# A Corpus-Based Study of the Discourse of Recreational Marijuana Legalization

Susana Sotillo
sotillos@mail.montclair.edu
Emerita Professor of Linguistics, Montclair State University

## Abstract

Recreational Marijuana Legalization (RML) is strongly supported in business, legal, and political circles, but opposed mostly by the medical professions. A nationwide rush to legalize recreational marijuana (a.k.a cannabis) at all levels is underway in the United States. Proponents claim that there are millions of dollars to be made in tax revenues that will create jobs and address social justice concerns. Opponents from law enforcement, medical and related fields highlight the risks to public health of smoking or ingesting cannabis with high concentrations of delta-9-tetrahydrocannabinol (THC). This study examines the discourse of RML in texts from two corpora: a corpus of articles from scientific fields, and a reference corpus consisting of texts from newspapers, online cannabis news, economic and legal journals. Using the Words tool in LancsBox 6, word frequencies and keywords were obtained. Next, hundreds of concordance lines of selected keywords as nodes from texts in each corpus were analyzed and evaluated to see whether word clusters around or near nodes could reveal hidden facts about the RML discourse. The results show that slightly more positive discourse patterns characterize texts in the reference corpus, highlighting revenues to be generated by taxing cannabis industry participants, but information about regulations and hefty license fees is not publicly shared with voters. Keywords as nodes in medical research texts yield mostly negative associations, stressing harmful effects on cognition, motor skills, and adolescents' brains. This has implications for the dissemination of important factual and scientific information to the average voter through print and online sources.

*Keywords:* recreational marijuana legalization, cannabis, discourse word clusters, LancsBox, revenues

## A Corpus-Based Study of the Discourse of Recreational Marijuana Legalization

Recreational marijuana legalization (RML) in the United States is a hotly debated topic. The potential for millions or billions of dollars in tax revenues for states where marijuana has been legalized is repeatedly stressed in business and financial circles and media outlets. Colorado was the first state to legalize marijuana in 2012. Since January 2014, total sales by Colorado retailers have exceeded $9.9 billion (Davis, 2021). Sixteen other states followed Colorado's lead (Gomez, 2020). In January 2021, the state of Illinois reported a record sales of cannabis purchases of nearly $89 million, with more than $25 million coming from out-of-state visitors (Jaeger, 2021). At present, 38 states, the District of Columbia, the territories of Puerto Rio, Guam, the Northern Mariana islands and the US Virgin islands have legalized medical marijuana to varying degrees. Nineteen (19) states have legalized both recreational and medical marijuana (see disa.com/map-of-marijuana-legality-by-state, July 2022).

Marijuana Moment (MM), an online morning newsletter, provides the latest news about policy, industry trends, and the culture of cannabis. MM helps activists, investors, businesses, and policymakers navigate the contentious field of cannabis politics and publishes information on pending bills in state legislatures.

Lobbying efforts and campaign contributions to US Senators and other politicians have been in place for at least the past 19 years. The National Organization for the Reform of Marijuana Laws (NORML), for example, has been advocating on behalf of marijuana users since 2002. Likewise, the Marijuana Policy Project (MPP) has funded state and federal candidates who support legalizing medical and recreational marijuana, decriminalizing its use, and, as with alcohol consumption, regulating marijuana. Since 2014, generous donations to US Senators and legislators have been made by various organizations representing the interests of the medical-recreational marijuana industry (Angell, 2020; opensecrets.org-Center for Responsive Politics, 2019).

Advocates in state legislatures claim that marijuana legalization is essential for funding social justice projects (e.g., drug treatment centers, counseling, and providing employment for those unjustly targeted by The War on Drugs, initially declared by President Nixon in 1971, expanded by President Reagan in the 1980s, and continued under the Clinton Administration in the 1990s). In 1992, with the philanthropic support of George Soros, Ethan Nadelmann expanded his work on marijuana reform and founded the Lindesmith Center, a drug policy institute in 1994, which later merged with the Drug Policy Foundation and eventually became the Drug Policy Alliance (Black, 2019). Between 2009 and 2013, approximately 40 states began modifying their drug laws by reducing mandatory minimum sentences and lowering penalties (Desilver, 2014, April 2). Presently, 60% of American voters of all ages support marijuana legalization for both medical and recreational purposes (van Green, 2021).

Previous research on RML based on a national survey using an online panel of 979 participants was conducted by McGinty et al., (2017). Political party affiliation and attitudes among individuals living in states that had approved RML and those that had not were included. Answers were sought to questions concerning Americans' perception of the relative strength of competing arguments about RML. The results showed that pro-legalization respondents were primarily persuaded by economic and criminal justice arguments, whereas anti-legalization

arguments focusing on conflicts between states and federal marijuana laws, inability to eradicate the black market, or public health risks were not convincing.

This exploratory study seeks to add to the growing research on RML from a different perspective by using methodology and tools from corpus linguistics (CL) to analyze the RML discourse of texts in two different corpora. At the time these corpora were compiled, during the COVID pandemic, there were no studies from a CL perspective that addressed the RML movement. This study thus seeks to account for why the RML discourse of proponents is more accessible and convincing to American voters than the discourse of medical researchers and scientists opposing RML. LancsBox 6.0 modules were used in this study, specifically Words KWIC, Text and Whelk. Two research questions are addressed: 1. What do positive keywords differences when compared with negative keywords disclose about the discourses of RML? 2. Can these discourse differences be interpreted by analyzing concordance lines of keywords as nodes and evaluating word clusters around or near them?

The following Literature Review and Theoretical Framework sections provide background information for the present exploratory study that addresses two different discourses of RML as instantiated in texts from a medical-scientific research (MSR) corpus and texts from a variety of sources in the reference corpus. RML has been carefully investigated by the scientific community but there are no CL studies on this topic.

## Literature Review

### Discourse of Proponents and Opponents of RML

Proponents of RML advance claims that economic growth and employment opportunities will benefit large segments of the American population, especially those who were victimized by draconian laws regarding the use and distribution of marijuana. Advancing an economic model, drug policy researchers Carnevale et al. (2017) propose a practical framework for a relatively restrictive regulatory approach to recreational marijuana, supporting a single market for medical and recreational marijuana in all 50 states based on the for-profit commercial model. They offer five recommendations in areas that need to be regulated: (1) cultivation, production, and processing; (2) sales, consumption, and possession; (3) taxes and finances; (4) public health and safety; and (5) governance (p. 71). However, from a legal perspective, Mandiberg (2019-2020) proposes a hybrid approach to marijuana federalism after examining various models and recommends applying different models to legalized marijuana activities that maximize both federal and state regulatory interests.

Concerning price elasticity, Amlung et al. (2018) used behavioral economics to investigate substitutability of legal and illegal cannabis in legalized regions of the United States. Their findings show that cannabis users treat legal cannabis as a superior product compared to illegal cannabis and recommend optimizing pricing policy and enhancing the benefits of a legally regulated marketplace. However, Virgil Grant, an operator of three licensed dispensaries and co-founder of the California Minority Alliance, claims that unlicensed shops greatly outnumber licensed stores and pose serious public safety threats to residents of South Los Angeles because of the lack of enforcement (Zamost, Lee, & Schlesinger, 2019). Thus, not all licensed cannabis stores are earning the profits predicted by the industry.

As to employment opportunities, a study by Isa (2017) based on data from the North American Industry Classification System (NAICS) sought a causal relationship between RML and changes in economic activity using multiple multivariate regressions among different industry codes, including a difference-in-differences test. Except for one case, no statistically significant changes in employment or business growth were found.

RML proponents are also active in social media, posting positive marijuana-related tweets. A study by Cavazos-Rehg et al. (2015) based on a random sample of 7000 tweets by popular influencers found that positive pro-marijuana tweets increased marijuana use and associated harms in adolescents through social contagion. Likewise, Park and Holody's (2018) systematic review of public discourses about marijuana based on peer-reviewed journal articles and one thesis revealed a shocking dissemination of positive messages that target and affect young people and increase their likelihood of using marijuana.

MM, Cannabiz Media, and other related outlets strongly advocate for RML at all levels by disseminating information to policy makers, businesses, and state legislatures. In 2021-22, new bills were introduced in state legislatures, and there is currently an effort to legalize marijuana at the federal level (Angell, 2021). The powerful lure of the for-profit motive is further revealed by the participation of Charles Koch, an American billionaire, who has joined forces with the Cannabis Freedom Alliance, and expressed support for legalizing not just marijuana but all drugs (Zhang, 2021).

**Opposition to RML**

Not all legislators in several states remain convinced of the legalization rush at the state or federal levels as there is opposition by the American Medical Association, mental health professionals, and the Association of Chiefs of Police in various states (e.g., Illinois, Louisiana, New Hampshire, New Jersey, and Pennsylvania). The Chiefs of Police have expressed concerns about traffic safety, mental health, and harm to adolescents based on the results of a survey conducted by Police1 (2020). A significant increase for driving under the influence of marijuana was reported, with 78% in legal states and 50% in illegal states. Forty-five percent of the respondents reported an increase in illegal distribution or sale of marijuana in their jurisdictions.

Nine years earlier, Caulkins & Lee (2012) warned that marijuana legalization could generate unforeseen and harmful results despite well-meaning policies. Along these lines, Smart Approaches to Marijuana (S.A.M.), a non-profit organization, opposes marijuana legalization but not decriminalization. Sabet (2021a, pp. 61-65) points out that legalization advocates claim that the rush to legalize marijuana stems from a desire to increase racial equality and social justice, but they successfully conceal the ultimate profit motive. He argues that this legalization frenzy is driven by the millions in revenues expected from cannabis purchases by regular users in all the legalized states.

An in-depth report based on publicly available state-level data, on-site investigations, and peer-reviewed experimental and survey studies shows disquieting public health findings: Following RML in Colorado and Alaska in 2012 and 2015, respectively, hospitalization cases increased by 101% in Colorado and 45% in Alaska. In Colorado, traffic fatalities rose from 55

deaths in 2013 to 115 deaths in 2018 and active THC was found in the blood sample of drivers involved in fatal collisions (Colton, 2020).

Other public health concerns outlined in the 2020 S.A.M. Impact-Report include a thriving underground or illegal market, a growing trend of workplace marijuana use among middle-class, suburban professionals which impacts employers, an almost non-existent participation of minorities in the marijuana industry (fewer than 2%), watersheds contaminated from pesticides, and threats to the environment from carbon dioxide emissions of energy-intensive marijuana cultivation (Colton, 2020).

**Scientific Opposition to RML**

Medical researchers and neuroscientists have also voiced opposition to RML. Nader & Sanchez's (2018) review of 56 neuropsychological and neuroimaging studies of the effects of regular cannabis use on cognition, brain structure, and function revealed mild cognitive deficits and structural and functional changes in the brain of adult users such as abnormalities in hippocampus volume and gray matter density. Other health effects on adults include several smoke-related toxic chemicals in their blood and urine (Dana-Farber Cancer Institute, 2021).

Neuroscientists examining the effects that THC plays on the brain's reward system recognize that it has a more serious impact on the immature brain of adolescents than on the brain of adult users. For instance, James et al. (2013) chose 24 studies for review among 141 neuroimaging studies based on strict criteria. Their findings revealed that there is greater utilization of brain areas under cognitively demanding tasks in healthy adolescent cannabis users, and greater memory loss and hippocampal volume changes. Currently there are no pharmacological studies of cannabis administration or PET studies with adolescents due to ethical considerations and legal constraints.

Survey studies on marijuana use among adolescents have yielded alarming results. Paschall et al. (2021) examined data based on over three million 7th, 9th, and 11th graders who had participated in the California Healthy Kids Survey in the 2010-2011 and 2018-2019 school years. Following RML in California, results show that there was a significant increase in adolescent marijuana use in 2017-2018 and 2018-2019, especially for 11$^{th}$ graders, Asian Americans, and African Americans. Similarly, the 1975-2019 Monitoring the Future national survey results indicate that daily marijuana use rates among 8$^{th}$, 10$^{th}$, and 12$^{th}$ graders rose in all three grades in 2019, especially in the lower two grades, reaching 1.3%, 4.8%, and 6.4%, respectively. Marijuana vaping also increased substantially in 2018 and 2019, with 22.1% among 12$^{th}$ graders, 19.1% for 10$^{th}$ graders, and 8.1% among 8$^{th}$ graders (Johnston et al., 2020). Moreover, Cerdá et al. (2020) found a 25% increase in Cannabis Use Disorder (CUD) among children ages 12-17 in states that had legalized marijuana.

Pertaining to children, Richards et al. (2017) reviewed the findings of 44 case series, studies, and reports on the unintentional ingestion of cannabis by children aged 12 or younger. Common signs and results of physical examination of cannabis toxicity include lethargy, ataxia (lack of movement coordination), hypotonia (muscle weakness), mydriasis (pupil dilation), tachycardia, and hypoventilation. They conclude that cannabis ingestion by children represents a serious public health threat because young children tend to explore and test things that look

appetizing by placing them in their mouth, such as cannabis-infused edibles in the form of brownies, muffins, or gummies carelessly left by adult family members.

THC effects of regular marijuana use also have an impact on cognitive-motor skills and brain mechanisms that regulate coordinated movement and driving. Weinstein et al. (2008) used two positron emission tomography (PET) scans with 12 regular users of Marijuana while performing a virtual reality maze task. One was carried out while participants were under the effects of 17mg of THC and the other without THC. The results show that when subjects were under the effects of THC, they hit the walls more often than when the task was performed with no THC.

Since most of this information is not accessible to the public, it is important to draw attention to the results of scientific and survey studies, though this review might seem too detailed for this type of study. It should also be made clear that most published articles cited on the effects of THC on health and brain functions show statistically significant results and positive correlations that support hypotheses but not causality due to confounding factors in their design, small sample sizes, and the use of a wide range of methods and analytic tools. This makes replication of studies very difficult.

**Traffic Fatalities**

Medical researchers Aydelotte et al. (2019) compared changes in fatal crash rates in Washington, Colorado, and nine control states with anti-marijuana laws or medical marijuana laws over the five years before and after the opening of commercial marijuana dispensaries. Difference-in-difference analyses were performed on data from the Fatality Analysis Reporting System (FARS) that showed more increases in fatal crash rates in Colorado and Washington than expected. The effect was greater and statistically significant following the 2014 opening of commercial marijuana dispensaries, which supports previous findings by Pollini et al. (2015) and Salomonsen-Sautel et al. (2014). However, Hake's (2019) quasi-experimental study using rational choice theory and perceptual deterrence theory examined the relationship between traffic fatalities involving cannabinoids in Washington State before and after marijuana legalization and compared these against 43 non-RML control groups using data from FARS of the National Highway Traffic Safety Administration (NHTSA). No significant differences were found between Washington State and other non-legalized states concerning traffic fatalities involving cannabinoids

A similar study by Miller (2018) analyzed changes in measures of excessive drinking, driving-under-the influence (DUI) fatalities, crime, and DUI arrests in Colorado and Washington using theoretical arguments from Sociology. The effects of legalizing marijuana possession in 2013 were examined, followed by an analysis of the effects of opening recreational marijuana dispensaries in 2014. No significant effects were obtained, but there was a significant finding for DUI policing following RML. Miller suggests that the lack of negative consequences may well be a reason for proponents to argue in favor of legalization policies. But a recent Rocky Mountain HIDTA 2021 report shows that marijuana-impaired driving is continuing to have a disastrous impact in Colorado: "The percentage of all Colorado traffic deaths involving drivers who tested positive for marijuana increased from 11 percent in 2013 to 20 percent in 2020" (p. 6). According to Grabenauer (2020), problems remain with blood sampling methods for assessing cannabis

impairment because there are no reliable instruments. This makes it difficult for law enforcement to measure cannabis impairment on the spot.

This overview of studies by researchers in medical and other fields is based on a sample of texts from both corpora and current marijuana-related news articles. The present investigation seeks to analyze the discourse of texts from two different corpora informed by contemporary studies using methodology and tools from CL.

**Theoretical Framework**

In CL, corpus construction is an important undertaking that requires access to large amounts of different types of texts in electronic format. Biber et al. (1993) conceive of a corpus as a large and principled collection of natural texts, and Bauer and Aarts (1999) propose a systematic selection of an adequate amount of material that represents the whole, but stress that statistical random sampling does not necessarily apply in textual analysis. McEnery & Wilson (2001), Baker (2006) and Brezina (2018) also view corpus construction as a principled, representative, and systematic collection of large texts. Corpus construction in other languages follows similar guidelines (Dash, 2010). Recently, Egbert, Biber & Grey (2022) have questioned the prominence of size in corpus construction.

CL is regarded by many primarily as a methodology that can be used in many branches of linguistics. For example, researchers have successfully combined CL tools with various theoretical approaches in investigating socio-political, cultural, and other aspects of language. Mautner (2007), for instance, used the 57-milllion Wordbanks Online corpus and combined quantitative and qualitative approaches to the study of conventional social constructions of age and aging. Semantic preferences of the discourse on the elderly included disability, illness, dependence on others for care, helplessness in the face of crime, and neglect, which are associated with negative semantic prosodies. Mautner drew attention to the limitations of concordances and encouraged the analysis of large stretches of text. Baker et al., (2018), addressed these concerns by combining rigorous CL methods using the 140-million-word corpus of British news articles with an in-depth qualitative analysis.

The use of CL was first proposed as a tool for discourse analysis (DA) and later for critical discourse analysis (CDA) to achieve explanatory adequacy. The researcher needs to show what this approach entails by analyzing the discourse in question and explaining its meaning (Antaki et al., 2003; Partington and Marchi, 2015; Stubbs, 1996, 2001; van Dijk, 1998, 2005).

Linked to the study of words and collocation research in CL are the concepts of semantic preference and semantic prosody. Semantic prosody was first discussed by Louw (1993) and expanded by other researchers. Partington (2004) regards semantic prosody as an aspect of evaluative meaning, and Sinclair (2004b) sees it as a unit of meaning, consisting of a long sequence of co-occurring items. The concepts of semantic preference and semantic prosody will not be examined in this study because there appears to be no consensus as to how to define semantic prosody (Bednarek, 2008; Hunston, 2007; McEnery & Hardie, 2011; Partington, 1998, 2004; Sinclair, 2004a). A recent syntactico-semantic analysis called prosody concord has been proposed to solve existing contradictions by Tang and Liu (2018).

In this study, the keyword analysis will be followed by an analysis of keywords as nodes in concordance lines to identify patterns or word clusters around or near the node following Baker's (2006, pp. 121-149) discussion of keyness and his analysis of keywords using a different metric. An effect-size metric, the simple maths parameter (SMP) proposed by Kilgarriff (2009), is used in this study, though there are various metrics for keyword analysis. SMP looks at the ratio between the relative frequencies of words in the texts of two corpora while adding the constant k=100 (Brezina, 2018, p.85). It is an effect-size measure and not a test of significance. Another one, %DIFF, was proposed by Gabrielatos & Marchi (2012). Gabrielatos (2018, p.5) stresses that statistical significance "does not reveal the size of the frequency difference," but is a means of establishing the level of confidence or reliability of observations made. Effect-size metrics, in contrast, can identify frequency similarities between corpora (see Gabrielatos, 2018, for an explanation of the difference between effect-size metrics and statistical significance).

With new free software such as LancsBox version 6.0, CL researchers can analyze functional aspects of written or spoken discourse using tools such as KWIC to obtain concordance lines, GraphColl for collocations, Whelk for displaying concordance lines of a node and seeing its distribution in the whole corpus, and Words for producing frequencies and keywords (Brezina et al., 2015; Brezina, et al., 2020).

## Methodology and Data

Two corpora were compiled from scratch using the researcher's university databases such as JSTOR, Google Scholar, Nexis Uni, ProQuest Central, and Science Direct. To avoid allegations of cherry picking, texts were randomly selected (every fifth article) from various databases using simple and advanced queries such as *marijuana* OR *cannabis*, *THC effects*, *driving and cannabis*, *recreational marijuana*, *cannabis regulations*, and *cannabis revenues.* This random search for texts from databases for the MSR corpus initially yielded 80 files. Samples of approximately 1700 words from the beginning, middle, and end of articles and a thesis were selected while enclosing author (s), source, date, and title in tags < >, </>. Unless embedded in the narrative, figures, tables, and references were removed when converting pdf files to plain text using Notepad++. Original titles and articles were kept in a separate folder. These texts/articles focused on opposition to commercialization due to health and safety concerns.

After revisions, the MSR corpus consisted of 75 files, 379,551 tokens, 20,536 types, and 18,228 lemmas. Table 1 shows the samples of texts obtained and approximate number of words/tokens in texts.

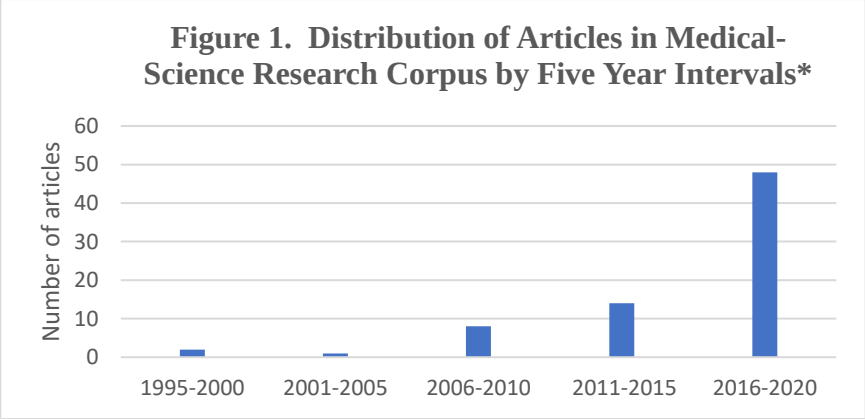
The Reference corpus was slightly larger, consisting of 97 files, 414,930 tokens, 22,007 types, and 19,328 lemmas. It was compiled following the same procedures but focusing on information for public consumption in an accessible language on marijuana commercialization and regulations.

**Table 1**

*Medical-Scientific Research Corpus and Reference Corpus*

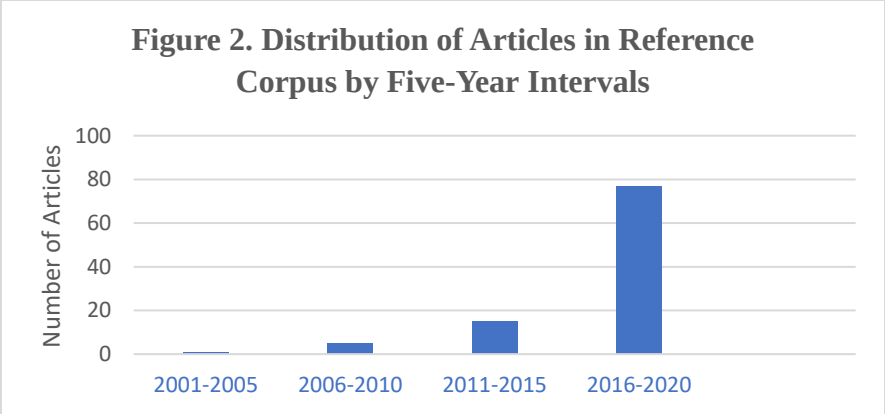
| Text Categories | No. of Texts | Approx. Num. of Words/Tokens | Text Categories | No. of Texts | Approx. Num. of Words/Tokens |
|---|---|---|---|---|---|
| Scientific articles: Neuroscience, Pediatrics, Prevention Medicine, Psychiatry, Psychopharmacology, etc. | 64 | 5550 | Journal Articles and Quarterlies: Economic behavior, American Indian Quarterly, Public health, Milbank Quarterly, etc. | 17 | 5600 |
| Doctoral thesis | 1 | 12,400 | Doctoral Thesis | 1 | 13680 |
| Journal Editorials | 3 | 1200 | Law Journals | 11 | 18000 |
| Letters to the Editor | 2 | 800 | Business Newsletters | 9 | 1500 |
| | | | Newspapers | 11 | 2000 |
| Government Reports: Congressional Digest, Fatality Analysis (FARS), NHTSA, NIH, S.A.M. | 5 | 1350 | Government Reports | 16 | 2800 |
| | | | Magazines | 6 | 1800 |
| | | | Online News | 25 | 650 |
| | | | NJ Blog | 1 | 700 |

Figures 1 and 2 below show the distribution of articles in these corpora. Figure 1 displays the distribution of articles in the RSM corpus from 1995 to 2020. An increase in published articles follows a favorable shift in public opinion concerning the legalization and commercialization of recreational marijuana.




*Two letters to editor in journal articles were not included in this graph.

Figure 2 displays the distribution of articles in the reference corpus from various sources, including legal and economic journals. As more and more states legalized marijuana in the U.S., the number of articles retrieved increased. Coverage of cannabis news increased from 2016 through 2020.




Modules of LancsBox 6 such as Words, KWIC, Text, and Whelk were selected. The Words module was used to first obtain absolute and relative frequencies using type and then switching to lemmas, which permitted filtering for function words in both corpora using a pop-up box (.*n|v|adj|). Next, the MSR and reference corpora were merged to obtain keywords. Positive and negative keywords, as well as lockwords (not shown or analyzed), were obtained using SMP. The results were automatically calculated by LancsBox 6. Manually checking for errors using KWIC or the Text feature was important since converting pdf documents to plain text sometimes resulted in symbol substitutions or splitting of words (e.g., ¼ instead of the equal sign; al cohol). It was also important to check for POS tagging errors.

Five out of the first 20 positive keywords in the MSR corpus were selected for close reading and analysis in concordance lines since they represented the aboutness or intentionality of texts in the corpus (Baker, 2006, pp-121-149). Positive or negative evaluations of clusters of words around or near keywords were examined. The analysis of 371 concordance lines of selected keywords as nodes within a span of seven words to the left and seven to the right was carried out by two raters working separately and following Brezina's guidelines (2018, pp 79-91). Longer stretches of text were used when it was difficult to identify whether a keyword as node had a negative or positive evaluation. Five negative keywords in the reference corpus were also analyzed using the same procedures. The researcher and an MA graduate in linguistics sought to separately analyze these keywords as nodes in concordance lines to determine their polarity and find a possible interpretation for differences in the discourse of proponents and opponents of RML.

## Results

The results section that follows first presents an explanation of keywords obtained (4.1) and then displays both positive and negative keywords in Figure 3. This is followed by a table that displays the first ten keywords in the MSR corpus. Section 4.2 describes the procedures for analyzing keywords as nodes in concordance lines by two raters and presents our findings in Tables 3 through 12.

### Positive and Negative Keywords in both the MSR and Reference Corpora

Columns one and two of Figure 3 below show the first 20 positive and negative keywords in both corpora. These were generated via the Words module, which computes a comparison of frequencies between the corpora using the default metric SMP with k=100. Positive keywords

occur more frequently in the MSR corpus, but not as frequently or not at all in the reference corpus. Negative keywords are those that occur infrequently in the MSR corpus but more frequently in the reference corpus.

Among the first 20 positive keywords in the MSR corpus we find terminology typical of experimental research, meta-analysis, and survey studies. These include *mj*_n (marijuana), *p*_n (stands for significance levels), *rml*_n (recreational marijuana legalization), *es*_n (endocannibinoid system), *rat*_n, *receptor*_n, *al*_(other), *et*_con (as in et al), *activation*_n, *connectivity*_n, *dui*_n (driving under the influence), *WIN*_n (WIN 55, 212-2, a chemical aminoalkylindole derivative), *male*_n, *task*_n, *pet*_n (positron emission tomography), *onset*_n, *caffeine*_n, *rcbf*_n (regional cerebral blood flow), and *mml*_n (medical marijuana legalization). Negative keywords that are found with more frequency in the reference corpus include *percent*_n, *say*_v, *bill*_n, *federal*_adj, the nouns *court*, *teen*, *business*, *industry*, *pot* (marijuana, weed), *school*, *tax*, *commerce*, *company*, *wa* (Washington), *enforcement*, (New) *Jersey*, *license*, and *revenue*. Business, commerce, economics, and industry are emphasized, as well as legal terminology.

**Figure 3**

*First Twenty Positive Keywords and Negative Keywords- MSR and Reference Corpora (LancsBox 6.0)*

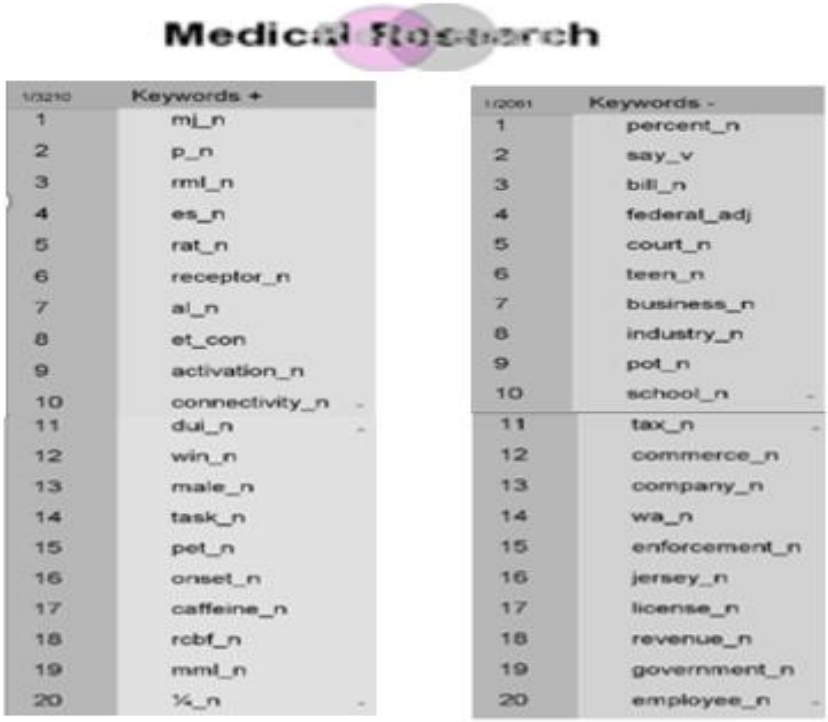



Table 2 below displays absolute and relative frequencies for the first 10 positive keywords in the MR corpus, as well as dispersion of keywords in texts, using the default metric SMP.

**Table 2**

*First 10 Positive Keywords in the MSR Corpus*

| Lemma | Ab. Frequency | Rel. Frequency | Dispersion | Metric Used SMP |
|---|---|---|---|---|
| mj_n | 641 | 16.89 | 3.572396 | 17.88838 |
| p_n | 875 | 23.05 | 1.427521 | 12.39926 |
| rml_n | 337 | 8.88 | 4.168941 | 9.424635 |
| es_n | 293 | 7.72 | 4.870463 | 8.719648 |
| rat(s)_n | 215 | 5.90 | 3.743924 | 6.215221 |
| receptor(s)_n | 319 | 8.42 | 2.170425 | 6.050703 |
| al_n | 2506 | 66.03 | 1.026166 | 5.916465 |
| et_con | 2506 | 66.03 | 1.142221 | 5.873953 |
| activation(s)_n | 195 | 5.14 | 2.309387 | 5.855412 |
| connectivity_n | 200 | 5.27 | 3.59992 | 5.846664 |

The Keyword analysis module shows differences in the relative frequencies of keywords in each corpus as shown in Figure 3. The MSR corpus highlights terminology from experimental research aimed at a specialized audience, including endless citations (et al.) as keywords. In contrast, negative keywords in the reference corpus reveal the regulatory or controlling language of institutions or entities (e.g., federal, court, industry, enforcement, tax, etc.) used in business, education, law, and government.

## Analysis of Keywords as Nodes

A total of 371 concordance lines with seven words to the left and seven to the right of the keyword as node were generated using the KWIC module in LancsBox 6.0, and analyzed by two raters to determine whether they conveyed a positive or negative evaluation in the corpora texts. Disagreements, as well as a revision of keywords following corpus data reanalysis were discussed first in person and later online. The raw agreement score obtained by identifying whether a keyword as node was evaluated as positive or negative was 0.91 or 91%, and Cohen's Kappa yielded 0.85, which indicates very good agreement between raters (Brezina, 2018 pp. 87-91; Landis & Koch, 1977). Keyword nodes selected for this analysis from the MSR corpus were: *MJ* (marijuana), *rml*, *task*, *connectivity*, and *onset*. Keywords such as et al. were not selected as they represented thousands of citations. Due to space constraints, only five concordance lines from texts of each corpus are shown in the tables below.

Table 3 shows concordance lines for MJ or marijuana, which the Random House Dictionary of the English language defines as a narcotic, or hallucinogen, which folk etymology traces to a personal name in Mexican Spanish (Maria Juana). This has negative connotations though this may be changing given its legalization in many states and countries. MJ in these concordance lines, retrieved from a larger concordance spread sheet, is often preceded by *past year* and *with*, and by noun phrases (NPs) such as *dose-dependent relationship/association*, and adjectives modifying nouns (e.g., *significant* abnormalities, *major* depressive disorder (MDD,) and *cognitive* deficits). MJ use/exposure as a NP is followed by comparative adjectives modifying nouns with negative connotations in medical or addiction research (e.g., *smaller* Medial Orbito-

Frontal Cortex or mOFC and *poorer* sustained attention and psychomotor speed). In these concordance lines, negative collocations with MJ intensify their negative associations with undesirable results.

**Table 3**

*Concordance of MJ (mj)-MJ use/exposure*

| | | | |
|---|---|---|---|
| 1 | Dataset, which examined relationships between **past year** | **MJ** | **use** and attention/executive function in emerging adults |
| 2 | support a dose-dependent relationship between **past year** | **MJ** | **use** and smaller mOFC volumes, there was |
| 3 | reported a dose-dependent association between **past year** | **MJ** | **exposure** and poorer sustained attention, psychomotor speed |
| 4 | demonstrate significant abnormalities in this region **with** | **MJ** | **use** or MDD alone, but showed greatly |
| 5 | depressive disorder (MDD) and cognitive deficits **with** | **MJ** | **use**, particularly early-onset use. We investigated cognitive |

Table 4 shows that the key node RML, a noun phrase, is preceded by *the/a/an effect(s) of*, *substitution effect*, forming the cluster *the effect(s) of RML* followed to the right by outcomes such as *on marijuana use*, *on the odds of marijuana use*, and *among first-year students*. A substitution effect of RML (line 5) on college students' binge drinking implies that legalization may encourage marijuana as an alternative to binge drinking, but the words and phrases in these concordance lines and larger stretches of text analyzed are not neutral and have negative connotations and evaluations that influence their meaning in this scientific discourse.

**Table 4**

*Concordance of RML (Adj+N+N)*

| | | | |
|---|---|---|---|
| 1 | interpretation and given that **the effect of** | **RML** | on marijuana use was similar in three |
| 2 | any heavy alcohol use, **the effect of** | **RML** | on the odds of marijuana use was |
| 3 | It is possible that **the effects of** | **RML** | we noted among first-year students who denied |
| 4 | was sufficient to detect **an effect of** | **RML.** | Based on the prevalence of the least |
| 5 | Al., 2017). Thus, **a substitution effect of** | **RML** | on college students' binge drinking could have |

Connectivity is a noun that denotes relatedness, association, or in a technical sense, the ability to connect systems, programs or networks. In a medical sense it can have negative connotations. In Table 5, when examining the context of connectivity as node, adjectives and nouns to the left of the node stress harmful effects that are associated with *disturbed brain connectivity*, *chronic-marijuana users* and *changes in brain connectivity*, *disruption* in the neural connectivity, and *disrupted frontoparietal connectivity*. To the right of the node, negative effects include *cognitive impairmen*t, *risk alleles in six genes*, and *changes* associated *with marijuana-related behavior*. These extended collocations with brain connectivity have strong negative associations in the discourse of these medical texts.

**Table 5**

*Concordance of (brain) Connectivity*

| | | | |
|---|---|---|---|
| 1 | may minimize such harmful effects. **Disturbed brain** | **connectivity** | in cannabis users may underlie **cognitive impairment** |
| 2 | were measured. Default mode network (DMN) **brain** | **connectivity** | was determined. **Risk alleles** in six genes |
| 3 | THC-enriched areas in chronic marijuana users that will be associated with **changes** in **brain** | **connectivity** | and **marijuana-related behavior**. Significance: The existing literature (Extended text search to the left of connectivity,) |
| 4 | Reports indicating the disruption in the **neural** | **connectivity** | in the specific regions of the brains of regular marijuana users. |
| 5 | attention, possibly secondary to **disrupted frontoparietal** | **connectivity** | with a compensatory right prefrontal over activity |

Table 6 below displays concordance lines with task as a node. Task means an amount of work or duty assigned to a person. In the idiomatic phrase, *to take someone to task*, it implies censure, which has a negative evaluative meaning. Task in these concordance lines forms the cluster *working memory (WM) task*. *Decreased performance* and *deleterious effects* precede the WM task node, and to the right of task, verbs describe undesirable changes following marijuana smoking, including *reduced alpha band EEG reactivity*, a *decrease in amplitude*, *smaller task load differences*, *increased reaction time*, and particularly *deleterious effects on performance* on a more difficult version of the *WM task*. Collocations in context on both sides of the node task with other nouns, verbs, and adjectives have negative associations that diminish or make smaller (reduce or decrease). Increased reaction time has no positive association since this involves a working memory task that takes longer to complete. Task in these concordance lines has persistent negative evaluations.

**Table 6**

*Concordance of (WM) Task(s) (AF: 562, RF: 14.81, not shown in Table 1)*

| | | | |
|---|---|---|---|
| 1 | power was not observed. In the **WM** | **task,** | marijuana **reduced alpha band EEG reactivity** in |
| 2 | tasks and the P300 in the **WM** | **task**, | **decreased in amplitude** after marijuana smoking. Such |
| 3 | subjects exhibited **decreased performance** in the **WM** | **task**, | and **smaller task load differences** on alpha |
| 4 | P>0.10. **Reaction time (RT)** in the **WM** | **Task** | **increased** after smoking marijuana [F(1,9)=16.15, P |
| 5 | **deleterious effects** on performance of the **WM** | **task**, | particularly **on the more difficult version**, and |

Onset means beginning or start, but it can also refer to an assault or attack. In the reanalysis of *onset* as node in Table 7 below, these concordance lines have only three examples of what we evaluated as having negative connotations since the studies deal with adolescents using a hallucinogen. Here *age of onset* of use is followed by the right superior frontal, which stands for the functioning of the frontal lobe that contributes to cognitive functions. To the left of the node, *age of marijuana use onset*, we learn that this influences *the magnitude of the changes* and thus has a negative evaluation, which is also the case with the last line that describes *the trajectories of*

*adolescents* before and after the onset of marijuana use. These three lines we evaluated as having negative associations, but lines two and four are seen as having no polarity because they describe in one case, imprecise results and in the other, studies merely focused on adolescents.

**Table 7**

*Concordance of Onset (AF: 207, RF: 5.45)*

| | | | |
|---|---|---|---|
| 1 | urinary cannabinoid levels and between **age of** | **onset** | **of use** and the right superior frontal |
| 2 | is not clear whether this predates the | **onset** | of regular use or is the result |
| 3 | Analyses confirm that age of marijuana use | **onset** | influences the magnitude of these changes. |
| 4 | studies focused on adolescents, even though the | **onset** | of marijuana use typically occurs during this |
| 5 | trajectories of adolescents before and after the | **onset** | of marijuana use. Our design considered prior |

Table 8 below illustrates the evaluation of the adjective federal in the reference corpus. Its normal usage implies a central government or a union of states, but distinct from the individual government of separate states. In these concordance lines, federal is linked to authority, government regulation of commerce, laws that supersede state laws, and prosecution of marijuana crimes, though this last line indicates that the Department of Justice has relaxed its prosecutorial authority. Federal in these contexts has negative evaluations since it implies control and regulation of state activities (interstate commerce). This word (adj) forms part of the discourse of legal and governmental entities in the reference corpus texts.

**Table 8**

*Concordance of Federal (AF: 717, RF: 18.39)*

| | | | |
|---|---|---|---|
| 1 | commerce was considered a critical component of | **federa**l | **authority**. After the American Revolution, serious trade |
| 2 | the Constitutional Convention, the role of the | **federal** | **government in regulating commerce** received tremendous attention. |
| 3 | were unconstitutional since they were repugnant to | **federal** | acts created by Congress **in regulating interstate** |
| 4 | use, cultivation and possession remain illegal under | **federal** | **law.** However, in response to several states' |
| 5 | the DOJ has relaxed its policy on | **federal** | **prosecution of marijuana crimes."** The Department of |

Table 9 below shows concordance lines of the repetitive phrase *an industry*, which is normally associated with a group of productive enterprises in a particular field. In concordance lines analyzed, VPs and NPs to the left of the node, such as making *money like grass* is a goal of investors in the marijuana industry, and words to the right of industry convey positive associations, such as *growing opportunities* around marijuana, *a revenue to be generated*, and *value* for minority communities that is brought about by this *new industry*. The words that collocate with industry in extended stretches of text communicate positive evaluations in the discourse of those supporting marijuana legalization (from online news and newspapers).

**Table 9**

*Concordance of (an) Industry (AF: 456, RF: 10.99)*

| | | | |
|---|---|---|---|
| 1 | Headlines such as "Investors hope to make money **like grass"** (DN May 5, 2014), "An | **Industry** | **is growing** around marijuana" (DN November 12, 2014), (Extended context). |
| 2 | of the drug war. "We'll build **an** | **industry**, | it would be a **revenue-generator**," New Jersey Gov. Phil Murphy said in an interview about the state's legalization referendum last week. "I think at first it would be modest, but ultimately will grow, I think, into several hundred million dollars in the state budget. Along with social justice, that's a pretty good, winning combination." (Extended context) |
| 3 | collected $262.9 million from the **regulated marijuana** | **industry** | in fiscal year 2019, which made up |
| 4 | people of color. Black folks create **an** | i**ndustry** | that has **value**, whether through legal or |
| 5 | get a chance to be in **an** | **Industry** | this **new**," he says. "This is like |

The single word tax has negative evaluations since it means that a sum of money or resources are demanded by a government of individuals or businesses to support its bureaucracy, facilities, services, and institutions (e.g., the military). Table 10 displays the following context lines to the left of the keyword node tax, *economic development*, *impacts on patients*, *support for policy options*, and *benefits tied to medical cannabis*, all indicate that tax revenues will be generated, thus securing positive outcomes such as increases in revenue. Also, *deductibility* in insurance costs and *advertising restrictions*, probably to prevent vulnerable populations from buying cannabis products, indicate positive economic and social outcomes. However, line five has negative associations because it implies coercion of members regulating the cannabis industry and is a burdensome requirement.

**Table 10**

*Concordance of Tax (AF: 540, RF: 13.01)*

| | | | |
|---|---|---|---|
| 1 | and **see the economic development opportunity** and | **tax** | **revenue** opportunity they offer and the jobs," |
| 2 | marijuana– seeing its **impacts on patients** and | **tax** | **revenues**– full legalization often follows. California fully |
| 3 | a position to **support policy options**--such as | **tax** | **increases and advertising restrictions**--that would promote the |
| 4 | **benefits tied to medical cannabis**, such as | **tax** | **deductibility**, potential insurance coverage and employer accommodation |
| 5 | making the **regulatory enforcement body dependent** on | **tax** | **revenues** from the very industry it oversees |

In Table 11 below, revenue, which stands for the income of a government from taxation, excise duties, customs or other sources that pay for public expenses, has a positive evaluation in these concordance lines. It highlights the billions of dollars generated that double or almost triple in two years or how RML encourages a boost in *revenue*, *job creation*, *benefits*, and *10 million a year in revenue* with potentially hefty license fees (for those applying for a dispensary license).

These concordance lines predict positive financial outcomes with strong favorable connotations and are thus positively evaluated.

**Table 11**

*Concordance of Revenue (AF: 269, RF: 6.48)*

| | | | |
|---|---|---|---|
| 1 | revenue in 2017 to **$16.9 billion in** | **revenue** | in 2019 to more than **$31 billion** |
| 2 | industry will grow from **$6.9 billion in** | **revenue** | in 2017 to **$16.9 billion** in revenue |
| 3 | the state will see **a boost in** | **revenue** | from marijuana, and the likely **job creation** |
| 4 | the benefits it also **brings in** | **revenue**." | "There's much more that needs to be |
| 5 | million to **$10 million** in a year **in** | **revenue,** | potentially **making a license worth $5 million** |

In contrast to the positive evaluations of keywords as nodes in concordance lines that highlight financial benefits for states, cities and municipalities, Table 12 below illustrates the use of legalese language when discussing the act of enforcement. Enforcement has a negative connotation since it involves the act or process of compelling someone or an entity of doing something. In these concordance lines, the environmental protection agency (EPA) has enforcement responsibilities regarding the cultivation and distribution of marijuana, but it also shares enforcement responsibilities with the states. Line four can be interpreted as a positive evaluation of enforcement since it protects children, and line five conveys the need for not granting unlimited licenses since it would put a strain on enforcement officials and pose a threat to public health. In general, we evaluated enforcement as having negative associations in the discourse of those writing these texts or reporting on legal problems because it shows strained relations between the states and federal government over many aspects of RML.

**Table 12**

*Concordance of Enforcement (AF: 377, RF: 9.09)*

| | | | |
|---|---|---|---|
| 1 | actively involved in its own administrative and | **enforcement** | activities. EPA becomes the lead enforcer on |
| 2 | which the federal government shares administrative and | **enforcement** | responsibility with states that meet minimum requirements. |
| 3 | numbers also speak to Commission licensing and | **enforcement** | staff working around the clock to make |
| 4 | on age verification. If continuous monitoring and | **enforcement** | are not in place, however, children might |
| 5 | public health. Unlimited licensing could strain regulatory | **enforcement** | and may have negative public health implications. |

To sum up, analyses of selected keywords as nodes in concordance lines indicate slightly more positive evaluations of the discourse in the reference corpus texts than in those of the MSR corpus. Keyword as nodes in the MSR texts had strong negative associations and in some cases, had no polarity since study procedures were being described in sections of the concordance lines examined. We also noticed that certain critical words such as decriminalization, which occurs 21 times in the MSR corpus and 51 times in the reference courpus, did not appear among the first ten or twenty positive and negative keywords.

## Discussion

This study used CL tools to analyze the discourse of texts in two corpora and sought answers to two research questions. Regarding RQ 1, as to what frequency differences of positive keywords when compared with negative keywords reveal about the discourses of RML, the results show that texts in the MSR corpus highlight experimental research findings in medicine, neurology, psychiatry and other related fields that focus on the use of marijuana and its effects on the brains of adults and adolescents. These are texts not easily accessible to the general public in a language that can be easily understood. The reference corpus, in contrast, stresses tax revenues, job creation, guidelines for cannabis cultivation and distribution, licensing fees for dispensaries, along with enforcement of regulations to prevent access to cannabis products by children and adolescents. In other words, the discourse of these texts appeals to the average individual who is asked to vote on RML public questions. This has important implications for scientists in medical and related fields who often lack access to mainstream media outlets for disseminating important findings to the general public. Opponents, on the other hand, have sufficient funds from the cannabis industry to blanket the airwaves, disseminating their messages in print and online news outlets.

With respect to RQ 2 or whether these discourse differences could be interpreted by analyzing concordance lines of keywords as nodes and evaluating word clusters around or near them, the results show that clarifying these differences between the discourse of opponents and proponents was difficult since not all the concordance lines examined showed positive or negative connotations. In many concordance lines, there was an apparent neutral position. However, some phrases and words in context were negatively loaded in the overall discourse of scientists writing about marijuana and its effects, such as:

- Age of onset
- Brain connectivity
- Depressive disorder and cognitive deficits
- Reaction time
- Working memory and decreased performance

Since the reference corpus consisted of more diverse texts from a variety of news sources, government reports, as well as economic, financial, and legal journal articles, it had more positively evaluated outcomes:

- The growth and normalization of the marijuana/cannabis industry
- The revenues to be generated by taxing partakers in the cannabis industry
- Guidelines to monitor sales of cannabis at dispensaries as a preventive measure to protect vulnerable populations

Some obstacles to RML had negative evaluations:

- Clashes between states and the federal government concerning interstate commerce in marijuana/cannabis
- Challenges in the regulation, cultivation, distribution, and consumption of cannabis

- Imposition of hefty dispensary fees that would prevent participation of minority groups in this expanding industry

Negative evaluations of clusters of words in the discourse of texts in both corpora provide some degree of support for McGinty et al.'s (2017) findings concerning Americans' perception of competing arguments about RML that result in overwhelming public support at the ballot box. A discourse that appears in numerous media outlets and other publications with more positive associations, such as productive economic outcomes and the setting aside of licenses for entrepreneurs in minority communities, is more appealing to the general public or voters who obtain their information from social media, newspapers, TV, or are convinced by politicians while ignoring objections from health experts, safety concerns by law enforcement, or warnings from other scientists (Sabet, 2021a, Sabet, 2021b).

## Conclusion and Limitations

The results of this exploratory study using tools from CL show that the keyword technique can reveal specific frequency differences in the discourse of texts from two different corpora. This was followed by an analysis of keywords as nodes in concordance lines that sought to explain what these discourse differences meant by evaluating word clusters around or near the nodes. The results did not yield definitive answers, but showed that to an extent, more positive evaluations were apparent in the discourse of RML proponents.

Limitations of this study include the small size of the two corpora and a not sufficiently systematic approach to text compilation since random selection of texts, separated by opposition or support for RML, resulted in uneven text sizes, with dispersion scores above 1.0 in both corpora. Additionally, the MSR corpus represents an academic genre aimed at a specialized audience, while the reference corpus texts cover a wide range of RML-related topics, and includes new genres (blogs, and online news). Those interested in pursing this topic would need to compile corpora from available reports or summaries of studies written for public consumption by skilled professionals, and compare these to mainstream, regional and local newspaper coverage of this contentious RML topic. CL inspired researchers could also investigate the discourse of social justice by politicians in transcribed data from regional and local nespapers, and reports of the extent to which cannabis tax revenues are spent on jobs or projects in minority communities, as proponents promised. Lastly, CL studies are needed that involve transcribed in-depth interviews with equal numbers of those voting for and against the legalization of marijuana and other hallucinogens, in order to assess the extent of reliable scientific information individuals possess for making informed decisions on important topics.

**References**

Adlin, B. (October 30, 2020). New Jersey Governor steps up marijuana legalization rush as new ad touts economic benefits days before election. *Marijuana Moment-Politics*.

https://www.marijuanamoment.net/new-jersey-governor-steps-up- marijuana-legalization-push-as-new-ad-touts-economic-benefits-days-before-election/

Amlung , M., Reed, M.M., Morris, V., Aston, E. R., Metrik, J. & MacKillop, J. (2019). Price elasticity of illegal versus legal cannabis: a behavioral economic substitutability analysis. *Society for the Study of Addiction*, *114*(1), 112-118.

Angell, T. (2020, October 29). Cory Booker urges New Jersey voters to legalize marijuana as data shows supporters outraising opponents. *Marijuana Moment: Politics*.

https://www.marijuanamoment.net/cory-booker-urges-new-jersey- voters-to-legalize-marijuana-as-data-shows-supporters-outraising-opponents/

Angell, T. (2021, January 15). Feds to compile tax revenue data. *Marijuana Moment*.

https://www.marijuanamoment.net/feds-to-compile-cannabis-tax-revenue-data-newsletter-january-15-2021/

Antaki, C., Billig, M., Edwards, D., & Potter, J.A. (2003). "Discourse analysis means doing analysis: A critique of six analytical shortcomings." *Discourse Analysis Online*, *1*(1), 1-24. http://www.shu.ac.uk/daol/previous/v1/n1/index.htm.

Aydelotte, J.D., Mardock, A. L., Mancheski, C. A., Quamar, S. M., Teixeira, P.G., Brown, C. V. R., & Brown, L. H. (2019). Fatal crashes in the 5 years after recreational marijuana legalization in Colorado and Washington. *Accident Analysis & Prevention*, *132*, 105284.

Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.

Baker, P., Gabrielatos, C., Khosravinik, M., Krzyzanowski, M., McEnery, A., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and Society*, *19*(3), 273-306.

Bauer, M. W. and Aarts, B. (2000). Corpus Construction: a Principle for Qualitative Data Collection. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative Researching with Text, Image and Sound*, (pp. 19-37). Sage Publishing.

Bednarek, M. (2008). Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory*, *4*(2), 119-139.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Black, L. (2019, April 15). The Little-Known Story of Marijuana Legalization. *The Stranger*.

https://www.thestranger.com/green-guide-spring-2019/2019/04/15/39894305/the-untold-story-of-marijuana-legalization legalization

Brezina, V. (2018). *Statistics in corpus linguistics*: *a practical guide*. Cambridge University Press.

Brezina, V., McEnery, T. & Wattam, S. (2015). Collocations in Context. *International Journal of Corpus Linguistics*, *20*(2), 139-73.

Brezina, V., Weill-Tessier, P., & McEnery, A. (2021). #LancsBox v. 6.x. [software package]

Caulkins, J.P. and Lee, A.C. (2012). The drug policy roulette. *National Affairs*, (12), 35-49.

Carnevale, J. T., Kagan, R., Murphy, P. J., & Esrick, J. (2017). A practical framework for regulating for-profit recreational marijuana in US States: Lessons from Colorado and Washington. *International Journal of Drug Policy* (42), 71–85.

Cavazos-Rehg, P. A., Krauss, M., Fisher, S.L., Salyer, P., Grucza, R. A., Bierut, L. J. (2015). Twitter Chatter about Marijuana. *Journal of Adolescent Health*, *56*(2), 139-45.

Cerdá, M., Mauro, C., Hamilton, A., Levy, N.S., Santaella-Tenorio, J., Hasin, D., Wall, M. M., Keyes, K.M., & Martins, S.S. (2020). Association between recreational marijuana legalization in the United States and changes in marijuana use and cannabis use disorder from 2008 to 2016. *JAMA Psychiatry*, *77*(2)165.

Colton, G. (2020, September 28). Fourth Annual Report on Harms of Marijuana Legalization. *Smart Approaches to Marijuana (SAM)*.

https://learnaboutsam.org/sam-releases-fourth-annual-report-on-harms-of-marijuana-legalization/

Dana-Farber Cancer Institute (2021, January 11). New evidence of health threat from chemicals in marijuana and tobacco smoke. *ScienceDaily.*

https://www.sciencedaily.com/releases/2021/01/210111084226.htm

Dash, N. S. (2010). *Corpus linguistics: A general introduction.* CHL.

Davis, R. (2021, January 14). Colorado's record-breaking marijuana sales top $2 billion in 2020. *Marijuana Moment-Politics.* https://www.marijuanamoment.net/colorados-record-breaking-%09marijuana-sales-top-2-billion-in-2020/

Desilver, D. (2014). Feds may be rethinking the drug war, but states have been leading the way. *Pew Research Center*, Washington D.C.

Disa.com (July 2022). https://disa.com/map-of-marijuana-legality-by-state

Egbert, J., Biber, D., & Gray, B. *Designing and Evaluating Language Corpora*. Cambridge University Press.

Gabrielatos, C. (2018). Keyness analysis: nature, metrics and techniques. In C. Taylor & A. Marchi, A. (Eds.), *Corpus approaches to discourse: A critical review* (pp. 1-31). Routledge.

Gabrielatos, C. & Marchi, A. (2012). Keywords: appropriate metrics and practical issues. CADS International Conference, Bologna, Italy, 13-15, September 2012. https://repository.edgehill.ac.uk/4196.

Gomez, S. (2020, November 12). All the States that legalized marijuana. *Addiction Center*.

https://www.addictioncenter.com/news/2020/11/states-legalized- marijuana/

Grabenauer, M. (2020). *Differences in cannabis impairment and its measurement due to route of administration*. Office of Justice Programs, National Criminal Justice Reference Service.

Hake, M. L. (2019). *Marijuana Legalization and Traffic Fatalities Involving Cannabinoids*. [Unpublished Doctoral Dissertation. Walden University, College of Social and Behavioral Sciences, Public Policy Administration.]

https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=7609&context=dissertations

Isa, N. A. (2017). *Legalization of Recreational Marijuana and Its Impact on Economic Activities*. [Unpublished Master of Public Policy. Faculty of the Graduate School of Arts and Sciences of Georgetown University.]

https://repository.library.georgetown.edu/bitstream/handle/10822/1043953/Isa_georgetown_0076M_13598.pdf?sequence=1&isAllo wed=y

Jaeger, K. (2021, February 3). Illinois sets new marijuana sales record in first month of 2021. *Marijuana Moment-Business*

https://www.marijuanamoment.net/?s=Illinois+sets+new+marijuana+sales

James, A., James, C., & Thwaites, T. (2013). The brain effects of cannabis in healthy adolescents and in adolescents with schizophrenia: A systematic review. *Psychiatry Research: Neuroimaging*, *214*(3), 181-189.

Johnston, L. D., Miech, R. A., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Patrick, M. E. (2020). Monitoring the Future–National Survey Results on Drug Use 1975–2019: Key Findings on Adolescent Drug Use. *The National Institute on Drug Abuse at the National Institutes of Health.*

Kilgarriff, A. (2009). Simple maths for keywords. In M. Mahlberg, V. González-Diaz, & C. Smith (Eds.). *Proceedings of the Corpus Linguistics Conference, CL2009*. Liverpool: University of Liverpool. http://ucrel.lancs.uk/publications/CL2009/171

Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 157-176). Amsterdam/Philadelphia: John Benjamins.

Mandiberg, S. F. (2019). A Hybrid Approach to Marijuana Federalism, *Lewis & Clark Law Review* 823.

Map of Marijuana Legality by State (July 2022). https://disa.com/map-of-marijuana-legality-by-state

Mautner, G. (2007). Mining large corpora for social information: The case of elderly. *Language in Society*, *36*(1), 51-72.

McEnery, T. & Hardie, A. (2011*). Corpus linguistics: Method, theory and practice.* Cambridge University Press.

McEnery, T. & Wilson, A. (2001). Corpus linguistics: An introduction, 2nd ed. Edinburgh University Press.

McGinty, E. E., Niederdeppe, J., Heley, K., & Barry, C. L. (2017). Public perceptions of arguments supporting and opposing recreational marijuana legalization. *Preventive Medicine*, *99*, 80-86.

Miller, T. (2018). *The consequences of legalizing recreational marijuana: Evidence from Colorado and Washington.* [Purdue University. ProQuest Dissertations, Publishing, 10844298]

Nader, D. A. & Sanchez, Z. M. (2018). Effects of regular cannabis use on neurocognition, brain structure, and function: a systematic review of findings in adults. *The American Journal of Drug and Alcohol Abuse, 44*(1), 4-18.

NORML (July 2022). *Medical Marijuana Laws.* NORML and The NORML Foundation. https://norml.org/laws/medical-laws/

Opensecrets.org/dollarocracy/index.php (2019)**.** *The top 10 things every voter should know about Money-in-politics.*

https://www.opensecrets.org/resources/dollarocracy/index.php

Gavin, W. (April 4, 2022). The marijuana industry spent millions lobbying on legalization in 2021. https://www.opensecrets.org/news/2022/04/the-marijuana-industry-spent-millions-lobbying-on-legalization-in-2021/

Park, S-Y, Holody, K. J. (2018). Content, exposure, and effects of public discourses about marijuana: A systematic review. *Journal of Health Communication*. *23*(12), 1036-1043.

Partington, A. (1998). *Patterns and meanings*. John Benjamin.

Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, *9*(1), 131-156.

Partington, A. & Marchi, A. (2015). 'Using corpora in discourse analysis.' In D. Biber and R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics*, (pp. 216-234). Cambridge University Press.

https://doi.org/10.1075/ijcl.9.1.07par

Paschall, M.J., Garcia-Ramirez, G., & Grube, J. W. (2021). Recreational Marijuana Legalization and Use among California Adolescents: Findings from a Statewide Survey. *Journal of Studies on Alcohol and Drugs*, *82*(1), 103-111.

Police1 Staff. (2020, September 30). '*Policing in an era of legal marijuana' survey results*.

https://www.police1.com/policing-era-legal- marijuana/articles/policing-in-an-era-of-legal-marijuana-survey-results-VpFRIKz5phYMtLP0/

Pollini, R. A., Romano, E., Johnson, M., B., & Lacey, J. H. (2015). The impact of 148 marijuana decriminalization on California drivers. *Drug and Alcohol Dependence*, *150*,135–140.

Richards, J. R., Smith, N. E., & Moulin, A. K. (2017). Unintentional cannabis ingestion in children: A systematic review. *Journal of Pediatrics*, *190*, 142-52

Sabet, K. A. (2021a). *Smokescreen*. Forefront Books.

Sabet, K. A. (2021b). Lessons learned in several states eight years after states legalized marijuana. *Current Opinion in Psychology*, *38***,** 25– 30.

Salomonsen-Sautel , S., Min, S-J, Sakai, J. T., Thurstone, C. & Hopfer, C. (2014). Trends in fatal motor vehicle crashes before and after marijuana commercialization in Colorado. *Drug and Alcohol Dependence*, *142*, 137-44.

Sinclair, J. (2004a). The lexical item. In J. Sinclair & R. Carter (Eds.). *Trust the Text. Language, Corpus and Discourse*, (pp. 149-172). Routledge.

Sinclair, J. (2004b). The search for units of meaning. In J. Sinclair & R. Carter (Eds.). *Trust the Text, Language, Corpus and Discourse* (pp. 138-148). Routledge.

Stubbs, M. (1996). *Text and Corpus Analysis: Computer Assisted Studies of Language and Culture.* Blackwell.

Stubbs, M. (2001). *Words and Phrases: Corpus Studies and Lexical Semantics.* Blackwell.

Tang, X. & Liu, G. (2018). Solving contradictions in semantic prosody analysis with prosody concord. *International Journal of Corpus Linguistics*, *23*(4), 437-466. https://doi.org/10.1075/ijcl.17057.liu

van Dijk, T.A. (1998). Discourse and Ideology [Editorial]. *Discourse and Society*, *9*(3), 307-308.

van Dijk, T. A. (2006). Discourse and manipulation. *Discourse & Society*, *17*(3), 359-383.

van Green, T. (2021, April 16). *Americans overwhelmingly say marijuana should be legal for recreational or medical use*. Pew Research Center, Washington, D.C. https://pewrsr.ch/3doCMAq

Weinstein, A., Brickner, O., Lerman, H., Greemland, M., Block, M., Lester, H., Chisin, R., Mechoulam, R., Bar-Hramburger, R., Freedman, N., & Even-Sapir, E. (2008). Brain imaging study of the acute effects of *Δ9*-tetrahydrocannabinol (THC) on attention and motor coordination in regular users of marijuana. *Psychopharmacology*, *196*,119-131.

Zamost, S., Lee, M., & Schlesinger, J. (2019, July 12). A look inside the black market for weed shows the huge threat to legal businesses**.** *CNBC Interview*.

https://www.cnbc.com/2019/07/11/las-black-market-for-weed-threatens-the-growth-of itslegalbusiness.html?&qsearchterm=A%20look%20inside%20the%20black%20market %20for%20weed%20shows%20the%20huge%20threat%20to%20legal%20businesses.

Zhang, M. (2021, April 6). Koch-backed group joins marijuana push after Zoom with Snoop Dogg: The Cannabis Freedom Alliance could change the dynamics of the marijuana legalization debate. *Politico*.

https://www.politico.com/news/2021/04/06/charles-koch-snoop-dogg-marijuana legalization-479148