An individual subjectivist critique of the use of corpus linguistics to inform pedagogical materials

Kendall Richards[1]
Edinburgh Napier University, UK

Nick Pilcher
Edinburgh Napier University, UK

Abstract

Corpus linguistics, or the gathering together of language into a body for analysis and development of materials, is claimed to be an assured, established method (or field) that valuably informs pedagogical materials and knowledge of language (e.g. Ädel 2010; Gardner & Nesi, 2013). The fundamental validity of corpus linguistics is rarely, if ever, critiqued. In this empirical paper we critically consider the foundations of corpus linguistics as being based on an abstract objectivist view of language. We critique this foundation through the lens of an individual subjectivist view of language. Our introduction outlines abstract objectivist and individual subjectivist views of language described by Voloshinov (1973). We then present what is claimed regarding corpus linguistics, and consider contemporary critiques of these claims . We then critique the foundations of corpus linguistics from an individual subjectivist view of language. We illustrate this critique by drawing on data from interviews and focus groups with content material lecturers and students in the subject areas of 'Business'; 'Nursing'; 'Design' and 'Computing'. These data question the fundamental assumption about how corpus linguistics operates: that what is counted is indeed countable. The data show how ostensibly similar words are understood in very different ways with very different underpinning psychological elements. We argue that corpus linguistics thus informs pedagogical materials with a merely passive understanding of the language. This view can only gain access to the inert crust of previous language, because it removes language from its individual subjective context. This context is fundamental to giving language the conscious and psychological elements that underpin its use. We argue the language should be taught through dialogue in this subject context and not removed from it.

Keywords: Corpus linguistics; Critique; Individual subjectivism.

Kendall Richards is a lecturer in the School of Computing at Edinburgh Napier University (UK) with a focus on addressing academic support, retention, progression and widening access. He works with a significant proportion of non-traditional, international, mature and direct entrants and his research interests are in education as social justice and language. He has presented globally and is published in the International Journal of Qualitative Studies in Education, Higher Education Research and

---
[1] Correspondence:
Kendall Richards, School of Computing, Edinburgh Napier University, Edinburgh, EH10 5DT, UK., Tel: (00)44 131 455 2659 email: K.Richards@napier.ac.uk
Nick Pilcher, The Business School, Edinburgh Napier University, Edinburgh, EH14 1DJ UK., Tel: (00)44 131 455 4731 email: N.Pilcher@napier.ac.uk

Development, Dialogic Pedagogy, and elsewhere such as in the book Researching Intercultural Learning (edited by Lixian Jin and Martin Cortazzi).

Nick Pilcher is a lecturer in the Business School at Edinburgh Napier University (UK). He is the programme leader for the MSc in Intercultural Business Communication and also helps students with writing in academic subjects. His research interests centre around education, language and qualitative research methods. He has published and contributed to work published in journals such as Qualitative Research, Psychology of Music, The Qualitative Report, Research in Transportation Business and Management, The International Journal of Shipping and Transport Logistics, and the Journal of Education and Work.

<p align="center">೪೦೪೬</p>

Introduction

> "… linguistics has always taken as its point of departure the finished monologic utterance – the ancient written monument considering it the ultimate realium. All its methods and categories were elaborated in its work on this kind of defunct, monologic utterance or, rather, on a series of such utterances constituting a corpus for linguistics by virtue of common language alone" (Voloshinov[2], 1973, p.73).

> "English is becoming an aggregate of vocabularies only loosely in connection with one another, which yet have many words in common, so that there is much danger of accidental ambiguity, and you have to bear firmly in mind the small clique for whom this author is writing" (Empson, 1930, p. 236).

In 1929, Voloshinov published *Marxism and the philosophy of language*. In it, he argued that two key trends can be identified in linguistic thought: individual subjectivist approaches to language, and abstract objectivist approaches to language. The latter is grounded in the idea that language is objective and can be detached for analysis and teaching, whereas the former is grounded in the idea that language is subjective and individual, living only within its specific context. Voloshinov outlines each of these approaches as having four key trends which are antithetical to each other. We quote these four trends of the abstract objectivist approach (Voloshinov, 1973, p. 57) and those of the individual subjectivist approach (ibid, p.48) and we place them side by side (Table 1) for the purposes of illustration of their antithetical nature. We also code them and use these codes to illustrate the relevance of the examples of literature and data we use throughout the article. Although we use these four trends in our coded examples throughout the article, in the table we briefly note their significance in terms of what they imply for the possibilities of language analysis.

We draw on both this literature and also empirical data from interviews and focus groups to argue that corpus linguistics is based upon, and justified by, an abstract objectivist view of language. The whole approach of corpus linguistics (see immediately below) is based upon this view as it collects together

---

[2] We use the name Voloshinov throughout but are aware that much debate surrounds the issue of whether the author of works ascribed to Voloshinov may in fact be Mikhail Bakhtin (Morris, 1994).

language from different contexts and deposits it into a bank which it uses for later analyses. It is based on the fundamental assumption that the counted words are in fact countable. However, here we raise the following questions: Does this view only gain access to the inert crust of previous language? Consequently, does this remove language from its individual subjective context which gives it the conscious and psychological elements underpinning its use?

Abstract objectivist approaches	Individual subjectivist approaches
1. *Language is a stable, immutable system of normatively identical linguistic forms which the individual consciousness finds ready-made and which is incontestable for that consciousness;* [AO1] i.e. words form a system that can be presented for a scientific study and they can be counted as having one or more stable meanings, even if they come from different places (and contexts).	1. *Language is activity, an unceasing process of creation (*energeia*) realized in individual speech acts;* [IS1] i.e. words are individual, unique to context, and cannot be removed for objective study.
2. *The laws of language are the specifically linguistic laws of connection between linguistic signs within a given, closed linguistic system. These laws are objective with respect to any subjective consciousness;* [AO2] i.e. individuals use words according to a closed objective system, thus words represent equal items.	2. *The laws of language creativity are the laws of individual psychology;* [IS2] i.e. words are used individually, their meaning is unique to each individual, and they do not represent equal items in a system.
3. *Specifically linguistic connections have nothing in common with ideological values (artistic, cognitive, or other). Language phenomena are not grounded in ideological motives. No connection of a kind natural and comprehensible to the consciousness or of an artistic kind obtains between the word and its meaning;* [AO3] i.e. language is neutral, objective and can be scientifically studied.	3. *Creativity of language is meaningful creativity, analogous to creative art;* [IS3] i.e. language is individual and subjective, and used in a unique creative way by individuals at the time of usage.
4. *Individual acts of speaking are, from the viewpoint of language, merely fortuitous refractions and variations or plain and simple distortions of normatively identical forms;* but precisely these acts of individual discourse explain the historical changeability of linguistic forms, a changeability in itself, from the standpoint of the language system, is irrational and senseless. *There is no connection, no sharing of motives, between the system of language and its history. They are alien to one another;* [AO4] i.e., individual variations are anomalies or errors in individual usage of the objective system, only important when embedded in diachronic variation over time.	4. *Language as a ready-made product (ergon), as a stable system (lexicon, grammar, phonetics), is, so to speak, the inert crust, the hardened lava of language creativity, of which linguistics makes an abstract construct in the interests of the practical teaching of language as a ready-made instrument;* [IS4] i.e. words cannot be counted and accorded importance by frequency, since words only represent a hardened crust of the live acts of speaking.

Table 1. Abstract objectivist and individual subjectivist approaches to language

Thus, from an abstract objectivist perspective, language is fixed, immutable and disconnected from consciousness, whereas from an individual subjectivist viewpoint, language is continually evolving, is individual and rooted in ideology and consciousness. Arguably, these two views of language are paradigmatically different (cf. Kuhn, 1970) in that they represent two opposite frames of seeing language. The individual subjective view also approaches language through its ontological (cf. Lafford, 2007) or *emic* and individual perspective, while the abstract objectivist view approaches language through its *etic* and structural characteristics. Such paradigmatically different views in turn govern how language is approached and taught. In other words, such views will underpin how certain fields or methods both view language and how it can be analysed and taught. From an individual subjectivist view of language, language must be taught through dialogue within its context of use, as the meaning and usage of the language is individual and underpinned by creative ideological and psychological elements. Conversely, from an abstract objectivist perspective of language, as the system of language is closed, stable and immutable, language can be taken from a range of different contexts and assembled for study of frequency and usage, and then packaged to be delivered in any context, such as the language classroom. This latter view is very much true of corpus linguistics, which assumes that what it counts is indeed countable.

Corpus linguistics and what it claims to do

Corpus linguistics has been described as the method (cf. McEnery & Hardie, 2012) or research field (Andor 2004) whereby a corpus of texts is gathered together for linguistic analysis. It has a number of major schools in Lancaster, Birmingham, Nottingham, and Arizona; and specific academic journals (e.g. *The International Journal of Corpus Linguistics* and *Corpus Linguistics and Linguistic Theory*). Regarding how it is defined, generally speaking, "we could reasonably define corpus linguistics as dealing with some set of machine-readable texts which is deemed an appropriate basis on which to study a specific set of research questions" (McEnery & Hardie, 2012, p.1). Such research questions can generally be divided into ones that ask about language diachronically [i.e. over a period of time], and about language synchronically [i.e. at the same point in time] (McEnery & Hardie 2012). With a suitable corpus, it is argued, "we would have at our disposal many thousands of words with which to conduct our study" (McEnery & Hardie, 2012, p. 27-8) and that with current tools and computer analysis, it is claimed that, "there is no good excuse for failing to test theoretical work against corpora" (Wasow, 2002, p.163). Corpora consist of texts, rather than words as it is argued that "experience of the human world is largely a textually mediated experience, and, to that extent, human beings live in a textually mediated world" (McEnery & Hardie, 2012, p.230). As a result of its reliance on texts rather than words, "corpus linguistics questions the position of the word as the core unit of language" (Teubert & Cermáková, 2007, p.49-50). Thus, corpus linguistics is clearly founded on the idea that language is stable and immutable [AO1] and can be removed from its original contexts for the purposes of objective study.

Although corpus linguistics gives prominence to texts rather than words, the key way that corpora answer questions is through their affordability to generate frequency counts of words, multi-word phrases or stylistic occurrences within these texts. This is believed to allow linguists to count the frequency of certain words and phrases in order to create lists of words, phrases, or grammatical items, to study or to inform pedagogical materials. Further, lines of text can be extracted from a corpus to form concordance lines, again for study or to inform pedagogical materials. Such approaches are used extensively in English language teaching (e.g. Campoy, Cubillo, Belles-Fortuno, & Gea-Valor, 2010) to create lists of words or lexical bundles for general (West, 1953), academic (e.g. Coxhead, 2000, 2010, McCarthy &

O'Dell, 2008, Gardner & Nesi, 2013) and specific (e.g. Lee & Swales, 2006, Hyland 2008; Cheng, 2010) purposes. Further, it is also used increasingly in other fields. For example, literary analysis is undertaken with corpus stylistics (Fischer-Starcke, 2010), and corpora are used extensively in applied linguistics (Hunston, 2002). Again, this could only be done guided by the fundamental abstract objectivist belief that language is comprised of a closed linguistic system [AO2] with the only connections being linguistic ones between the words and the text.

Many claims are made regarding what corpora, and the lists of frequently occurring words or phrases they generate, can do. It is claimed, for instance in terms of the meanings of words, that "if we want to find out what words, sentences and texts mean, we should opt for corpus linguistics" (Teubert & Cermáková, 2007, p.37) as only after gathering all usages of a word in texts can we see what its meaning is, and in fact, "there is nothing else that could tell us what this word means. And all of it is verbal communication" (ibid, p.38). Thus, there is little connection between the word and individual consciousness [AO3] and the word only represents an inert crust [contra IS4] which must be compiled in its numerous occurrences to fully understand its meaning.

Overall, it is generally accepted as *a priori* that corpus linguistics is correct, valuable, assured, and established (e.g. Chang & Kuo, 2011; Gilquin, Granger & Paquot, 2007; Hafner & Candlin, 2007; Walker, 2011; Gardner & Nesi, 2013). There does indeed need to be no dialogue on the matter. As Ädel (2010, p. 39-40) notes in the introduction to a chapter related to how corpus linguistics can help with the teaching of writing: "One thing which will not be discussed is why we might want to use corpora in teaching in the first place. Others have already presented very good reasons (e.g. Gavioli & Aston, 2001; Sinclair 2004; Yoon & Hirvela 2004; O'Sullivan & Chambers, 2006)." Further, as noted elsewhere by corpus linguistic researchers, "There are few areas of linguistics where there is no possible role for corpus methods …, it is safe to say that the place of corpus analysis in linguistics is assured" (McEnery & Hardie 2012, p 236-7), and that corpus linguistics, "has to be accepted as an outstanding research field of empirical importance" (Andor 2004, p.97, cf. Baker, 2013). Such claims can only be made if language is assumed immutable [AO1], and existing within a closed linguistic system [AO2] that has no connection with ideology [AO3] or history [AO4].

Nevertheless, despite its widespread acceptance, corpus linguistics has not been without its detractors or its critics. One key critic who is considered in recent corpus linguistics literature is Noam Chomsky. Chomsky's critique is levelled from a strong foundational basis of linguistics and he argues that, "corpus linguistics doesn't mean anything" (cited in Andor, 2004, p.97) from a language perspective as it is concerned not with how language works, but with how language is observed. It is, therefore, external remnants of language [cf. IS4] rather than illustrations of its inner workings (cited in Andor, 2004). Thus, it is not scientific. In response, corpus linguists have argued that "Chomsky's view is at best somewhat naïve and at worst deliberately misleading… just as observation of the universe through astronomy can help to prove the hypotheses of physicists such as Einstein, so observation of language through corpora can help linguists to understand language" (McEnery & Hardie, 2012). Indeed, it has even been argued that "corpus linguistics is concerned with language use in real contexts [and] therefore, it is often contrasted with Chomskyan linguistics, which emphasizes language competence and often involves made-up examples as the basis of its exploration of language" (Adolphs & Lin, 2013, p. 597; cf. McEnery & Wilson, 2001). In order for such observations to allow an understanding of language, the language being observed must be abstract, and must be part of a stable and immutable system. Thus, the language use takes precedence over the language structure, and this use is stable and has a fixed meaning [AO1, AO2]. Such an underpinning assumption would afford the argument that the 'real

contexts' (Adolphs & Lin, 2013) of the language can be used by corpus linguistics to remove language for analysis.

Another critique of corpus linguistics, despite some corpus linguists claiming it is concerned with 'language use in real contexts', is that the process of compiling texts and then extracting words from them 'decontextualises' the language in that it removes the language from its 'context' (e.g. Widdowson, 2000) and also removes it from an ethnography of communication (Hymes, 1964). Yet, from an abstract objectivist view of language, this *context* involves elements of a closed immutable system [AO1, AO2] which can be analysed by the linguist, such as communicative context, typical writer-reader roles, cultural values and intertextual knowledge (Ädel, 2010). From an abstract objectivist view of language, the *context* is social and cultural (Koester, 2010), or possibly linguistic, as illustrated by McEnery & Hardie's (2012, p.35) reference to "key word in context (KWIC) concordancing" showing the "context in use" (ibid) of the nominalizing suffix 'ness' by showing a number of lines with words ending in 'ness' situated amongst other words. Only in a closed system [AO2] removed from ideology [AO3] and containing normatively identical forms [AO1] can such a context reveal the inner workings of the language.

In addition to such lines providing a 'context in use' of words, context itself can be provided, according to corpus linguists, by reducing the size of the corpus, so the smaller corpora provide more specific data for analysis (Koester, 2010; Hyland, 2008). In such small corpora, where "the analyst is more often than not also the compiler, the analyst often has recourse to the original context, and the corpus can be marked up for these different contextual features, e.g. setting, text type, communicative purpose, etc" (Flowerdew, 2011, p.31, see also Flowerdew, 2005). Furthermore, context in the genre can be provided by the students if they have the genre knowledge of the subject being studied (cf. Lee & Swales, 2006). It has even been argued, that such ostensible 'de'contextualization can help with contextualization, as "students are stimulated to look further into the context, to expand the lines as far as necessary or to consult the source text" (Charles, 2007, p. 298). Thus, according to the claims of Charles (2007), the paradox exists that by 'de'contextualizing, contextualization is actually enhanced, in that students are encouraged to return to the original source of the language after being shown its salience through its removal. Thus, from the abstract objectivist point of view, the language can be studied, but only on the basis that the linguist can provide the context, and this can only be done where the language is believed to be analyzable. This in turn is only possible if language is believed to be disconnected from individual ideology [AO3] and part of a closed linguistic system [AO2] of normatively identical forms [AO1]. Only if such a view of language exists can it be removed for study and the linguist be able to analyze it, and students be able to learn about it in such a context. Such a view forms the foundation for all structuralist research, including corpus linguistics.

Not surprisingly then, in the field of corpus linguistics, even when it is acknowledged as a drawback, decontextualization, "should not cause us to abandon corpus-informed teaching" and that corpora work best as a "*complement*" (Ädel, 2010, p.28). What is more, it is often the case that corpus linguists wish to access a large body of corpora to reach wider coverage, for example, when developing the PHRASE list, Martinez and Schmitt (2015) note that "the full 100-million word BNC [British National Corpus] was deemed the best choice from among the publicly available large corpora for a number of reasons, including its size, diversity, and reputation" (Martinez & Schmitt, 2015, p. 452). Further, when compiling the Academic Word List, a total of 3.5 million running words were used (Coxhead, 2000) and the student textbook *Academic Vocabulary in Use* (McCarthy & O'Dell, 2008) was prepared "by identifying academic vocabulary in the Cambridge International Corpus of written and spoken English and the CANCODE corpus, as well as checking the Cambridge Learner Corpus for common learner errors" (Coxhead, 2010, p. 463). For such large corpora to be accepted as valid and useful, and for it to be

accepted that the words can be counted as being the same from such a wide range of contexts (i.e. countable), the words in them can only be assumed to be from a closed system [AO2] of normatively identical forms [AO1] removed from individual ideology [AO3].

Other critiques of corpus linguistics are also dealt with through recourse to an abstract objectivist view of language. For example, what is said to be one of the most serious critiques of corpus linguistics is, "the inevitable focus on surface forms in corpus work" (Ädel, 2010, p.48) [cf. IS4] which means that there is "the risk of focusing exclusively on the word and the phrase level when using computer-assisted methods" (ibid, p.49). This results in the challenge for corpus linguists that they need to connect these "surface forms (which are easy to search for by computer) to meaning (which tends to require human analysis) – whether lexical, collocational, pragmatic or discursive" (ibid, p.49). It is, however, argued that such connection can be provided by the linguist, who can help analyze hedging and modal verbs, for example, and that "the challenge for corpus work is to find mappings between functional categories (such as politeness, evaluation or metadiscourse), which are very important in writing, and surface forms" (ibid, p.50). Thus, according to corpus linguists, linguistic analysis can reveal such connections through knowledge about the language held by the linguist, for example, through the linguist's ability to identify homographs of words (Bambrook, 1996, cited in Flowerdew, 2011). Here again, such a connection can only be made by the linguist if language is viewed through an abstract objectivist lens, only if language is seen as a closed system [AO2] of normatively identical items [AO1] removed from individual ideology [AO3] can the linguist make such connections. If it was, in contrast, accepted that the usage and meaning of the language was creative [IS2], individual [IS1], and only represented the inert hardened crust of the language [IS4] then the linguist would be unable to analyse it isolated from the context in which it was used. Instead, it could only be understood and used in its unique individual context, and such seemingly similar words would in fact be completely different in specific contexts, and, therefore, could not be counted, as they would be individual and psychologically unique [IS3].

In addition to such critiques, there are also a number of qualifications made by linguists about corpora and their uses. Yet, as language is viewed as a closed system [AO2] of normatively identical forms [AO1] unconnected with individual consciousness [AO3], these critiques are framed within this abstract objectivist view. For example, one critique is that the indirect use of corpora, where they inform materials, is generally agreed to be of value, but the direct use of corpora where students do the analysis, is more controversial (Ädel, 2010). Thus, it is not the foundations of the approach that are wrong, merely how it should be used. Further, as McEnery and Hardie (2012, p.26) note, not only do many other tools to study language exist, but it should also be remembered that although "corpora allow us to observe language… they are not language itself." Therefore, we need to analyze the language, but we can do this through corpora. Consequently, the foundations of the approach are sound. In another qualification, computer corpora are noted by linguists to be "incomplete" in that they "say nothing about how many people have read or heard a text or utterance, or how many times…some phrases pass unnoticed precisely because of their frequency, others strike and stay in the mind, though they may occur only once. And because different individuals notice different things, such saliency can never be included in a corpus" (Cook, 1998, p. 59, cited in Flowerdew, 2011, p. 16). It is also possible that, "a word or expression may be possible, but not recorded in the corpus" (Flowerdew, 2011, p. 32). Yet, such qualifications are more problems with specific corpora rather than with the fundamental basis of corpus linguistics as a field and method. Thus, despite these qualifications, and the acknowledgement and recognition of critiques of corpus linguistics, as its foundation is grounded in an abstract objectivist paradigm of language, corpus linguistics gathers pace, expands, and continues unabated. It is assumed that what can be counted is indeed countable.

Corpus linguistics seen through an individual subjectivist lens

In an individualist subjective view of language, rather than being an immutable set of normatively identical linguistic forms [contra AO1] language is an unceasing process of creation realized in individual speech acts [IS1]. Fundamental to the meaning of language in such a view, and therefore how it is approached pedagogically, are dialogue and context. The importance of dialogue (Bakhtin 1981, 1986) means that language consists of a stream of unfinished utterances that is continually evolving and is never completed. Context is fundamentally important to language, as what context provides is a 'password' known only to the individual users of the language [cf. IS2]. The specific context shows the meaning of the language as part of a chain of utterances in a dialogue: "any utterance is a link in a complexly organized chain of other utterances" (Bakhtin, 1986, p. 69). In this context, a fundamental role is played by elements such as intonation in the underpinning creative nature of language [cf IS3]. This is illustrated by Voloshinov in '*Literature as ideological form*' (cited in Morris, 1994) with regard to how the intonation around the word 'Well' is of fundamental importance to its meaning. Such intonation is, critically, only known by the individuals using the word (cited in Morris, 1994), and used by these individuals to imbue the word with individual meaning unique to a specific context.

There are, therefore, non-linguistic elements that are only seen in the context the word is used in, and such contexts may differ so greatly that each word will have a different meaning in each context, which will be linked very much to individual psychology [IS2] and be at times highly creative [IS3]. As Empson wrote in 1930, "English is becoming an aggregate of vocabularies only loosely in connection with one another, which yet have many words in common, so that there is much danger of accidental ambiguity, and you have to bear firmly in mind the small clique for whom this author is writing" (p. 236). Such ambiguities could be ones of meaning, of grammar, or changes in meaning over a period of time, and others, a total of seven types (Empson, 1930). What this means is that the structure of words is indeed highly complex (Empson, 1951) and highly individual in nature [cf. IS1]. For Bakhtin (1986), the word is of fundamental importance to giving meaning to language, and each word has three owners: the addresser, the addressee, and nobody (1986). For Bakhtin, nobody owns the word when it is removed from its context, and it becomes neutralised, and similar to a dictionary definition (1986). If this is the case, then it can be expected that there will be elements within the context and dialogue of the word in usage that are critical to its meaning, and that cannot be accessed through an abstract objectivist approach, even though this is precisely what structural linguistics does. Such examples if they exist would have to be only examples and cannot be taken as being reified tangible occurrences to be repeated, rather, they would simply be excerpts glimpsed from another context. Once again, this is assuming that what can be counted is in fact countable. However, a key question from an individual subjectivist view that asks whether such words are countable by corpus linguistics is: **How will future usage of a word be known from studying its usage in the past?**

From an individual subjectivist perspective, language has an essentially individualised nature [IS1] only realizable in the specific context it is used. As Vygotsky wrote, "the meaning of a word represents such a close amalgam of thought and language that it is hard to tell whether it is a phenomenon of speech or a phenomenon of thought" (1962, p.120, [cf IS2]). Although a dictionary definition of a word is both neutral and neutralised, ostensibly giving common generic features for all to recognise, "the use of words in live speech communication is always individual and contextual in nature" (Bakhtin 1986, p.87). Thus, words are, from an individual subjectivist perspective, only the 'inert crust' [IS4] of language activity, a system of forms (Saussure's langue (1959)). Words are, as Borges wrote, "symbols that assume a shared memory" (1979, p. 33), they are highly ambiguous (Empson, 1930), complex (Empson, 1951), individual in nature, and inextricable from their context (Bakhtin, 1981, 1986; Voloshinov, 1973). Critically, they resist being 'fixed' or set in stone. Samuel Johnson, when reflecting on

the first dictionary he had written, nine years later commented on how he realized it was folly to attempt to fix the language and that in fact language was in a lively state of mutability (Mullan, 2010). As Fecho (2011) observes, "to expect that just because you and I are using the same term or phrase that we have a consensus understanding of its meanings is to deny that context and experience having anything to do with our understandings" (Fecho, 2011, p. 19). As Saussure notes, "all the mistakes in our terminology, all our incorrect ways of naming things that pertain to language, stem from the involuntary supposition that the linguistic phenomenon must have substance" (Saussure 1959, p. 89). Thus, from an individual subjectivist perspective, a key question to ask is: **How does word usage in context show its individual and subjective nature?**

From an individual subjective perspective of language, the linguistic point of view cannot gain access to these contextual details as "the linguistic point of view deals with abstract, bare words and their equally abstract components (phonetic, morphological, and so on). Therefore, the total import of discourse and its ideological value – the cognitive, political, aesthetic, or other – are inaccessible to it" (Voloshinov, cited in Morris, 1994, p.169). Such an understanding of the language can only be passive, not active (Voloshinov, 1973). From an individual subjectivist view of language therefore, what corpus linguistics is doing when it removes texts and words from their context to compile a list of ostensibly 'frequent' words is actually compiling not a frequency of words at all, rather, it is compiling a frequency of 'signs', what Saussure (1959) would call 'parole', what Chomsky (1988) would call 'E-language' and what Voloshinov (1973) would refer to as the inert crust of the language [cf IS4]. Thus, instead of representing equivalents, these signs represent decontextualized signifiers, which in turn are only evidence of past thoughts. In other words, what corpus linguistics claims to count, is, from this perspective, uncountable. In context, individuals will have a non-neutral, individualized definition of many words, one which is continually changing through dialogue and based in the context. Thus, from an individual subjectivist perspective, a key question to ask is: **Is word usage in context influenced by psychological and ideological elements to allow for individual expression in the specific context?**

Based on this critique of corpus linguistics from an individual subjectivist view of language, the following questions should be asked of corpus linguistics:

- **How will future usage of a word be known from studying its usage in the past?**
- **How does word usage in context show its individual and subjective nature?**
- **Is word usage in context influenced by psychological and ideological elements to allow for individual expression in the specific context?**

We now describe the source of our data and methodological approaches before we discuss these questions in our Results and Discussion section.

Methodology

Our ontological and methodological approach involved exploring language from a non-text-based, and deliberately simplistic, perspective that involved asking students and lecturers what was needed for understanding and success in the subject. This approach was simplistic because there were very few questions and we did not ask for any texts, we just let participants speak. Our rationale for this approach was to complement existing text-based studies in the area of student support. We therefore approached our collection of data from a perspective of stimulating dialogue around the subject based understandings and usages of the 'English' needed by students to succeed, and of assessment task words. In this way,

our methodological approach was qualitative, subjective, and dialogic, in that it explored the language through dialogue and context. Our dialogues were with students and lecturers in interviews (in 2012) and focus groups (in 2013) that took place in classrooms and offices and looked at the role of 'English' in the contexts of different subjects and perceptions of key assessment terms. This entailed asking participants about what speaking, listening, reading and writing students needed to do in English, and then gathering together lecturers in focus groups (average 90 minutes) to compare and contrast the results from these interviews. The focus groups discussed aspects we had identified such as the role of silence in Nursing, and the use of video assignments in Mandarin in Interactive Media Design. We interviewed 21 subject lecturers (average 30 minutes) in the subject areas of Nursing, Design, Computing, and Business. These subjects were chosen simply because we considered them to represent a broad range and also because we helped students studying them. Previously, in a related area, focus groups with students and lecturers from the UK and China had been set up (in 2010) where participants discussed their perceptions of key assessment terms such as 'critically evaluate' and 'analyse' (for average 90 minutes). All interviews were reflexive and active (Holstein & Gubrium, 1995) to allow for discussion. Focus groups (12 in total) contained a maximum of twelve participants and a minimum of three (Shamdasani & Stewart, 1990; Barbour, 2007). All procedures were approved by the relevant ethics committees and all data is anonymised (Christians, 2011). All interviews and focus groups were audio recorded digitally and then transcribed by the authors. These were usually in English but sometimes in Chinese. When conducted in English these were facilitated by the authors, and when in Chinese by native Chinese speaking colleagues. The Chinese focus groups were transcribed and translated manually by a professional translator using a 'skopos', or goal oriented approach (Vermeer, 2004) that aimed to achieve as readable a translation in the target language as the source language. To achieve this one of the authors discussed and compared the Chinese and translated transcripts with the translator.

Our approach to analysing the data was also dialogic in that we as researchers engaged in close dialogue with the interview and focus group transcripts. This dialogue involved two researchers discussing and continually revisiting the data to analyse it diffractively (Mazzei, 2014) using a constructivist grounded theory approach (Charmaz, 2011) of continual re-reading for emergent themes. This was in contrast to a more objectivist grounded theory approach that would have specified the themes to identify in advance of the analysis. Such diffractions allowed us to go beyond the surface of the words, to consider them in their context and to identify and isolate individual and psychological elements which underpinned the usage of the words and language (cf. Pilcher & Richards, 2016). We then took these results to focus groups with the lecturers we had interviewed and used the data as the basis of a further dialogue. The initial aim of this further dialogue was to generate words and phrases students would need to use in their subject context. Yet, what we found (as we further detail and discuss below) was that in fact, when removed from the subject context and placed in the decontextualized focus groups, lecturers were unable to provide us with any words. Instead, they merely agreed with the ideological and psychological elements we had isolated, such as the importance of the emotional in Nursing and of the visual in Design (see below).

Results and Discussion

An outline of our understanding and approach to these results and discussion

Before we respond to the questions as outlined above, using the data we have, and approaching it from an individual subjectivist position, we stress two points: firstly what we are *not* doing and what we believe the data do *not* show; and secondly what we *are* doing and what we believe the data *does* show.

In terms of the first, what we are not doing: we are not conducting a linguistic analysis, and we do not believe the data we have allow us to gain access to the meanings of the words and the meanings of language we discuss. If we were to argue this were possible, then we would have to do so on the basis of an abstract objectivist view of language. This, however, as we have argued above, cannot convey the psychological and ideological elements underpinning language usage that an individual subjectivist view of language shows to exist, as demonstrated through our data below. An abstract objectivist view can only convey a passive linguistic understanding of the inert crust of words, when they are removed from their context. Instead, and in terms of the second point, what we are doing, and what we believe the data do show, is the operation of individual and psychological aspects behind the language. These elements underpin its usage and underline the importance of its usage in context, as part of a dialogue, to help give the language life. We now respond to the questions we outlined above through recourse to our data. We present and discuss the results here together in one section as we want to discuss how the results relate to our arguments in response to each question during the narrative.

How will future usage of words be known from studying past usage of words?

For a student, or teacher, to use corpus linguistics concordance lines to study a word for future use, they would, from an individual subjectivist perspective, need to know which context the word was used in. However, they will not know which corpora to choose. This is because the corpora are detached from their original context, and a huge range of possible choices of corpora exist. For example, The Lextutor site (Lextutor, 2015) has 22 different corpora to choose from. So which should be chosen? Furthermore, how can the selected corpora reflect correct future usage of the word if it is removed from its original context? For example, if it is a Nursing student who wants to know how to use the word 'sick' – they would then choose (or perhaps be directed to) the 'Med corpus', but the variety in the usage of 'sick' in the sentences is so wide it would not help. They could undoubtedly see how it had been *previously* used, but not how it *should* be used in a future context, nor the importance of the individual nature of language.

As an illustration of changing usages from our data, in Mental Health Nursing, one lecturer we interviewed noted that their Job title had frequently changed over time: *"I am a 'Registered Nurse for Mental Handicap', the qualification now is a 'Registered Nurse in Learning Disabilities' and in the future it will be a 'Registered Nurse in Intellectual Disabilities'. Because 'Mental Handicap' is now [a] derogatory term and you've got other labels you know, like older people with disabilities may have been diagnosed at the time as 'Cretin' and 'Mongol'. These were medical labels that now evoke very different feelings. But it is in writing and is in older people's files."* Notably, to type in 'Mongol' and 'Cretin' to the BNC Med corpus generates no matches, but this does not mean these words are not important. They are very important, but they are not used anymore to refer to patients, so they are not in a corpus. Nurses therefore, know not to use these words, they know when to hold their tongue (see below) and of the importance of the role of silence (see below). Corpus linguistics could not relay this, nor could a linguistic analysis through a corpora reveal the power behind these words and their usage, simply because they are no longer used.

Similarly, in Design, certain words were 'banned' due to having been overused. In the words of one lecturer: *"When I first started working here… we used to ban words like nice, minimal; minimal you might think… that's minimalism which is an approach and term but when you end up with it being overused you get nothing out of that at all."* Here therefore, studying the past usage of the word through a corpora would not reveal its meaning nor how it would be used in the future. What a corpus may show was how frequently the 'word' was used, but not how its meaning was understood, nor how in fact, its high frequency actually is a negative aspect.

Our data also contained examples to show how language was not fixed and how it changed over time for individuals, thus making it impossible to determine future usage from past. In the words of one student from China about the assessment task words they were expected to use: *"Do you think our understandings of these words changed since we learnt them? I feel they must be changing naturally and gradually. You wouldn't notice it. The most obvious one is, for example "justify", our initial understanding was "prove", later we know it has more meanings, it is a gradual process."*

Thus, with regard to the question **'How will future usage of words be known from studying past usage of words?'** our data shows how language is underpinned, from an individual subjectivist position, by flux, by change, by its link in a chain of time-determined dialogue (Bakhtin, 1981, 1986). We could not, we stress again here, actually witness the usage of the language itself from our focus groups and interviews, rather, we could only witness the fact that it had underpinning individual psychological elements and meanings. We could not study this usage in context, and the individual subjectivist perspective underlines how the language could not be removed from this context. This underlines how language was not abstract and objectivist, and how language must be delivered and used in context for its meaning to be known. Thus, studying the past usage of words does not inform how they should be used in the present, or the future, simply because such study removes the words from their context, and places them in contexts that are different from ones of their future use. As corpus linguistics views language to be abstract and objectivist, it believes this can be done, yet, this fails to account for individual and psychological elements.

How does word usage in context show its individual and subjective nature?

From an individual subjectivist view of language, the specific contextual nature of words is fundamental to what they mean, what they convey, and how they are used. In our data there are many illustrations of this in terms of how specific contexts would influence how words were used. For example, regarding the word 'analyse', one student commented this would contextually relate to the task involved, that, *"when I am doing quantity surveying it's more analysing results, cost ranges or... facts and figures… rather than statements"*. For 'discuss' another student commented that, *"it all, all depends on the question, or what it actually is you've been asked to do"* and also for 'examine' that it "*depends on the question."* For a lecturer in engineering, to show 'analysis', students need to *"intercept what happens at naught, negative infinity, positive infinity… [put] error bars on it, at least that's a first, only a very small incremental step to analysis."* Further, to 'describe', this lecturer commented that, "*if I say 'describe' the soil, my lawyer sister would say well it's wet and muddy and grey or brown or maybe white if it's a China clay, but to actually get an engineer's 'description' of soil requires certain tests, requires certain calculations and you plot that on the chart and that gives you your description."*

The critical role of the subject context is highlighted by another lecturer for the term 'discuss': *"when you 'discuss' something… the student needs to be able to place it within its subject domain… and… include things like 'define' if we ask a student to 'discuss'… the role of ERP[3] systems integrating data in organizations, first of all I would expect them to 'define' an ERP system."* Similarly, the way that 'critical' operates in a Design context is *"being able to test the relationship between Design and culture and economics and politics and so on. Why is it that Design had the form it did in the 1950's?... I would expect them to look at post-war issues of austerity, relationships, international relationships and today I would expect global relationships… It's that ability to contextualise and see shifts in time. And offer interpretation."* Thus, for 'words' context is absolutely fundamental, and often, words cannot be bounded or isolated, but incorporate other words as part of a chain of utterances (Bakhtin, 1981). To remove these

---

[3] ERP = Enterprise Resource Planning

words from individual contexts would therefore be taking them away from this context. We argue that a linguistic analysis done under the assumption that language is abstract and objectivist cannot gain access to these contexts. Thus, corpus linguistics cannot gain access to how language operates individually underneath the inert hardened crust of the word.

What is more, lecturers said they developed understandings of words integrally with the context and the subject, and that a dictionary definition (a neutral definition for Bakhtin (1986)) of the term 'value' for instance would give only *"a very narrow interpretation of the term. We find that students from different parts of the world come to these concepts in different ways and I find myself developing an understanding of the vocabulary in the class. We do that through discussing the terms and what they mean, operationalising or actually using the term in a business case study. Putting them in context, putting them into operation. Carrying out a value analysis… working at establishing that 'value' is not just economic - there are different levels of service value and quality aspects as well."* Lecturers stressed that words were specific to the level and degree both of the year and specific module students were studying: *"don't forget… every single module is different… it's so specifically situational."* The importance of subject content knowledge was also noted by a student, that using the words without such knowledge was impossible: *"I'm told to 'critically evaluate' what I've learned so far and I haven't been able to transfer that because I haven't learnt anything so far."* Indeed, as another lecturer commented: *"I'm teaching students a lot of stuff. Some of that they can acquire but you know what? They really, really need to know their subject."*

In Nursing as well, terms required contextualisation beyond their neutral dictionary definitions. For example, with the terms 'attachment', 'vulnerability', 'resilience': *"there are clearly defined definitions and theories… so they may have had an understanding of what the term means but not necessarily the concept of the theory within the profession … they would come up with a sort of dictionary type definition… so attachment means to be attached… and then you say yes, but actually; then you introduce the theory and there are key concepts within that theory that then they will develop an understanding of… vulnerability and resilience…. they may have an understanding of… what's meant by vulnerability but there's a clearly defined definition around child development."*

One Design lecturer commented on how a student's experience of language usage in Design was 'completely different' to engineering, saying that the *"use of language in Art Design has been completely different to the experience they have in engineering... it's how you explain and discuss a project and how that relates to people and the scenario that's been created… so sometimes words change, not completely changed meaning, but change as a result of the perspective of their learning."*

The complexity and critical importance of context was also highlighted by a Nursing lecturer, describing their need to speak completely differently with different sets of parents, that *"if I was talking to parents in a deprived area I had to speak at their level… I couldn't put on my [posh] accent because… they wouldn't have listened to me… I had to [be]… somebody that they could understand and felt that understood them, but then… in… other areas, I would have more educated parents coming to speak to me…. so I would have to then adapt to that so it was quite a shift really you know in all aspects."* The importance of this context was also the case with Learning Disability (LD) Nursing: *"they're having to speak and construct effective relationships with a huge amount of people and… that language is different for different people so if you're speaking to a speech and language therapist… you'll be using Nursing terminology… to the parent you might be describing the same condition but in a very different way, it's not "Hydro-cephalis" it's "there's some fluid on the brain"… then to the person with the learning disability it'll be "you got a sore head"."*

When different subject lecturers came together in a focus group context, such differences could be even more marked, as the following dialogue between a Design lecturer and a Nursing lecturer regarding the word 'empathy' illustrates:

> Design – *I find it quite interesting you use the word 'empathy', I was just thinking the way empathy is probably thought about or linked to from subjects would be completely different as well… empathy within Design is usually with the idea of having some sort of resolution at the end of it, and yet… in especially Nursing it might be more to do with merely being willing to understand and listen…*

> Nursing – *Yeh, absolutely and I think within Nursing we would see 'empathy' as a professional quality and a skill, something that could be learnt… you have to employ that as part of your work and is very core and maybe again that's very different in other disciplines.*

Critically, this Nursing lecturer later emphasized the importance of context: *"that issue of context is really important isn't it, I'd agree absolutely that context changes all the time you've got to get the concept in context if you're gonna really understand what's going on and what the person is saying."*

All the above data support the individual and subjective view on the nature of language for the speakers. Again we stress that what we were being told was not so much how the words were used, rather it was evidence of the underpinning individual psychological and ideological elements. What we attempt to show here though, is that this data underlines how a linguistic analysis that is underpinned by an abstract objectivist view cannot gain access to these meanings. Thus, not only is it *not* possible for corpus linguistics to convey these meanings, but, that given it is a purely text-based method and approach it is also impossible for it to gain access to understandings of these individual elements. In contrast, considering language through an individual subjectivist lens, shows that it is only when used in their true context can active conscious meaning be given to the words. Pedagogically therefore, the teaching of language and words needs to be approached and practiced through dialogue (Bakhtin, 1981) in the individual subject context they will be used in (cf. Pilcher & Richards, 2016).

Is word usage in context influenced by psychological and ideological elements to allow for individual expression in the specific context?

From an individual subjectivist perspective, language is underpinned with non-verbalised psychological elements of individual consciousness that are expressed in para-linguistic elements such as intonation (Voloshinov, 1973). We found many examples of individual psychological and ideological elements (cf. Pilcher & Richards, 2016) underpinning the words and usage of language in specific subject contexts. Of Design, one lecturer noted that *"I think in Design… students sign up for a kind of non-verbal, non-written degree actually."* Further, a key non-verbalised element of both Business and also Computing was, unsurprisingly, mathematics. In Computing, one lecturer noted that *"you're talking about the theory of databases… about set theory… statistics and physics of games and so on now that is mathematical and that is something students would have to be familiar with to be able to… to wrap the local vocabulary around it."* Critically, this idea of *wrapping* the words of the subject around the mathematics of the subject was commented on by an accounting lecturer, who said that although not totally numerical, accounting's numeracy was tied in with the words: *"not every question's numerical… we aren't totally numeric driven… it probably only makes up 25% maybe… obviously the numbers can help the students… they can actually, sort of, then, sort of, tie their words into it".*

Another fundamentally important non-verbalised element to Computing and to Design was the 'visual'. One interactive media lecturer asked students to make a video in their original language and provide subtitles and explanations as *"we are not interested in whether your English is perfect. It's not what this module is about, it's about your ability to take media and tell stories through it."* This idea of the 'visual' was fundamental. One lecturer noted that *"it's very cinematic, very visual and that's where you go over the language. You can explain visually what utopia is and what dystopia is and then it allows you to start to talk about the culture, the politics of utopias and dystopias."* Another Design lecturer underlined the importance of the visual: *"we are teaching a visual subject, referencing lots of visual language, if you are referencing great cinema or literature."* This same lecturer, critically, intertwined the idea of the visual with the verbal, saying that, *"a visual English, yes. You've got the semiotics of that."*

In Nursing, an essential psychological and conscious element underpinning any verbal usage (or non-usage) of language was the emotional, the reflective, feeling focused 'total communication'. One Learning Disability (LD) lecturer described communication as *"about active listening… about the whole… body language and attending to people… people have to actively listen and interpret what the person's saying… LD nurses in particular, communication is 90% of our job… we have to be skilled in actually attending to people and interpreting information and again because a lot of our guys have communication problems… it's integrated you know in a kind of total communication."* In Mental Health Nursing, listening would involve the ability to almost listen to feelings: *"you know skilled listening is a real art… to listen not just to the words but to the feelings that are being communicated."*

Further, in Mental Health Nursing, the words were intertwined with emotion, and students would *"have to be able to deal with kind of abstract concepts… describe human emotion, kind of thinking processes and I guess you would also look for an ability to language that can describe yourself… so language that's got to do with reflection… self-awareness."* For Learning Disability (LD) nurses as well, the verbal and non-verbal were intertwined: *"it's the tone of voice… but it also, with our guys, the people we care for, we have to give as many clues as possible as to our meaning so it's not just the speech, it's the face, it's the proximity, it's the gestures you know… and it's getting that across to people that it's total communication, and speech is really only one part of it… that's hugely important… that's crucial there they have to learn to integrate all those elements."*

With regard to the actual power of such psychological and emotional elements in communication, the lecturer gave the following example: *"one of the guys I worked with was profoundly learning disabled. I love this man for years and years and initially, the first two years he would never give me eye contact. And then one day he gave me eye contact for about a minute and see the power in that, totally non-verbal nothing… nothing to do with English or anything like that it was a communication at a very basic level but a very powerful level as well. Incredible."*

Such elements were not just important for Mental Health Nursing or Learning Disability (LD) Nursing, they were also key for Nursing in general: *"your non-verbal skills, communication skills are almost as important as verbal communication skills in Nursing and that's the thing that we kind of concentrate on… before we let them out into practice. When they are on the ward especially, you know, visitors will be in and … watching… how they interact with each other if they're standing around… if they have got their hands in their pocket, all of these non-verbal, you know, ways of communicating are being observed by everyone… even the basic one like how they are dressed, how they portray themselves to the patient… that is a form of communication, you are actually communicating to the patient that I am a professional and I am going to be here to look after you and you are safe with me… For nurses it is not so much the language they use verbally it is their non-verbal communication."*

Also, for nurses, critical to the ability to communicate non-verbally, was the role of silence, of *not* using words, and that this too was very much linked with facial communications: *"in Nursing there are lots of instances when you have got… to hold your tongue. When working with sick kids for example you have got children coming in with no-accidental injuries… you suspect the parents have done it, everybody suspects the parents… but it's not up to us to be judgmental… it is not up to us to discuss and it is very difficult to show in your face that you are not disgusted by it or, you know, upset by it. You have just got to act as if everything is normal and that is quite difficult."*

Thus, the answer to the question **Is word usage in context influenced by psychological and ideological elements to allow for individual expression in the specific context?** is 'Yes'. When talking about the living context of the subject, what these lecturers said, reveals, through an individual subjectivist view of language, a range of psychological, non-verbalised, conscious elements that both underpinned and were integral to verbal language usage. An abstract objectivist view of language would not be able to access these elements, nor would a linguistic analysis be able to reveal them. An abstract objectivist analysis could only reveal a passive understanding of the inert crusts of the words (Voloshinov, 1973). Thus, using corpus linguistics to inform pedagogical praxis will not be able to access or convey these elements. Instead, only by placing and teaching these words in the individual context of the subject can these psychological and ideological elements be conveyed.

Conclusion

In this article, we have argued that using corpus linguistics to inform pedagogical materials for language use is based on an abstract objectivist view of language which assumes that what it is counting is countable. We argued that such a view is based on the foundation that language consists of a normatively identical closed linguistic system, one that is removed from any individual consciousness. Further, that such a view is based on the assumption that language can be taken out of its context, removed from its dialogue, and compiled into an objective body of texts from which lists of normatively identical words can be compiled for study and to inform pedagogical materials. Critically, such an approach and view is unable to convey or appreciate the fundamentally important psychological and ideological elements underneath the language that are key to pedagogical praxis. Instead, when viewed from an individual subjectivist perspective, language is inseparable from an individual consciousness: both its psychological and ideological elements. We presented data to illustrate these arguments. We showed that in the four subject areas of 'Nursing'; 'Business'; 'Computing'; and 'Design' the language used is inextricably linked with context. Further, we showed that removing words and language from the subject contexts takes them away from a number of key elements that give life and meaning to the language. We stressed that the data we had, rather than show the meanings of words *per se*, instead showed the following: the words and language these lecturers used were inextricably connected to the individual psychological aspects of dialogues, i.e. ideological and conscious elements that could be talked about but not necessarily absorbed or observed, and certainly not rendered visible through any purely linguistic analysis.

Instead, our data showed that words and language are intertwined with specific contexts, and consist of many subject specific elements such as the visual for Design discourses, the emotional for Nursing discourses, including the role of silence, and the mathematical for the Accounting discourses. As Saussure noted; "all the mistakes in our terminology, all our incorrect ways of naming things that pertain to language, stem from the involuntary supposition that the linguistic phenomenon must have substance" (Saussure 1959, p. 89). Yet, the use of corpus linguistics to inform pedagogical materials and praxis is

based on an abstract objectivist paradigm of language that does precisely this. It assumes that there must be substance to the linguistic phenomenon, that what it counts is indeed countable. However, by removing these 'words' from their contextually based usage and 'teaching' them outside the living context of the subject means that not only will we be teaching students neutralised passive meanings which will differ from those they will be taught and need to know for their subjects, we will be inevitably cutting the words and language from their living non-verbal, and sometimes silent, context. In other words, what counts in a living language is in fact not countable.

We highlight here that we have only looked at a few subjects: Nursing, Business, Computing and Design. We would imagine that in other subjects such as Physics, Chemistry, Music, Art, and Literary Criticism, there would be both specific contextual language usage and also specific contextual non-verbal communicative elements that a corpus linguistics approach again could not convey when used to inform pedagogical materials and practice. At the outset of this paper we presented two quotes from quite different fields, one from Language Philosophy (Voloshinov, 1973) and one from Literary Criticism (Empson, 1930). We would highlight here that despite being from different fields we believe both quotes clearly illustrate the underpinning individual subjectivism approach to language and how language should be studied in specific contexts. It is notable that both quotes are from works that were from a similar period of time, and that that period of time is over 80 years ago. In spite of that, since that time, corpus linguistics has been used to develop and inform pedagogical materials based on an abstract objectivist view of language, assuming that a passive linguistic analysis can reveal the true 'context' or 'meaning' of words. We hope the above has shown how this is not possible. Rather than use corpus linguistics to inform pedagogical materials to teach language, we should instead be teaching the language in the subject context and dialogue it is intended to be used in. Only through pedagogical materials designed to foster dialogue within the specific contexts of usage will students learn the language they need to know, and learn it in relation to its individual psychological and ideological elements.

References

Ädel, A. (2010). Using corpora to teach academic writing: Challenges for the direct approach. *Corpus-based approaches to ELT. London, UK: Continuum*, pp.39-55.

Adolphs, S and Lin, P., M., S., (2013) Corpus Linguistics. In J. Simpson, (Ed). *The Routledge handbook of applied linguistics*. (pp 597 – 610) London: Routledge.

Andor, J. (2004). The master and his performance: An interview with Noam Chomsky. *Intercultural Pragmatics*, *1*(1), 93-111.

Baker, P. (Ed.). (2012). *Contemporary corpus linguistics* (Vol. 16). Bloomsbury Publishing.

Bakhtin, M. M. (1981). The dialogic imagination: Four essays by MM Bakhtin (M. Holquist, Ed.; C. Emerson & M. Holquist, Trans.).

Bakhtin, M. M. (1986). Speech genres and other late essays. *Trans. Vern W. McGee. Ed. Caryl Emerson and Michael Holquist. Austin, TX: The University of Texas Press*.

Barbour, R. S. (2007). *Doing focus groups. London, UK: SAGE.*

Borges, J. L., Di, G. N. T., Reid, A., & Borges, J. L. (1979). The book of sand; translated [from the Spanish] by Norman Thomas de Giovanni; [and], The gold of the tigers: selected later poems; translated [from the Spanish] by Alastair Reid/[by] Jorge Luis Borges. London: Allen Lane.

Campoy, M. C., Cubillo, M. C. C., Belles-Fortuno, B., & Gea-Valor, M. L. (Eds.). (2010). *Corpus-based approaches to English language teaching*. A&C Black.

Chang, C. F., & Kuo, C. H. (2011). A corpus-based approach to online materials development for writing research articles. *English for Specific Purposes*, *30*(3), 222-234.

Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes*, *6*(4), 289-302.

Charmaz, K (2011) Grounded Theory Methods in Social Justice Research. In Denzin, N. K., & Lincoln, Y. S. (Eds). *The Sage handbook of qualitative research*. Thousand Oaks, Calif: Sage. Pp 359 – 380

Cheng, W (2010) Empowering professionals in training to learn the language of their profession in Campoy, M. C., Cubillo, M. C. C., Belles-Fortuno, B., & Gea-Valor, M. L. (Eds.). *Corpus-based approaches to English language teaching*. A&C Black. Pp 67 - 78

Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures* (Vol. 16). MIT press.

Christians, C. G. (2011). Ethics and politics in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 61 – 80) Thousand Oaks CA: Sage.

Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, *34*(2), 213-238.

Coxhead, A (2010) 'What can corpora tell us about English for Academic Purposes?' In O'Keeffe, A., & McCarthy, M. (Eds). *The Routledge handbook of corpus linguistics.* London: Routledge. Pp. 458 – 470

Empson, W. (1930). *Seven types of ambiguity*. London: Chatto and Windus.

Empson, W. (1951) *The Structure of Complex Words.* London: Chatto & Windus

Fecho, B. (2011). *Teaching for the students: Habits of heart, mind, and practice in the engaged classroom*. Teachers College Press.

Fischer-Starcke, B. (2010). *Corpus linguistics in literary analysis: Jane Austen and her contemporaries*. Bloomsbury Publishing.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for Specific Purposes*, *24*(3), 321-332.

Flowerdew, L. (2011). *Corpora and language education*. Palgrave Macmillan.

Gardner, S., & Nesi, H. (2013). A Classification of Genre Families in University Student Writing. *Applied Linguistics*, *34*(1).

Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, *6*(4), 319-335.

Hafner, C. A., & Candlin, C. N. (2007). Corpus tools as an affordance to learning in professional legal education. *Journal of English for Academic Purposes*, *6*(4), 303-318.

Holstein, J. A., & Gubrium, J. F. (1995). *The active interview*. Thousand Oaks, CA: Sage.

Hunston, Susan. *Corpora in Applied Linguistics*. Cambridge University Press, 2002.

Hymes, D.H. (1972) 'On communicative competence' in J. Pride and J. Holmes (Eds) *Sociolinguistics.* Harmondsworth: Penguin Books.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes*, *27*(1), 4-21.

Koester, A (2010) Building Small Specialised Corpora. In O'Keeffe, A., & McCarthy, M. (2010). *The Routledge handbook of corpus linguistics*. London: Routledge pp 66 - 79

Lafford, B. A. (2007). Second language acquisition reconceptualized? The impact of Firth and Wagner (1997). *The Modern Language Journal*, *91*(s1), 735-756.

Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, *25*(1), 56-75.

Lextutor (2015) Corpus Concordance English. http://www.lextutor.ca/conc/eng/ last accessed April 24th, 2015

Martinez, R and Schmitt, N (2015) 'Vocabulary' in Biber, D., & Reppen, R. (Ed.) (2015). *The Cambridge handbook of English corpus linguistics*. Cambrideg University Press. Pp 439 – 459

Mazzei, L.A. (2014). Beyond an easy sense: A diffractive analysis. Qualitative Inquiry, 20(6), 742–

746.

McCarthy, M., & O'Dell, F. (2008). *Academic vocabulary in use: 50 units of academic vocabulary reference and practice; self-study and classroom use*. Ernst Klett Sprachen.

McEnery, T. & Wilson, A (2001). *Corpus linguistics: An introduction*. Edinburgh, UK: Edinburgh University Press.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Morley, J., & Bayley, P. (Eds.). (2009). *Corpus-assisted discourse studies on the Iraq conflict: Wording the war*. New York, NY:Routledge.

Morris, P. (1994). The Bakhtin Reader. *London: Edward Arnold*.

Mullan, J. (2010, August 6). The folly of preserving English in aspic. G2 Magazine, Guardian, p. 3.

Pilcher, N., & Richards, K. (2016). The paradigmatic hearts of subjects which their" English" flows through. *Higher Education Research and Development*. DOI: 10.1080/07294360.2016.1138455

Romer, U (2010) 'Using general and Specialized corpora in ELT' In Campoy, M. C., Cubillo, M. C. C., Belles-Fortuno, B., & Gea-Valor, M. L. (Eds.). *Corpus-based approaches to English language teaching*. A&C Black. Pp18-35

Saussure, F. (1959). Course in general linguistics. New York, NY: Philosophical Library.

Shamdasani, P. N., & Stewart, D. W. (1990). Focus groups: Theory and practice. Applied social research methods series, Vol. 20. Newbury Park, CA: Sage.

Teubert, W., & Cermáková, A. (2007). *Corpus linguistics: A short introduction*. Continuum.

Vermeer HJ (2004) Skopos and commission in translational action. In: Venuti L (ed.) *The Transla-tion Studies Reader* (2nd Edition). London: Routledge, 221–232.

Vološinov, V. N., Matejka, L., & Titunik, I. R. (1973). *Marxism and the philosophy of language*. New York [usw.: Seminar Pr. (original 1929)

Vygotsky, L. S. (1962). Thought and language. *Trans. E. Hanfmann and G. Vakar. Cambridge: MIT Press*.

Walker, C. (2011). How a corpus-based study of the factors which influence collocation can help in the teaching of business English. *English for Specific Purposes*, *30*(2), 101-112.

Wasow, T. (2002, p.163). *Postverbal behavior*. Center for the Study of Language and Inf.

West, M. (1953). A general service list of English words. London: Longman, Green.

Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied linguistics*, *21*(1), 3-25.