International Journal of
LANGUAGE & LAW

Hashimoto & Heilpern, JLL 12 (2023): 48–72
www.languageandlaw.eu

Solving the Cherry-Picking Problem in Legislative History Use

– A Corpus-Based Approach for Empirical Intentionalist Legal Interpretation Analysis

Brett Hashimoto and *James Heilpern**

Abstract

The use of legislative histories under intentionalist/purposivist theories of statutory interpretation is frequently criticized because it can be easily biased (see, e.g., Scheppele, 2012). To date, corpus-based statutory interpretation has relied almost exclusively on textualist theory of legal interpretation. However, corpus linguistic methods are not necessarily bound to any one theory (e.g., Biber & Reppen, 2015). The present study analyses two legislative histories as corpora and compares them against a general corpus of English to determine if interpretative theory makes a meaningful difference in two example cases (Costello v. United States, 2012; Taniguchi v. Kan Saipan Pacific, 2012). Senses of relevant terms were manually annotated by two independent human coders with high interrater reliability in the two types of corpora. The results indicate that a legislative history corpus can reveal multiple patterns of lexical meaning and produce unbiased and distributional results rather than a single biased data point as most legislative history analyses do. These two case studies show significant and meaningful differences in both cases using Fisher's Exact Test (Costello, $p < 0.0001$, Cramer's $V = 0.70$; Taniguchi, $p < 0.0001$, Cramer's $V = 0.53$) between using a legislative history corpus versus a general language corpus. These results indicate that intentionalist/purposivist methods can be improved by using corpus-based analyses as well as the fact that intentionalist/purposivist and textualist theories produce practical semantic distinctions in legal interpretive settings due to the differences in relevant texts, registers, and speech communities.

Keywords

statutory interpretation, legislative history corpus, intentionalism, corpus-based legal interpretation, law and corpus linguistics, legal linguistics

Submitted: 31 December 2022, accepted: 26 July 2023, published online: 18 August 2023

* *Hashimoto*: Brigham Young University, brett_hashimoto@byu.edu; *Heilpern*: Judicial Education Institute, jheilpern@judicialeducation.org.

1. Introduction

The use of corpus linguistics as a tool to assist in the interpretation of law in the United States has been on the rise in recent years. For example, corpus-based evidence has been a part of several briefs submitted to Supreme Court cases (see, e.g., Baron et al., 2019; Strang, 2021; The Center for Constitutional Jurisprudence, 2021). In addition to these briefs, corpus-based reasoning has appeared in opinions in several other cases in a variety of lower courts, including various Circuit Courts of Appeal and state supreme courts (see, e.g., Caesar's Entertainment Corp. v. International Union of Operating Engineers, 2019; Richards v. Cox, 2019; State v. Lantis, 2019; State v. Misch, 2021; United States v. Scott, 2021; United States v. Woodson, 2021). There have also been a growing number of law review articles that use corpus linguistics (see, e.g., Cunningham & Egbert, 2020; Heilpern, 2018; Lee & Mouritsen, 2021; Lee & Phillips, 2019). Another major indication of corpus linguistics' growing importance in the field of legal interpretation is the fact that Justice Alito of the US Supreme Court has recently asked for corpus-based investigations of the canons of construction, widely applied maxims applied by courts in the United States to aid in the interpretation of legal documents (Facebook Inc. v. Deguid, 2021).

Corpus-based legal interpretation studies have almost universally relied on the same type of logical foundation. That is, these studies have relied on the premise that laws should be interpreted according to their ordinary meaning. Ordinary meaning – while ubiquitous in judicial opinions – is an underdeveloped concept in the law (Lee & Mouritsen, 2017). As two commentators have observed, there is "no ordinary meaning of ordinary meaning" (Lee & Mouritsen, 2017: 798). At least some of the time, judges imply that ordinary meaning refers to how an ordinary speaker of English would understand the words of the law. This is clear from the words of Justice Holmes of the U.S. Supreme Court who argued that "it is reasonable that a fair warning should be given to the world in language that the common world will understand, of what the law intends to do if a certain line is passed" (McBoyle v. United States, 1931). Because this definition of ordinary meaning has been central to the motivation of previous corpus-based legal interpretation studies, these investigations have virtually all made use of corpora designed to be representative of the language of ordinary American people (see, e.g., Lee & Mouritsen, 2017: 828–836 for more discussion). For example, in their 2017 Yale Law Review article, Lee and Mouritsen analyzed a series of legal cases – Muscarello v. United States (*carry a firearm*); Costello v. United States (*harbor*); and Taniguichi v. United States (*interpreter*) – using the collocate and KWIC functions of the Corpus of Historical American English (COHA) and News on the Web (NOW) corpora (Lee & Mouritsen, 2017). Gales and Solan also primarily used COHA, the Corpus of Contemporary American English, and the Google Books N-gram Viewer to analyze the Holy Trinity Church v. United

States case (*labor or service*), a 19th century case typically associated with a non-textualist approach to statutory interpretation (Gales & Solan, 2020)[1].

However, sometimes what is of interest to jurists is not just how the public would interpret the law, but rather, what legislators intended to accomplish by enacting it (see, e.g., Metropolitan Omaha Property Owners Association, Inc. v. City of Omaha, 2021; United States v. Acedvedo-De La Cruz, 2017). Laws are passed for a reason, and the purpose or intention behind a piece of legislation may be important for some in interpreting the language in a particular statute (Eskridge, 1987). Historically, one of the primary methods used in providing evidence for what the legislature's intentions by a statute is legislative history. Legislative history is the "background and events leading to the enactment of a statute", especially those documents produced by the legislature during the bill's drafting and debate, including "hearings [transcripts], committee reports, and floor debates" (Black & Garner, 2004: 919). However, the use of legislative history has been criticized because one can cherry-pick examples that make their arguments look most favorable and choose to ignore other examples that may indicate other potential meanings of a contested term.

A few scholars have suggested that corpus linguistics could be used as a means of approximating Congressional intent. In an amicus brief submitted to the United States Supreme Court, Heilpern used a corpus of Supreme Court opinions to approximate "a distinct dialect-of-art we might call *legalese*." (Heilpern, 2019). Likewise, Gales and Solan created a corpus of United States Statutes-at-Large to look for patterns in statutory language generally (Gales & Solan, 2020). In their most recent article, Lee and Mouritsen used the Corpus of Current U.S. Code to do the same (Lee & Mouritsen, 2021).

The purpose of the present study is to outline a method for a more rigorous, transparent, replicable, and valid method of using a legislative history to provide less biased and more useful and comprehensive evidence about the legislature's intended meaning in statutory language. We plan to accomplish this by creating corpora of legislative histories. By randomly sampling hundreds of instances of how a contested term is used, we hope to be able to provide additional nuance about the intended meaning of those contested terms. We also hope to provide evidence that using a legislative history corpus can provide an understanding about the meaning of contested terms in statutes that a more general corpus could not. This will be accomplished by comparing the results of the legislative history corpus analysis against an analysis of a corpus of general language.

[1] It is worth noting that Gales & Solan (2020) also relied on a custom corpus of United States Statutes-at-Large to look for patterns in statutory language generally.

2. Literature Review

The debate over the use of legislative history has historically boiled down to a fundamental disagreement over the nature of the judicial inquiry. That is, when interpreting a statute, should judges be searching for the "intent" behind the legislative body's enactment of the law, or should they focus on giving effect to the actual text that was passed? Proponents of the former – called "intentionalists" or "purposivists" – assume that "legislation is a purposive act, and that judges should construe statutes to execute that legislative purpose" (Katzmann, 2014: 31). They therefore rely on legislative history to illuminate the "mischief" that Congress was trying to remedy (Driedger, 1981). At intentionalism's height in the 1960s and 1970s, reliance on legislative history was actually *more* important to judges than the actual text of the statute. For example, in Citizens to Preserve Overton Park v. Volpe (1970: Headnote 5), the Supreme Court noted that it was only *because* "the legislative history [of the relevant statutes] is ambiguous" that the Court "must look primarily to the statutes themselves to find the legislative intent". Also, Judge Wald noted in 1981, almost every statutory case the Supreme Court decided relied on legislative history: "Not once last Term was the Supreme Court sufficiently confident of the clarity of the statutory language *not* to double check its meaning with the legislative history" (Wald, 1982: 197).

However, this approach began to be challenged in the 1980s. For instance, Judge Easterbrook of the U.S. Court of Appeals for the Seventh Circuit took issue with the idea that a multi-member, political divided, diverse body like Congress could have a collective intent at all (Easterbrook, 1988). There is no guarantee that all of the legislators who voted for a particular law did so for the same reason or were even trying to remedy the same problem. The only thing that the majority agreed on was the *language* of the statute itself. Consequently, Easterbrook (1988: 65) argued that "[t]he meaning of statutes is to be found not in the subjective, multiple mind of Congress but in the understanding of the objectively reasonable person". Others questioned the good faith of those judges employing legislative history. Justice Scalia of the US Supreme Court called it "the last hope of lost interpretive causes, that St. Jude of the hagiology of statutory construction" (United States v. Thompson, 1992: 505). Likewise, D.C. Circuit Judge Levanthal accused judges of cherry picking from the legislative record to reach their preferred outcomes, comparing the use of legislative history to "looking over a crowd and picking out your friends" (quoted in Samaha, 2017: 556).

These criticisms worked; over the last three decades, intentionalism has faded as a legal interpretive theory, gradually being replaced by textualism as the dominant interpretive philosophy favored by judges. Accordingly, citations to legislative history in judicial opinions has decreased dramatically, as well. Gone are the days where judges began their inquiry with a review of the legislative history. If legislative history is consulted at all, it is to try to clarify an ambiguity in the text. In other words, the text rather than

the intent is still what matters. As Justice Kagan stated, "we're *all* textualists now" (Harvard Law School, 2015).

However, textualism is not without its own critics. A number of scholars have argued that textualism as a theory is "overly malleable, and thus not much different than its main competitor [intentionalism]" and insist that the theory is simply a "smokescreen by conservative judges to reach ideologically acceptable outcomes" (Grove, 2020: 265). While is it possible that this is true, there is little evidence that textualist judges are not sincere in their jurisprudential commitments. Linguists have also noted issues with textualist methods as well. The tools that judges traditionally rely on to analyze the text and discover the ordinary meaning of the law – including analyses using intuition, dictionaries, etymology, morphology, and the so-called "canons of construction" for sense disambiguation – have been shown to all be deeply flawed (Gries, 2020). For example, Mouritsen (2010: 1926) has noted that when utilizing dictionaries, courts often fall victim to the "sense-ranking fallacy" – the assumption that more common senses of a word are listed first in the dictionary. Judges frequently interpret contemporary meanings of words by examining their meanings or the meaning of its morphemes in Latin or Greek (Gries, 2020). However, what is most troubling perhaps is that when faced with ambiguity or vagueness in the meaning of terms in the law, judges more frequently rely on their intuitions (Gries, 2020; Lee & Mouritsen, 2017; Mouritsen, 2010). Thus, even the most genuine textualist judges might interpret a legal document incorrectly because the available tools cannot accomplish the kind of semantic analysis that they wish to perform.

Combatting these operationalization errors is precisely the reason that many textualist judges have been quick to adopt corpus linguistics in recent years; they view it as a tool – in the words of Justice Lee of the Utah Supreme Court – to better "deliver on the promises of textualism" (Federalist Society, 2017: 23:20). They laud its transparency and replicability (Wilson v. Safelite Group Inc., 2019). This is especially true of those judges who believe strongly that the law should provide the regulated public with a "fair notice" of what conduct is permitted and prohibited. As a result, it is unsurprising that virtually *all* of the judicial opinions that have utilized corpus linguistics thus far have employed more general language corpora or corpora more so designed to represent the language of ordinary American citizens such as the Corpus of Contemporary American English (Davies, 2009), the Corpus of Historical American English, or the Corpus of Founding Era American English (Hashimoto, 2023).

Thus, most corpus-based legal interpretation studies have adopted a textualist lens. However, corpus linguistics has the potential to be equally helpful to intentionalists as well. It is simply a method, and it although it has been applied almost exclusively in textualist-informed analyses, it could be applied by intentionalists as well with the same benefits of transparency, replicability, and validity that corpus methods have contributed to textualist analyses. In fact, corpus linguists often describe corpus methods as independent of any particular interpretative or linguistic theory (Biber & Reppen, 2015).

Although primary reliance on intentionalism as the predominant legal interpretive theory has fallen out of vogue, many judges continue to cast their textual analysis in intentionalist terms. For example, the U.S. Court of Appeals for the Third Circuit has repeatedly stated that "[o]ur task is to give effect to the will of Congress, *and where Congress's will has been expressed in language that has a reasonably plain language*, that language must ordinarily be regarded as conclusive" (see, e.g., Byrd v. Shannon, 2013: 122, emphasis added). Such formulations of the ordinary meaning canon suggest a greater interest on the part of these judges in legislative use of the language of the law, especially in the discourse concerning that law than the public's understanding of the language of the law. In other words, these judges focus on a different speech community than their "fair notice" colleagues – the legislature itself.

We hypothesize that due to the demographic differences between legislators (especially those in Congress) and the body politic as a whole, their language usage will differ in some respects as well. More specifically, the population of Congress has been shown to be very different from the population of the United States in that they have a much higher average age and level of education as well as very different distributions of religious affiliation, occupational backgrounds, religious affiliations, and ethnicities (Manning, 2017). All of these variables have been demonstrated to contribute to differences in language use (see, e.g., Baker & Bowie, 2010; Forrest & Dodsworth, 2016; Fought, 2006; Haeri, 1997; Rampton et al., 2008; Zilles, 2005).

Accordingly, a general language corpus such as COCA or COHA may not be appropriate for these judges. Instead, legislative history – consisting of documents produced by Congress and their staff specifically about a certain statute – may be a better proxy for analyzing Congressional language usage. After all, even Judge Easterbrook has conceded that legislative history can serve as a type of lexicon of legislative speech:

> Clarity depends on context, which legislative history may illuminate. The process is objective; the search is not for the contents of the authors' heads but for the rules of language they used. (In re Sinclair, 1989: 1343)

Corpus linguistics can empower judges to do this in a systematic way that is both transparent and a replicable, effectively ameliorating the problems identified by Justice Scalia and Judge Levanthal.

To test this hypothesis, two case studies were conducted based on statutory interpretation cases that have been discussed at length in the existing law and corpus linguistics literature which have been analyzed repeatedly using general purpose corpora: *Costello v. United States* (2012) and *Taniguchi v. Kan Pacific Saipan, Ltd.* (2012).

Costello v. United States (2012), hereafter *Costello*, was a criminal law case. The defendant in *Costello* was a woman who allowed her boyfriend – who she knew was an undocumented immigrant – to move in with her. When her boyfriend was arrested and convicted on both drug charges and returning to the United States after having been removed, she too was charged for having violated a federal statute, namely, 8 U.S.C. § 1324(a)(1)(A)(iii) which made it illegal for anyone to

> knowing or in reckless disregard of the fact that an alien has come to, entered, or remains in the United States in violation of the law, conceals, harbors, or shields from detection, or attempts to conceal, harbor, or shield from detection, such alien in any place, including in any building […]. (Bringing in and harboring certain aliens, 2012)

Because there was no evidence that she had attempted to hide or conceal her boyfriend, the judges determined that the case turned on the definition of the word *harbor*. Specifically, the outcome of the case was determined by whether the scope of the word *harbor* meant to simply provide shelter or if it involved an effort to conceal a person. While the district court judge found the defendant guilty, the Seventh Circuit reversed the decision by distinguishing between harboring and providing shelter without concealing or shielding from protection.

Taniguchi v. Kan Saipan Pacific, Ltd. (2012), hereafter *Taniguchi*, arose out of a personal injury lawsuit brought by a professional baseball player from Japan against the owner of a resort in the Northern Mariana Islands. In preparation for its defense, the resort paid to have a series of documents translated from Japanese into English. When the resort prevailed in the district court, it asked the court to require the baseball player to cover these translation fees, pursuant to the federal costs statute, namely 28 U.S.C. § 1920(6), which allows the district court to order the losing party to reimburse the winner for any "compensation of interpreters" (Taxation of costs, 2012). The question for the Court to decide was whether the word *interpreter* extended to a person who translates written documents or if it is restricted to spoken communication only. In this case, the district court chose to include costs associated with the translation of written documents into the compensation for interpreters, which was later upheld by the Ninth Circuit. However, the case was appealed, and the Supreme Court reversed the decision with a 6:3 majority, indicating that the ordinary meaning of *interpreter* is someone who translates orally from one language to another, and thus, document translation should not be included.

3. The Present Study

As mentioned previously, the purpose of this study is to propose a more robust methodology for analyzing legislative histories for statutory interpretation using corpus methods. In the present study, we conducted two analyses with parallel designs. In each analysis, we compared differences between results from using two types of corpora for statutory interpretation: a legislative history corpus and a corpus designed to represent general language. In the first analysis, we compared a legislative history corpus for *Costello* against a ten-year cross-section from the Corpus of Historical American English from 1912–1922 (COHA) (Davies, 2012). In the second analysis, we compared a legislative history corpus for *Taniguchi* compared again a ten-year cross-section from COHA from 1973–1983. Through these two analyses, we plan to respond to the following research questions.

Research questions (RQ):

- RQ 1: To what extent is additional information gained by examining many examples from a legislative history as opposed to only one example?
- RQ 2: To what extent is there a difference between the legislative discourse surrounding the formation of a piece of legislation and general language of the public for corpus-based legal interpretation?

4. Methods

4.1. Corpora

The corpus used as the general corpus, which we used to represent the language of ordinary American citizens was COHA. Presently, COHA contains over 475 million words across more than 100,000 texts. Details about the composition of the corpus can be found at english-corpora.org/coha. However, for the present study, not the entire corpus was used. In order to represent language from around the relevant time period of each of the relevant statutes, a five-year window on either side of the year of enactment/addition of the relevant term was used. For example, as the contested term *harbor* was first added to the US Code in 1917, we subsampled from COHA from the years 1912–1922. COHA contains five registers. The TV/Movies register begins in the 1930s, which is why there are no texts for the 1912–1922 COHA subcorpus. The TV/Movies register was derived from sampling from the OpenSubtitles database (Lison & Tiedemann, 2016: opensubtitles.org). Fiction texts were sampled from several sources: Project Gutenberg, Making of America, scanned books, and movie and play scripts. Magazines were also sampled from several sources, including Making of America and scanned magazines. Non-fiction is likewise from several sources: Project Gutenberg, archive.org, and scanned books. A breakdown of the composition of the two corpora derived from COHA used in the present study for the relevant time periods are shown in Table 1 and additional detail as well as a full list of sources can be found at English-corpora.org.

Table 1: Description of subsamples of COHA used as general corpora in the present study.

Years	# of	TV/ Movies	Fiction	Magazines	News	Non-fiction	Total
1912–	words	0	20,621,260	5,813,594	2,098,166	3,905,521	32,438,541
1922	texts	0	449	1,249	2,447	113	4,258
1973–	words	1,935,249	12,788,629	6,527,401	3,991,346	2,855,679	28,098,304
1983	texts	249	366	6,139	4,114	93	10,961

In addition to COHA, two additional corpora were constructed for this study, which were legislative histories for *Costello* and *Taniguchi*. The United States Code Service on LexisNexis (LexisNexis, 2022) contains a history section for each US statute which contains a list of public law numbers related to the code. Once the relent public law numbers were identified, the legislative histories were constructed by using the Legislative Insight tool in Proquest (ProQuest, 2022). Proquest claims the validity of their legislative histories as follows:

> Each history includes the full text of the public law itself, all versions of related bills, law-specific Congressional Record excerpts, committee hearings, reports, and prints. Also included are presidential signing statements, CRS reports, and miscellaneous congressional publications that provide background material to aid in the understanding of issues related to the making of the law.

Each relevant statute was queried according to its public law number and all related documents resulting from the query were included. For the Costello Legislative History Corpus (CLHC), PL 82–414, PL 95–582, PL 97–116, PL 99–603, PL 100–525, PL 103–322, PL 104–208, PL 106–185, PL 108–458, and PL 109–97 were queried, and for the Taniguchi Legislative History Corpus (TLHC), PL 80–773, PL 95–359, and PL110–406 were queried and the results were downloaded as PDF documents. Duplicate documents within each corpus were removed from consideration, so that they would not be counted twice.

After gathering the relevant documents, they were all converted to a plain text format using Adobe OCR (Optical Character Recognition) software. The overall quality of the OCR process was sometimes quite poor because many of the documents were scanned PDF images. The documents underwent some systematic cleaning to remove or replace non-UTF-8 characters as well as to fix commonly reoccurring OCR errors using Python v.3 (Van Rossum & Drake, 2009). This resulted in 1,664 texts (67,118,848 words) for the Costello Corpus and 109 texts (2,717,309 words) for the Taniguchi Corpus.

4.2. Query Methods and Concordance Sampling

The queries in COHA were relatively straightforward. In each case, the flemma for the contested term (*harbor* and *interpreter*, respectively) was queried using the English-corpora.org interface (Davies, 2012). Flemmas are similar to lemmas except that they also group in identical orthographic forms that are different parts of speech along with the inflections of those forms (Pinchbeck, 2014). For example, the word *bottle* can be both a noun and verb. So, a flemma for this word would include singular and plural forms as well as the verbal forms of *bottling*, *bottles*, and *bottled*. Flemmas were selected over lemmas to allow for the possibility that the words were mistagged to an incorrect part of speech or were not converted into characters correctly. For instance, the *i* or *t* in *interpreter* was often converted as *1* or *l*. OCR quality of the legislative history corpus severely affected the accuracy of automated part-of-speech tagging programs, and therefore, methods of lemmatization. The accuracy of multiple programs that were trialled (viz.,

CLAWS7, Stanza, and TreeTagger) was so inaccurate that it was more work to try and fix the tags than to simply modify our queries and filter out irrelevant observations. Because the overall OCR quality of the legislative history corpora was quite poor, it was important to adapt the querying approach in the legislative history corpora to be as flexible as possible. Using a custom Python script, each corpus was queried for approximate matches for the target flemmas. This was done by querying every possible sequence of three or more characters in the target word. For instance, for the word *interpreter*, the queries were *int, nte, ter, erp, ret, ete, ter*. Concordances were generated for each occurrence 300 characters on either side of the query result. Redundant concordances were deleted based on flexible matches of the concordances, files, and location within the file. Each concordance was then manually reviewed to get rid of any false positive results (i.e., any case where the query returned a result that was not the target flemma). After the results were manually reviewed, 500 concordances were randomly sampled using the random library in Python, which were then manually coded for their meaning.

4.3. Data Coding

Each concordance line was assigned a code by two trained and independent coders. The scheme used to evaluate the meanings in the *Costello* analysis is described in Table 2; the scheme used to evaluate the meanings in the *Taniguchi* analysis is described in Table 3. These coding schemes were developed over several rounds of piloting and coder calibration. After a round of coding, if coders did not achieve a 95 % raw agreement and .9 kappa statistic, all lines were recoded. Reliability between the coders was calculated using raw percentage agreement and unweighted Cohen's Kappa using the *R* package *irr* (Gamer et al., 2012; R Core Team, 2013). In the event of disagreement between the coders, the concordance line was discussed together to see if the disagreement could be resolved. When coders did not agree, discrepancies were discussed to see if modifications to the coding scheme were necessary. Coders were instructed to read through the whole concordance line, to use best judgment when it comes to poorly OCRed texts, and to consider whether alternative interpretations were possible for each line. For the final analysis, the coding of the first author was relied upon because the second coder achieved the desired level of agreement in all cases. In order to achieve this level of agreement, the data underwent three iterations of coding.

Table 2: Coding scheme for the Costello analysis.

Code	Description
Concealing a person	You feel strongly that it could only mean that a person was being hidden from others.
Sheltering a person	You feel strongly that it could only mean that a person was being sheltered.
Both concealing and sheltering a person	You feel strongly that it could only mean both that a person was being hidden from others and sheltered.
Either concealing or sheltering a person	You believe that it could mean either concealing or sheltering a person. There is not enough information for you to be sure about which of the two meanings is intended, but you feel confident that it is one of these two meanings and not another meaning.
Another sense that does not have to do with either concealing or sheltering a person	You feel strongly that it does not mean either concealing or sheltering a person.
No category	No category in the coding scheme fit well, it was not being used a verb, or the line was not readable/understandable/did not have enough context.

Table 3: Coding scheme for the Taniguchi analysis.

Code	Description
Written translation only	You feel strongly that it could only mean that the language was being written by the interpreter. This could mean that the translation was from spoken to written or written to written.
Spoken translation only	You feel strongly that it could only mean that the language was being spoken by the interpreter. This could mean that the translation was from spoken to spoken or written to spoken. This includes signing.
Both written and spoken translation	You feel strongly that it means both written and spoken translation.
Either written or spoken translation	You believe that it could mean either written or spoken. There is not enough information for you to be sure about which type of translation is occurring.
Another sense that is not translation	You feel strongly that it is another meaning that does not involve rendering one language to another.
No category	No category for a meaning in the coding scheme fit well, it was not being used as a noun, or the line was not readable/understandable/did not have enough context.

4.3. Analysis

To summarize the results obtained from each corpus, raw counts of each code were taken and proportions of the total for each code of all concordances lines was calculated. Raw counts and proportion distributions were first compared within each corpus and then compared between the general and their respective legislative history corpora. "No category" results were not considered in the final analyses since they were not relevant or not interpretable.

Dispersion of the senses was also measured using range (i.e., the number of different texts in which a sense occurred at least once from the coded concordance lines). Dispersion is useful for answer RQ 1 in helping to ascertain whether instances of a given sense were all from one text or whether they were derived from many different writers/speakers across many events.

The general and their respective legislative history corpora were then compared using a Fisher's Exact Test to determine if there is a statistical difference in result depending on the type of corpus one selects. Fisher's Exact Test was selected over a Chi-squared test because the data did not meet the assumptions of a chi-squared test. The assumptions of Fisher's Exact Test of independence and fixed totals were checked and met. Cramer's *V* was also calculated as a measure of effect size to determine the magnitude of the differences between corpora. We rely on Cohen's (1988) recommendations for interpreting effects sizes. The concordance lines were qualitatively analyzed in order to explain the patterns observed in the corpora by examining whether there was any correspondence between syntax, cotext, register, or topic that could easily predict the meaning of the contested term that could shed additional light into interpreting its meaning in the statute.

5. Results

5.1. Descriptive Results

Three hundred and twenty-nine observations were removed from consideration from CLHC, and 13 were removed from the corresponding COHA corpus (1912–1922) because they did not match the correct part of speech. Eight observations were "no category" in TLHC all because of OCR issues and were removed from consideration from the rest of the analyses.

The results of the descriptive analyses showed marked differences in the distribution of senses between the legislative history corpora and the general language corpora (see Tables 4 and 5). In the case of the *Costello*, using a legislative history corpus made the likelihood of observing senses relevant to the case much more likely than in the general corpus. Of the 137 instances from COHA, 97 were not relevant to this case. These included instances of *harbor* being used with thoughts, feelings, pathogens, or other non-animate/non-human entities as in the following examples (*harbor* in bold):

> Runner was inclined to **harbor** resentment and refused to answer me. Hacker, however, readily informed me: (FIC_A Virginia Scout_1922_COHA)

> More strangely still, it is found that some breeds and some animals, and even some persons, may **harbor** the germs of a given infection without any disagreeable effects. (MAG_The reservoirs of contagion_1912_COHA)

Table 4: Counts and range from concordance line coding in Costello analysis.

Code	TLHC Freq.	COHA Freq.	TLHC Disp.	COHA Disp.
Concealing a person	1	2	1	3
Sheltering a person	64	69	10	48
Both concealing and sheltering a person	34	0	4	0
Either concealing or sheltering a person	393	58	28	46
Another sense that does not have to do with either concealing or sheltering a person	0	24	0	22

When considering only the relevant senses from the *Costello* analysis, the legislative history corpus referred proportionally more to a sense indicating that *harboring* involved exclusively concealing (and not sheltering) a person and had many more instances of an ambiguous context. On the other hand, the general use of the word in COHA more frequently connotated exclusively 'sheltering' and never exclusively 'concealing' a person.

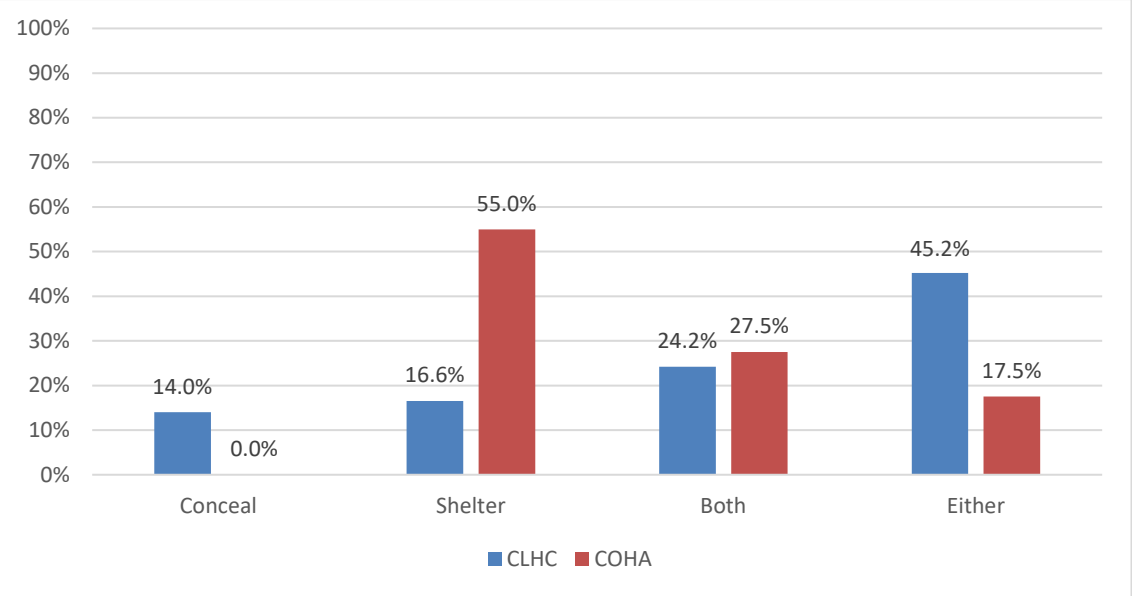

Figure 1: Proportions of relevant senses from Costello analysis based on frequency. Note that this figure reflects the proportions of only those categories contained within the figure.

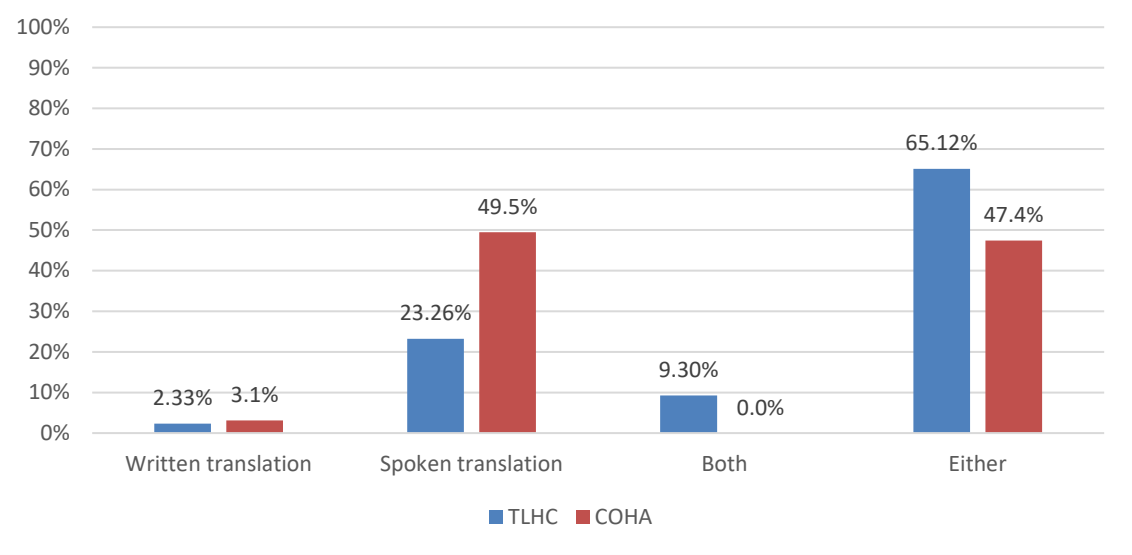

Figure 2: Proportions of relevant senses from Costello analysis based on range. Note that this figure reflects the proportions of only those categories contained within the figure.

In the case of the *Taniguchi* analysis, an *interpreter* in the legislative history corpus always referred to a relevant sense of the word, whereas in the general corpus, there were instances of interpreters of art, including music, dance, literature, and cinema; of the law; or other senses not related to the transferring of a message from one language to another as in the following examples (*interpreter* in bold).

> Manhattan's Philharmonic Hall was crammed as Pianist Earl ("Fatha") Hines, Singer Mabel Mercer, Saxophonist Gerry Mulligan and other **interpreters** jazzed the songs of composers like George Gershwin and Cole Porter. (MAG_Jazz, by George_1973_COHA)

> Welcome to Elysia. Devna, **Interpreter** of Laws. Gentlemen, you now stand before the Ruling Council. (TV/MOV_Star Trek_1973_COHA)

Table 5: Counts and range from concordance line coding in the Taniguchi analysis.

Code	CLHC Freq.	COHA Freq.	CLHC Disp.	COHA Disp.
Written translation only	22	0	18	0
Spoken translation only	26	22	22	19
Both written and spoken translation	38	11	29	10
Either written or spoken translation	71	7	49	7
Another sense that is not translation	14	97	11	77

In the *Taniguchi* analysis, the legislative history corpus had proportionally more instances of the use of the word being ambiguous or definitively both, whereas the general language corpus had many more instances of the word meaning exclusively spoken translation.

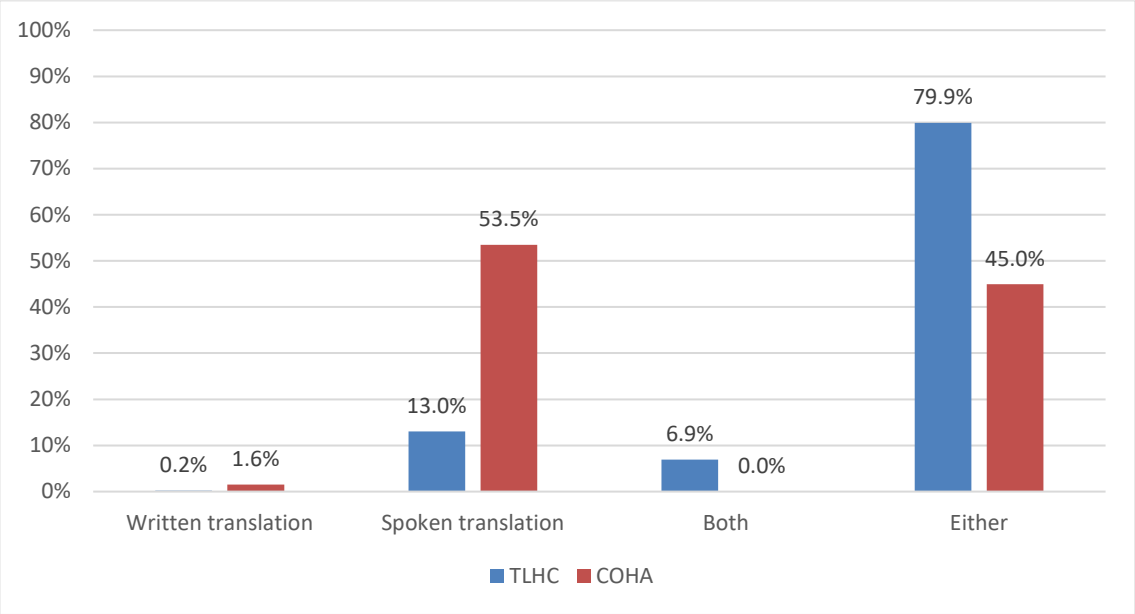

Figure 3: Proportions of relevant senses from Taniguchi analysis based on frequency. Note that this figure reflects the proportions of only those categories contained within the figure.

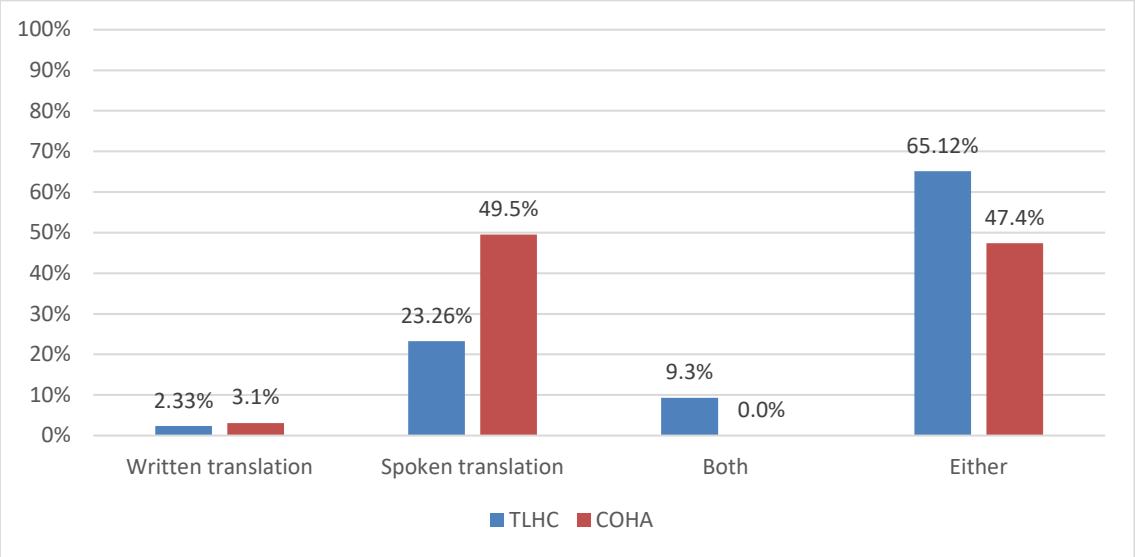

Figure 4: Proportions of relevant senses from Taniguchi analysis based on range. Note that this figure reflects the proportions of only those categories contained within the figure.

In both cases, the legislative history corpora contained all possible relevant sense codes. Also in both cases, the legislative history corpus resulted in markedly different results to the general language corpus in the distributions of the codes overall.

5.2. Statistical Tests

In conducting more formal statistical tests, the differences between the results of the legislative and general language corpora are significant in all cases. In both *Costello* Fisher's Exact Tests, a statistically significant result and moderately strong effect size was obtained (see Table 6), indicating a meaningful difference between the two types of corpora overall. In both *Taniguchi* Fisher Exact Tests, a statistically significant result with a moderate effect size was obtained (see Table 6), indicating a meaningful difference between the two types of corpora overall.

Table 6: Fisher's Exact Test results

	p	*v*
Costello Frequency	< 0.001	0.70
Costello Dispersion	< 0.001	0.67
Taniguchi Frequency	< 0.001	0.53
Taniguchi Dispersion	< 0.001	0.40

6. Discussion

6.1. RQ 1: To what extent is additional information gained by examining many examples from a legislative history as opposed to only one example?

The results from the analysis of senses from both legislative corpora indicate that within the context of a single legislative history, there can be great variation in the use of a word. Tables 4 and 5 show that within a legislative history all senses of interest could be found. In other words, any jurists with motivated reasoning could find examples of a definition of the word in question that suited their interest in either of these two cases. In fact, in the *Costello* analysis, the legislative history corpus revealed many instances of each of the senses. Consider these illustrative examples of contrasting uses of the word *harbor* that could have been used to support either side of the argument (*harbor* in bold).

> I am advised by the staff that what **harboring** means […] – what **harboring** means is concealing persons from detection, persons who are here without proper documentation. (CR19840613PL99603H.txt)

> That any person, including the master, agent, owner, or consignee of any vessel, who shall bring into or land in the United States […] or shall conceal or **harbor**, or attempt to conceal or **harbor**, or assist or abet another to conceal or **harbor** in any place, including any building, vessel, railway car, conveyance, or vehicle, any alien not duly admitted by an immigrant Inspector […] (CMP1952SJS0003.txt)

In the former example, the speaker is clearly indicating that their understanding of the word is that it means 'concealing'. In the latter example, the alternative connective *or* is

repeated three times separating *conceal* and *harbor,* strongly implying that the two are separate actions or, rather, that one does not entail the other.

Likewise, in the *Taniguchi* analysis, contrasting senses were also found as in the following examples (*interpreter* in bold).

> After any such determination has been made, each party to the proceeding shall be entitled to utilize the services of the **interpreter**, certified pursuant to section 604 2 (a) of this title, to provide a simultaneous translation of the entire proceeding […] or to the translation of such document, from such non-English language to English (HRG1975HJH0017.txt)

> some **interpreters** will be instructed to interpret only the testimony of the non-English speaking person to Judge, Jury, court reporter, and litigating counsel. When the non-English speaking person Is not testifying, the Interpreter ls frequently excused from the courtroom

In the former example given, interpreters are clearly meant to provide both spoken and written services. In the latter, the interpreter is instructed to only interpret spoken testimony or else they are excused from the courtroom, indicating that their work is exclusively spoken.

So, theoretically, one could come up with many examples to help them whatever their cause, even though the sense that was being advocated for might not have been the most prevalent one, a prevalent one, or even a common one. In the *Taniguchi* analysis, the 'written translation only' sense of the word was found only one time. Thus, if that one instance were solely relied upon for interpreting the statute, it could be very different from what the legislature actually intended by using the word *interpreters*. What is worse is that it could lead to "strategic insertions designed to produce a judicial interpretation that did not have enough votes to be written into the statutory text" (Eskridge, 1990: 381). For example, the D.C. Circuit "condemned the well-recognized phenomenon of deliberate manipulation of legislative history at the committee level to achieve what likely cannot be won before Congress as a whole" (Federal Election Commission v. Rose, 1986), and Dickerson (1984: 82) notes that the reliability of the use of legislative history is "undermined by the widespread practice, at least in Congress, of allowing legislators to amend or supplement their remarks in the published version of the Congressional Record". Such instances of legislators "planting" desired meanings in the legislative history have been discovered as being of consequence in practice as well. Meltzer (1979: 441) elaborates on this issue in discussing United Steelworkers v. Weber (1979), where the majority of Supreme Court Justices relied heavily on only two lines of legislative history in their ruling that were "so far out of step with the rest of [the statute]'s voluminous legislative history that [they are] not entitled to significant weight". Clearly, the method used in the present study cuts through these kinds of issues with how legislative histories have been used in the past.

What is also apparent in this analysis is that Congress is not a monolith. Eskridge (1990) calls this issue the "problem of aggregating intent". Eskridge (1990) points out that Congress is a collection of hundreds of members across two different bodies (the House and Senate), and as such, these various people have differing backgrounds, ideologies,

concerns, objectives, and constituencies, which all in turn affect their intent in passing legislation. To reduce the intent of so many with only a handful of examples from a legislative history could potentially greatly misrepresent what the general consensus of the two legislative bodies was or else at least falsely propagate the idea that there was only one intent behind the passage of a piece of legislation rather than the much more likely reality, which is that there was a diversity of reasons.

What the frequency analysis reveals in the *Costello* case is that according to the legislative history, either interpretation was plausible. While the most likely probability was that the use of the word was ambiguous (45.2 %), 16.6 %, one out of six uses of the word strongly indicated that the meaning was exclusively 'sheltering' and did not necessarily imply 'concealing' and 24.2 % of the occurrences strongly indicated that 'sheltering' was included in its meaning in addition to 'concealment' (Figure 1). Thus, 40.8 % of the uses indicated that 'concealment' was not necessary to constitute *harboring*. On the other hand, 14.0 % of the occurrences strongly indicated that *harboring* was exclusively 'concealment' and did not necessarily extend to 'sheltering', so both interpretations are plausible. The *Costello* dispersion analysis reveals a quite similar pattern as the frequency analysis although with a slightly lower proportion of ambiguous cases and a higher proportion of the other categories (Figure 2).

The *Tanguchi* frequency analysis results are equally elucidating. While most occurrences were ambiguous (79.9 %), whenever it is clear, it is overwhelmingly meant 'speech' (13.0 %) and not 'writing' (0.2 %) (Figure 3). However, if we take the total of the occurrences of those which strongly indicated 'both speech and writing' as well as only 'writing', 7.1 % of the occurrences indicated that writing should be included in the cost of the interpreters. Here also, one sense was dominant, but the alternative is also plausible. These two cases illustrate that patterns and distributions emerge when looking at many examples, confirming the oft-quoted adage that "language looks rather different when you look at a lot of it at once" (Sinclair & Sinclair, 1991: 100). The *Taniguchi* dispersion analysis reveals a similar pattern but with a far smaller proportion of ambiguous codes (65.12 %) and much higher proportion of the spoken sense (23.26 %) (Figure 4).

The advantage of using a corpus-based approach is that is allows for the analysis of the entirety of expressed intentions of Congress through an unbiased sampling of every available part of the legislative history. Rather than a clear-cut answer, the analysis confirms that the confusion of the courts in understanding these words is justified, that one sense is more prevalent than the other, but also that both senses are possible. Even with this kind of analysis, it remains possible for a single person or even group of legislators to strategically insert their intent into law through legislative history, it mitigates this bias. Furthermore, if a more sophisticated dispersion were incorporated, this potential area for bias could be greatly reduced (see, e.g., Keller & Egbert, 2020).

6.2. RQ 2: To what extent is there a difference between the legislative discourse surrounding the formation of a piece of legislation and general language of the public for corpus-based legal interpretation?

In all comparisons, (for both frequency and dispersion), the legislative history and the general corpus analysis revealed significantly different distributions of senses with moderate or strong effect sizes, meaning that the results of the analyses from the two types of corpora lead to quite different results and, therefore, potentially lead to different conclusions in statutory interpretation. In the case of Costello, a number of differences between the two datasets can be observed. When considering only relevant instances of the word in the frequency analysis, general language use has proportionally many more instances where the verb *harbor* meant only 'sheltering' a person (55.0 %) than the legislative history (16.6 %). Another interesting difference is: Whereas 14.0 % of the occurrences of the word meant exclusively conceal in the legislative history in the frequency analysis (18.6 % in the dispersion analysis), not a single occurrence of that sense was observed in general language.

Thus, in this case, if a judge were a textualist who followed the ordinary meaning doctrine, that person would likely come to quite different conclusions about what the contested term meant than an intentionalist would if the same empirical approaches were applied. The textualist would be inclined to believe that sheltering a person was almost always included in the ordinary meaning of the term because that is what 82.5 % of the occurrences from COHA strongly implied (i.e., the categories of 'sheltering a person' and 'both concealing and sheltering a person' combined) and the term never exclusively means to conceal a person. Therefore, the textualist might determine that the ordinary meaning of the word should definitely include shelter and rule against the defendant. On the other hand, the intentionalist would be inclined to believe that *harboring* only sometimes means sheltering because that sense is strongly implied less at less than half the rate as ordinary language (40.8 % of the occurrences) and may therefore come to a different conclusion about the meaning of the term than the textualist. It is worth noting, though that there were similarities in the one category. Both the frequency and dispersion analyses had a somewhat similar proportion of occurrences that implied both senses. Other than that, however, there proportions were quite different in every other category between the legislative history and general corpora.

The results of the *Taniguchi* analysis also revealed some interesting and meaningful differences between the two types of data. Both datasets indicate that written conversion from one language to another represent a minority of occurrences. However, the contrast is more stark in the general language analysis of COHA, where only 1.6 % of the time does it mean written (i.e., the categories of 'written translation only' and 'both written and spoken translation' combined). Thus, although it might be possible in ordinary language, it is perhaps unlikely that that costs of an interpreter necessarily implies writ-

ten translation. The legislative history analysis, on the other hand, shows a less definitive disparity with 13.0 % of the cases indicating that only 'spoken translation' was strongly implied to the 7.1 % of the cases where 'written translation' was strongly implied.

Here also, jurists might come to different interpretations of the data depending on their jurisprudential theoretical orientational. Based on the results of the corpus analysis, the textualist judge would probably understand that while the written translation is rarely implied if at all, especially when compared to spoken translation. The intentionalist judge would probably be led to understand that both interpretations are highly plausible with one interpretation being less than 6 % more likely than the other. Therefore, from both of these cases, it is clear that the language of legislative history can be markedly different than ordinary language use, which empirically confirms a substantial difference in outcomes based on jurisprudential ideology.

It is unsurprising that these differences manifest themselves in this data. After all, the nature of the texts within the two corpora differ in registers along many important situational characteristics, which of course leads to linguistic differences (see, e.g., Biber & Conrad, 2019). Two major situational differences are (1) the demographics of the speakers/writers and audiences and (2) the purposes of the texts. After all, Congress is fundamentally different from the general citizenry in practically every major demographic category known to cause socio-linguistic differences among speakers, including socio-economic status, age, ethnicity, gender, level of education (Manning, 2017). This fundamental difference in the social dimension of language use is bound to have an effect on the language use within legislative histories to differentiate them from ordinary language use. Also, the purposes of the language contained within legislative histories is much more focused and singular than general language. Congressional committee reports, for instance, are focused on summarizing bills that are being proposed to brief the rest of congress and may contain persuasive arguments in favor of or against the bill (Parrillo, 2013). Colloquys are transcriptions of floor debates where members of congress argue for changes to a bill, propose reasons to vote in favor of or against a bill, indicate their personal stance, and persuade others to agree with them (Costello, 1990). Most registers in COHA represent more ordinary language, such as fiction novels or news articles. These registers are not nearly so focused on the purposes of persuasion and establishing stance.

DOI: 10.14762/jll.2023.048

7. Conclusion

The present study's findings advance research into statutory interpretation in several ways. First, in answer to RQ 1, it empirically confirms a concern that many have expressed that cherry-picking jurists can find many different senses of a word in both legislative histories or general language use if they simply look hard enough. Thus, anybody with motivated reasoning can make compelling arguments with actual examples to support their argument almost regardless of whatever that perspective may be as long as it is reasonable or plausible. Second, in response to RQ 2, this study empirically demonstrates that textualist and intentionalist approaches can result in differing objective understandings of what a disputed term means. Furthermore, just as corpus linguistic methods have helped in "delivering on the promise of an objective inquiry into ordinary meaning" for textualists (Lee & Mouritsen, 2017: 127) by making it more systematic, transparent, empirical, replicable, valid, and therefore, more generally scientific, this study demonstrates that the same is also possible in making deliver on the promise of intentionalism as well. If judges really want to subscribe to intentionalist or purposivist theories of legal interpretation without being accused of "picking out [their] friends" (quoted in Samaha, 2017: 556), there is a methodologically rigorous way forward to overcome that issue.

Of course, intentionalist judges may view these methods as impractical in an actual case. Certainly, there are those that have argued that to be the case (e.g., Hessick, 2017: 1523). However, Gries (2020: 639) points out that expert analysis is also required for ballistics, genetics, and other forensic sciences that the legal community has little expertise in and also that "if the stakes are high enough, why would one *not* seek expert testimony". Clearly, information gained in these types of analysis may be highly relevant. Thus, in practical terms, judges who wish to have corpus analysis performed for either textualist or intentionalist purposes can request brief as was the case in *Wilson v. Safelite* (Wilson v. Safelite Group Inc., 2019). Lawyers could also hire corpus linguists as expert witnesses to provide scientific evidence of interpretations of a statute that support their argument even conducting both kinds of analyses performed in this study to best advocate for their clients.

7.1. Limitations

There are several notable limitations to this study. For instance, the OCR methods struggled to accurately convert many of the documents that were scanned images resulting in time-consuming cleaning of data and sorting through concordance lines to filter out false positive hits, a practical issue if this type of analysis were to be used in an actual case. This issue becomes even more problematic the farther back in time one goes as

fewer and fewer documents are digitized. This study could have also looked at more concordance lines and likely found even more interesting patterns within the data. Also, although COHA is probably the best corpus currently in existence to represent general historical American English, it is limited in several ways in that it does not contain conversational speech and likely overrepresents both the language of higher socio-economic status and urban language since it contains a great deal of published and edited language.

7.2. Future Directions

Of course, future analyses can address many of the limitations of this study by advancing OCR methods and creating better historical corpora. More immediately, future studies could conduct corpus-based legislative history analysis on cases where cherry picking was suspected to have occurred. Also, corpus analysis could be performed on the congressional record since that is another dataset that intentionalists sometime rely upon. Several text types also exist in legislative histories (e.g., committee reports, drafts of bills, hearings, colloquys) that have not yet been well described. A deeper linguistic understanding of these registers could help in understanding their role as strata in a legislative history corpus. Finally, as Keller and Egbert (2020) suggest, more sophisticated methods of measuring dispersion should be incorporated into future corpus-based statutory interpretation analyses.

References

Baker, Wendy & Bowie, David (2010). Religious affiliation as a correlate of linguistic behavior. *University of Pennsylvania Working Papers in Linguistics*, *15*(2), 2–10.

Baron, Dennis E., LaCroix, Alison L., Gries, Stefan Th. & Merchant, Jason (2019). *Brief of Amicus Curiae Corpus Linguistics Professors and Experts Supporting Respondents*, *No. 20–107*.

Biber, Douglas & Conrad, Susan (2019). *Register, Genre, and Style*. Cambridge: University Press.

Biber, Douglas & Reppen, Randi (2015). *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: University Press.

Black, Henry C. & Garner, Bryan A. (2004). *Black's Law Dictionary* (3rd ed.). West Group.

Bringing in and harboring certain aliens (2012). 8 U.S.C. § 1324(a)(1)(A)(iii).

Burrus, Trevor & Meyer, Randal J. (2016). Between the Scylla of Disparate Impact and the Charybdis of Disparate Treatment. *Cato Institute*. Available at: policycommons.net/artifacts/1307223/between-the-scylla-of-disparate-impact-and-the-charybdis-of-disparate-treatment/1910515 (accessed 8 August 2023).

Cohen, Jacob (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Erlbaum.

Costello, George A. (1990). Average voting members and other benign fictions. The relative reliability of committee reports, floor debates, and other sources of legislative history. *Duke Law Journal*, 1990(1), 39–73.

Coulthard, Malcolm, May, Alison & Sousa-Silva, Rui (Eds.) (2020). *The Routledge Handbook of Forensic Linguistics*. New York: Routledge.

Cunningham, Clark D. & Egbert, Jesse (2020). Using empirical data to investigate the original meaning of 'emolument' in the Constitution. *Georgia State Law Review*, *36*, 465–490.

Davies, Mark (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, *14*(2), 159–190.

Davies, Mark (2012). Expanding horizons in historical linguistics with the 400-million word Corpus of Historical American English. *Corpora*, *7*(2), 121–157. DOI: 10.3366/cor.2012.0024.

Dickerson, Reed (1984). Statutory Interpretation in America: Dipping into Legislative History – I. *Statute Law Review*, *5*(3), 76–86.

Driedger, Elmer A. (1981). Statutes: the mischievous literal golden rule. *Canadian Bar Review*, *59*(4), 780–786.

Easterbrook, Frank H. (1988). The role or original intent in statutory construction. *Harvard Journal of Law and Public Policy*, *11*(1), 59–66.

Eskridge Jr., William N. (1990). Legislative history values. *Chicago-Kent Law Review*, *66*(2), 365–440. Available at: search.proquest.com/docview/1297632488 (accessed 9 August 2023).

Eskridge Jr., William N. (1987). Dynamic statutory interpretation. *University of Pennsylvania Law Review*, *135*(6), 1479–1556.

Federalist Society (2017). Corpus Linguistics and Legal Interpretation. 19th Annual Faculty Conference, Video. Available at: fedsoc.org/conferences/19th-annual-faculty-conference?#agenda-item-corpus-linguistics-and-legal-interpretation (accessed 8 August 2023).

Forrest, Jon & Dodsworth, Robin (2016). Towards a sociologically-grounded view of occupation in sociolinguistics. *University of Pennsylvania Working Papers in Linguistics*, *22*(2), 31–40.

Fought, Carmen (2006). *Language and Ethnicity*. Cambridge: University Press.

Gales, Tammy & Solan, Lawrence M. (2019). Revisiting a classic problem in statutory interpretation: Is a minister a laborer? *Georgia State Law Review*, *36*, 491–534.

Gamer, Matthias, Lemon, Jim, Fellow, Ian & Singh, Puspendra (2012). *Package 'irr'. Various Coefficients of Interrater Reliability and Agreement*, *22*.

Gries, Stefan Th. (2020). Corpora and legal interpretation: Corpus approaches to ordinary meaning in legal interpretation. In Coulthard, May, & Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (pp. 628–644). New York: Routledge.

Grove, Tara Leigh (2020). Which textualism? *Harvard Law Review*, *134*(1), 265–307. Available at: search.informit.org/documentSummary;res=AGISPT;dn=20201201040526 (accessed: 8 August 2023).

Haeri, Niloofar (1997). *The Sociolinguistic Market of Cairo: Gender, Class, and Education*. New York: Routledge.

Hashimoto, Brett. (2023). Corpus of Founding Era American English: Designing a corpus for interpreting the United States Constitution. *Corpora*, *18*(1), 1–14.

Harvard Law School (2015). *The Antonin Scalia Lecture Series: A Dialogue with Justice Elena Kagan on the Reading of Statute*. Video, YouTube. Available at: youtube.com/watch?v=dpEtszFT0Tg (accessed 8 August 2023).

Heilpern, James A. (2018). Dialects of art. *Jurimetrics*, *58*(4), 377–410. Available at: jstor.org/stable/27009972 (accessed 8 August 2023).

Hessick, Carissa B. (2017). Corpus linguistics and the criminal law. *BYU Law Review*, *2017*(6), 1503–1530.

Katzmann, Robert A. (2014). *Judging statutes*. Oxford: University Press. DOI: 10.1093/acprof:osobl/9780199362134.001.0001.

Keller, Daniel & Egbert, Jesse (2020). Hypothesis testing ordinary meaning. *Brooklyn Law Review*, *86*(2), 489–532.

Lee, Thomas R. & Phillips, James Cleith (2019). Data-driven originalism. *SSRN Electronic Journal*, *167*(2), 261–335. DOI: https://doi.org/10.2139/ssrn.3036206.

Lee, Thomas R. & Mouritsen, Stephen C. (2017). Judging ordinary meaning. *Yale Law Journal*, *127*(4), 788–879.

Lee, Thomas R. & Mouritsen, Stephen C. (2021). The corpus and the critics. *The University of Chicago Law Review*, *88*(2), 275–366. DOI: jstor.org/stable/26986409.

LexisNexis (2022). *United States Code Service (USCS)*. Available at: store.lexisnexis.com/products/united-states-code-service-uscs-skuSKU7560/details (accessed 8 August 2023).

Lison, Pierre & Tiedemann, Jörg (2016). OpenSubtitles 2016: Extracting Large Parallel Corpora from Movie and TV Subtitles. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 923–929.

Manning, Jennifer E. (2017). *Membership of the 115th Congress: A Profile*. Congressional Research Service.

Meltzer, Bernard D. (1979). The Weber Case: The Judicial Abrogation of the Antidiscrimination Standard in Employment. *University of Chicago Law Review*, *47*(3), 423–466.

Mouritsen, Stephen C. (2010). The dictionary is not a fortress: definitional fallacies and a corpus-based approach to plain meaning. *Brigham Young University Law Review*, *2010*(5), 1915–1980. Available at: search.proquest.com/docview/866207951 (accessed 8 August 2023).

Parrillo, Nicholas R. (2013). Leviathan and interpretive revolution: The administrative state, the judiciary, and the rise of legislative history, 1890–1950. *Yale Law Journal*, *123*(2), 266–529.

Pinchbeck, Geoffrey G. (2014). *Lexical Frequencies Profiling of Canadian High School Diploma Exam Expository Writing: L1 and L2 Academic English*. *Roundtable Presentation*.

ProQuest (2022). *Legislative Insight*. Available at: li.proquest.com/legislativeinsight (accessed 8 August 2023).

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. Available at: R-project.org (accessed 8 August).

Rampton, Ben, Harris, Roxy, Collins, James & Blommaert, Jan (2008). Language, class, and education. In May & Hornberger (Eds.), *Language Policy and Political Issues in Education* (pp. 71–81). Boston: Springer.

Rossum, Guido van & Drake, Fred L. (2009). *Python 3 reference manual*. CreateSpace, Scotts Valley, CA.

Rossum, Guido van (1995). Python reference manual. *Department of Computer Science [CS]*(R 9525).

Samaha, Adam M. (2017). Looking over a crowd. Do more interpretive sources mean more discretion? *New York University Law Review*, *92*(2), 554–621.

Scheppele, Kim L. (2012). Judges as architects. *Yale Journal of Law & the Humanities*, *24*(1), 345–396.

Sinclair, John & Sinclair, Les (1991). *Corpus, Concordance, Collocation*. Oxford University Press, USA.

Taxation of costs (2012). 28 U.S.C. § 1920(6).

The Center for Constitutional Jurisprudence (2021). *Brief of Amicus Curiae Center for Constitutional Jurisprudence, No. 20–107*.

Wald, Patricia M. (1982). Some Observations on the Use of Legislative History in the 1981 Supreme Court Term. *Iowa Law Review*, *68*(2), 195–216.

Zilles, Ana M. S. (2005). The development of a new pronoun: The linguistic and social embedding of a gente in Brazilian Portuguese. *Language Variation and Change*, *17*(1), 19–53.

Cases

Byrd v. Shannon (2013). 715 U.S. (3d Cir. 2013).

Caesar's Entertainment Corp. v. International Union of Operating Engineers (2019). Local 68 Pension Fund, No. 18-2465 (3d Cir. 2019).

Costello v. United States (2012). 666 F.3d 1040 (7th Cir. 2012).

Citizens to Preserve Overton Park v. Volpe (1970). 401 U.S. 402.

Facebook Inc. v. Deguid. (2021). 592 U.S. ___ (2021).

Federal Election Commission v. Rose (1986). 608 F. Supp. 1.

Heilpern, James A. (2019). *Brief of amicus curiae*. No. 17–1625.

In re Sinclair (1989). 870 F.2d 1340.

McBoyle v. United States (1931). 283 U.S. 25.

Metropolitan Omaha Property Owners Association, Inc. v. City of Omaha (2021). 8:19CV341.

Richards v. Cox (2019). 2019 UT 57.

State v. Lantis (2019). 165 Idaho 427.
State v. Misch (2021). 256 A.3d 519 (Vt. 2021).
Steelworkers v. Weber (1979). 443 US 193.
Strang, Lee J. (2021). *Brief of amicus curiae*. No. 19–1392.
Taniguchi v. Kan Saipan Pacific, Ltd. (2012). 566 U.S. 560.
United States v. Acedvedo-De La Cruz (2017). No. 15–10418 (9th Cir. 2017).
United States v. Scott (2021). No. 20–1514 (3d Cir. 2021).
United States v. Thompson (1992). 504 U.S. 505.
United States v. Woodson. (2021). No. 20–1142.
Wilson v. Safelite Group Inc. (2019). No. 18–3408 (6th Cir. 2019).