Qual Quant
DOI 10.1007/s11135-017-0517-4

Measuring populist discourse with semantic text analysis: an application on grassroots populist mobilization

Paris Aslanidis[1]

Abstract Populism is a concept employed to qualify the political behavior of a large number of actors at a worldwide scale, with scientists classifying the latter into populists and non-populists according to dimensions such as ideology, strategy, discourse, economic policy, and even style. This article analyzes existing schools of thought on the nature of populism and argues that conceptualizing populism as a specific type of anti-elite discourse in the name of the People is both conceptually and methodologically the most coherent and useful way to understand the phenomenon. Additionally, it suggests discarding crude, dichotomous classification in favor of a gradated view of populist mobilization by means of quantifying populist discourse and observing its spatial and temporal variation. It adds value to current methods of measurement by demonstrating why and how clause-based semantic text analysis can provide optimal quantitative results while retaining qualitative elements for mixed-methods analysis. Aiming, moreover, at expanding the scope of populism studies by overcoming a narrow view that focuses exclusively at party system developments, it applies semantic text analysis to the study of grassroots mobilization during the Great Recession. Results point to the wide use of populist discourse on the part of movement activists seeking an inclusive language when framing disparate social grievances in a given constituency, a finding with important implications with regards to how populism can facilitate straddling the divide that purportedly distinguishes institutionalized party system behavior from the social movement milieu.

Keywords Populism · Discourse · Text analysis · Social movements · Quantitative methods

✉ Paris Aslanidis
paris.aslanidis@yale.edu

[1] Department of Political Science and Hellenic Studies Program, Yale University, 34 Hillhouse Avenue, New Haven, CT 06520-8206, USA

Published online: 06 June 2017

1 Introduction

A great number of worldwide political phenomena gets treated today as populist, yet the analytical and conceptual structure supporting the term's broad application in academic and media circles is not entirely clear. Frequently, analysts fail to communicate a sufficiently rigorous understanding of populism prior to analyzing empirical instances classified as such (Moffitt 2016). In other cases, definitions are supplied but their operationalization is sufficiently vague to allow *ad hoc* classificatory decisions (Pappas 2016). Nevertheless, the concept's extension is taken to cover a vast area, from the xenophobic right in Europe (Mudde 2007), to the electoral surge of the left in Latin American since the late 1990s (Conniff 2012), and the highly polarizing 2016 US presidential primaries (Oliver and Rahn 2016).

Maintaining a coherent view of populism that travels seamlessly across this extensive empirical terrain is admittedly not a straightforward undertaking. The field is ripe with contrasting interpretations, ranging from seeing populism as ideology (Mudde 2007), to style (Moffitt 2016), regime (Caramani 2017), discourse (Laclau 1977), economic policy (Dornbusch and Edwards 1991), and strategy (Weyland 2001).[1] Frustration with conceptual battles that seem increasingly futile tempt many to altogether abandon the quest for proper analytical rules and opt for a 'you know it when you see it' attitude or unquestionably endorse opinions of 'national experts' on whether X or Y party or leader is populist. Yet others do stand firm by their chosen definitions and follow their implications to the end, even if that means denying the populist label to political episodes widely seen as such[2] or attaching it to ones that are usually spared the association.[3]

This conceptual morass over defining attributes, as well as its corollary, the discord over classification, is of course not unique to populism. Basic political concepts such as democracy (Munck and Verkuilen 2002), authoritarianism (Stenner 2005), and liberalism (Bell 2014), are equally hard—if not harder—to disambiguate and employ for social-scientific analysis. However, this article argues that populism can be spared the dismal fate of other 'essentially contested concepts' (Gallie 1956). Compromising on a minimal understanding of populism as a discursive frame that emphasizes the political primacy of the People over any elite claims on political authority may allow us to infuse our classificatory efforts with increased rigor and, even more importantly, to measure populism in quantitative terms rather than analyzing it in a purely qualitative fashion.

This endeavor will require adherence to three core ideas, to be defended in the next section: (a) that, even if rhetoric is perhaps not the *sole* defining element of populism, it is impossible to grasp populist mobilization without some exhibition of anti-elite discourse in the name of the sovereign People; (b) that since populism *is* discourse, it can emanate from any agent capable of communicating political claims (political parties, leaders, movements, the media, etc.) and its traces should thus be accessible within various types of text produced by these agents (speeches, interviews, manifestos, tweets, slogans, editorials, etc.); and (c) that populist discourse inevitably manifests in degrees of intensity: it can be employed at different levels of frequency by different agents and its ensuing cross-case and cross-temporal variation undermines—or at least questions—rigidly dichotomous classifications. In sum, the orienting thesis is that populism, as a discursive frame that delivers an

[1] The list is non-exhaustive.

[2] Müller (2016), for instance, argues that the US People's Party in the 1890s was not populist.

[3] For example, Mair (2002) discusses Tony Blair as a populist leader in the UK.

anti-elite message in the name of the sovereign People, can be utilized in varying degrees by political agents, producing measurable output within political text.

In the next sections, this article develops a case for the discursive approach to populism, presents the state of existing quantitative approaches, argues in favor of choosing semantic text analysis over competing alternatives of measuring populism, and demonstrates how semantic analysis can be applied to political texts produced by movements active during the Great Recession in order to assess their populist character.

2 Conceptualization

Ever since its inception, populism has undergone analysis in predominantly qualitative and dichotomous terms, allowing its association with variables that later proved partial or redundant as means of classifying populist phenomena. Decades of study have now taught us that populism is not unequivocally tied to the mobilization of specific social classes (Müller 2016), to unique socio-historical circumstances of development (Panizza 2005), to economic policies of fiscal irresponsibility (Canovan 2005; Roberts 1995), or to strictly top-down electoral phenomena (Jansen 2011).[4] All these dimensions, once considered necessary and/or sufficient, have been gradually discarded in face of empirical and theoretical developments (Mudde 2007). Only a few ideas have thus survived this evolutionary process of elimination.

To begin, most scholars would today find it hard to disprove Canovan's (1981: 294) claim that 'all forms of populism without exception involve some kind of exaltation of and appeal to "the people," and all are in one sense or another antielitist.' The fact that populists fundamentally rely on the construction of a social dichotomy of people versus elites to mobilize their supporters has become increasingly obvious. Even before Canovan, Laclau (1977) had analyzed populism as the discursive interpellation of a '"people"/power bloc contradiction' (p. 196). Later, Kazin (1995), in his influential historical analysis of American populism, pointed to language as 'the most basic and telling' (p. 1) characteristic of populism: a 'language whose speakers conceive of ordinary people as a noble assemblage not bounded narrowly by class, view their elite opponents as self-serving and undemocratic, and seek to mobilize the former against the latter' (p. 1). Studying Latin American populism, de la Torre (2000: 4) similarly defined the term as 'a rhetoric that constructs politics as the moral and ethical struggle between el pueblo and the oligarchy' and Hawkins (2009: 14) saw it as 'a Manichaean discourse that identifies Good with a unified will of the people and Evil with a conspiring elite.' In more recent work, Oliver and Rahn (2016: 190) also understand populism as 'a type of political rhetoric that pits a virtuous "people" against nefarious, parasitic elites who seek to undermine the rightful sovereignty of the common folk.'

Other important strands see populism as ideology (Mudde 2004; Albertazzi and McDonnell 2008), strategy (Weyland 2001; Philip and Panizza 2011), or style (Moffitt 2016). Yet, while a political ideology usually denotes a set of ideas with a fair level of coherence that carry clear programmatic implications on areas of policy and the general organization of society (Gerring 1997), populism is regularly attached to a colorful assortment of politicians, ranging from the far right to the far left and everything in-between. Philip and Panizza (2011: 71) therefore find it 'clear that [populism] has no distinctive ideology, as it effectively incorporates a wide range of diverse and often

[4] For an overview of the academic development of the term, see Moffitt (2016), Chap. 2.

contradictory political beliefs.' Tellingly, ideological claims for populism are unable to escape discursive implications, and are almost always couched in terms of rhetoric. Mudde (2004: 543), for instance, sees populism as a 'thin-centred' ideology depicting a society 'ultimately separated into two homogeneous and antagonistic groups, "the pure people" versus "the corrupt elite"', adding that this ideology 'argues that politics should be an expression of the *volonté générale* (general will) of the people,' while conceding that populism is 'moralistic rather than programmatic'. Similarly, Albertazzi and McDonnell (2008: 3) see populist ideology as pitting 'a virtuous and homogeneous people against a set of elites and dangerous "others" who are together depicted as depriving (or attempting to deprive) the sovereign people of their rights, values, prosperity, identity and voice.' In both these cases, nothing specifically ideological is offered outside the Manichean form already endorsed by those who see populism as 'mere' discourse. Promoting this discursive feature to the status of ideology has been inadequately problematized, given the many theoretical and empirical gaps inherent in the process (Aslanidis 2016a; Moffitt 2016).

Hence, dropping the superfluous ideological clause brings the previous perspective squarely in line with the discursive approach analyzed earlier. In turn, defining populism as strategy likewise exhibits compatibility with the discursive view. If politicians are generally understood as strategic agents, then their discourse—alongside other elements of their behavior—is bound to follow strategic calculations. Consequently, Bonikowski and Gidron's (2016: 1593) definition of populism as a 'discursive strategy that juxtaposes the virtuous populace with a corrupt elite and views the former as the sole legitimate source of political power' again points to the primacy of rhetoric. The same is true for Philip and Panizza (2011: 10) who see it as 'a political strategy based on the discourse of popular unity and the stigmatization of unpopular elites.' Even though it is impossible to distinguish populist 'ideologues' from mere 'tacticians', there are indeed politicians who openly admit using populist discourse strategically in order to maximize electoral support (Iglesias 2015).

Conceiving populism as political style acknowledges the importance of discourse (Moffitt 2016), while emphasizing the significance of performance and aesthetics in 'doing populism'. As such, it can be seen as an addition to studying populism's linguistic manifestations, with the idea being that populists may employ particular pieces of attire, props, or gestures, in order to demonstrate their affinity towards ordinary people and their distance vis-à-vis 'elites'. However, while populist style *may* accompany populist discourse, at least in a visually mediatized form, it is rather safe to assume that mute performance alone cannot sufficiently broadcast the populist message to the audience; language remains crucial. Moreover, apart from the near impossibility of operationalizing 'style', this approach assigns undue primacy to political leadership, an association that does not survive empirical scrutiny and unnecessarily excludes instances of populism as a grassroots phenomenon (Formisano 2008; Gerbaudo 2012).

The arguments above defend discourse as the proper conceptual terrain to study populism. The latter is understood as a type of anti-elite rhetoric in the name of the sovereign People (Aslanidis 2016a), broadcasted by various types of political entrepreneurs who champion the People as the ultimate source of political authority and condemn powerful elites as scheming to usurp—or having already usurped—popular sovereignty in pursuit of their own narrow interests. Nevertheless, to facilitate a more hospitable introduction to the methodological contribution of this paper, this substantive view can alternatively be read as an instrumental one, suggesting to scholars holding non-discursive views that—conceptual issues aside—operationalizing populism as a discursive frame facilitates quantification, even if dissenters would hold that we are merely measuring shadows cast by a

more elegant genus. The next section reports on the efforts of those who have endorsed either of these two perspectives for quantitative work.

3 Methodological overview

The main benefits of quantifying populism for comparative analysis are three-fold: (a) increasing measurement reliability and validity, thus helping to avoid researcher bias; (b) transforming the dependent variable from dichotomous (populist/non-populist) to incremental (more/less populist); and (c) constructing large datasets amenable to statistical and causal analysis. Irrespective of specific definitions, virtually all attempts at quantification have treated populism as discursive output, measured in various types of (written or oral) text. A summary of published quantitative research on populism is provided in Table 1, divided into three broad categories: dictionary-based computerized content analysis, holistic grading, and traditional thematic text analysis, discussed sequentially in the following subsections.

3.1 Dictionary-based computerized content analysis

Pauwels (2011) and Rooduijn and Pauwels (2011) assembled special-purpose dictionaries and utilized the processing power of computers to measure populism in political texts produced by West European parties. Operationalizing Mudde (2004), they distill two necessary and sufficient dimensions: people-centrism and anti-elitism. The frequency of these two latent themes is calculated by counting occurrences of specific words, chosen to indicate populist traces.[5] Bonikowski and Gidron (2016) employ their own dictionary to gauge populism in American presidential campaigns between 1952 and 1996 by capturing the moral opposition of people and elites in more than 2400 speeches. Oliver and Rahn (2016) closely follow Bonikowski and Gidron (2016) but opt to apply three distinct dictionaries: two for capturing 'political' and 'economic populism' and a third one for blame attribution.

The strongest point of computer-assisted analysis is the guarantee of perfect reliability, an achievement unapproachable to human coding. However, validity is an equally important performance benchmark that depends crucially on the specific content of the dictionaries employed. Unfortunately, dictionary quality is notoriously hard to establish (Grimmer and Stewart 2013). In effect, human interpretative bias is still at work, concealed within the preparatory stage of choosing words to populate dictionaries (Shapiro and Markoff 1998).[6] Researchers habitually refer to the burdensome preliminary qualitative reading of their subject matter before constructing their final dictionaries. The outcome

[5] Harrison and Bruter (2011) worked along similar lines in their analysis of the European extreme right.

[6] Pauwels (2011) includes *absurd*, *admit*, *politic**, *propaganda*, *ruling**, *tradition**, and *truth*; (Harrison and Bruter 2011) include *nonsense*, *shambles*, *chaos*, *contempt*, *favouritism, blind, innovation,* and *regret*; Oliver and Rahn (2016) include *politician(s)*, *IRS*, *donors*, and *CEOs*. The validity of such spurious choices is evidently contested.

Table 1 Original scholarship on quantitative content analysis of populism by chronological order

Author(s)	Operational definition	Sampling unit(s)	Coding unit	Themes/index	Reliability scores
Oliver and Rahn (2016)	Based on Bonikowski and Gidron (2016)	2016 US primary campaign speeches	Words and phrases	Dictionary-based	N/A
Bonikowski and Gidron (2016)	Based on Laclau (1977) and Mudde (2007)	US presidential campaign speeches	Words (n-grams)	I. Dictionary-based II. Unspecified (see footnote 6)	I. N/A II. $\kappa = 0.70$
Bernhard et al. (2015)	March (2012), based on Mudde (2004)	Press releases, press conferences, party newspaper, political ads	Eq. sampling unit	1. People-centrism 2. Anti-elitism 3. Popular sovereignty	Unavailable
Rooduijn (2014)	Mudde (2004)	Newspaper opinion articles	Paragraph	1. People-centrism 2. Anti-elitism	$\alpha = 0.81$ $\alpha = 0.72$
Rooduijn et al. (2014)	Mudde (2004)	Party manifestos	Paragraph	1. People-centrism 2. Anti-elitism	$\alpha = 0.72$ $\alpha = 0.69$
Vasilopoulou et al. (2014)	Vasilopoulou et al. (2014)	Party leader parliamentary speeches	Core-sentence	1. Blame-shifting index 2. Exclusivity index	Unavailable
Rooduijn and Pauwels (2011)	Based on Mudde (2004)	Party manifestos	I. Paragraph II. Word	I. 1. People-centrism 2. Anti-elitism II. Dictionary-based	I. $\alpha = 0.73$ $\alpha = 0.70$ II. N/A
Pauwels (2011)	Based on Mudde (2004)	Party manifestos, party magazines	Word	Dictionary-based	N/A
Reungoat (2010)	Reungoat (2010)	European election manifestos	Excerpt	1. People-index 2. Anti-elite index 3. Claim for democracy index	Unavailable
Hawkins (2009)	Hawkins (2009)	Speeches of chief executives	Eq. sampling unit	None (holistic grading)	Stage I. $\kappa = 0.68$ Stage II. $\kappa = 0.44$

Table 1 continued

Author(s)	Operational definition	Sampling unit(s)	Coding unit	Themes/index	Reliability scores
Jagers and Walgrave (2007)	Jagers and Walgrave (2007)	10 min. TV broadcasts	Excerpt	1. People-index 2. Anti-establishment index 3. Exclusivity-index	Unavailable

There is a set of further works which use content analysis to study populism in political texts, without however having the aim to produce a ratio scale metric of populist discourse, but rather to investigate various qualitative facets of populist themes; see Deegan-Krause and Haughton (2009), Ruzza and Fella (2011), Caiani and della Porta (2011) and Bruhn (2012). Moreover, some authors (e.g. Moffitt and Tormey 2014) consider Armony and Armony (2005) as an early example of quantitative content analysis of populism; this is mistaken, since it is simply a content analysis of answers to the question "In spite of the current situation, do you feel proud to be Argentine? Why?" taken from an Argentine online forum

varies: Pauwels (2011) uses twenty-eight terms, Harrison and Bruter (2011) about two-hundred-and-thirty, and Bonikowski and Gidron (2016) thirty-six.[7]

No golden rule exists on which and how many terms the ideal dictionary should incorporate to capture populist discourse. The considerable resources invested in this step blunt the argument celebrating dictionary-based methods as less labor intensive than rival approaches (Franzosi 2008). Dictionaries carry a broad brush when identifying populism and pose hurdles to comparative research across different languages and political cultures.[8] Most importantly, decontextualizing and sterilizing key terms such as 'people' or 'elites' is bound to produce errors. Harrison and Bruter (2011) and Bonikowski and Gidron (2016) found it necessary to manually evaluate the outcome of their automated analyses in an *ad hoc* fashion, rooting out false positives and other discrepancies in their data.[9] Surprisingly, Bonikowski and Gidron (2016) opted to altogether discard the term 'people' from their dictionary; Rooduijn and Pauwels (2011) declared the measurement of people-centrism unfeasible, settling with counting words for anti-elitism alone. In both cases the analytical value of findings is therefore compromised and even if particular results exhibit some level of validity, generalization becomes questionable (Pauwels 2014).

[7] The numbers reflect both single words and n-grams employed by the authors. Words can be stemmed or remain whole. Oliver and Rahn (2016) only provide examples from their dictionaries rather than their whole content.

[8] Rooduijn and Pauwels (2011) had to include 'context-specific words' in their cross-national study, since many idiosyncratic political expressions fail to transcend linguistic barriers. The significance of the word *regenten* for Flemish and Dutch populism, or *katestimeno* for Greek populism are telling illustrations. In terms of proportion, the Dutch dictionary in Rooduijn and Pauwels (2011) contains fourteen 'core' words and seven 'context-specific' words, proving that contextual impact is far from negligible when analyzing text (Krippendorff 2004).

[9] Bonikowski and Gidron (2016) read 40.1% of their total dataset, hand-coding excerpts from 890 speeches, and manually assessing all positive instances identified in a first iteration of dictionary analysis. They report a Cohen's *kappa* of 0.70 for intercoder reliability when performing manual coding (footnotes 5–6).

Due to these inherent limitations Shapiro and Markoff (1998) evaluate automated methods as hobbling rather than encouraging research. Bauer (2000) proclaims them simply absurd. Yet these are rather hasty dismissals of an otherwise useful approach. Apart from their stand-alone value, dictionaries can, for instance, be used as a quick way to triangulate other methods (see Pauwels 2014) or contribute to preliminary case selection. Future developments in information theory may soon improve automated methods, overcoming some of the reservations noted above.

3.2 Holistic grading

Computerized text analysis methods work by splitting text into words, assigning a binary value to each word-as-coding-unit according to dictionary membership. They usually produce a percentage of relevant words compared to the total word count, informing us that a sample unit (e.g. a political speech or manifesto) is *x*% populist. This signifies an advantage over approaches that report nominal or ordinal values, since dictionary-based methods benefit from the highest possible dynamic range when gauging the intensity of populist discourse.[10]

Bonikowski and Gidron (2016) diverge from this norm. They ignore the frequency of populist keywords in the sample unit and use their dictionary to grade their sampling unit (campaign speeches) *dichotomously*, as populist or non-populist.[11] Hence, they follow a hybrid approach, situated between dictionary-based computerized content analysis and holistic grading. With the latter, 'whole-text analysis' (Hawkins 2010: 8) of political texts is performed, relying on informed interpretation by trained coders in evaluating the populist degree of the sample unit as a whole. Coders assign a 'populist grade' along a three-point scale of either 0 (non-populist), 1 (mixed), or 2 (populist). Bernhard et al.'s (2015) analysis of the discourse of the Swiss SVP fits in the same subcategory since they also analyze non-segmented sampling units.[12]

With proper training, holistic grading can indeed yield results of enhanced validity compared to automated methods. The project is however predicated upon two limitations. First, reliability. Hawkins (2009) reports an acceptable Cohen's *kappa* of 0.68 for the first coding stage, but a much lower score of 0.44 for the second stage.[13] Longer training sessions can perhaps improve this performance, and it should be noted that Cohen's *kappa* fails to take into account the ordinal nature of the grading scale.[14] Coding populist dimensions in Bernhard et al. (2015) was performed exclusively by a contributing author and therefore its reliability cannot be assessed.[15]

Without doubt, impeccable reliability scores for latent concepts such as populism are a rather nonsensical requirement and some degree of leniency is justified. The second limitation is of a different nature, arising from the use of a limited ordinal scale of only three levels of differentiation. Even though this is still superior to categorical qualitative

[10] A higher range would require splitting text into syllables or letters (nonsensical in this context).

[11] Their rationale is using this evaluation as a boolean variable for statistical analysis.

[12] Bernhard et al.'s (2015) coding scheme draws from March (2012) with the overall score for each sample unit resulting from the combination of scores in three different dimension.

[13] Hawkins (2009) attributes this moderate performance on the large number of coders, the small set of speeches, and the limited experience of coders in reading political texts.

[14] For a critique against using Cohen's *kappa* for content analysis and a way to calculate the reliability of ordinal data, see Hayes and Krippendorff (2007).

[15] Personal communication with authors.

appraisals or the scale employed by Bonikowski and Gidron (2016), it is still rather insufficient in capturing enough variation among cases, especially in small samples. Moreover, holistic grading fails to offer secondary quantitative or qualitative information such as frequency of specific themes or references to social actors, so as to facilitate a more comprehensive analysis of populist discourse. Holistic grading is arguably the best compromise between validity and speed in our current toolset, but problems of reliability and resolution somewhat undermine its standing.

3.3 Thematic text analysis

Standard quantitative content analysis refers simply to the quantitative analysis of qualitative data (Bernard and Ryan 2010). Text analysis, its most typical application, uses a coding frame as a standard instrument of comparison, with humans applying a set of codes on sections of textual material to derive inferences for further analysis (Bauer 2000). This process requires precise definitions, explicitness of procedure, and proper training of coders to produce reliable, valid, and replicable findings (Franzosi 2008). Such provisions guarantee a check on the impact of unrestrained human interpretation, a requirement verified by statistical intercoder reliability tests as indicators of quality (Neuendorf 2002). Save for dictionary-based and holistic grading approaches, the remaining contributions in Table 1 rely on some variant of 'traditional' thematic text analysis. Typically, they start by splitting text into meaningful and specific segments (coding units) which coders then undertake to classify into one or more thematic categories according to specific rules derived from operationalizing a certain definition of populism. Dividing occurrences with the total number of segments usually provides the final score.

In their ground-breaking analysis, Jagers and Walgrave (2007) define populism as a political communication style employed strategically to mobilize support, using transcripts of televised messages by Belgian political parties as their dataset. Despite their interesting insights and methodological innovations, their use of an 'exclusivity' index as a defining feature (alongside people-centrism and anti-elitism) results in a bias towards capturing exclusively right-wing populism (Rooduijn 2014). Moreover, people-centrism tends to get overcoded, since all references to 'the People' are coded indiscriminately into this category. Jagers and Walgrave (2007) vaguely refer to arbitrary 'excerpts' as coding units and issues of reliability are not dealt with (Pauwels 2011).

Reungoat (2010) applied a revised version of this original coding scheme on French euromanifestos, swapping the problematic exclusivity index for a claim-for-democracy index. Coding people-centrism is also overly inclusive[16] and the claim-for-democracy index captures a lot more than appeals for unrestrained popular sovereignty that distinguish populism from standard democratic appeals. Finally, general criticism of the French left is included in the anti-elite index, conflating anti-incumbent rhetoric with populist, morally-laden, anti-establishment discourse. As with Jagers and Walgrave (2007) and Reungoat (2010) employs 'excerpts' as the coding unit and does not discuss reliability ratings.

The contribution by Vasilopoulou et al. (2014) employs a quite robust segmentation approach, dissecting Greek parliamentary speeches into 'core-sentences'. However, the authors opt for an idiosyncratic conceptualization of populism that diverges considerably from the accumulated wisdom of the field, missing the basic dimension of people-centrism. While employing a novel coding scheme to construct blame-shifting and exclusivity

[16] The 'people index' incorporates references to any 'group[s] of people having explicit constant features in common' (Reungoat 2010: 311).

indexes, it is mostly anti-elitism actually captured, rather than populism. The authors explicitly refrain from coding 'we' sentences that regularly exhibit the populist tendency of speaking on behalf of 'the People', and, as with Reungoat (2010), they do not distinguish between standard oppositional rhetoric and anti-elite discourse expressed in the name of 'the People'.[17]

Rooduijn and Pauwels (2011) and Rooduijn et al. (2014) rely on paragraphs as coding units and election manifestos as sampling units to identify West European populist parties and investigate whether mainstream parties have become more populist over the years. They trained coders to identify people-centric and anti-elitist themes, yielding Krippendorff's *alpha* reliability scores of approximately 0.70.[18] Operationalizing Mudde (2004), they identify populism only when a paragraph contains both people-centric and anti-elitist claims.

The rigorous thematic text analyses performed by Rooduijn and his associates constitute a significant improvement over competing approaches in terms of operationalization, speed, and reliability. The choice of paragraph as coding unit occupies a valuable position between the overly segmented nature of dictionary coding and its opposite, holistic grading. Nevertheless, this choice inevitably imposes the semantic segmentation intended by the texts' original authors, undermining the validity of findings. Coding paragraphs as they appear in political texts is inherently arbitrary since paragraph length varies considerably among party manifestos. Moreover, these documents frequently contain bullet-pointed lists and short motivational sentences or quotes, features that further undermine comparability. Sources of political rhetoric such as speeches, interviews, and parliamentary debates are problematic in a different respect, absent a standard and systematic method of segmenting transcribed text into paragraphs. Additionally, while far superior to holistic grading, segmentation into paragraphs still involves a discount in semantic resolution, since mildly populist paragraphs receive identical scores with intensely populist ones that carry greater informative content.

Table 2 summarizes this survey, assessing existing methods with respect to validity, reliability, and resolution. Notwithstanding certain setbacks, this original body of work establishes the feasibility and usefulness of quantifying populism, highlighting its advantages over strictly qualitative assessments. The gradated measurement of populism teaches us a lot more that coarse, dichotomous classifications. At the same time, persisting limitations show that there is considerable room for improvement. Striking a balance between reliable—yet crude—dictionary-based approaches and holistic assessments that introduce valuable interpretative judgment at the cost of resolution seems like the most promising avenue. The novel approach submitted next caters for these issues, incorporating best practices while avoiding new pitfalls.

4 Clause-based semantic text analysis

When aiming to capture latent political meanings such as populism, many scholars see automated methods as particularly unrefined. At the same time, as outlined above, there are significant gains in quantification that cannot easily be dismissed. Therefore, content

[17] Reliability cannot be assessed since coding was performed exclusively by the first author (personal communication with first author).

[18] The method is also applied to newspaper articles in Rooduijn (2014), with improved reliability scores. In the same paper, scores given to 'introductory paragraphs' are doubled and paragraphs in long manifestos are weighted differently than those in shorter ones, thus employing an interesting non-linear scoring scheme.

Table 2 Assessment of quantitative literature on measuring populism in terms of validity, reliability, and resolution

Content analysis studies	Validity	Reliability	Resolution
Oliver and Rahn (2016)	Moderate	High	High
Bonikowski and Gidron (2016)	Moderate	High	Low
Bernhard et al. (2015)	High	Low	Low
Rooduijn (2014)	High	High	Moderate
Rooduijn et al. (2014)	High	Moderate	Moderate
Vasilopoulou et al. (2014)	Low	Low	High
Rooduijn and Pauwels (2011)	High	Moderate	Moderate
Pauwels (2011)	Moderate	High	High
Reungoat (2010)	Low	Low	High
Hawkins (2009)	High	Moderate	Low
Jagers and Walgrave (2007)	Low	Low	High

analytical methods can gain validity by assuming a hybrid form: providing a reliable ratio metric to expose variation in the dependent variable, while simultaneously allowing for nested, qualitative analysis. In measuring populism, a method can only aspire to accomplish this if it: (a) posits a minimal set of dimensions capable of capturing populism, (b) uses a coding unit that strikes a balance between incorporating context and allowing for significant semantic resolution, and (c) employs a scoring scheme that captures frequency reliably while retaining the qualitative content of the coding unit. Since current methods fail to satisfy all three requirements I suggest turning to a lesser-known variant of quantitative text analysis: *clause-based semantic text analysis* (CBSTA).

CBSTA (Popping and Roberts 2015) diverges in a number of ways from simply counting word occurrences within heaps of textual elements. Its most distinctive feature is the way it handles and segments the sample unit. By use of a *semantic grammar*, original text is reformulated into a set of clauses termed *semantic triplets* (Franzosi 2004, 2010), a superior coding unit in many respects (Shapiro and Markoff 1998), comprised of the elementary syntactic components of language: *subject*, *verb*, and (optionally) *object.* Originally developed by linguists, the significant value of the semantic grammar approach for text analysis is now widely acknowledged in the social sciences (Roberts 1997; Johnston and Oliver 2000; Franzosi 2004). Advantages for measuring populism are threefold: inherent reliability, increased resolution, and structural commensurability.

Semantic triplets are inherently reliable coding units. They follow objective, structural, grammatical rules, restoring control of the text back to the researcher. They are preferable to the arbitrary segmentation dictated by typographical elements (sentences or, worse, paragraphs) found in readily supplied political texts, or the necessarily unsystematic choices researchers make when transcribing oral artifacts such as speeches and interviews. The strict rules of grammar guarantee rigorous and systematic handling of textual material, producing perfectly comparable coding units in *subject-verb-object* format, in any language.[19]

Second, triplet clauses offer increased resolution compared to both paragraphs and sentences, since even natural sentences usually yield more than one semantic triplet clause, due to the inclusion of multiple verbs, subjects, or objects. This multiplies the information

[19] Apart from the option of manually distilling semantic triplets, software packages exist (e.g. IDEA, KEDS) to parse text for syntactical elements. With improved tools, we can expect increased reliability and even lower processing costs in the future. See discussion in Franzosi (2010: 62–63).

extracted from each sample unit, allowing for a wider dynamic range while enhancing validity. Triplets are also more compatible with short texts, where paragraph coding tends to become almost equivalent to holistic approaches. Therefore, CBSTA is not constrained by length. The coding scheme can treat slogans, Facebook posts, tweets, short passages, and long texts equally well.[20]

Third, semantic triplets are particularly fitting instruments for measuring populism due to their perfect structural commensurability with the formal features of populist discourse. Since the 'the perennial cry of populists' is that 'power has been stolen from the people by politicians and special interests' (Canovan 2005: 5), the core populist message is essentially structured around references to two symbolic categories of social actors (people, elites) and their interactions. This formal nature of populist discourse (Laclau 2005) fits perfectly with the syntactic form of semantic triplets, where people (*people-centrism*) and elites (*anti-elitism*) are expected to function as the subject (and/or object) part of semantic triplet clauses, with their interaction denoted by the verbal segment. Hence, semantic triplets are excellent vessels for accommodating the full essence of populist form and preparing it for further analytical dissection.

Apart from advantages in terms of quantification, semantic text analysis carries the extra merit of optionally allowing for the preservation of the qualitative features of populist discourse, facilitating mixed-methods research. This further stage of analysis can contribute to the validity of quantitative findings, since qualitative content—and not only the sum of occurrences yielding a 'populist index'—can be captured during the coding process into datasets comprised of the semantic elements (subject, verb, object) employed. The most impressive application of semantic analysis in this respect is Shapiro and Markoff's (1998) voluminous study of the *Cahiers de doléances* archives from 1789 France where the coding of grievances expressed by citizens towards King Louis XVI produced a remarkable database of French public opinion. Similarly, Franzosi (1999) and Franzosi et al. (2012) use semantic text analysis to derive quantitative and qualitative inferences from postwar Italian strike activity and newspaper accounts of lynching episodes in Georgia between 1875 and 1930. Roberts (1997) applies the same methodology to comparatively study the coverage of military conflicts by East and West Berlin radio stations, while Roberts et al. (2010) use semantic analysis to distinguish between Islamic and Hindu cultural traits, and Johnston and Alimi (2013) use semantic triplets to study Palestinian nationalism. In every instance, coders capture semantic triplets into databases subsequently probed to identify qualitative features and/or frequencies of social actors registered as subjects and objects as well as their interactions expressed in verbal form. They are also able to derive inferences about general themes emerging in the clauses, such as immigration, economics, security, and so on. Some authors (e.g. Franzosi 2004) have used large databases of this sort to derive informatively rich network graphs of social interaction.

Scholars of populism can similarly build databases of populist semantic clauses where social actors denoting 'the People' interact with 'elite' actors. Subjects and objects within populist clauses produce a {people-centrism} actor-set containing all types of invocations of 'the People', and an {anti-elitism} actor-set containing all references to 'elites'. The size of these sets contributes quantitative information and their content provides qualitative information on which social actors typically constitute 'the People' and their enemies. The

[20] To compensate for any potential loss in contextualization, coders in CBSTA projects are instructed to draw judgment from larger context units (Krippendorff 2004), or even the entire sample unit, prior to appending codes to individual semantic triplets.

verb part produces an {interaction} set containing information on the type of interplay between the two actor-sets.

As an example, consider a sample text (e.g. a speech) parsed into a sequence of semantic triplets which includes the clause *'Politicians have betrayed the people'*. If correctly interpreted as constitutive of populism, the coder registers the following entries in the respective datasets:

$$\{\text{people-centrism}\} += \{\text{'people'}\}$$
$$\{\text{anti-elitism}\} += \{\text{'politicians'}\}$$
$$\{\text{interaction}\} += \{\text{'betray'}\}$$

where the notation += signifies addition to the major set of the first column. Coding further populist semantic clauses adds more elements into the three datasets, until the sample is exhausted. Apart from producing overall frequency scores of populist discourse we can now utilize findings to answer secondary quantitative and qualitative questions, such as *Which social categories are depicted enemies of the People*?', or '*Which grievances produce tension between People and elites?'*. Datasets incorporating several national party systems or time periods can furnish interesting comparative conclusions, improving the predominantly descriptive approach usually taken in the literature.

Caiani and della Porta (2011) have published a very first application of a semantic grammar for populism, with the objective of studying populist frames in the discourse of Italian and German extreme right organizations. While their analysis leads to interesting comparisons, its methodology is not targeted at providing an overall metric for populist discourse within a specific sampling unit. Their approach uses a story grammar to locate subject-actors in a heap of data from online forums and report their relative frequencies and qualitative characteristics. Therefore, separate indexes for 'identity actors' and 'oppositional actors' do not overlap with *people-centrism* and *anti-elitism* as expected from a quantitative analysis that employs a specific scoring scheme.

Treating their corpus as a priori constituting a body of populist discourse, Caiani and della Porta (2011) are not interested in applying semantic text analysis to gauge the level of populism therein.[21] Additionally, the use of the 'statement' as a coding unit is vaguely explained, and they opt for a standardized codebook of actors rather than allowing for an inductive, emergent process of populating their database, as semantic text analysis prescribes. Despite these limitations, Caiani and della Porta (2011) illustrate that systematic research on populism is compatible with the use of a semantic grammar. In the next section, I contribute a more comprehensive and focused empirical application of this approach.

5 Measuring populism in the Great Recession movements

In existing literature, populism predominantly qualifies the behavior of institutional actors such as political parties and their leaders. Quantitative research has largely followed this trend. However, this limited perspective neglects a wide array of populist phenomena taking place at the grassroots level, outside formal institutional channels (Aslanidis 2016b). For instance, recent episodes such as the Tea Party (Formisano 2012; Van Dyke

[21] The dataset was coded by two individuals (one of which one was the first author) with no actual overlap in content, and therefore, no reliability test was performed (personal communication with first author).

and Meyer 2014) have been studied as a reflection of populist themes harking back to a long history of grassroots politics in the US (Kazin 1995; Formisano 2008). Similarly, in Latin America, anti-IMF and anti-austerity mobilizations between the late 1970s and the early 1990s have been found to exhibit a strong populist element (Walton and Seddon 1994). In Eastern Europe, populism consisted a pivotal element of bottom-up mobilization that challenged communist rule and led to the capitulation of authoritarian regimes in the region during the 1980s. Canovan (2005) and di Tella (1997) have studied *Solidarity* as a distinctly populist movement, while Pfaff's (2006: 263–264) analysis reaches the conclusion that 'East Germany's revolution was largely a populist rebellion'. And in India, regional populist social movements have come to the political fore under the banner of the *jan andolan* (Kumar 2011).

More recently, the global economic slowdown that followed the 2008 Lehman Brothers collapse precipitated an avalanche of social grievances in the West that, at least superficially, seems ripe with populist interpellations. Partly inspired by the Arab Spring, grassroots episodes of mobilization such as the Spanish *indignados*, the Greek *aganaktismenoi* and the US *Occupy Wall Street* brought thousands of people to the streets in 2011. The *Geração à Rasca* in Portugal and the *J14* in Israel were also highly successful during the same period. The shared use of protest repertoires in the form of occupying central city squares, setting-up of tent camps, establishing General Assemblies as decision-making mechanisms, and making extensive use of social media, were characteristic elements of this transnational wave of protest (della Porta and Mattoni 2014). The crisis eventually catalyzed the rise of populist electoral contenders such as SYRIZA in Greece and Podemos in Spain (Kriesi and Pappas 2015), with the 2016 campaign of Bernie Sanders echoing similar themes.

At a discursive level, these grassroots movements claimed to represent and speak on behalf of the 'People' or the '99%' in favoring the restoration of popular sovereignty or 'real democracy' purportedly thwarted by self-serving, corrupt elites of the '1%'. However, researchers using qualitative tools of classification are torn over whether some or all of these movements can be seen as partaking of a wave of populist mobilization, with some even denouncing the existence of populist movements *tout court* (cf. Aslanidis 2016b; Gerbaudo 2012; Urbinati 2014; della Porta 2015; Calhoun 2013; Gamson 2011). The application of CBSTA to the discourse of these movements can help us discern whether and to what degree populism was actually employed as a mobilizing frame in the squares of major Western cities in 2011.

Following a purposive (relevance) sampling approach (Krippendorff 2004), a dataset of movement manifestos was assembled, composed of those considered as the most important instances of Great Recession mobilization (Gerbaudo 2012). Manifestos are moderately sized texts written in narrative format, encompassing the diagnostic, prognostic, and motivational frames usually employed by social movement entrepreneurs (Snow and Benford 1988). The manifestos of the Great Recession movements are important summaries of the essence of the protests, employed as official or 'constitutional' documents approved unanimously in the General Assemblies of the squares, the sole collective decision making institutions. The most significant movements, those in Portugal, Spain, Greece, and the United States, all produced such texts, disseminated through social media and unequivocally acknowledged by their activists as representing their core ideas. Additionally, on October 15, 2011, activists from several movements jointly organized a worldwide *United for Global Democracy* protest. The respective manifesto was also included in the dataset.[22]

[22] The J14 movement in Israel did not produce a manifesto and therefore is not part of the case selection.

Table 3 Case selection

Movement	Manifesto title	Date issued	Words
Geração à rasca (Portugal)	Manifesto–precarious generation protest	Feb 18, 2011	305
15M Indignados (Spain)	Manifesto "real democracy now"	May 2, 2011	570
Aganaktismenoi (Greece)	Resolution of the popular assembly of syntagma Sq.	May 27, 2011	156
Occupy wall street (USA)	Declaration of the occupation of New York City	Sep 29, 2011	642
United for global democracy	Manifesto for global democracy	Oct 14, 2011	313
Populists (USA)	The populist manifesto	Feb 22, 1895	460
Greenpeace (USA)	Greenpeace declaration of interdependence	Nov 1976	628
Immigrant movement Int'l	International migrant manifesto	Nov 5, 2011	464

Observing some degree of populism in these manifestos would not suffice to reach a valid verdict on the substantial role played by populist discourse without employing a control group as a means of validation. Therefore, the dataset includes the *Populist Manifesto* of 1895, addressed to members of the American People's Party, widely considered as the paradigmatic case of populist mobilization in the literature (Ionescu and Gellner 1969; Canovan 1981). A second (negative) type of control is provided by including manifestos from movements typically classified within other genres of mobilization with no obvious association to populism: the *Greenpeace Declaration of Independence*, and the *International Migrant Manifesto* were selected to fulfill this role. The hypothesis is that the manifestos of Great Recession movements, if indeed populist, will illustrate a degree of populist discourse much closer to the former rather than the latter benchmark. A summary of cases can be seen in Table 3.

Regarding operationalization, as explained in the second section of this paper, populism is identified as a discursive frame that contains an anti-elite message in the name of the sovereign People. The populist frame is minimally operative along two dimensions:

(a) *People-centrism* The invocation of an overwhelming majority of 'the People' as rightful political sovereigns within a given polity, and
(b) *Anti-elitism* References against a slim minority of unaccountable power holders engaging in the misappropriation of popular sovereignty.

Moreover, it is considered that a *full populist frame* needs to incorporate (as subject *and* object) both the positively valenced group, 'the people' (or a metonym) as well as the negatively valenced one, 'the elites' (or a metonym), within a given coding unit (semantic triplet). The separate existence of *people-centrism* or *anti-elitism* in a coding unit signifies a *partial populist frame*.

Each sentence in the manifestos, or each part of a sentence hosting an inflected verb and (optionally) a subject and a predicate, underwent a process of transformation into a semantic triplet-clause. Hence, the original text was turned into a set of semantic clauses, each with a unique sequence number to preserve the original flow and integrity of contextual meaning. The transformation followed standard rewrite rules, most basic of which were (1) turning passive voice clauses into active voice, and (2) segmenting clauses that contain coordinating conjunctions.[23] Hence, the number of extracted semantic triplets

[23] Preparation of each manifesto took an average of 2 h.

exceeds the number of natural sentences, enabling us to capture rhetorical and syntactical elements vital for populist discourse such as repetition, conditional phrases, and adjectival expressions, enhancing overall resolution and providing for nuanced scoring outcomes.

Following segmentation, three coders were trained to recognize populist invocations and identify empirical examples of *people-centrism* and *anti-elitism* according to the theoretical framework presented by the researcher.[24] They were also introduced to a computer application developed by the author to facilitate the assignment of semantic clauses into one of eight independent, exclusive, and exhaustive (Popping 2010) codes (Table 4).[25] Coding required approximately 1 h per sample unit (manifesto) and was executed as follows.

Prior to assigning codes to clauses, coders were asked to read their assigned material at least twice to acquaint themselves with the context. In this first, holistic stage, they were instructed to consider the main subjects addressed by the manifesto and locate emergent collective identities, adversarial or otherwise. Then, taking up each clause in succession, they mentally constructed the set of {actors} in the clause and consulted a detailed decision scheme (Fig. 1) in order to follow the correct steps in assigning codes.[26]

Three codes (U01 for *people-centrism*, T01 for *anti-elitism*, and V01 for *full populist frame*) contribute to identifying the dimensions of populist discourse, while the remaining are classified as residual, employed to cater for research questions that lie outside the scope of this article. The overall intercoder reliability score was recorded at a Krippendorff's *alpha* of 0.745 on a random sub-sample of the data consisting of approximately 20% of the overall dataset. The respective scores for the various stages of coding can be seen in Table 5.[27] In the next section I present my scoring scheme and discuss final results.

6 Scoring scheme and results

An overall ratio index of populism is obtained by dividing occurrences of 'populist' semantic triplets with the total number of clauses extracted from the original text. The scoring scheme differentiates between *full populist frames* and *partial* populist frames that only contain elements of either *people-centrism* or *anti-elitism* in order not to miscode anti-incumbent discourse or any appeal to popular sovereignty as populism. Admittedly however, a populist text longer than just a few sentences will inevitably contain *partial* populist frames along instances of *full* populist frames, and these extra *people-centric* and *anti-elitist* appeals can be seen as contributing to the overall character of the text. A compromise is reached by assigning a full score to *full* populist frames (V01) and a weight of one-third[28] to *partial* populist frames coded as *people-centrism* (U01) and *anti-elitism*

[24] The researcher was not part of the coding team.

[25] The software package was used to streamline the coding process and avoid typing or other unnecessary coding mistakes. Any spreadsheet or word processing application can be used with equal merit.

[26] The use of decision schemes is another significant advantage of this method, enhancing reliability by organizing the coding process along a series of dichotomous decisions, increasing coding speed and minimizing errors owed to coder fatigue (Krippendorff 2004).

[27] The moderate intercoder agreement in identifying the presence of an actor in the clause (NAC) is largely due to disagreement over the eligibility of vague notions such as 'society' or 'humanity'. Further training can disambiguate such observations and lead to highly increased intercoder scores altogether.

[28] The value of factor *c*, indicating weight, is of course a matter of perspective.

Table 4 Codes and descriptions

Code	Description	Rule
U01	People-centrism	{actors} contains a populist "we"
U02	Us-residual	{actors} contains a "we" that is non-populist
T01	Anti-elitism	{actors} contains a populist "they"
T02	Them-residual	{actors} contains a non-populist "they"
V01	Full populist frame	{actors} contains a populist "we" and a populist "they"
V02	Us-vs-them residual	{actors} contains a "we" and a "they" but not both are populist
NAC	Non-actor	{actors} is empty
RES	Actor-residual	{actors} contains no "we" or "they" elements

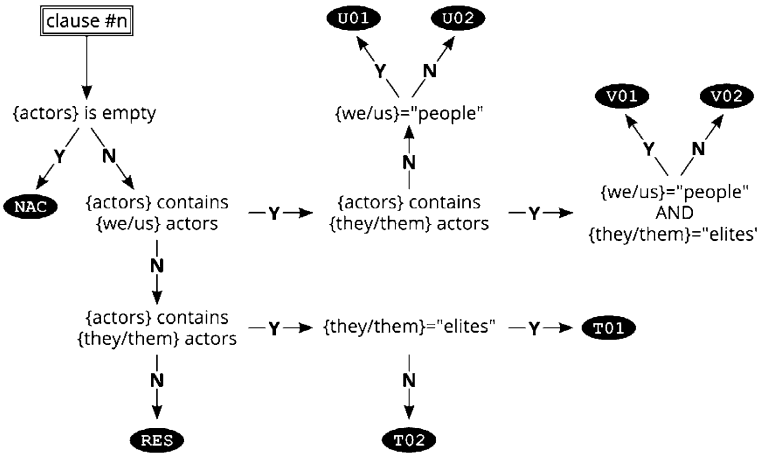

Fig. 1 Decision scheme for coding SVO clauses into one of eight categories

Table 5 Intercoder reliability scores for the various stages of coding

Coding stage	*Alpha*
Agreement on whether clause contains actor (NAC)	0.629
Agreement on whether clause contains we/us actors (U01/U02/V01/V02)	0.948
Agreement on whether clause contains they/them actors (T01/T02/V01/V02)	0.887
Agreement on *people-centrism* for we/us actors (U01/V01)	0.799
Agreement on *anti-elitism* for they/them actors (T01/V01)	0.780
Agreement on *full populist frames* (V01)	0.816

(T01).[29] The sum of populist-related clauses is divided with the total number of n clauses in the sample unit. The formula for the final *populist discourse index* (*PDI*) is then simply:

$$PDI = \frac{\sum_1^n V01 + c \times \left(\sum_1^n U01 + \sum_1^n T01\right)}{n}$$

Table 6 presents the results of the process. The manifesto of the American People's Party sets a benchmark of 0.50 for the exemplary case of populist mobilization.[30] Those from *Greenpeace* and the *Immigrant Movement International* exhibit very small traces of partial populist elements and no *full* populist frames at all. In all manifestos issued by Great Recession movements we find an average of about 20% *full* populist frames. Although there is variation in the intensity of populist discourse among movements of the Great Recession, there is an obvious overrepresentation of *full* and *partial* populist frames. Stipulating a strict threshold to dichotomously distinguish between populist and non-populist discourse is certainly a debatable decision, and the intricacies of political competition require that a degree of partially populist appeals is almost always to be expected. However, the lower weight attached to partial frames prevents the inordinate scoring of mild cases, while *full* populist frames are considered safe indicators of a populist character and their existence must be taken at face value. Therefore, in this setup, it is reasonable to assess a performance of $PDI > 0.20$ as signifying considerable usage of populist discourse, and a score of $PDI > 0.35$ as a safe indicator of populism.[31]

Regarding cross-case variation, the populist mix in Greece invests heavily in *people-centrism* while *Occupy Wall Street*, on the contrary, engages considerably in *anti-elitist* discourse. Portugal's *Geração à rasca*, while over the 0.20 threshold, constitutes a borderline case that warrants further qualitative investigation to reach a solid verdict on its populist character.

Apart from providing quantitative information, CBSTA also allows for the systematic extraction of qualitative information. In this case, it is informative to observe the two basic actor-sets that emerge from the coding process, {people-centrism} and {anti-elitism}, revealing how the crucial subjectivities of 'the People' and their enemies are constructed. Methodologically, this can be done by producing tables of actors or, more graphically, through word-clouds. To provide a simple example, we can compare the actor sets in the manifesto of the Spanish *indignados* (Table 7) and *Occupy's Declaration* (Table 8).

A range of conclusions on the particular type of populist discourse employed can be drawn from these tables. For instance, *Occupy Wall Street* names only one explicit category as enemy, 'corporations', while the *indignados* provide a host of actors who conspire against the People. Spaniards also engage in portraying 'the People' through various synonyms, a tactic that the Americans generally refrain from, presenting a sharper image of the populist frame.

Findings from applying CBSTA on the manifestos of the Great Recession movements illustrate the strikingly high level of populist discourse employed by these groups. In most cases, their PDI score is in the vicinity of paradigmatic cases of populist mobilization such as the *People's Party*, an outcome that provides quantitative triangulation to similar

[29] Residual categories receive no score.

[30] The unusually rounded score for the manifesto of the benchmark case was a matter of coincidence.

[31] For instance, a text with 10% full populist frames, 15% people-centrism, and 15% anti-elitism yields PDI = 0.20. A text with no full populist frames would require a combined 60% of partial populist frames to pass this threshold, a rather unlikely scenario for non-populist discourse. Hence, the possibility of a false positive is very slim.

Table 6 Content analysis results

Sample unit	Full populist frames (%)	People-centrism (%)	Anti-elitism (%)	Residual (%)	PDI
Geração à rasca (Portugal)	3.4	56.9	0.0	39.7	0.22
15M Indignados (Spain)	18.0	48.6	16.2	17.1	0.40
Aganaktismenoi (Greece)	21.4	73.8	0.0	4.8	0.46
Occupy wall street (USA)	20.7	16.1	50.6	12.6	0.43
United for global democracy	31.8	53.4	8.0	6.8	0.52
Populists (USA)	38.5	13.5	21.2	26.9	0.50
Greenpeace (USA)	0.0	20.0	0.0	80.0	0.07
Immigrant movement Int'l	0.0	11.4	5.7	82.9	0.06

Table 7 We/they actors (people-centrism and anti-elitism) in the Spanish indignados

Indignados (Spain)			
We-actors (people)		They-actors (elites)	
Actor	Occur.	Actor	Occur.
We-generic	33	Political class	10
(The) people	12	A few	7
Society	2	Politicians	2
Citizens	1	Businessmen	2
Like you	1	Bankers	2
The rest [against the few]	1	A minority	2
Anonymous	1	Partidocratic dictatorship (PPSOE)	1
Majority	1		

Table 8 We/they actors (people-centrism and anti-elitism) in occupy wall street

Occupy (USA)			
We-actors (people)		They-actors (elites)	
Actor	Occur.	Actor	Occur.
(The) people	16	They-generic	54
We-generic	7	Corporations	7
NYC GA	4		
Taxpayers	1		

conclusions stemming from purely qualitative research (e.g. Aslanidis 2016b; Gerbaudo 2012; Calhoun 2013). Hence, the characterization of Great Recession movements as populist carries considerable merit. This illustrates that populism is able to straddle the divide that purportedly separates political parties from social mobilization, while emphasizing the potency of an inclusive populist discourse in expressing societal grievances by means of constructing a popular identity that resonates with diverse constituencies. It becomes evident that both social movement scholars and political scientists studying party system change should study the role populism plays in transforming diffuse popular disaffection into an antagonistic dichotomy between 'people' and 'elites' that may later find institutional representation by political parties eager to adopt it for their own agendas.

7 Conclusion

After decades of hard work, qualitative analysts of populism have picked all the low hanging fruits. Moving forward requires sharpening our focus upon measurable dimensions, embracing methodological developments and aligning with wider advancements in the field of social science. Dichotomous perspectives fail to account for much diversity found in empirical reality; populism is not a matter of 'either-or'. A graded reasoning can address more interesting research questions and test a range of hypotheses that have so far remained in the dark.

In this article, after arguing in favor of the conceptual and operational merits of the discursive view of populism, a review of existing quantitative literature exhibited the feasibility of measuring populism in a wide range of textual data sources such as electoral manifestos, speeches, newspaper editorials, and TV commercials. Furthermore, CBSTA was introduced as an improvement over earlier methodological limitations, allowing for the construction of a reliable index of populist discourse while maintaining the ability to accommodate qualitative inferences. As an illustration, CBSTA was applied on non-institutional political phenomena, gauging the level of populist discourse found in the manifestos of the Great Recession movements in order to adjudicate on their disputed populist character. Results uncovered a strongly populist discursive content in the cycle of mobilization of the Great Recession, on a par with paradigmatic cases of populist mobilization and divergent from other types of collective action.

While the use of a semantic grammar in CBSTA strikes a balance between methods measuring purely manifest content through *ad hoc* dictionaries and those performing holistic assessments of latent meanings, it inevitably carries its own limitations. First, the overhead introduced by manually segmenting text into semantic triplets prior to coding. Second, the need to carefully train human coders in order to reach adequate levels of reliability. While the latter point is a standard liability in all manual coding methods, it is a necessary step if we are to overcome the low validity introduced by crude automated methods. The first point could be considered negligible for short or medium-sized texts, yet it can become a concern with large datasets containing longer texts. Here, automated methods retain an advantage. However, large and well-funded research projects can cope with large datasets irrespective of the heavy workload, as the well-known *Manifestos Project* has successfully proven (Budge et al. 2001). The ongoing technological innovation in automated methods of extracting semantic triplets is also bound to lower the bar in the near future.

Scholars of populism now possess an impressive arsenal of techniques to measure their empirical objects. Depending on the research question, the nature and size of available data, the required scoring resolution, or the availability of academic resources, we can complement qualitative analysis by employing dictionaries, performing holistic grading, or coding text with various thematic analysis approaches to derive quantitative inferences. With proper funding, research teams can further increase the validity and reliability of measurement and produce large datasets of populist discourse that cover wide geographical regions and time periods. The availability of such data will undoubtedly encourage comparative work and build bridges towards collaborating with scientists working on fields that directly interact with the multifaceted activity of populist political actors.

Acknowledgments The author would like to thank Nikos Marantzidis and Hanspeter Kriesi for their guidance, the two anonymous reviewers of this journal for their constructive critique, and his team of coders for their patience and perseverance.

References

Albertazzi, D., McDonnell, D.: Introduction: the sceptre and the spectre. In: Albertazzi, D., McDonnell, D. (eds.) Twenty-First Century Populism: The Spectre of Western European Democracy, pp. 1–14. Palgrave Macmillan, Basingstoke (2008)

Armony, A.C., Armony, V.: Indictments, myths, and citizen mobilization in argentina: a discourse analysis. Lat. Am. Polit. Soc. **47**(4), 27–54 (2005)

Aslanidis, P.: Is populism an ideology? a refutation and a new perspective. Polit. Stud. **64**(IS), 88–104 (2016a)

Aslanidis, P.: Populist social movements of the great recession. Mobil. Int. Q. **21** (3), 301–321 (2016b)

Bauer, M.W.: Classical content analysis: a review. In: Bauer, M.W., Gaskell, G. (eds.) Qualitative Researching with Text, Image and Sound, pp. 132–152. Sage, London (2000)

Bell, D.: What is liberalism? Polit. Theory **42**(6), 682–715 (2014)

Bernard, H.R., Ryan, G.W.: Analyzing Qualitative Data: Systematic Approaches. Sage, Thousand Oaks (2010)

Bernhard, L., Kriesi, H., Weber, E.: The populist discourse of the swiss people's party. In: Kriesi, H., Pappas, T.S. (eds.) European Populism in the Shadow of the Great Recession, pp. 123–138. ECPR Press, Colchester (2015)

Bonikowski, B., Gidron, N.: The populist style in American politics: presidential Campaign discourse, 1952–1996. Soc. Forces **94**(4), 1593–1621 (2016)

Bruhn, K.: "To hell with your corrupt institutions!": AMLO and populism in Mexico. In: Mudde, C., Kaltwasser, C.R. (eds.) Populism in Europe and the Americas: Threat or Corrective for Democracy?, pp. 88–112. Cambridge University Press, Cambridge (2012)

Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tanenbaum, E.: Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998. Oxford University Press, Oxford (2001)

Caiani, M., della Porta, D.: The elitist populism of the extreme right: a frame analysis of extreme right-wing discourses in Italy and Germany. Acta Polit. **46**(2), 180–202 (2011)

Calhoun, C.: Occupy wall street in perspective. Br. J. Sociol. **64**(1), 26–38 (2013)

Canovan, M.: Populism. Harcourt Brace Jovanovich, New York (1981)

Canovan, M.: The People. Polity Press, Cambridge (2005)

Caramani, D.: Will vs. Reason: The Populist and Technocratic Forms of Political Representation and Their Critique to Party Government. Am. Polit. Sci. Rev. **111**(1), 54–67 (2017)

Conniff, M.L. (ed.): Populism in Latin America, 2nd edn. The University of Alabama Press, Tuscaloosa (2012)

Deegan-Krause, K., Haughton, T.: Toward a more useful conceptualization of populism: types and degrees of populist appeals in the case of Slovakia. Polit. Policy **37**(4), 821–841 (2009)

de la Torre, C.: Populist Seduction in Latin America: The Ecuadorian Experience. Ohio University Center for International Studies, Athens (2000)

della Porta, D.: Social Movements in Times of Austerity. Polity Press, Cambridge (2015)

della Porta, D., Mattoni, A. (eds.): Spreading Protest: Social Movements in Times of Crisis. ECPR Press, Colchester (2014)
di Tella, T.: Populism into the twenty-first century. Gov. Oppos. **32**, 187–200 (1997)
Dornbusch, R., Edwards, S. (eds.): The Macroeconomics of Populism in Latin America. The University of Chicago Press, Chicago (1991)
Formisano, R.P.: For the People: American Populist Movements from the Revolution to the 1850s. The University of North Carolina Press, Chapel Hill (2008)
Formisano, R.P.: The Tea Party: A Brief History. The Johns Hopkins University Press, Baltimore (2012)
Franzosi, R.: The return of the actor. Interaction Networks among social actors during periods of high mobilization (Italy, 1919–1922). Mobil. Int. Q. **4**(2), 131–149 (1999)
Franzosi, R.: From Words to Numbers: Narrative, Data, and Social Science. Cambridge University Press, Cambridge (2004)
Franzosi, R.: Content Analysis. Sage, Thousand Oaks (2008)
Franzosi, R.: Quantitative Narrative Analysis. Sage, Thousand Oaks (2010)
Franzosi, R., De Fazio, G., Vicari, S.: Ways of measuring agency: an application of quantitative narrative analysis to lynchings in georgia (1875–1930). Sociol. Methodol. **42**(1), 1–42 (2012)
Gallie, W.B.: Essentially contested concepts. Proc. Aristot. Soc. **56**, 167–198 (1956)
Gamson, W.A.: Arab Spring, Israeli summer, and the process of cognitive liberation. Swiss Polit. Sci. Rev. **17**(4), 463–468 (2011)
Gerbaudo, P.: Tweets and the Streets: Social Media and Contemporary Activism. Pluto Press, London (2012)
Gerring, J.: Ideology: a definitional analysis. Polit. Res. Q. **50**(4), 957–994 (1997)
Grimmer, J., Stewart, B.M.: Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Polit. Anal. **21**(3), 267–297 (2013)
Harrison, S., Bruter, M.: Mapping Extreme Right Ideology: An Empirical Geography of the European Extreme Right. Palgrave Macmillan, New York (2011)
Hayes, A.F., Krippendorff, K.: Answering the call for a standard reliability measure for coding data. Commun. Methods Meas. **1**(1), 77–89 (2007)
Hawkins, K.A.: Is Chavez populist?: Measuring populist discourse in comparative perspective. Comp. Polit. Stud. **42**(8), 1040–1067 (2009)
Hawkins, K.A.: Venezuela's Chavismo and Populism in Comparative Perspective. Cambridge University Press, Cambridge (2010)
Iglesias, P.: Understanding Podemos. New Left Rev. **93**(May–June), 7–22 (2015)
Ionescu, G., Gellner, E. (eds.): Populism: Its Meanings and National Characteristics. Weidenfeld and Nicolson, London (1969)
Jagers, J., Walgrave, S.: Populism as political communication style: an empirical study of political parties' discourse in Belgium. Eur. J. Polit. Res. **46**(3), 319–345 (2007)
Jansen, R.S.: Populist mobilization: a new theoretical approach to populism. Sociol. Theory **29**(2), 75–96 (2011)
Johnston, H., Alimi, E.Y.: A methodology analyzing for frame dynamics: the grammar of keying battles in palestinian nationalism. Mobil. Int. Q. **18**(4), 453–474 (2013)
Johnston, H., Oliver, P.E.: Breaking the frame. Mobil. Int. Q. **5**(1), 61–64 (2000)
Kazin, M.: The Populist Persuasion: An American History. Cornell University Press, Ithaca (1995)
Kriesi, H., Pappas, T.S. (eds.): European Populism in the Shadow of the Great Recession. ECPR Press, Colchester (2015)
Krippendorff, K.: Content Analysis: An Introduction to Its Methodology, 2nd edn. Sage, Thousand Oaks (2004)
Kumar, A.: The Making of a Small State: Populist Social Mobilisation and the Hindi Press in the Uttarakhand Movement. Orient Blackswan, New Delhi (2011)
Laclau, E.: Politics and Ideology in Marxist Theory: Capitalism–Fascism–Populism. LNB, London (1977)
Laclau, E.: On Populist Reason. Verso, London (2005)
Mair, P.: Populist democracy vs party democracy. In: Meny, Y., Surel, Y. (eds.) Democracies and the Populist Challenge, pp. 81–98. Palgrave Macmillan, Basingstoke (2002)
March, L.: Towards an understanding of contemporary left-wing populism. In: Political Studies Association (PSA) Annual International Conference, Belfast, 3–5 April 2012
Moffitt, B.: The Global Rise of Populism: Performance, Political Style, and Representation. Stanford University Press, Stanford (2016)
Moffitt, B., Tormey, S.: Rethinking populism: politics, mediatisation and political style. Polit. Stud. **62**(2), 381–397 (2014)
Mudde, C.: The populist zeitgeist. Gov. Oppos. **39**(4), 542–563 (2004)

Mudde, C.: Populist Radical Right Parties in Europe. Cambridge University Press, Cambridge (2007)
Müller, J.-W.: What Is Populism?. University of Pennsylvania Press, Philadelphia (2016)
Munck, G.L., Verkuilen, J.: Conceptualizing and measuring democracy: evaluating alternative indices. Comp. Polit. Stud. **35**(1), 5–34 (2002)
Neuendorf, K.A.: The Content Analysis Guidebook. Sage, Thousand Oaks (2002)
Oliver, J.E., Rahn, W.M.: Rise of the Trumpenvolk: populism in the 2016 election. Ann. Am. Acad. Polit. Soc. Sci. **667**(1), 189–206 (2016)
Panizza, F.: Introduction: populism and the mirror of democracy. In: Panizza, F. (ed.) Populism and the Mirror of Democracy, pp. 1–31. Verso, London (2005)
Pappas, T.S.: The specter haunting Europe: distinguishing liberal democracy's challengers. J. Democr. **27**(4), 22–36 (2016)
Pauwels, T.: Measuring populism: a quantitative text analysis of party literature in Belgium. J. Elect. Public Opin. Parties **21**(1), 97–119 (2011)
Pauwels, T.: Populism in Western Europe: Comparing Belgium, Germany and the Netherlands. Routledge, Oxon (2014)
Pfaff, S.: Exit-Voice Dynamics and the Collapse of East Germany: The Crisis of Leninism and the Revolution of 1989. Duke University Press, Durham (2006)
Philip, G., Panizza, F.: The Triumph of Politics: The Return of the Left in Venezuela, Bolivia and Ecuador. Polity Press, Cambridge (2011)
Popping, R.: Ag09: a computer program for interrater agreement for judgments. Soc. Sci. Comput. Rev. **28**(3), 391–396 (2010)
Popping, R., Roberts, C.W.: Semantic text analysis and the measurement of ideological developments within fledgling democracies. Soc. Sci. Inf. **54**(1), 23–37 (2015)
Reungoat, E.: Anti-EU parties and the people: an analysis of populism in french euromanifestos. Perspect. Eur. Polit. Soc. **11**(3), 292–312 (2010)
Roberts, C.W.: A generic semantic grammar for quantitative text analysis: applications to East and West Berlin radio news content from 1979. Sociol. Methodol. **27**(1), 89–129 (1997)
Roberts, C.W., Zuell, C., Landmann, J., Wang, Y.: Modality analysis: a semantic grammar for imputations of intentionality in texts. Qual. Quant. **44**(2), 239–257 (2010)
Roberts, K.M.: Neoliberalism and the transformation of populism in Latin America: the peruvian case. World Polit. **48**(1), 82–116 (1995)
Rooduijn, M.: The mesmerising message: the diffusion of populism in public debates in Western European media. Polit. Stud. **62**(4), 726–744 (2014)
Rooduijn, M., Pauwels, T.: Measuring populism: comparing two methods of content analysis. West Eur. Polit. **34**(6), 1272–1283 (2011)
Rooduijn, M., de Lange, S.L., van der Brug, W.: A populist Zeitgeist? Programmatic contagion by populist parties in Western Europe. Party Polit. **20**(4), 563–575 (2014)
Ruzza, C., Fella, S.: Populism and the Italian right. Acta Polit. **46**(2), 158–179 (2011)
Shapiro, G., Markoff, J.: Revolutionary Demands: A Content Analysis of the Cahiers de Doléances of 1789. Stanford University Press, Stanford (1998)
Snow, D.A., Benford, R.D.: Ideology, frame resonance, and participant mobilization. Int. Soc. Mov. Res. **1**(1), 197–217 (1988)
Stenner, K.: The Authoritarian Dynamic. Cambridge University Press, Cambridge (2005)
Urbinati, N.: Democracy Disfigured: Opinion, Truth, and the People. Harvard University Press, Cambridge (2014)
Van Dyke, N., Meyer, D.S. (eds.): Understanding the Tea Party Movement. Ashgate, Farnham (2014)
Vasilopoulou, S., Halikiopoulou, D., Exadaktylos, T.: Greece in crisis: austerity, populism and the politics of blame. J. Common Mark. Stud. **52**(2), 388–402 (2014)
Walton, J., Seddon, D.: Free Markets & Food Riots: the Politics of Global Adjustment. Blackwell Publishers, Cambridge (1994)
Weyland, K.: Clarifying a contested concept: populism in the study of Latin American politics. Comp. Polit. **34**(1), 1–22 (2001)