move from passion for a cause to a plan for a project? How are they to assemble a methods toolkit that is fit for purpose and manageable? And how should they monitor their own analytical practice as they go along? In short, they may still not be entirely sure ‘how one actually does discourse’. The bravest among them — and I am lucky to have met some — will persistently ask basic questions about method, but even they tend to be apologetic about their lack of certainty. Unnecessarily so, I believe, for two reasons. First, whatever doubts they may have are usually shared by their elders and betters (who may or may not be brave enough to own up to their own insecurities). And second, the difficulties involved in ‘doing discourse’ are often due less to any deficiency in the researcher than to the opaque and multi-layered nature of the link between language and society. Social cause and linguistic effect — or linguistic cause and social effect — are fiendishly difficult to match up, and rarely can a single discovery procedure do the trick. Yet in their heart of hearts, that single procedure is precisely what many researchers would wish to find, despite protestations to the contrary.

It is not surprising, then, that junior researchers often feel let down. One such colleague recently remarked on ‘the tendency for research methods and related theories to be viewed as either completely right or completely wrong’. He added that ‘this helps create an impression of an impenetrable academic excellence’, which in his view ‘can have a potentially debilitating effect on the confidence of budding researchers’.¹ This unfortunate tendency, in my experience, is rarely played out in refereed publications, but it may be more prevalent in informal and quasi-tutorial settings, where personal power differentials can stifle genuine critique and debate more easily than in the course of double-blind peer review. Be that as it may, the thoughts I have gathered here are among other

1 Personal e-mail communication, 30 July 2017. The author prefers to remain anonymous.

things intended to bolster the confidence of junior researchers and reassure them that their trials and tribulations in developing research designs are in fact common currency.

Of course, as in any professional domain, it would be foolish to deny that some ideas are more firmly grounded than others, and that not all research designs yield equally pertinent results. Nor can there be any doubt that the researcher's experience plays a big part in determining whether the research findings are dull or exciting, and indeed whether the whole project can be considered a success or failure. So no: where research designs are concerned, 'anything goes' is not the answer either.

Before moving on, a quick word about what the following reflections will *not* cover — for reasons of focus as much as of space. For the most part, I will refrain from commenting on the mechanics of corpus building and mark-up, for example, or the transcription of spoken language. Although these issues can be very complex indeed, they represent a different kind of challenge from the one I outlined earlier. When researchers worry about 'where to start' and ask 'how does one *actually* do corpus-based discourse analysis', their insecurity is usually related less to mastering individual techniques than to finding ways of seeing the bigger picture. In some ways, learning how to use the tools of the trade is easier than developing a vision of what goals one can sensibly achieve with them.

3. Five key challenges in research design

When a research project is first conceived — whether within CADS or indeed any other framework — two basic questions arise. *What* is it you want to find out, and *how* are you going to go about it? In other words, the researcher needs to formulate research questions (rather than just identifying a 'topic'), and then choose the analytical tools most appropriate to the task. Important though these initial stages are, questions of method need to be addressed throughout the research process. In what follows, I will discuss five ongoing concerns which experience suggests have a particularly strong impact on the final outcome. As we shall see, all five challenges are partly a matter of hands-on empirical technique, and partly a question of theory, relating as they do to operational choices as much as to foundational assumptions about how language and social life are assumed to be related. The five are:

- (1) *Finding the right fit between the research question(s), data, method and the underlying theory.*
- (2) *Preventing analytical tools from dictating the research process.*
- (3) *Maintaining the distinction between quantitative and qualitative tools.*
- (4) *Ensuring that empirical claims are commensurate with the representativeness of the data.*
- (5) *Avoiding the misinterpretation of findings.*

It ought to be noted that unpacking the research process in this way is not the same thing as prescribing a rigid algorithm. Ultimately, each researcher needs to forge their own

path through their projects, and overly detailed instructions would probably be counter-productive, reducing rather than increasing quality. Although my five challenges correlate to a certain extent with stages in the research process (to take an obvious example, the choice of tools comes before interpretation), they are probably best seen as permeating it. In other words, these issues are less a matter of procedure, and more of research philosophy, principle and policy. I will now discuss each of them in greater detail.

3.1. *Finding the right fit between the research question(s), data, method and the underlying theory*

The relationship between these four elements can make or break a project. Unless they are all well matched, the project's feasibility will be jeopardised, or its findings will lack relevance, or both. Exciting data can yield boring results, or none at all, if they are tackled with unsuitable methods. Likewise, results often *become* exciting only when they are explained in relation to a robust theoretical framework. To take a simple example, using computer-supported methods of exploring the minutiae of spoken language, such as hesitation or discourse markers, will only be possible if those methods are applied to a fine-grained transcript in which the minutiae in question are actually recorded. And the project's findings will only break new ground if they are contextualised and interpreted on the basis of theoretical assumptions, among other things, about the rule-governed nature of spoken interaction.

The issue of fit between theory, research questions, data and method also relates to the very mundane, but absolutely fundamental question of how data are collected, processed and stored. Deceptively simple cut-and-paste jobs can result in messy corpora so full of typographical errors and other 'noise' that the quantitative results produced by corpus-linguistic software are too unreliable for comfort. Such questions of corpus quality appear to be particularly sensitive in interdisciplinary cooperation with researchers who may have collected texts with entirely different, not primarily linguistic, research questions in mind. To then take that collection of texts and turn it into a corpus fit for linguistic analysis can be a difficult task (with implications for timelines, manpower and funding).

Perhaps the most crucial match is that between methods and data. To put it in culinary terms: even the best ingredients will yield poor results unless they are cooked properly; the finest fillet steak will be ruined if you boil it. Avoiding such catastrophes may seem simple, but it can be surprisingly difficult. Perhaps ironically, if you are partial to the idea of making your project 'corpus-assisted', one of the biggest challenges is to determine when *not* to resort to the computer's assistance. Remember that the definition of CADS by Partington (2010) cited earlier contained the phrase 'where appropriate' — a brief but crucial qualifier that leads us conveniently into my second challenge.

3.2. Preventing analytical tools from dictating the research process

Put bluntly, the tail should never wag the dog. The temptation to allow that to happen seems particularly great when the tools in question are computer-based. However, in linguistics as much as in carpentry, tools not only solve problems (if we are lucky), but also influence what we identify as a problem in the first place. As Abraham Maslow (1966, p. 15) famously remarked, ‘I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail’. Over and above the general temptation caused by this so-called ‘law of instruments’, computer-based tools seem to hold a particular allure. They have an aura of efficiency and infallibility rarely associated with manual, qualitative methods. There is a danger, in other words, that we miss things merely because they cannot be captured easily by the software we have at our disposal. Word frequencies, collocations, keywords and n-grams, to name just a few standard features, all have their uses, but we should also remain alert to phenomena that are beyond their reach.

What is more, in spite of ready-made programs being commercially available, considerable time and effort is still required before a corpus-based discourse analysis can really take off. So, once researchers have a system up and running, they have a strong interest in making the most of it, rather than accepting its shortcomings and spending even more time and effort on an additional method. It is almost as if, once you have ‘bought into’ a particular approach, you need to remain faithful to it, come what may. Of course, I am far from the first to point out the ‘hammer and nail’ problem. In theory, most researchers acknowledge its significance and are quick to emphasize the need for complementary, mixed-methods approaches. In practice, however, the ‘affordances’ of the chosen software — what it can and cannot do — frequently do end up as key drivers behind research design, often at the expense of alternative perspectives.

There is a related problem. Unless you write your own corpus-analytic software — an approach strongly advocated by Laurence Anthony, for example (Anthony, 2013) — you are at the mercy of the tools that the original software developers have devised, including a raft of background computational operations that most discourse analysts are likely to find unfathomable. Yet even those who have both the talent and inclination to turn themselves into computer programmers may not be willing to muster the time and energy needed to do so. As a result, most discourse analysts have no choice but to accept whatever the software offers them. But that is all it should be: acceptance, not blind trust.

3.3. Maintaining the distinction between quantitative and qualitative methods

Some corpus-linguistic tools are quantitative in nature, such as frequency lists or the keyword technique, whereas others, such as KWIC concordances, facilitate qualitative analyses. Arguably it is the combination of both that yields the most robust results, particularly when qualitative discourse analysis is thrown into the ring as well. Hence the ap-

peal of mixed methods (Bednarek, 2009) and triangulation (Baker & Egbert, 2016; Taylor and Marchi, 2018). Yet, in this context, ‘mixing’ should not be misread as ‘blurring’. However useful the combination, it is essential that a clear distinction be maintained between quantitative and qualitative statements about the data. Thus, in presenting evidence gleaned from the corpus, it is advisable to avoid vague quantifiers such as *many*, *most*, or *predominantly*. In papers submitted for review, such quasi quantifiers can be irritatingly frequent. They are alright when backed up with figures, but not on their own. In essence, you either count or you don’t; there is no room for a fuzzy in-between.

There is yet another angle to the ‘quant-qual’ dilemma. In a sense, the decision to go for a mixed-methods approach is only the first step. Important though it is, it needs to be followed by careful consideration of how the two strands are to be sensibly combined. Otherwise, two sets of apparently unrelated results are simply placed side by side, with links between them asserted rather than demonstrated. Significantly, Partington’s definition of CADS does not talk about combining discourse studies with techniques and tools from corpus linguistics, but about ‘*integrating* [them] into the analysis’ (Partington, 2010, p. 88; my emphasis). And integration requires constant oscillation between quantitative and qualitative viewpoints, moving back and forth between computer-based discovery procedures and traditional, human hermeneutics.

3.4. *Ensuring that empirical claims are commensurate with the representativeness of the data*

Bringing corpus-linguistic methods on board is meant to put discourse studies on a sounder empirical footing. Instead of cherry-picking — focusing on a handful of particularly interesting texts which you know beforehand prove your point — you look at a larger, usually computerised, collection of texts, a corpus. However, in most cases, it will be impossible to include *all* the available texts in your corpus; so certain texts have to be selected, on the basis of stringent criteria or by random sampling, from the wider ‘universe’ of available texts. That said, it is often impossible to determine what precisely that comprehensive universe of texts comprises. If you are interested in, say, the discursive constructions of poverty, you would first have to narrow down the scope of your inquiry considerably — for example, to a particular time period, a particular medium, a country and/or an ethnic community and so on. Even after trimming your research design in this way, the chances are that the range of documents qualifying for inclusion in your corpus is still far too wide to be manageable, even if you enlist computer support. So another round of culls would be in order, further reducing the breadth of your inquiry, while at the same time sharpening its focus. Thus, although corpus-linguistic methods have vastly increased our options in managing data, we must still acknowledge that corpus building involves selection, and that the choice of selection criteria is a judgement call. Inevitably, therefore, a corpus is an artefact in its own right, not a perfect mirror of ‘language’ or ‘so-

cial reality' — which inevitably conditions the strength of any empirical claims we may base on its analysis.

Whether you use a CADS approach or traditional, qualitative methods without computer support, your claims about the data must be commensurate with how representative — in other words, how typical — they are of the wider universe of texts 'out there'. An analysis of a single text — such as a politician's speech about eradicating poverty, or a newspaper editorial — is just that: an analysis of one text. Obviously, the more representative your corpus is, the more legitimate it is to generalise, and the bolder your claims can become. Although most researchers are in principle aware of these limitations, it is not uncommon for written-up research to contain tell-tale slippage into unwarranted and over-confident generalisations.

So, in terms of corpora's representativeness, is bigger always better? Not necessarily. Representativeness is a malleable, highly contingent concept. If you study the annual reports of a company that has only existed for five years, then its five annual reports to date are all it takes to make the corpus 100% representative. If the company is 50 years old, then including only five annual reports in the corpus, one for each decade, appears to leave too much material unaccounted for.

There is thus no hard-and-fast rule about ideal corpus size. Like so many questions related to research design, the most appropriate answers are those that are sensitive to the relationship between theory, research questions, data and method discussed in Section 3.1 above. To illustrate the point, take a concept typically invoked in sampling: saturation. Essentially it means that you stop adding more data when it becomes obvious that more data will simply be more of the same. Yet, if your theory and method are inspired by the principles of corpus linguistics, then size does matter, because 'more of the same' is precisely what yields the most interesting and reliable results about word frequencies, set phrases and collocational patterns.

3.5. Avoiding the misinterpretation of findings

One area particularly subject to overambitious claims is that of supposed causal links between the macro level of social structures and the micro level of linguistic choices. Be they lexical or syntactic, such choices convey meaning only in conjunction with the specific co-text and context in which they are made. Frequent use of the passive voice, for example, is not necessarily about 'hiding agency'; infrequent use of 'I' does not necessarily point to the speaker's self-effacing nature, and so on. Such claims are the stock-in-trade of popular books on communication, yet they sometimes also creep into linguists' accounts. Hence a brief reminder that it is essential to avoid taking speculative and overly ambitious interpretative leaps from form to function or, more broadly still, from the micro level of language to the macro level of society.

That the two levels are connected, systematically and in a mutually constitutive way, is one of the axiomatic foundations of discourse analysis (and indeed of much of what has been happening in applied linguistics generally over the past half-century or so). The question is *how*. And pinpointing the precise nature of the link in a specific instance continues to be a challenge that one is well-advised to approach on tip-toe rather than with a confident stride. Why are speakers who share Demographic A more likely to use Linguistic Form X than speakers with Demographic B? Is it really A or B that ‘makes’ them use X more or less frequently, or is the difference perhaps due to a third, as yet unidentified, factor? The most frequent pitfall in this context is to mistake the co-occurrence or correlation of two phenomena as a sign of causality — what ancient rhetoric called the ‘*post-hoc ergo propter-hoc*’ fallacy.

4. No easy answers

I would argue that a CADS project whose design does not give careful thought to the five challenges discussed above is likely to flounder at some point. However, since it is impossible to anticipate every specific challenge that a project may pose, I do not claim that they are the only issues to be faced, and that if you have them sorted, everything will be alright. In other words, meeting these five challenges ought to be seen as a necessary, though not sufficient, condition for sound research design.

For those new to discourse studies, and grappling with problems in research design, that acknowledgment will perhaps merely compound what is already a major disappointment: the realisation that there is no canonical algorithm that texts can be ‘run through’ and which would guarantee worthwhile results. Instead, there are only loose assemblages of analytical categories that we know have worked in the past — the usual suspects, as many of you will be aware, are pronoun usage, the naming of social actors, nominalization, agent deletion and the apparatus of systemic-functional grammar (including notions such as transitivity, which captures the semantic differences between types of verb). There are no hard-and-fast rules about which of these to look at first, which ones to focus on and which ones to ignore, either temporarily or permanently. Equally, it is difficult to decide when and how to factor various aspects of context into the analysis.

The idea of following ‘protocols’ for dealing with context, as Leitch and Palmer (2010) propose, is definitely alluring when one feels adrift on the high seas of methodological choices. However, as the critique of their paper by Chouliaraki and Fairclough (2010) showed, protocols are not uncontroversial. The point that the latter set of authors made (as convincingly as the former made theirs) was that intuition had a big part to play, and that the very idea of a protocol, with its promise of certainty, was unduly constraining. Clearly, both duos have a point. Leitch and Palmer’s approach to dealing with context introduces a welcome note of rigour, while their respondents’ approach puts greater emphasis on the ‘more versatile and porous methodologies that make space for novel, inter-

disciplinary research designs in the field' (Chouliaraki & Fairclough, 2010, p. 1214). Not only do both positions have their merits, but it would also appear that the greatest mileage can be obtained from combining them: using methods that are as rigorous as necessary *and* as 'porous' as possible.

Debates such as these are worth having not only because they are fascinating in their own right, but also because in the long term they pave the way for both methodological and substantive progress in the study of discourse. However, to novice researchers looking for clarity and certainty in the short term, such controversies are not necessarily all that helpful, as their academic appeal in theory may be greater than their usefulness in practice. My own account here has not come down on the side of either 'protocol' or 'intuition'. Both have their place, I would argue, and each should be tempered with the other. For protocol without intuition can easily become an uninspired bureaucratic exercise, while intuition without protocol can quickly turn into an impressionistic mess.

Ultimately, the key methodological questions have to be addressed afresh for each piece of research, on the basis of past practice, of course, but with a healthy distance from it. What we see, once research gets published, are the answers but rarely the questions, doubts, false starts, dead ends, and the reams of perfectly accurate but meaningless results that most methods also produce. In a Q&A session after a talk, a young colleague recently complained that he would benefit a lot more from reading about what did *not* work than what did. Point taken.

5. The quest goes on

In the past two decades, CADS has undoubtedly made impressive progress, and is set to make many more in years to come. There is clearly ground for optimism, though to varying degrees depending on which elements of the approach we focus on. We can be very confident that corpus-linguistic techniques, including those useful for studying discourse, will become ever more refined. We can be reasonably confident that more of the necessary software will become available widely and affordably. And we can be cautiously optimistic that the type of interdisciplinary collaboration that is the spice of CADS work will become more frequent than it is at present.

Nonetheless there are questions to which it is unrealistic to expect answers. At the interface of discourse and society, the really big questions seem to fall into that category. For example, how *exactly* are the minutiae of linguistic detail linked to macro level social phenomena, and how systematic is that link? How much strategic leeway do speakers *actually* have when they choose one linguistic item over another, in a given language and in a given context? How can we be *really* certain that our written-up results say at least marginally more about the discourse data at hand than about the nature of our analytic tools and our own preconceived notions? Most projects manage to say something useful about at least one of these questions, but without settling them once and for all, however much

we may secretly long for that to be possible. In this and all other areas of human knowledge, a grandly conceived quest for the Philosophers' Stone would be misguided from the outset.

However, every single project, if inspiringly conceived, carefully executed and thoughtfully argued, does help us edge a little closer to finding answers to the big questions. And that, indeed, is all we can ever expect to do: poke a little hole here or there in the thick fence that surrounds the vast territory of social structures and relationships. That sobering insight can lead to very different decisions when it comes to planning research designs. Some researchers prefer to narrow down the scope of their inquiry and, metaphorically speaking, put their objects of study under a microscope. Others are more inclined to broaden the scope and view the bigger picture through a wide-angle lens. Again, ideally, the close-up and wide-angle views of discourse ought to be integrated for mutual benefit. The *Journal of Corpora and Discourse Studies* has the potential to become a home for both approaches, and encourage interaction between them. Watch this space.

Competing interests

The author has no competing interests to declare.

References

- Angermuller, J., Maingueneau, D., & Wodak, R. (2014). The Discourse Studies Reader: An introduction. In J. Angermuller, D. Maingueneau & R. Wodak (Eds.), *The Discourse Studies Reader. Main Currents in Theory and Analysis* (pp. 1-14). Amsterdam and Philadelphia: John Benjamins.
- Anthony, L. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, 30(2), 141-161. doi:10.17250/khisli.30.2.201308.001
- Baker, P. & Egbert, J. (Eds.). (2016). *Triangulating Methodological Approaches in Corpus-Linguistic Research*. New York and London: Routledge.
- Bednarek, M. (2009). *Corpora and discourse: A three-pronged approach to analyzing linguistic data*. Paper presented at the HCSNet Summerfest '08, Somerville, MA, USA: Cascadilla Proceedings Project. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.212.6312&rep=rep1&type=pdf>.
- Chouliaraki, L. & Fairclough, N. (2010). Critical discourse analysis in organizational studies: Towards an integrationist methodology. *Journal of Management Studies*, 47(6), 1213-1218. doi:10.1111/j.1467-6486.2009.00883.x

- Gee, J. P. & Handford, M. (Eds.). (2012). *The Routledge Handbook of Discourse Analysis*. London and New York: Routledge.
- Leitch, S. & Palmer, I. (2010). Analysing texts in context: Current practices and new protocols for Critical Discourse Studies in Organization Studies. *Journal of Management Studies*, 47(6), 1194-1212. doi:10.1111/j.1467-6486.2009.00884.x
- Maslow, A. (1966). *The Psychology of Science*. New York: Harper & Row.
- Partington, A. (2010). Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK newspapers: An overview of the project. *Corpora*, 5(2), 83-108. doi:10.3366/cor.2010.0101
- Putnam, L. L. & Fairhurst, G. T. (2001). Discourse analysis in organizations. Issues and concerns. In F. M. Jablin & L. L. Putnam (Eds.), *The New Handbook of Organizational Communication. Advances in Theory, Research, and Methods* (pp. 78-136). Thousand Oaks, London & New Delhi: Sage.
- QSR International. (2018). nVivo (Version 12) [Computer software]. Retrieved from <https://www.qsrinternational.com/nvivo/home>
- Taylor, C. & Marchi, A. (Eds.). (2018). *Corpus Approaches to Discourse. A Critical Review*. London and New York: Routledge.