

John Benjamins Publishing Company



This is a contribution from *Pragmatics and Society* 9:2
© 2018. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy on <https://benjamins.com/content/customers/rights>

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Does a corpus informed analysis provide any insights as to why Robert Phillipson's theory of Linguistic Imperialism is labelled by some as a conspiracy theory?

Sean Thornton
Toyo University

This paper uses the corpus tools DOTA and WordSmith to see if they can provide any indication as to why some label Phillipson's theory of Linguistic Imperialism as a conspiracy theory. The tools were applied to multiple corpora composed of texts drawn from: Phillipson's works, conspiracy theory books, and a control corpus of general academic papers. The quantitative data generated was subjected to a corpus informed qualitative analysis with the tools being applied to facilitate a corpus-assisted discourse study of Linguistic Imperialism.

Keywords: Corpus-Assisted Discourse Studies (CADS), conspiracy theory, Dogmatism Text Analysis (DOTA), Linguistic Imperialism, Robert Phillipson, WordSmith

Introduction

Upon first encountering both Linguistic Imperialism, and the accusation that it is a conspiracy theory, I was puzzled. At that time, my image of what constituted conspiracy theories was perhaps extreme, and to suggest that Phillipson's ideas were comparable to them seemed unconvincing. That the accusation exists made me wish to investigate it and attempt to find out why Linguistic Imperialism has been dubbed a conspiracy theory. However, I was hesitant to consume conspiracy material because it seemed challenging to do so objectively. For this reason I enlisted a corpus approach to the study with the aim of being guided by the lexical details of the ideologies rather than their arguments.

This paper begins with a literature review that introduces the use of corpora in the analysis of texts; Phillipson's theory of Linguistic Imperialism; an outline

of other academic responses to it; and of some of the works chosen for inclusion in the conspiracy theory corpus, and why their inclusion seemed appropriate. It then goes into detail regarding the methods used to create and analyse the corpus data with particular reference to the DOTA and WordSmith tools that were employed. The results are then presented and discussed, ahead of a conclusion, which is informed by the data gathered, rather than the snap judgements that instigated the research initially.

Literature review

Linguistic imperialism

The basic theory of Linguistic Imperialism is that dominant nations take overt and covert action to use their language as a tool to consolidate their dominance. The exact nature and manifestation of Linguistic Imperialism has been debated in applied linguistics over the last twenty years. Amongst the perspectives on this topic, Phillipson (1992) champions it, Spolsky (2004) accuses it of being a conspiracy theory, Davies (1996) questions whether it might be a hoax or a parody, while Canagarajah (2003) accepts it, and focuses on solutions.

Phillipson (1992) relates prior examinations of imperialism to linguistics where previously the focus had been on aspects such as: culture, economics, and politics. He defines English Linguistic Imperialism thus: “the dominance of English is asserted and maintained by the establishment and continuous reconstruction of structural and cultural inequalities between English and other languages.” (1992:47). He sees it as a modern form of cultural imperialism. He identifies the post World War II period as a time of rapid and dramatic neo-colonialism through English Language Teaching. The theory claims that the British and American governments cultivated the ELT profession, alongside media products, to create a gradual shift towards ideological imperialism in the same way there had previously been a shift from militant to economic imperialism (*ibid.*: 53). Phillipson is not anti-English, and implies that France was attempting much the same thing with French as the British and Americans are with English.

The theory admonishes the ELT profession in particular for its early history of simply trying to transplant UK and US educational standards and methods around the globe with no consideration for the cultural context or need for localization. Phillipson (1992:185) elaborates on this by identifying what he perceives as the tenets of the ELT business:

- English is best taught monolingually, and by a native speaker.
- The best results come from English being taught frequently, and from an early age.
- If other languages are used much, standards of English will drop.

He challenges these tenets as a series of fallacies, and suggests that they have more to do with the advantage they give to the export market for ELT teachers, than with pedagogical concerns. He pre-empts the possibility that some may accuse him of conspiratorial thinking by emphasising the role of hegemony and competitive policies, and rejecting the notion that there is a shadowy cabal of elites trying to rule the world through English (1992: 72–74).

Spolsky's book discusses in extensive detail language policy, and not only English language policy, but other languages also. He elaborates on the impact and implementation of language policy, which he identifies as existing on all levels from the individual making choices about which words to use or not use, up to the development of UN policies on linguistic rights. In the chapter of his book that deals with the spread of English in particular, there is a section (2004: 78–81) subtitled "Conspiracy theory", in which he characterises Phillipson's theory of Linguistic Imperialism as a conspiracy theory on the basis that it views the spread of English as orchestrated with the intention of domination through language policy. Spolsky concludes that Linguistic Imperialism does not seem a rational or empirical explanation for the spread of English. Instead, it is essentially an argument driven by the notion that because the English speaking countries (arguably) benefit the most from the spread of English, then they must have caused it (*ibid*, 79). He goes on to support his perspective by drawing contrast between the overt colonial policies of England and France. For about three centuries France engaged in extensive language management policies, both domestically and overseas, including attempting to have all education in their colonies taught in French. The British often ruled that primary education was better left to mother tongues and local languages, and had a lack of aggressively pursued language policies. Spolsky sees the primary driving forces of the spread of English as socio-economic and hegemonic.

Conspiracy theories

There has been a proliferation of conspiracy theories over the last two decades. The Internet has made it possible for anyone to champion any idea. This is perhaps what Knight is referring to when he writes, "Despite the pervasiveness of this culture of conspiracy, it is often difficult for scholars, students, and general readers to gain accurate and dispassionate Information" (Knight 2009: xi)

Davis, in Knight (2003), explains that the range of conspiracies includes both legitimate concerns that might actually be true, right through to ideas so extreme that it would be shocking if they turned out to be true at all. This is supported by Cooke (2009) who declares that while most of the ideas outlined in his book have no basis in fact, some do, or could.

Cooke's (2009) lays out his perspective on the basics of the conspiracy field. He identifies a number of tendencies that a piece of discourse could possess that would indicate it being conspiratorial: the need to attribute agency; a negative mind set: "You'll never hear about a conspiracy theory that says things are better than they seem." (Cooke 2009:5); effort spent on cover-ups and secrecy; an over-emphasis on symbols and/or coincidence; and a reliance on semantics and wordplay.

Wilson (1998) outlines the manner in which he believes that conspiracy theories arise. This could be summarised as a cycle of the public reacting badly to something that an organization does. This bad reaction is often countered by a combination of damage control – which can end up being perceived as a cover-up – and research into public opinion – which can end up being perceived as spying. This cycle repeats and escalates with increasing risk of paranoia and over-reaction on both sides, and conspiracy theories emerge from this zeitgeist. Wilson feels that conspiracy theories stem from the human tendency towards wanting to explain things, especially bad things, and the brain's tendency towards pattern recognition.

Wilson (1998) and Knight (2003) remark on the extent to which conspiracy thinking has penetrated contemporary culture. Wilson cites a survey indicating that 74% of Americans believe that the US government actively participates in conspiracies. Cooke does not necessarily agree with this claim. His dealings with people in the conspiracy field seem to give him the impression that they are still fringe people with views about politics and history that are hard for most people to accept.

Some distinction needs to be made between conspiracy theory in the sense used above and other uses. Byford (2015:20–21) defines the labelling of an idea as a 'conspiracy theory' as being distinct from the legal definition. A legal ruling may identify a crime as having a factual element of conspiracy wherein multiple perpetrators are found to have colluded criminally (Byford 2015:20). A 'conspiracy theory' extends beyond the more personal legal definition and the term "tends to be reserved for conspiracy-based explanations which deal with large scale, dramatic social and political events" (Byford 2015:21). In an academic context, such as in the case of Spolsky's dubbing of Linguistic Imperialism as a 'conspiracy theory', Chomsky (in Byford 2015:22–23) notes that the term is typically employed in attempts to marginalise and delegitimise an idea.

There now follows a summary of the authors included in the conspiracy theory corpora. These summaries are included to provide some context to attempt to illustrate where Spolsky, or others, are placing an idea when they posit that it is a conspiracy theory. There is no debate regarding the veracity of the ideas, it is the intention of this section to express them briefly and neutrally; as such even the more controversial or extreme ideas are written in a matter of fact way.

David Icke is the author of numerous books. His early focus was on self-realization, later shifting towards conspiracy theories. Most of his work is concerned with the global elite controlling the world and all humans in it. Initially, he emphasised a conspiracy theory that the ruling elites of the world are not human; instead, they are reptilian humanoids, descendants of the serpent from the Biblical Garden of Eden, who shape-shift into human form and infiltrate the ruling elites. He also expounds a less fantastical version of a global ruling elite focusing on the New World Order as a modern manifestation of the Illuminati and comprised of key business, political and aristocratic families. In both versions, the elites wish to maintain the subjugation of the human race for their own convenience. To this end, they use the media to engage in fear-mongering around issues such as terrorists, drugs, and paedophiles, so that the wider public will give up their rights and freedoms in exchange for protection. Members of this elite supply the drugs, fund the terrorists, and are the paedophiles.

David Irving was at one time a reputable World War II historian. At points, his publications were contentiously sympathetic towards the Nazi perspective on WWII. This has led to him being labelled a holocaust denier, and his prior legitimacy as a historian has eroded because of this. Holocaust denial includes such ideas as:

- The Nazis did not start the war with the specific agenda of the extermination of Jews.
- The total number of Jews killed by the Nazis was at most 250,000.
- There was not enough fuel in the Third Reich to have burned 6 million people in the gas chambers.
- Zionists have been exploiting European guilt over the Holocaust to empower their own agenda.

Eustace Mullins was a staff member in the US Congress. His theories are primarily concerned with institutional corruption in America. In particular, the nature of the financial and banking system, which he perceives as an illusionary monetary system, with the primary aim of pro-actively enslaving people through debt. He also accuses the medical industry of favouring profit by suppressing effective health care, and sometimes extending this into deliberately making people worse rather than better.

Antony Sutton was a British scholar in the field of economic history and with a particular interest in relationships between financial systems and the military. His conspiracy theories largely centre around the notion that American money finances the creation of enemies for America, in order to fuel the military-industrial complex.

Michael Tsarion's conspiracy theories begin with his version of the Atlantis myth, according to which Atlantis was inhabited by extra-terrestrial genetic engineers who created superior forms of human life. Atlantis, which was near Ireland, was wiped out by a great flood, caused by the destruction, 50,000 years ago, of a planet, which had orbited the Sun between the Earth and Mars. The remnants of that culture relocated to Egypt and were far more advanced and civilised than the rest of the world. Over the centuries, various elite (non-Atlantean) groups sought to consolidate their power by suppressing the ancient wisdom to mere mythology and fairy tales, to prevent people from tapping into their true power and potential. To this end, they also sought to eradicate people and groups (such as druids and shamans) who wielded this ancient power, and resisted subjugation.

Methodology

Corpus Linguistics involves the construction and analysis of banks of language data. These collections of texts can be subjected to quantitative analysis yielding empirical results about the patterns of language use employed within them. This approach permits the swift processing of a large amount of content. A corpus can be scrutinised in isolation or in relation to other corpora. One application of these techniques is outlined in Adolphs (2006:80–84) as the examination of ideologies.

Adolphs (2006) identifies a range of research possibilities, of which the most relevant to this study is discussed below as the relationship between language use and ideology. The prejudice or bias of a text can be examined in terms of lexical items, their semantic prosodies and their context (Adolphs 2006:10). This has the potential to provide insights into what constitutes a conspiracy theory. Impressions may be formed on the basis of single word frequencies (*ibid.*: 40), recurrent sequences (*ibid.*, 42), and by comparison to other texts on a keyword basis (*ibid.*: 44). Such analysis should help direct the research towards areas being further examined on a concordance level (*ibid.*: 56).

Adolphs (2006:83) does caution that: "The study of ideology with the use of corpora requires a very detailed rationale for the chosen texts, as their origin is likely to influence the outcome to a considerable degree." The manner of this scrutiny can be clarified with reference to ideas drawn from Baker (2009); Biber, Conrad, and Reppen (2006); and McKenny (2005). Biber, Conrad, and Reppen

(2006: 3, 8) caution that researchers using a corpus approach need to make a point of considering both the typical and the unusual, and resist the human tendency of letting an intuitional spotlight on the abnormal lead to a conclusion that doesn't make full use of the advantages corpora can bring.

Lexical frequency could highlight words shared across the corpora with a comparable frequency, as ideal targets for concordance level metaphoric consideration (Baker 2009:17). Frequency and ideology come together in Ertel's DOTA (Dogmatism Text Analysis). DOTA is a content analysis instrument that claims to gauge the degree of dogmatism vs. open-mindedness evident in a given item of discourse. It focuses on the relative frequency of words from the six lexeme categories that Ertel considered as clear indicators of either dogmatism or open-mindedness. Ertel (1985: 230; quoted in McKenny 2003: 95) identifies these six groups as listed in Table 1.

Table 1. Dogmatism vs. open-mindedness according to Ertel (1985: 230)

Category	Dogmatic examples	Open minded examples
Frequency	<i>always; forever</i>	<i>not always; rarely</i>
Quantity/Amount	<i>all; none; every;</i>	<i>few; not all; some</i>
Degree/Measure	<i>totally; absolutely</i>	<i>hardly; considerably</i>
Certainty	<i>surely; naturally</i>	<i>apparently; perhaps</i>
Exclusion/Inclusion	<i>only; merely</i>	<i>in addition; moreover</i>
Necessity/Possibility	<i>must; impossible; have to</i>	<i>need not; may; possible</i>

The full English version of the DOTA dictionary features 465 words, collocations or clusters, identified as significant to assessing dogmatism. An adapted version of DOTA outlined in McKenny (2005) was employed for this study.¹ Ertel (1985) and McKenny (2005) both discuss the relationship between dogmatism, exaggeration, and cognitive disposition; this provides an avenue by which insights can potentially be gleaned by examining the corpora used in the present study. DOTA applies this filter both to gauge dogmatism quantitatively and to identify a point of entry where qualitative discourse level analysis may begin. DOTA could be considered an earlier manifestation of the kind of approaches and tools that would now be thought of as Corpus-Assisted Discourse Studies (CADS). Like DOTA, CADS, as described in Partington (in Stefanowitsch & Gries 2007: 271) and Jensen (2014: 130), is suitable for comparing ideologies across corpora at a discourse level. The ideologies assessed are those of conspiracy theorists, as com-

1. A more extensive study could attend to the full range of DOTA terms, rather than filtering by shared frequency.

pared with Phillipson's theory of Linguistic Imperialism and alongside a control corpus of general academic works.

In the building of all the corpora suggestions from Kennedy (1998:70–75) and Biber, Conrad, and Reppen (2006:246–250) were followed, including that the Phillipson corpus, above all else, had to be representative of his writings about Linguistic Imperialism. For the conspiracy corpus and the control corpus, a balance was needed between representativeness and variety.

In the case of the Phillipson corpus, the criteria for inclusion were: that Phillipson authored the content, and that Linguistic Imperialism is the subject matter. Eleven such articles were found. Additionally, Phillipson agreed to make a complete copy of his 2009 book, *Linguistic Imperialism Continued*, available as the core of the corpus. The resulting corpus comprised approximately 229,000 words, of which 139,000 could be attributed directly to, *Linguistic Imperialism Continued* (2009).

There were two elements to the acquisition of the conspiracy theory corpora. Initially, Wilson (1998), Knight (2003), and Cooke (2009) were used to identify either authors, or ideas, that could be considered conspiracy theories. Such texts were then gathered; the yield was substantial: were it all included in one large conspiracy corpus, the data collected would have totalled an estimated ten million words. A filter was applied: to restrict the conspiracy corpora to only texts that had been published as books. This narrowed the text selection down to a manageable size of works from five authors. It may be of note, particularly in relation to Kennedy (1998:62–63), that all of the works in the conspiracy corpus were authored by British or American middle-aged Caucasian male authors. This does not represent a deliberate decision; rather, it reflects those demographics as representing the authors that typically not only have written within the genre, but also succeed in having books published.²

Four of Icke's works were selected as being the ones that were most concerned with conspiracy issues (as opposed to self-realization). In addition, Irving's more recent and controversial books were deemed appropriate choices for inclusion in a conspiracy corpus, while his more plainly historical books were not. Six books by Mullins, four by Sutton, and three by Tsarion comprised the remainder of the conspiracy corpus, which totalled approximately 2.6 million words, roughly eleven times the size of the Phillipson corpus. Kennedy (1998:73–74) cautions that while a corpus needs to be big enough to be representative, if it is too large it can become unwieldy, and that a million word corpus is usually more than enough. Earlier versions of the conspiracy corpus had been much larger, but with Kennedy's advice in

2. There are non-white and/or non-male conspiracy theory content producers, but the filters of 'book' and 'available for use in the study' resulted in the above demographic bias.

mind, the text selection was filtered and the size reduced while maintaining comparability.

One other corpus was made: a control corpus of academic articles and master's degree dissertations totalling around 252,000 words. The only filter that was applied was a check to ensure that their topics were not related to conspiracy theories, language policy, or imperialism.³ None of the corpora were filtered by year of publication, therefore the study was diachronic rather than synchronic (Kennedy 1998:71).

The aim of the present study was to compare the Phillipson corpus to the others, to see if conspiracy theory elements could be identified within his works in this way. There seemed to be two paths: one of looking at terms that are shared across all the authors and the other of looking at terms that have very close frequency matches between Phillipson's and other authors' works.

After the corpora were built, it became necessary to choose a starting point for the investigations. As Sinclair (2004:10) points out, even though we should allow the data to lead us to new observations "it is impossible to study patterned data without some theory". Lexical frequency lists were built using WordSmith, one list for each author, plus a general conspiracy list and a general academic writing one (the control corpus). Each DOTA word was then looked at in each of the corpora, and the raw frequency and percentile frequency were recorded. From there, the lexes that had exact percentile frequency matches between Phillipson and another author, were examined on a concordance level. Two word collocations and three word clusters were identified, and the sentences containing them were examined at the sentence level. The preliminary focus on collocations and clusters was on the basis that, as discussed by Mahlberg (in Baker 2009:57), recurrent instances of these can reveal aspects of consistency in the world presented by an author across a number of texts.

McKenny (2005) presents DOTA as a potential bridge between corpus analysis and discourse analysis by focusing the corpus analysis on lexemes that may be "devoid of content, although, of course not devoid of meaning" (ibid.). A less Firthian reason is that conspiracy theories by their nature and content seem to feature some very open-minded ideas expressed in very dogmatic ways, and dogmatic thinking may seem prone to exaggeration (McKenny 2005). Initially examining Linguistic Imperialism in terms of dogmatism and open-mindedness could prove an efficient way of identifying lexis that warrants further investigation.⁴

3. Due to confidentiality issues the control corpus is unavailable. In retrospect it would have been better to use a public access corpus, such as British National Corpus or Lextutor, as the control.

Results

WordSmith tools revealed that the following DOTA words were shared across all the corpora, and they thus seemed the best place to begin a comparison. According to Ertel (1981) and McKenny (2005), the A terms listed (in Table 2) are those that indicate the level of dogmatism in a text, while the B terms the level of open-mindedness displayed.

Table 2. A terms and B terms (Sources: Ertel (1981) and McKenny (2005))

	Icke	Irving	Mullins	Phillipson	Sutton	Tsarion	Consp	Control
A Terms	%	%	%	%	%	%	%	%
all	0.31	0.28	0.2	0.19	0.19	0.31	0.26	0.18
cannot	0.02	0.01	0.02	0.02	0.03	0.04	0.02	0.03
every	0.06	0.05	0.06	0.02	0.03	0.05	0.05	0.03
must	0.04	0.06	0.05	0.03	0.04	0.07	0.05	0.03
only	0.13	0.16	0.14	0.11	0.14	0.11	0.14	0.12
Total	0.56	0.56	0.47	0.37	0.43	0.58	0.52	0.39
	Icke	Irving	Mullins	Phillipson	Sutton	Tsarion	Consp	Control
B Terms	%	%	%	%	%	%	%	%
also	0.21	0.07	0.21	0.16	0.16	0.22	0.16	0.25
can	0.19	0.07	0.07	0.19	0.13	0.15	0.12	0.24
may	0.05	0.17	0.08	0.09	0.12	0.08	0.1	0.09
possible	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.03
several	0.01	0.03	0.02	0.04	0.04	0.02	0.02	0.02
some	0.14	0.07	0.1	0.12	0.1	0.11	0.1	0.16
Total	0.62	0.42	0.5	0.61	0.57	0.59	0.52	0.79

What this data seems to indicate, is that of the individual authors, Phillipson's works, according to the shared terms frequency, are the least dogmatic by no small margin, and the dogmatism levels are very close to those of the control corpus. These observations do not extend to the frequency of shared open-minded terms, wherein Phillipson ranks higher than average, relatively close to the control corpus, but closer still to most of the conspiracy theorists.

A limited number of DOTA terms were found to have exact percentile frequency matches between Phillipson and at least one other author. Among the dogmatic words, *all*, *cannot*, and *only* had exact matches. From the open-minded words *also*, *can*, *often*, *possible*, and *several* all had exact matches. This guided the subsequent concordance level investigation.

- Beyond DOTA, WMatrix had also been intended to be applied, but technical problems ultimately prevented this.

only

The collocation *it is only* was found in equal measure in both the Phillipson and Tsarion corpora. In the majority of occurrences in Tsarion, the phrase is actually part of quotations from other sources that seem to support his stance. Conversely, in Phillipson, *it is only* is sometimes used to reject the ideas of other scholars. Each author happens to have two instances of using *it is only* to express their own ideas: “There is competition among the most dominant ‘international’ languages, though currently it is only English which is a constituent part of a neoliberal empire.” (Phillipson 2008: 20); “Fundamentally, it is only our own basic thoughts that possess truth and life, for only these do we really understand through and through.” (Tsarion 2010: 54). Phillipson seems to use the term to speculate on perceived evidence, whereas Tsarion uses it for introspection. As with the support-refute dynamic noted above, this has the texts going in distinctly different directions in their use of *it is only*.

can only be is often used by Phillipson in quotes by other scholars to support his stance, and also to express his perspective, particularly around the issue of linguistic equality. For Tsarion, *can only be* is predominantly used in reference to matters of faith and the dismissal of the need for evidence in preference for intuition.

The final *only* cluster to be considered is *is not only the*. Phillipson employs this phrase to emphasise the intentions he perceives organisations to have: “Consulting the website of a successful and possibly representative British university with campuses in Malaysia and China gives the clear impression that what is being exported is not only the British English medium but also British content.” (Phillipson 2008: 10). In all but one case, it appears to be used in a critical fashion to condemn agendas that he does not support. Tsarion on the other hand uses it once in that fashion, but in all other instances, he uses it in a positive manner.

can

Phillipson and Icke shared the same overall percentile frequency of *can*. The first cluster in question is *which can be*. Phillipson’s instances of *which can be* manifest as statements that draw attention to something that he wishes the reader to consider. He uses this term to imply collusion between the leaders of the US and the UK in the exercise of Linguistic Imperialism; for him, *which can be* is a hedging device. Icke uses *which can be* a number of times to denote ability, draw attention to observations, and lay out the genealogies of the elites. In this, there is some overlap: “The American dog has wagged the British tail since 1945 but with a ‘special relationship’ which can be traced through Churchill-Roosevelt Thatcher-Rea-

gan and Blair-Bush II.” (Phillipson 2011:10); “you will find they come from the same bloodlines which can be charted back to the British and European royal and aristocratic (reptile-Aryan) families” (Icke 1999: 187). Both authors use *which can be* to draw attention to the relationships between the perceived agents of their respective theories. Phillipson uses it more as a casual observation, Icke more as if he is presenting evidence.

This trend continues with the collocation *can be traced*, which Phillipson uses when giving accounts of the history of Linguistic Imperialism's role in colonialism. Icke is also concerned with tracing the history of the agendas he perceives over the course of the last few thousand years. For Phillipson, what *can be traced* is no small matter, and points towards Linguistic Imperialism having been brewing long before it became apparent. For Icke, tracing feels central to his conspiracy theories, it is the strings that lead to the puppet-masters.

Icke uses *that can be* numerous times to identify the methods by which the elites' agendas are carried out, often identifying things *that can be* done by 'them' to 'us'. His other recurrent use of *that can be* is to directly address his audience with clarifications or rhetorical questions. Phillipson often applies *that can be* in reference to the manner in which language policy manifests in education. While both use *that can be* to point out the chain of human decisions, Phillipson's use seems more observational than conspiratorial.

The final *can* attended to is *can be seen*. Phillipson frequently uses *can be seen* in reference to how hegemony manifests in any given place, particularly Europe. He uses it to draw attention to what is apparent, but perhaps not obvious. When he is talking about universities, he uses *can be seen* in a more critical manner: “The results of this can be seen in the way universities are being turned into mass production machines that do not generate critical thinking.” (Phillipson 2009: 22). His implication seems to be that something *can be seen* but is deliberately ignored – part of the conspiracy of silence about which he has written. Phillipson's *can be seen* is about results and behaviours. Icke's *can be seen* is in many cases much more literal and is a direct reference to what you will see if you go and look at a certain building or location. He does occasionally use *can be seen* to explain behaviours or policies: “It can be seen in their articulated understanding that those in authority are pedalling [sic] lies to sell a secret agenda to the people.” (Icke 2003: 285). In the instances where their underlying use is similar, there is at the same time a stark difference in the degree of the assertion – authorities actively telling lies, verses universities not actively cultivating people who will challenge the status quo.

cannot

One recurrent collocation of *cannot* was found in Phillipson's and Mullins' *cannot be*. Phillipson's use of *cannot be* is spread across a number of topics and concerns and where *cannot be* occurs, it is either in reference to things which he is certain about, or in statements he wishes to resolutely stand by. *cannot be* seems to mark some of the core issues that Phillipson wishes to tackle and attend to, and among these the things he really wants his readers to think about, such as:

- Language policy (or a lack of it) is a factor in domination/inequality.
- Europe should not give up and let English dominate it, or even the world.
- Language and culture go hand-in-hand, English is never truly a lingua franca.
- While English is not a panacea, it still is important.
- Linguistic Imperialism is not a conspiracy theory.

He expresses these opinions both directly and through citation, often within conclusions. A number of Mullins' instances of *cannot be* are practical and factual statements about what the actual rules are within the systems he discusses. Yet many other examples are much more ideological in nature with multiple declarations about what humans must not fail to do if disaster is to be averted. Like in the case of Phillipson, *cannot be* is employed here to assert key concerns of (his) ideology in a firm manner.

several

Several had shared frequency between Phillipson and Sutton. The collocation *by several* is used in Phillipson to indicate an amount that is either not specified, or does not need to be specified. The same is true for Sutton. Both use *by several* for vague numerical purposes. Phillipson's use covers the doings of scholars, the presence of errors, and the amount of contributing factors. Sutton's *by severals* are in relation to numbers of people or groups involved in a situation.

The other shared collocation, *several decades*, finds Phillipson placing his ideas in a historical context, sometimes in a merely observational way: "As a result of several decades of this policy, over half of Singaporeans now use English as the language of the home." (Phillipson 2016: 4). At other times, *several decades* seems part of a more sinister reading of history. He makes mention of the long-term struggle between French and English Linguistic Imperialism. He also implies a *several decades* long relationship between the CIA and Linguistic Imperialism, the tone of which suggests that cultivating Linguistic Imperialism would present little logistical or ethical problem. Sutton's use has a more pronounced sense of doom: "Over the past several decades, quietly, without media attention, many Ameri-

cans in diverse fields of activity have been pressured into silence” (Sutton 1986: 15). Some of the overlap here seems more likely to relate to the word *decades* than to the word *several*.

possible

Possible was found in shared frequency in the works of Phillipson and Irving. The first collocation to be looked at is *as possible*, which Phillipson uses when talking about the desires and intentions of other parties within his discussions. Irving often uses *as possible* in a spatial and temporal manner with reference to distances, time-frames, and practical limits. A frequent collocation in Irving is *as soon as possible*, often used in quotations where Nazi officers are giving written orders to their units. At no point does Irving use *as possible* in the same manner as Phillipson.

In the case of *possible to* Phillipson uses it to talk about what is or is not likely or realistic, particularly regarding multilingualism. Though Irving's use of *possible to* includes a few equivalent cases, it also extends into numerous other different uses including: suggestions, requests, speculations, denials, and accusations.

The final *possible* to be considered is *possible that*. Speculation is how Phillipson uses *possible that* particularly concerning a potentially bleak future for certain people: “It is therefore perfectly possible that the global linguistic map may change violently in the coming decades, and that in the intervening period, monolingualism in English may be a serious liability in the job market.” (Phillipson 2008: 19). Irving uses the term for general speculation, but in a less pronounced or consistent fashion.

Discussion

Table 3 shows the raw frequency and corresponding percentage of both dogmatic and open lexis across each author and in both reference corpora.

Table 3. Dogmatic vs. open lexis

	Icke Freq	Irving Freq	Mullins Freq	Phillipson Freq	Sutton Freq	Tsarion Freq	Consp Freq	Control Freq
Dogmatism	4820	5084	2809	953	1621	1948	16435	1254
Openness	5538	3971	2799	1755	1960	2033	16301	2803
Total	10358	9055	5608	2708	3581	3981	32736	4057
Dogma %	46.5	56.1	50.1	35.1	45.3	48.9	50.2	30.9
Open %	53.4	43.8	49.9	64.8	54.7	51	49.7	69

The Phillipson corpus is closest, both in its level of dogmatism and openness, to the control corpus, while it is less dogmatic and more open-minded than the conspiracy theory corpora. At the other end of the scale, the Irving corpus is the most dogmatic and the least open-minded. His theories hold some superficial similarity to Linguistic Imperialism in that, as evidence for their claims, both use government documents and interviews with people who have worked in government related fields, while both are concerned with actual people and situations that have a strong basis in fact. Thus, it would have been a reasonable prediction to suppose that there would be greater statistical commonality between Phillipson and Irving. However, this is not the case. Irving's scores could be in part attributed to his prolific use of quotes and citations of Nazi officers. A major source in DOTA's formation was Nazi propaganda. As such, any book that talks so much about discourse content that the Nazi high command produced is likely to feature a high degree of dogmatism.

While completing the analysis of the section dealing with the use of *can*, there seemed to emerge a similarity between how Linguistic Imperialism and Icke's conspiracy theories are expressed. Although there was a stark difference in the content specifics, the underlying sentiments seemed much closer: *can* highlighted matters regarding agency, the history of agency, and the reaction – or lack of it – by the general population. Also, *can* seems to be one of the foundations upon which both Phillipson and Icke build their arguments on, regardless of whether the theories are conspiracy theories or not. Further investigation of *can* across the other authors might help to validate this claim.

The use of *only* demonstrates very little common ground at all between the two corpora compared on this basis. At least as far as their respective use of *only* reveals, *only* pretty much serves to highlight the differences between the texts. What marginal overlaps there are prove irrelevant, given the otherwise totally unrelated or even contradictory use of *only*. What this seems to reveal is more that Tsarion doesn't employ *only* in key sentences to assert his ideas, but more as a softening device: "It can only be shunted aside, cast in shadow, arrested, and unexpressed." (Tsarion 2010: 49). Within the Linguistic Imperialism corpus, *only* is often employed to consolidate a position either by being used in the identification of evidence, or used in supporting or rejecting other scholars' ideas: "globally this can only be done equitably by giving voice to speakers of a multiplicity of languages" (Phillipson 2010: 243).

Where there is overlap in the use of *possible* in the two corpora in question, it seems to have more to do with the general use of the term than with the core of the theories. This is particularly true of Irving, in whose works *possible* sees extensive and diverse use: "He would even produce documents to prove that he did everything possible to ensure that the labour force was well looked after." (Irving 1996:

297). This in itself is somewhat revealing because the Linguistic Imperialism concordance lines containing *possible* are almost always concerned with the theory rather than other matters: "Once the conspiracy theory is removed, it is possible to identify the paper's main claim" (Phillipson 2009:190). This could be a consequence of the corpus size, or of the authors' writing styles.

The use of *several* proves divisive.. In the context of this study, most of the occurrences seem of almost no interest: they are bland and typical uses of the term, showing us that both corpora have plenty of unremarkable language use in them. The collocation *several decades* is a departure from this, and seems to have a sinister undertone in some of its occurrences: "The French have been aware of the threat from cultural and linguistic imperialism for several decades." (Phillipson 2009:152). Perhaps this tendency would be reflected in the other corpora too, but it has more to do with a discussion of time-frames and the long-term cultivation and consequences of the theories than it does with the word *several*.

Having examined how certain terms manifest in both the Linguistic Imperialism corpus and the various conspiracy theory corpora, it seems relevant and necessary to look again at some of the terms in relation to the control corpus of academic texts.

Within this corpus *several decades* is found numerous times and largely in reference to the development and shifts in the media for example: "For several decades, print newspapers have not been wiped out by the radio or by television in the media environment." (Anonymous Control Corpus 2008: 29). While *several* here is being used as a non-specific counter of decades, as was often the case in the experimental corpora, the feel of the control corpus use is much more matter of fact – rather a case of accounting for the general history of something. Sutton and Phillipson both do use the term for general accounts, but also to insinuate the brewing of plots and the hatching of schemes: "Over the past several decades, quietly, without media attention, many Americans in diverse fields of activity have been pressured into silence" (Sutton 1986: 15); "An insider in the murky universe of CIA 'intelligence' over several decades wrote in 1998 that there was an underlying "devastating truth" (Phillipson 2008: 8).

All three corpora feature a sense of competition: between formats, between nations, between languages. The difference is that Sutton and Phillipson make some direct accusations about what has been happening in the last *several decades*, whereas the control corpus implies no interference or agenda as regards media development.

The sole instance of *can be traced* that is found in the control corpus is concerned with agency and agenda, as it relates to the relationship between the leaders of the Soviet Union and its media: "The theory can be traced back to the 1917 Russian Revolution" (Anonymous Control Corpus 2008: 25). Although this has

some superficial resemblance to the use of *can be traced* in the experimental corpora, the key difference is that there is no element of secrecy. The control corpus talks about the accepted fact that there was a traceable agenda to the use of the media by the communists' leaders in the USSR. Icke's use of *can be traced* is far more speculative, whereas Phillipson's use, while closer to the control corpus, is observational with a speculative edge.

Across all the corpora *cannot be* sees comparable use: the expression of assertions and expectations, with the occasional dip into the describing of rules. This manifests itself with the greatest consistency in the Linguistic Imperialism corpus, where it tends to be used for quite intense assertions and firm expectations: "English cannot be a neutral lingua franca if it is at one and the same time somebody's mother tongue and somebody else's foreign language." (Phillipson 2009:190).

That there is greater variation in the 'feel' of the use of *cannot be* in the other corpora may be attributed to the presence of multiple authors in the control corpus: "The politics of justice cannot be only national or only international." (Anonymous Control Corpus 2008: 25); "Thus babbling, while cute, cannot be considered language, but rather language practice, and gesture-based systems like Laura's can in fact be considered an effective protolanguage." (Irwin 2005:3). In the case of the Mullins corpus, a few factors may lead to the increased diversity of how, exactly, *cannot be* is deployed: "The Federal Reserve Act stipulates that the stock of the Federal Reserve Banks cannot be bought or sold on any stock exchange." (Mullins 1952: 225). This diversity could be attributed, among a number of things, to: the greater size of the corpus, the different focuses of the texts included in the corpus, and the fact that the texts were, in some cases, produced decades apart so that there may be some longitudinal variation in the authors' use of the term.

Conclusion

This paper sought corpus informed insights as to why Phillipson's theory of Linguistic Imperialism has sometimes been labelled as a conspiracy theory.

It began by outlining corpus linguistics and its potential use in analyzing texts both in general and with particular reference to DOTA and WordSmith: the tools that were used for this study. It then introduced Phillipson's theory of Linguistic Imperialism and some of the criticisms of it. It went on to describe research into conspiracy theories and some of the works that were included in the conspiracy theory corpora.

The methodology section outlined the construction of the corpora and how DOTA and WordSmith were used to investigate them, particularly in terms of

their use of dogmatic and open-minded collocations and clusters. The results of this analysis were then presented and discussed.

The data indicated that, in terms of dogmatism, the Phillipson corpus was distinctly less dogmatic than the conspiracy corpora, and very close to the levels of dogmatism found in the control corpus. The data also showed a wide range of variation in the levels of open-mindedness across the corpora. The control corpus stood out as having notably higher levels of open-mindedness than all of the other corpora. The gap between the Phillipson corpus and the control corpus was smaller in terms of open-mindedness compared to all but the Icke corpus. However, the Phillipson corpus exhibited levels of open-mindedness that were much closer to each of the conspiracy corpora (with the exception of the Irving corpus). On the basis of this data, it seems reasonable to conclude that if the DOTA analysis of the corpora has yielded any indication as to why Linguistic Imperialism has been accused of being a conspiracy theory, it appears to be because the level of open-mindedness it exhibits is much closer to that of the conspiracy theories than it is to that of the control corpus. This is not to say that Linguistic Imperialism is a conspiracy theory, only that this is one insight that the data offers as a possible explanation for the accusation.

The concordance level scrutiny of the corpora offers some other more qualitative insights. The manner in which Tsarion and Phillipson employed their shared key term *only* revealed no common ground. Conversely, there did seem to be significant overlap in how both Icke and Phillipson used *can*, with the exception of the particular collocation *can be seen*. *cannot be* was shared between Mullins and Phillipson and deployed in a similar vein. Although there were similarities around the key term *several* in both Sutton and Phillipson, there were also some indication that these were more due to the use of *decades* in *several decades* than to that of the key term itself. Finally, it was not possible to identify any pronounced correlation – positive or negative – regarding the use of *possible* between Irving and Phillipson. As with the data, the sentence level analysis focusing on collocations and clusters with key words from the DOTA list suggested that the open-minded terms, rather than the dogmatic ones, were the greater source of commonality between Linguistic Imperialism and the texts in the conspiracy theory corpora.

To summarise, the use of DOTA to study the corpora showed that Linguistic Imperialism is both less dogmatic than conspiracy theories tend to be, but also less open-minded than the academic writing control corpus. This finding was supported by the concordance level work conducted using WordSmith, which found greater parallels between the Phillipson corpus and the other corpora on the basis on open-minded key word collocations and clusters than it did with the dogmatic key words.

In conclusion, the insights offered by this corpus informed analysis are: the accusations of Linguistic Imperialism being a conspiracy theory do not seem to originate in its being overly dogmatic, but rather may arise on account of it showing levels of open-mindedness that are comparable with DOTA's indicators.

References

- Adolphs, Svenja. 2006. *Introducing Electronic Text Analysis*. Abingdon, Oxon.: Routledge.
- Baker, Paul (ed.). 2009. *Contemporary Corpus Linguistics*. London: Continuum.
- Biber, Douglas, Susan Conrad, and Randi Reppen. 2006. *Corpus Linguistics Investigating Language Structure and Use*. New York: Cambridge University Press.
- Bunce, Pauline, Robert Phillipson, Vaughan Rapatahana, and Ruanni Tupas (eds). 2016. "Confronting the Hydra". *Why English?* Bristol: Multilingual Matters.
- Byford, Jovan. 2015. *Conspiracy Theories A Critical Introduction*. London: Palgrave Macmillan.
- Canagarajah, A. Suresh. 2003. *Resisting Linguistics Imperialism in English Teaching*. Oxford: Oxford University Press.
- Cooke, Monte. 2009. *The Skeptic's Guide To Conspiracies*. Avon, Mass.: Adams Media.
- Davies, Alan. 1996. "Ironising the Myth of Linguicism". *Journal of Multilingual and Multicultural Development* 16: 485–496. <https://doi.org/10.1080/01434639608666297>
- Ertel, Suitbert. 1972. "Erkenntnis und Dogmatismus" [Insight and dogmatism]. *Psychologischer Rundschau* 23: 241–269.
- Ertel, Suitbert. 1981. "Prägnanztendenzen in Wahrnehmung und Bewußtsein" [Emerging tendencies in perception and consciousness]. *Zeitschrift für Semiotik* 3: 107–141.
- Ertel, Suitbert. 1985. "Content analysis: An alternative approach to open and closed minds". *High School Journal* 68 (4): 229–240.
- Jensen, Kim Ebensgaard. 2014. "Linguistics and the digital humanities: (computational) corpus linguistics". *Mediekultur: Journal of Media and Communication Research* 57: 117–136..
- Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. Harlow, Essex: Longman.
- Knight, Peter (ed.). 2003. *Conspiracy Theories in American History: An Encyclopaedia*. Santa Barbara, Calif.: ABC-CLIO.
- McKenny, John. 2003. "A double take on writer's stance: Content analysis of dogmatism compared with corpus analysis of student essays". corpus analysis of student essays". In Luuk Lagerwerf, Wilbert Spooren, and Liesbeth Degand (eds) 2003, *Determination of Information and Tenor in Texts: Multidisciplinary Approaches to Discourse*. Amsterdam: Vrije Universiteit, Stichting Neerlandistiek. pp. 93–104.
- McKenny, John. 2005. "Content analysis of dogmatism compared with corpus analysis of epistemic stance in student essays". *Information Design Journal + Document Design* 13 (1): 40–50. <https://doi.org/10.1075/idjdd.13.1.05mck>
- Phillipson, Robert. 1992. *Linguistic imperialism*. Oxford: Oxford University Press.
- Phillipson, Robert. 2007. "Linguistic imperialism: a conspiracy, or a conspiracy of silence?". *Language Policy* 6: 377–383. <https://doi.org/10.1007/s10993-007-9058-3>
- Phillipson, Robert. 2009. *Linguistic imperialism Continued*. Hyderabad: Orient Blackswan.
- Phillipson, Robert. 2011. "English: from British empire to corporate empire". *Sociolinguistic Studies* 5 (3): 441–464.

- Phillipson, Robert 2016. Myths and realities of 'global' English. *Language Policy*. (On line).
- Sinclair, John. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Spolsky, Bernard. 2004. *Language Policy*. Cambridge: Cambridge University Press.
- Stefanowitsch, Anatol, and Stefan Thomas Gries (eds). 2007. *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin: Mouton de Gruyter.
- Wilson, Robert Anton. 1998. *Everything is Under Control: Conspiracies, Cults, and Cover-ups*. New York: William Morrow.

Appendix

There follows a list of the texts included in each corpus. For the control corpus, some of the texts were censored to be anonymous, but where known the authors are also added.

Robert Phillipson

- (1992) *ELT the native speaker's burden?*
- (1996) *Linguistic imperialism: African perspectives*
- (2001) *English for Globalisation or for the World's People?*
- (2003) *The new linguistic imperial order: English as an EU lingua franca or lingua frankensteinia?*
- (2006) *Figuring out the Englishisation of Europe*
- (2007) *Linguistic imperialism: a conspiracy, or a conspiracy of silence?*
- (2008) *English, panacea or pandemic?*
- (2008) *The Linguistic Imperialism of Neoliberal Empire*
- (2009) *Some partner languages are more equal than others*
- (2010) *Realities and Myths of Linguistic Imperialism*
- (2011) *English: from British empire to corporate empire*

Michael Tsarion

- (2002) *Atlantis, Alien Visitation, and Genetic Manipulation*
- (2008) *Astrotheology and Sidereal Mythology*
- (2010) *What is Objectionism?*

David Icke

- (1999) *The Biggest Secret*
- (2001) *Children of the Matrix*
- (2002) *Alice in Wonderland and The World Trade Center Disaster*
- (2003) *Tales from the Time Loop*

David Irving

- (1989) *Goring: A Biography*
- (1996) *Goebbels: Mastermind of the Third Reich*
- (1996) *Nuremberg: The Last Battle*

Anthony Sutton

- (1976) *Wall Street and the Rise of Hitler*
(1981) *Wall Street and the Bolshevik Revolution*
(1986) *The Best Enemy Money Can Buy*
(2004) *America's Secret Establishment: An Introduction to the Order of Skull & Bones*

Eustace Mullins

- (1952) *Secrets of the Federal Reserve*
(1968) *The Biological Jew*
(1980) *The Curse of Canaan*
(1984) *The Secret Holocaust*
(1985) *The World Order, A Study in the Hegemony of Parasitism*
(1988) *Murder By Injection*

The control corpus

- Anonymous, 2008, *A Comparison of Culture between China and the West through the Analysis of Creation Myths*
Anonymous, 2008, *Blogging and Constructing Identity*
Anonymous, 2008, *Influence of Different Environment on Learning Style of Chinese and Western Students: A Cultural Approach International*
Anonymous, 2008, *Political, Economic and Cultural Factors That Influence Newspaper Publishing in Britain and China: A Comparison between The Times and The People's Daily*
Anonymous, 2008, *Reporting the same events? A comparative studying of the China Daily and the New York Times.*
Anonymous, 2008, *The future of print newspapers in the information age in China*
Anonymous, 2008, *The Novel Wears Prada An Analysis of Fashion Communication in Chick-Lit*
Anonymous, 2007, *The Road of Young Women: Female Images in the Films 'Forever Young' and 'Mona Lisa Smile'*
Bailey, Benjamin, 2004, *Misunderstanding*
Birdsong, David, 2006, *Age and Second Language Acquisition and Processing: A Selective Overview*
Brost, Molly, 2010, *Walking the Line: Negotiating Celebrity in the Country Music Biopic*
Featherston, Mike, and Couze Venn, 2006, *Problematizing Global Knowledge and the New Encyclopaedia Project*
Hayes, Michael, 1992, *On The Epistemology of Risk: Language, Logic and Social Science*
Hoshino, Ayako and Hiroshi Nakagawa, 2007, *Assisting cloze test making with a web application*
Huang, Philip, 1993, *"Public Sphere"/"Civil Society" in China?: The Third Realm between State and Society*
Irwin, Derek, 2005, *So What's on Georgia's Mind? A Systemic View of Language Development*
Marr, Andrew, 2005, *Unleashing Humphrys or Paxo is a democratic service*
Noel, Alain, and Jean-Philippe Therien. 2002, *Public Opinion and Global Justice*
Oakley, Kate, 2004, *Not so cool Britannia*
Robinson, William, 2009, *Saskia Sassen and the Sociology of Globalization: A Critical Appraisal*
Salgueiro, Teresa, 1997, *Consumption and City Fragmentation*

- Skaperdas, Stergios, and Bernard Grofman, 1995, *Modelling Negative Campaigning*
- Sondrol, Paul, 2008, *Totalitarian and Authoritarian Dictators: A Comparison of Fidel Castro and Alfredo Stroessner*
- Waite, C., 2009, *How to Study Communication: Notes on a Method*
- Ye, Lulu, 2006, *Opportunities and Challenges of CCTV Foreign Service*

Address for correspondence

Sean Thornton
Toyo University
1-1-1 Izumino, Itakura-machi
Gunma-ken, 374-1093
Japan
sean@toyo.jp

Biographical notes

Sean Thornton is a lecturer and researcher at Toyo University in Japan, where he primarily teaches English for Academic Purposes with a focus on Applied Science related courses. His academic background includes degrees in: religious studies, applied linguistics, and cultural astronomy. He has published and presented in the fields of ELT, applied linguistics, anthropology, and science history. His current research projects concern survival English, compulsory education, and Carl Jung's processes and methodologies. This wide range of interests is in part his own attempt to discover a suitable PhD topic that he may eventually pursue.