(Mis)interpreting the Commerce Clause: A Critical Look At Corpus Linguistics As A Solution to the Question of Originalism.

Yuval Kehila
University of British Columbia

Abstract

In recent years the use of corpus linguistics as a remedy to the problem of linguistic drift and United States legal language interpretation has proliferated, especially in the field of constitutional originalism of United States constitutional language. This paper investigates the Commerce Clause of the United States constitution as a case study for looking at the effects of legal language misinterpretation, through the Corpus of Founding Era American English (COFEA). This paper will then take a critical look at a modern case of legal language in order to question critical assumptions made by linguists about the employment of corpora for originalist analysis. This paper argues for the necessity of scepticism around the idea that 'original intention' behind written law can somehow be revealed by the syntax-oriented public-usage frequency analysis achieved through corpus analysis. Using corpus analysis to reveal intention assumes not only that legal language likely conforms to original public meanings of words, but more critically that these laws were intended to conform to one such interpretation. Rather, this paper will demonstrate that legal language is consistently written in a way which lends itself to multiple interpretations, and perhaps done so purposefully. By employing corpus linguistics to tackle the problem of original legal meaning, this paper argues, linguists could be ignoring this as a significant *feature* of legal language culture.

Introduction:

An integral part of modern state identity is the legal body which serves to regulate not simply the principles by which a state operates internally and internationally, but also the language of such interactions. Language, for legal systems, is the basis for not only the definition of such principles, but operates as an inbuilt notetaker and arbitrator; both documenting and assuring consistency in legality over time and space. This paper addresses one challenge in legal language culture, which is the interpretation and application of legal language – specifically of language which is in use hundreds of years after its writing. This paper critically addresses recent attempts using corpus linguistics to grapple the problems of original meaning in legal language interpretation in the United States Constitution, under a body of theoretical work termed 'originalism'. Furthermore, this essay serves to address how both the problem of interpretation, and the originalist solutions which have been employed in its remedy, speak to a deeper and more insidious set of problems in legal language culture.

Linguistic Drift and the Need for Originalism | *Introducing the Commerce Clause***:**

The history of law in Western tradition has seen numerous shifts in language and culture (May 2019). One challenge with the continued use of historic legal documents such as the United States Constitution is that the original legal language remains the *only* standard for use in any legal setting since their writing. This becomes a problem due to the linguistic drift – the theory that language shifts in meaning and usage over time (Lee and Phillips 2018) – which has occurred in American English since the 1780's. This drift means that the interpretation of constitutional legal language is not, by any means, straightforward, and its misinterpretation not without real world consequences.

Linguistic drift refers to a body of literature and assumptions regarding the interdependent change of language and culture over time. Inevitably, it seems, the hermeneutic resources of a society are embedded in their language, and thereby the change of the usage and meaning of language over time and space makes writing from distant historical or cultural settings challenging and complex (Lee and Phillips 2018). One prominent case to highlight the effect of such linguistic drift on legal practice is the case of the 'Commerce Clause' which has not changed in wording since the ratification of the United States Constitution in 1788, but is frequently used in contemporary judicial rulings, including as justification for seemingly unrelated policy. The Commerce Clause itself consists of only 16 words *"[That congress shall have the power] to regulate commerce with foreign nations, and among the several states, and with the Indian tribes;"*[1]. The history of the use of this clause, I argue, makes a cogent case for the importance of understanding some of the real world effects of linguistic drift on the cultural (mis)understanding and (mis)use of legal language.

During the Civil Rights movement of the 1960's, politicians fighting for desegregation faced a challenge with the 14th amendment. While the amendment should have provided all the necessary legal support for desegregation, it seemed overly focused with the actions of 'states' and could thereby not be used to force desegregation in smaller businesses such as restaurants – "*No <u>state</u> shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any <u>state</u> deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws*"[2]– Instead, the government passed the 1964 civil rights act on the basis of the Commerce Clause. In a subsequent case, Katzenbach v. McClung, a small business, Ollie's Barbeque, argued they should not be regulated as they did not constitute commerce *among the states* (WYNC 2018). In 1964, however, the United States Supreme Court ruled against Ollie's Barbeque, arguing for an interpretation of *commerce* under which a restaurant sourcing ingredients from out-of-state constituted interstate commerce and therefore could be regulated by congress.

Some have poignantly noted the absurdity implied in the need to pass civil rights legislation on the basis of commerce rather than areas of the law already intended for rights.

[1] U.S. Const. Art. 1. Sec. 8. Cl. 3
[2] U.S. Const. Amend. XIV. Sec. 1.

However, putting this aside, many legal scholars also argue that such interpretations of the Commerce Clause were undoubtedly not the intended use of its initial writing, even if they led to progressive outcomes (Lacour 2002). This sort of argument seems intuitive given that the civil rights movement and the question of segregation was not on the minds of the writers of the clause in the same way these questions occupied the minds of Americans during the civil rights movement. This change in cultural context which results from over *two-hundred* years of linguistic drift cannot be ignored. Namely, the socio-linguistic significance of *commerce* as it applies to the case of the Commerce Clause in the United States constitution, raises a significant question of *originalism* (Biancalana 2002).

Originalism in the Context of the United States Constitution:

Originalism in United States Law refers to the methods for the judicial interpretation of the United States constitution in order to ascertain its original meaning and intention (Lee and Phillips 2018). The group of theories which encompass originalism in the constitutional context agree that there is some fixed *communicative content* of the language in the constitution. This communicative content is ratified into the constitution alongside the lexical content of the clauses, which contains the actual relevant legal meaning which must be adhered to in the future use of the legal documents. Due to the prevalence of cases such as the aforementioned Commerce Clause, where legal language from a time both culturally and linguistically out of reach to the modern interpreter is used to make contemporary legal decisions, the question of originalism has understandably proved more and more necessary in the legal consciousness of America. Lee and Phillips argue that originalism, if mastered, is a tool which can not only streamline the application of constitutional clauses in legal settings, but also provide judges with "external and internal restraint" (268). Lee and Phillips split originalism regarding the United States constitution into three ideological camps: public meaning originalism, original intentions originalism, and methods originalism.

Public meaning originalism focuses itself on the garnering of understanding of constitutional-era semantic meaning in public interpretation. The meaning of the United States constitution is embedded in the regularities of use and the semiotic context of late 1700's America, and therefore what we are to strive for in their interpretation is how the 18th century public would have interpreted the writing in terms of their contemporary language conventions. Original intentions originalism, by contrast, searches not for how a work would have been interpreted, but for the intentions of those writing it. To illustrate how these differ, the public meaning originalist would be interested in any contemporaneous media discussion of the constitution, or the general semantic culture around the language used, whereas the original intentions originalist would opt for written record or discussion of the constitution by the people who originally wrote it. Methods originalists, by contrast, hold that the constitution should be treated as a work in a separate 'dialect' and that we can only understand the constitution by "immersing ourselves in the language community of that dialect" (Lee and Phillips 2018, 271). What all of these ideological camps have in common is that they call for the extensive analysis

of primary language data from the historical cultural context of the United States constitution. Therefore the question of originalism becomes an anthropological and linguistic question: How can we accurately bridge the cultural gap of time and tether ourselves to meanings which are hermeneutically inaccessible to us in the modern world? Many have thought corpus linguistics to be the answer to this question.

In a paper entitled *Corpus Linguistics & Original Public Meaning: A New Tool To Make Originalism More Empirical*, Phillips, Lee, and Ortner (2016) argue that given originalism is the predominant method for interpreting the constitution, it therefore becomes imperative that such interpretations are empirically driven so that we can attempt to bridge the historiographic chasm. With this call for empiricism, they envisioned a comprehensive corpora of texts from the period in which the constitution was written, for use in corpus analysis.

Corpus analysis refers to the compilation and systematic analysis of language from specific contexts whereby words and term-phrases could be sequenced out of this data, and organised by context and connotation (May 2019). Ergo a corpus ideally allows us to ascertain the specialised or 'ordinary public meaning' of language in a certain time and space, and then let us approach texts, such as the United States constitution, using a new context-oriented lens. Nonetheless, corpus linguistics is not without limitations. Lawrence M. Solan (2016) wrote in support of Lee and Ortner's call for a corpus, but noted fundamental problems with corpus linguistic analysis: I. A corpus may highlight multiple 'original public meanings', or none at all; II. There lacks a clear, and well supported theory for the selection of one public meaning over another – meaning that originalists, even when trying to be empirical in their methods, "must still exercise judgement to determine how the various occurrences of words or phrases should inform their meaning in the Constitution." (59); and III. For many abstract concepts, common in legal language – citing the example *'abridging the freedom of speech,'*– it is not always apparent whether the concept is meant in reference to particular examples of the concept, or rather if what is being referred to is the concept of freedom of speech itself.

Solan thereby situates the project within a pre-existing and multifaceted set of problems already inherent in other corpus analysis; problems which linguists and lexicographers must contend with constantly, in any work which must make meaning of one language in another (i.e foreign language dictionaries). Solan calls upon Lawrence Lessig (1993), who in *Fidelity in Translation* notes that "Translation is a practice that neutralises the effect of changed language on a text's meaning, where language is just one part of context, and changed language is just one kind of change in context". Simply put, linguistic drift is another contextual change, which requires a translation into our current cultural-linguistic context.

Lessig urges that the originalist, upon enlisting a corpora for constitutional interpretation, must take upon herself the role of a translator. She must view the language of the United States constitution as foreign, and must be even more diligent than the translator in that she must simultaneously fight the apparent lexical similarities, and the urge to affix to them modern judgements and hermeneutical perspectives. Solan (2016), however, argues that this may not be sufficient because even the lexicographer will consider the purpose and audience of their

dictionary, as well as the need for brevity or simply their financial resources. Similarly, in a legal setting, "there will always be lexicographic decisions to be made about how narrowly or broadly to define a term" (7). Meaning, even if taking the translator's perspective could help the originalist achieve neutrality and combat personal reflexivity bias by artificially distancing themselves from the language of the constitution, we have to recognise that even the perspective of the translator is not free from bias. It sacrifices 'accuracy' for 'purpose', and still does not free one from having to ultimately decide upon an 'original public meaning'. White and Phillips (2017) furthermore say that before even considering the interpretation of corpora, the selection of texts from which to produce a corpora is by no means straightforward, especially in such culturally and historically distant settings. Linguists must balance the need for a large sample size, which is fundamental to the validity of sampling methodology, and choosing a representative sample, which becomes problematic when considering the variables in play. A text must not only be from a similar time period, but a similar language usage-culture to the constitution, in order to be considered relevant in the comprehension of the constitution.

Nonetheless, in an effort to bridge the empiricism gap, the project discussed in these papers has since been completed: the Corpus of Founding Era American-English (COFEA), "tailored to cover the linguistic range of the Constitution and the Founding Era, beginning with the rule of King George III in 1760 and ending with the death of George Washington in 1799" (May 2019, 5-6). The first thing that this allowed scholars to do was to actually empirically study constitutional originalism for the first time. However, it also provided new insight into the effectiveness of corpus linguistics as a methodology for ascertaining original public meaning.

Garrett May (2019) used the COFEA to evaluate precisely the Commerce Clause aforementioned in the beginning of this essay. May's analysis of the corpus found four distinct senses of 'commerce', as well as two instances where meanings did not fit any of those four, labelled as *other* or *intermediate*, totalling six senses altogether. After identifying a significant sample of instances where there were duplicates (27 instances) or unclear meanings (20 instances), there emerged one sense, overwhelmingly used in 96 out of the 171 usable instances analysed. The second highest clear sense was far behind with only 16 instances. This most popular sense defines commerce as "the trading, bartering, buying, and selling of goods (and the incidents of transporting those goods within the definition)" (May 2019, 7). This sense does not seem to support the reading of the Commerce Clause utilised in the 1964 supreme court decision regarding Ollie's Barbeque, and calls into question many other interpretations of the clause which have influenced legal decisions in the last 200 years. May notes that given the corpus size of almost 120,000 texts, the appearance of the collocation 'regulate commerce' just 229 times seems wildly disproportionate and could indicate the general invalidity of any conclusions drawn from the analysis. May continues, however, that this is also a generalised argument which is often posed to originalism and to corpus analysis, and which does not really accomplish anything. Evidently, regardless of how much the regulation of commerce was discussed in the 1700's, and whether it bears the same weight as it does in contemporary public politics, these terms which nonetheless occupied a meaning, were used in the constitution. May reminds us

there are numerous things which did not exist at the time of the constitution, and which we still use constitutional law to regulate. Such novel items as the smartphone, modern car, or the globalised transnational corporation, are prime examples. “Regulated commerce was an item of discussion of the founding era…”, writes May, and “minimal though its representation in the corpus may be, it is represented, and our analysis of [it] stands attested.” (2019, 9).

Lee and Phillips (2019) also looked at the Commerce Clause using the COFEA for the same set of senses of commerce/regulating commerce. Interestingly, Lee and Phillips consider that these senses are not mutually exclusive. They describe that a manufacturing sense of commerce, “the production of goods for trade; manufacturing” (300) and a trade sense, “the trading, bartering, buying, and selling of goods (and the incidents of transporting those goods within the definition)” (300), would both fall neatly under a broader, market sense of commerce, “any market-based activity having an economic component (this would include trade, manufacturing, agriculture, labor, and services)” (300). Further, in an analysis of collocates of commerce, one collocate stood out in frequency, *amity*, perceivably the result of trade agreements the United States was making with other nations at the time, called ‘treaties of Amity and Commerce’ (303). This, as well as other notable collocates, also supports an overall trade-oriented understanding of commerce outlined in the leading sense found by both May, and Lee & Phillips’ corpus analyses. The originalist now has a solid case to argue that the 1964 passing of the civil rights act on account of the Commerce Clause, and the legal disputes against segregation with small local businesses such as Ollie’s Barbeque, constitute misuse of the original sense of commerce which existed in the language culture of the founders of the United States constitution. Ollie’s Barbeque partook in the sense of commerce as “the trading, bartering, buying, and selling of goods”, at the level of sourcing ingredients, but not at the level of the ‘goods’ which they created and served at their restaurant. Further, the original intentions originalist might now ask whether or not whom the restaurant served fell under the parts of ‘commerce’ which the clause allows congress to ‘regulate’. This by no means suggests that segregation should not have been made illegal in the United States, or anywhere else for that matter, but rather that it should have been made illegal under the power of human rights legislation as opposed to commerce legislation. Perhaps further corpus analysis of the 14th amendment would find that in 18th century legal discussion, there was less distinction between the power of individual states and that of a central government, and perhaps that we might be able to make the case of originalism to support an interpretation which allows for federal intervention in matters of the 14th amendment.

Regardless of these anecdotal cases, the emergence of this corpus as a tool clearly demonstrates that corpus linguistics serves as a powerful tool in both the interpretation and reformation of legal systems, in the battle against the challenges which arise out of linguistic drift. However, I will take this opportunity for a more general critique of the *modus operandi* of the use and interpretation of language in legal settings —or more precisely of the general assumptions made by linguists and legal scholars in this interpretation. When an originalist searches for original public meaning in the syntax and grammar of a text, they will thereby

assume that the meaning is wholly embedded in the words used. Ideally, this should be the case, especially in a legal context, because those words are all the lawmaker has at her disposal in defining the distinction between the legal and the illegal. However, one need not spend much time looking at legal writing or listening to courtroom law to recognise that the legal language culture is uniquely indeterminate and prevaricating. To exemplify this, I call the reader's attention away from the distant and out of reach language of a 230 year old constitution, and to a contemporary and clear modern case of linguistic research in law.

In a study of juror death penalty decisions in Texas capital punishment cases, Robin Conley (2013) notes that the structure of capital punishment law and the way it is enacted in jury trials is purposefully structured to systematically dehumanise defendants, and distance jurors from ethical decisions. Specifically, the language of the decision to sentence someone to death seems unendurably convoluted. Jurors are asked to answer two questions; 1. "Do you find from the evidence beyond a reasonable doubt that there is a prob-ability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society?" (505) and 2. "Do you find from the evidence, taking into consideration all of the evidence, in- cluding the circumstances of the offence, the defendant's character and back- ground, and the personal moral culpability of the defendant, that there is a sufficient mitigating circumstance or circumstances to warrant that a sentence of life imprisonment rather than a death" (505). In neither question is there explicit mention of the juror's responsibility towards sentencing the defendant to death. The juror is simply answering two questions, and in a different legal or cultural context, the answer to these two same questions might not result in death. Thereby, the juror is not answering the question *should the defendant be killed*, but rather the decision seems to be just a natural consequence of the legal system's interpretation of a certain combination of answers to the two questions. As if in an effort to prove itself even more reductionistic, the question has been divided into two, so that rather than the juror consider all aspects of the question of death in one moment, the juror is asked to separate the matter into distinct sections, and ignore the coherence which comes with viewing the problem as a whole. Thereby, they are also asked to not see the defendant as a whole, but rather to separate them into parts as well, so as to dehumanise them, and remove responsibility from a juror, to make the decision appear less daunting to the individual.

What this points to is a general tendency in the legal use of language — also observed in the aforementioned (mis)use of the Commerce Clause — to say things in ways which lend to multiple interpretations; they are overall inexplicit whilst having the forward appearance of being explicit. This problem cuts to the foundations of originalism, insofar as we assume that finding the 'original public meaning' of the words of a text like the constitution will help discover the meaning of the words as they are used in the legal treatise itself. One simultaneously must assume that the words in the treatise bear the ordinary or statistically likely sense of their language culture. However, it is the overwhelming tendency of law to use complicated language that has the appearance of being explicit due to its inaccessibility, in order to discreetly allow for open ended interpretations to use to their benefit. Just as it is the tendency of judges and lawyers

to use those possible interpretations of the legislation, like the Commerce Clause, to achieve certain legal outcomes, I argue that it is equally the tendency of the writers of these laws — the tendency of the entire linguistic culture of law — to write laws which allow for such multitudinous interpretations.

Conclusion:

Language is both an integral part of the legal identity of a modern nation-state like the United States, and its biggest challenge. Language promises to immortalise laws and ensure the constancy of their application across time and space, yet language equally locks the meaning of a law in a specific context of time and place; language explicitly demarcates the legal and the illegal, and yet language itself creates avenues for the (mis)interpretation and (mis)use of laws. Even in the employment of corpus linguistics to solve the problems of linguistic drift, language seems to be both the saviour and the corruptor of law — both law's biggest problem, and the tool which law uses to fix and reform itself. However, it seems that this dichotomous relationship is so deeply embedded in the very nature of legal culture that it becomes, itself, integral to defining what law is. It becomes insidious to the efforts of legal scholars and lawmakers who fight to ascertain the true meaning of legal language, and yet manipulate language to ensure that one singular clear-cut interpretation cannot be inferred. Corpus linguistics in the field of constitutional originalism must be treated as both a symptom of, and remedy for, the awkward waltz of legal language culture.

Bibliography:

Biancalana, Joseph. 2002. 'Originalism and the Commerce Clause'. *U. Cin. L. Rev.* 71: 383.

Conley, Robin. 2013. 'Living with the Decision That Someone Will Die: Linguistic Distance and Empathy in Jurors' Death Penalty Decisions'. *Language in Society* 42 (5): 503–526.

Lacour, Vanue B. 2002. 'The Misunderstanding and Misuse of the Commerce Clause'. *SUL Rev.* 30: 187.

Lee, Thomas R., and James C. Phillips. 2019. 'Data-Driven Originalism'. *U. Pa. L. Rev.* 167: 261.

Lessig, Lawrence. 1992. 'Fidelity in Translation'. *Tex. L. Rev.* 71: 1165.

May, Garrett. 2019. 'Immigration Law on the Line: Corpus Linguistics and the Commerce Clause'. *Brigham Young University Prelaw Review* 33 (1): 9.

Phillips, James C., Daniel M. Ortner, and Thomas R. Lee. 2016. 'Corpus Linguistics & Original Public Meaning: A New Tool to Make Originalism More Empirical'. *Yale LJF* 126: 21.

Phillips, James Cleith, and Sara White. 2017. 'The Meaning of the Three Emoluments Clauses in the US Constitution: A Corpus Linguistic Analysis of American English From 1760-1799'. *S. Tex. L. Rev.* 59: 181.

Solan, Lawrence M. 2016. ‘Can Corpus Linguistics Help Make Originalism Scientific’. *Yale LJF* 126: 57.

Solum, Lawrence B. 2017. ‘Triangulating Public Meaning: Corpus Linguistics, Immersion, and the Constitutional Record’. *BYU L. Rev.*, 1621.

WNYC Studios. 2018. ‘One Nation, Under Money’. More Perfect. Retrieved https://bit.ly/2XkKIv2

Edited by: Laura Derby