International Studies Review (2020) **22**, 1039–1067

THE FORUM

FORUM: CODING IN TONGUES: DEVELOPING NON-ENGLISH CODING SCHEMES FOR LEADERSHIP PROFILING

Klaus Brummer
Catholic University of Eichstätt-Ingolstadt

Michael D. Young
University at Albany, State University of New York

Özgur Özdamar
Bilkent University

Sercan Canbolat
University of Connecticut

Consuelo Thiers
University of Edinburgh

Christian Rabini
Catholic University of Eichstätt-Ingolstadt

Katharina Dimmroth
Catholic University of Eichstätt-Ingolstadt

Mischa Hansel
Development and Peace Foundation

AND

Ameneh Mehvar
University of St. Andrews

Over the last twenty years since the introduction of automated coding schemes, research in foreign policy analysis (FPA) has made great advances. However, this automatization process is based on the analysis of verbal statements of leaders to create leadership profiles and has remained largely confined in terms of language. That is, the coding schemes can only parse English-language texts. This reduces both the quality and quantity of available data and limits the application of these leadership profiling techniques beyond the Anglosphere. Against this background, this forum offers five reports on the development of freely available coding schemes for either operational code analysis or leadership trait analysis for

Brummer, Klaus et al. (2020) FORUM: CODING IN TONGUES: DEVELOPING NON-ENGLISH CODING SCHEMES FOR LEADERSHIP PROFILING. *International Studies Review*, doi: 10.1093/isr/viaa001

languages other than English (i.e., Turkish, Arabic, Spanish, German, and Persian).

Keywords: automated coding, foreign policy analysis, leadership profiling, leadership trait analysis, operational code analysis

Introduction: Decentering Leadership Profiling

Klaus Brummer

Catholic University of Eichstätt-Ingolstadt

AND

Michael D. Young

University at Albany, State University of New York

Leadership profiling is one of the hallmarks of foreign policy analysis (FPA),[1] and arguably the most prominent and widely used analytical tools for that purpose are leadership trait analysis (LTA) (Hermann 1980, 2005a) and operational code analysis (OCA) (George 1969; Walker, Schafer, and Young 2005). Both LTA and OCA are "at-a-distance" assessment techniques (Schafer 2000), which try to assess the traits or beliefs of leaders based on their verbal statements.[2]

While highlighting traits in its name, LTA covers a broader set of characteristics that also include leaders' cognitive abilities (for details, see table 1). The seven characteristics that are explored by the construct are leaders': belief in the ability to control events; need for power and influence; level of self-confidence; ability for differentiation (called "conceptual complexity"); task orientation; distrust of others; and bias toward their own group ("in-group bias"). Based on certain manifestations of those individual traits, eight distinct leadership styles can be inferred (e.g., "opportunistic," "evangelistic," or "actively independent"). Those traits, as well as already different manifestations of the individual traits as such, are associated with specific behavioral expectations (for details, see Hermann 2005a). To give but one example: leaders with a low conceptual complexity are considered to require stronger stimuli to engage in foreign policy change and are overall less likely to redirect their country's foreign policy compared to their high-complexity peers (Yang 2010; for other recent applications of LTA, see Van Esch and Swinkels 2015; Cuhadar et al. 2017).

OCA examines two sets of decision-makers' political beliefs, with each set containing five distinct elements. Philosophical beliefs address leaders' perceptions and diagnoses of situations, for instance with respect to the conflictual or cooperative nature of political life or their ability to control situations. Instrumental beliefs offer insights into leaders' strategies for goal selection and attainment and the specific means used for that purpose (see table 2 for details).

Leaders' beliefs are conceptualized as causal mechanisms, which "steer the decisions of leaders by shaping leaders' perceptions of reality, acting as mechanisms of cognitive and motivated bias that distort, block, and recast incoming information from the environment" (Walker and Schafer 2006: 5). Hence, whether a leader

[1] For general introductions to the field of FPA, see Breuning (2007) and Hudson (2014).

[2] On the validity of the respective approaches, see Hermann (2005a) and Walker (2003).

Table 1. Traits in leadership trait analysis

BACE	Belief in one's ability to control events	Perception of having control and influence over situations and developments
PWR	Need for power and influence	Aspiration to control, influence, or impact other actors
CC	Conceptual complexity	Ability to perceive nuances in one's political environment, differentiate things and people in one's environment
SC	Self-confidence	Sense of self-importance as well as perceived ability to cope with one's environment
TASK	Task focus/orientation	Focus on problem solving or group maintenance/relationships
DST	General distrust or suspiciousness of others	Tendency to suspect or doubt the motives and deeds of others
IGB	In-group bias	Tendency to value (socially, politically, etc.) defined group and place the group front and center

Source: Own depiction based on Hermann (2005a).

Table 2. Philosophical and instrumental beliefs in an operational code

Philosophical beliefs

P-1 What is the "essential" nature of political life? Is the political universe essentially one of harmony or conflict? What is the fundamental character of one's political opponents?

P-2 What are the prospects for the eventual realization of one's fundamental political values and aspirations? Can one be optimistic, or must one be pessimistic on this score; and in what respects the one and/or the other?

P-3 Is the political future predictable? In what sense and to what extent?

P-4 How much "control" or "mastery" can one have over historical development? What is one's role in "moving" and "shaping" history in the desired direction?

P-5 What is the role of "chance" in human affairs and in historical development?

Instrumental beliefs

I-1 What is the best approach for selecting goals or objectives for political action?

I-2 How are the goals of action pursued most effectively?

I-3 How are the risks of political action calculated, controlled, and accepted?

I-4 What is the best "timing" of action to advance one's interest?

I-5 What is the utility and role of different means for advancing one's interests?

Source: Own depiction based on George (1969).

perceives the environment as extremely hostile or friendly is consequential for the likelihood of her or him engaging in conflictual or cooperative behavior.

Both LTA and OCA were developed using human coding of relatively small volumes of source texts, which limited their use. Over the last twenty years since the introduction of automated coding schemes, FPA research using LTA and OCA has made great advances with an increasing volume of research from seven publications in 1998 to eighty-five in 2018.[3] This growth is due in part to the reduced coding costs of using automated coding schemes for LTA and OCA, which run on Profiler Plus (Levine and Young 2014) and profilerplus.org, along with concomitant increases in the reliability and comparability of data.[4] Indeed, over the last fifteen years or so, most of the LTA and OCA profiles have been generated through the automated coding of leaders' utterances, rather than hand coding.

However, this automation process is based on the analysis of verbal statements of leaders to create leadership profiles and has remained largely confined to

[3] Figures compiled by Young from scholar.google.com.

[4] The coding schemes are available from Social Science Automation (https://socialscience.net/).

English-language texts. This limits the scope of FPA research because most people do not speak English as their first language, if at all. One estimate suggests that only 378 million of approximately 7.5 billion humans speak English as their first language (Ethnologue 2019); that is, roughly five percent. English is also not the most widely spoken first language; it is in third place, well behind Chinese, although comparable to Spanish, Hindi, and Arabic. The problem for LTA, OCA, and, more broadly, FPA, is that many texts are not available in English, and neither machine translation, such as Google Translate, nor human translation provides an acceptable solution, due to issues that make machine translations problematic and to the cost of high quality human translation.

Both the LTA and OCA coding schemes running on Profiler Plus are fully automated like other systems such as DICTION (Hart 1985) and LIWC (Pennebaker, Francis, and Booth 2001). However, unlike those systems and more like PETRARCH2 and its predecessors (Norris 2016), they rely on part-of-speech information to correctly identify sequences of terms within, and sometimes across, sentences. This reliance on part of speech and potentially disparate terms renders them more sensitive to translation errors that may remain intelligible to a human. Each of our projects has encountered this problem in various forms. For example, many Turkish verbs become nouns or adverbs when translated by Google, as can be seen in President Erdogan's speech at the UN on September 24, 2019:

> **Turkish**: Üçüncü önemli konu, Suriye'nin dörtte birini işgal eden ve sözde Suriye Demokratik Güçleri adıyla meşrulaştırılmaya çalışılan Fırat'ın doğusundaki PKK-YPG terör yapılandırılmasının ortadan kaldırılmasıdır.
> **Google**: The third important issue is the elimination of the PKK-YPG terror restructuring in the east of the Euphrates, which occupies one quarter of Syria and is being legitimized under the name of the so-called Syrian Democratic Forces.
> **Human translation**: Thirdly, we have to destroy the PKK-YPG terrorist network, which occupies the one-quarter of Syria under the so-called umbrella term, Syrian Democratic Forces, to rebrand the terrorists as legitimate freedom fighters.

The difference in translations of kaldırılmasıdır results in a + 1 rather than the correct −3. An issue with Arabic-English translation is that Google translate gives more than one corresponding word for many Arabic words, including transitive verbs. For example, "تنحني لن" should be translated as "we will not bow to them (Israel and Egyptian military)," but Google translates it as "we will not lay squat kneel for them."

Other types of errors also occur. In the German and Persian examples below, meanings become quite mangled:

> **German**: Auch technische Details in Zusammenhang mit der Aufnahme der diplomatischen Beziehungen, zu denen der Umfang der beiderseitigen Botschaften gehört, bedürfen noch der Absprache.
> **Google translate**: Also technical details in connection with the establishment of diplomatic relations, to which the scope of the mutual messages belongs, still need to be agreed.
> **Human translation**: Technical details in relation to the establishment of diplomatic relations, such as the size of the Embassies on both sides, still need to be agreed upon.
> **Persian**: نظام سیاسی قبل از انقلاب که سر تا پایش ظلم بود.

Google translate: The political system before the revolution, which was cruel to the feet.
Human translation: The political system before the revolution was entirely cruel.

These problems have a quantifiable effect on the observations obtained from the texts (illustrated in table 3) and change the I1 or P1 score for three of four Spanish language speeches examined[5] by more than a standard deviation.

Beyond pure issues of translation, creating LTA and OCA coding schemes in other languages also increases the accessibility of the two approaches. There are many scholars who do not speak or write English but who are perfectly comfortable in other languages.

The current necessity of using English language texts reduces both the *quality* and *quantity* of available source texts and limits the application of these leadership profiling techniques beyond the Anglosphere. Regarding source text quality, leaders whose first language is not English are more at ease with their native language than with English. Hence, the better command of their native language compared to English means that leaders' utterances in their native tongue should lead to more nuanced, and thus also more accurate, expressions of "who they are" compared to statements in their second, or even third, language of English. Since the construction of LTA or OCA leadership profiles relies exclusively on verbal statements, the question of whether a leader's profile is built on utterances made in his or her first, second, or third language is highly relevant.

In addition to quality, quantity of source texts is also problematic. LTA and OCA both suggest that fifteen thousand words or more of source text is used for the construction of a profile (table 4). However, many non-English speaking world leaders do not make frequent statements in English, making it hard, and at times outright impossible, to compile a sufficient quantity of source texts that fulfill setting or spontaneity requirements. However, leaders "always" talk in their native language, and that text is available in much greater quantity. Overall, original speech acts, which are more plentiful and more nuanced, are preferred over translations. This alone increases the number of cases for which researchers can locate sufficient source texts.

The added value of non-English coding schemes is fourfold:

1. Non-English coding schemes significantly increase the volume of source text on which leadership profiles can be constructed.
2. They are likely to bring about more accurate profiles since they are based on leaders' utterances in their native tongue.
3. They help answer novel empirical questions as well as revisiting and maybe challenging established insights using a more rigorous methodology.
4. They broaden the scope of leadership profiling beyond the five percent of the global population in the Anglo-American core and, by extension, contribute to the decentering of FPA more generally.

The following five contributions to this forum report on the development of non-English coding schemes for either OCA or LTA. Those are: OCA coding schemes for Turkish (Özdamar, Canbolat, and Young), Arabic (Canbolat), and Spanish (Thiers), as well as a full LTA coding scheme for German (Rabini et al.) and an LTA conceptual complexity coding scheme for Persian (Mehvar). All schemes are

[5] Calculated using the 2007 dataset courtesy of Social Science Automation.

Table 3. OCA Observations obtained from high-quality human translation and Google Translate for four Spanish-language speeches

	self punish	self threaten	self oppose	self appeal	self promise	self reward	other punish	other threaten	other oppose	other appeal	other promise	other reward
1. Bachelet Google	1	0	2	10	0	0	5	2	5	25	3	17
1. Bachelet UN	2	0	1	9	0	2	5	0	5	25	4	13
Difference in observations	1	0	−1	−1	0	2	0	−2	0	0	1	−4
2. Bachelet Google	0	0	2	13	1	2	7	3	6	27	3	17
2. Bachelet UN	1	0	0	8	1	6	9	1	5	24	1	14
Difference in observations	1	0	−2	−5	0	4	2	−2	−1	−3	−2	−3
3. Bachelet Google	0	1	5	6	0	3	6	1	5	14	0	10
3. Bachelet UN	1	1	1	10	1	4	3	0	3	16	1	12
Difference in observations	1	0	−4	4	1	1	−3	−1	−2	2	1	2
4. Bachelet Google	6	0	3	16	1	5	11	0	9	25	2	21
4. Bachelet UN	4	0	3	18	1	8	8	2	7	21	5	22
Difference in observations	−2	0	0	2	0	3	−3	2	−2	−4	3	1

Source: Compiled by Consuelo Thiers.

Table 4. Source text requirements for LTA and OCA

LTA	OCA
Spontaneous speech acts only	Any verbal expression, including a complete speech, a press conference, or an interview.
100 speech acts (50 at the very least)	10+ speeches
150 words per speech act (100 at the very least)	1500+ words per speech act or 15–20 coded verbs per speech act
Delivered in different contexts/in front of different audiences	Agnostic to audience/match to research question
Covering different issue areas (foreign and domestic; though more targeted samples are not ruled out)	Targeted toward policy issue under examination
Spanning a leader's tenure (though more specific time periods are not ruled out)	Targeted in case time/timing is of relevance for research question (e.g., before/after a certain event)

Source: Own depiction based on Hermann (2005a, 2008) and Schafer and Walker (2006a).

fully operational and available online.[6] Hence, both other scholars and practitioners can use them—for instance, to systematically explore the impacts, respectively, of leadership traits or political beliefs on foreign policy processes (e.g., instances of low-quality decision making) and outcomes (e.g., foreign policy change), or to extrapolate how leaders are likely to respond to positive or negative incentives, as scholars have done with the English-language coding schemes.[7]

After a brief introduction, each contribution first discusses differences between the respective languages and English that render a mere translation of the original English coding schemes futile. Then, they discuss challenges for automated coding with respect to data quality, availability, and pertinence. Next, the contributions turn to the empirical domain by highlighting, for instance, which new research question can be addressed based on the new coding schemes or which established insights can now be systematically tested or challenged for the first time. All contributions end with suggestions for further development.

Profiling Leaders in Turkish

ÖZGUR ÖZDAMAR

Bilkent University

SERCAN CANBOLAT

University of Connecticut

AND

MICHAEL D. YOUNG

University at Albany, State University of New York

Leaders have been influential in Turkish politics and foreign policy since the early years of the Republic, ranging from Kemal Atatürk to Recep Tayyip Erdoğan. Heper and Sayari (2002, 7) argue that Turkish politics has always been "a stage for

[6] The coding schemes for Turkish, Arabic, Spanish, and German are available on profilerplus.org, and the one for Persian is on GitHub (https://github.com/DavidSymonz/PersianAnalyser.git).

[7] See http://socialscience.net/partners/research.aspx.

leader-based politics," as the Islamic creed extolling the role of a strong and charismatic leader in maintaining order enables personalities to shape domestic politics and foreign policy. Individual leaders, prime ministers, and now presidents have enjoyed both legal powers defined by the Turkish constitution and informal powers derived from their personality and charisma (Kesgin 2019a). Turkish media also plays a key role in personalizing politics as it zeroes in on certain individuals and their leadership characteristics in media coverage and framing of Turkish politics (Demir 2007; Kesgin, 2013, 2019a). Because individuals and their leadership style matter in Turkish politics, a nuanced and scientific explanation of Turkey's foreign and domestic policy-making requires a systematic approach to leadership analysis. We contribute to efforts in broadening the scope of FPA by creating a Turkish coding scheme for operational code analysis (OCA).

For the "translation" of the English OCA coding scheme to Turkish, there are two main linguistic challenges stemming from historical and cultural features of the language that might render simple translation of the original coding scheme unavailing. These hurdles are: (1) the agglutinative character of the Turkish language and phrasal verbs and (2) cultural and religious symbols embedded in the language.

First, Turkish is an agglutinative language, which is associated with phrasal (or compound) verbs containing multiple adjectives and nouns, and the agglutination complicates simple translation efforts. For example, while "darbetmek" (hit) and "affetmek" (forgive) are two Turkish verbs that have the same core verbs (i.e., suffix), "etmek" (make/render), their coded values are diametrically opposed (+3 and −3, respectively). Because the nonseparable suffix, as a form core verb, takes different coded values depending on the preceding nouns such as "darp" (a single act of hitting) or "af" (a single act of forgiving), coding rules must match the nonseparable core verb with the noun prefix.

Second, verbatim translation of Turkish text into English or all-out translation of the coding rules will fall short of detecting and accounting for context-specific words and hidden messages in many Turkish texts. For example, when talking about the Syrian civil war back in 2012, Erdoğan said: "Şu anda Suriye'de olanlar Kerbela olayı ile tamamen aynıdır." (What is happening in Syria now is exactly the same as Karbala).[8] One of the most powerful symbols in the Islamic religion, Karbala refers to a war fought between the righteous good (Turkey and anti-Assad rebels) and the ferocious evil (the Syrian regime and its allies).

Additionally, many words have different connotations in Turkish than the literal English translation. For example, the words "mücahit" (mujahedeen or jihadist) and "şehit" (martyr) have rather positive connotations in Turkish, in stark contrast to their negative implications in English. Conversely, certain words take a very negative meaning in Turkish when they are used in a political context such as "firavun" (pharaoh) (Özdamar 2017, 21).

Although the ability to use Turkish language sources presents greater opportunities for empirical data and hitherto hidden variables, data source persistence continues to be a challenge. There is a dearth of online data for certain Turkish political leaders who left active political life following a termination of their term in office or a disagreement with the top executive leadership resulting in removal of source texts. For example, systematic and reliable online data is not available for former prime minister Ahmet Davutoğlu or former president Abdullah Gül, who were once significant executive actors in Turkish politics and foreign policy but became somewhat dissident politicians with rather aloof relationships with the current president, Tayyip Erdoğan, and his AKP party.

[8] *Hürriyet Daily News*, "PM Erdoğan Likens Syrian Crisis to Karbala Massacre," September 8, 2012, http://www.hurriyetdailynews.com/pm-erdogan-likens-syrian-crisis-to-karbala-massacre-29671 (last accessed: February 2, 2020).

The creation of a Turkish operational code coding scheme has the potential to open new horizons in empirical and theoretical terms. There are at least three benefits of a Turkish coding scheme: (1) flagging context-specific indicators and symbols to improve data quality, (2) introducing and studying new political actors who have been put aside because of a lack of data, and (3) addressing novel research questions and (re)visiting critical Turkish foreign policy decisions and processes considering newfound data.

First, a Turkish coding scheme helps researchers find and highlight context-specific symbols and indicators, which are drawn from a corpus of Turkish texts. As the size of the Turkish text corpus expands, possible derivatives and synonyms of the indicators and symbols could be added to the scheme's dictionary. This would allow scholars to refine both coding rules and scored verb dictionaries and to further improve data quality. For instance, weeding out manipulative and domestic politics-oriented speeches of Turkish leaders from the corpus of Turkish foreign policy speeches by the help of context-specific indicators might be a stride for distinguishing the "signal from the noise" in leadership analysis data (Wohlstetter 1962, 56).

Second, a Turkish coding scheme extends OCA to leaders that could not be profiled because of a dearth of empirical data. For our project, we were able to collect the minimum of twenty eligible speeches for all Turkish prime ministers and presidents of the Republic of Turkey between 1946 and 2018.[9] However, comparable English source material is unlikely to match the size of the Turkish text corpora (see Kesgin 2013; Özdamar 2017) because the translation of leaders' statements from Turkish into English is a rather recent phenomenon, dating to the late 1990s, when Turkey became a candidate for the European Union membership.

Third, a Turkish coding scheme can help scholars formulate novel research questions or revisit some of the established research programs and insights in the context of Turkish foreign policy. For example, the current literature on domestic Turkish politics shows that entrenched populism and audience effect have an impact on the style and content of Turkish leaders' statements. Soner Cagaptay (2017, 180–81) asserts that Erdoğan could be the "inventor of 21st century populism whose speeches since he assumed the presidency, particularly after an attempted coup in 2016, have been the most consistently populist of his career." Dominating the political arena thanks to his presidential powers and partisan media support, Erdoğan has shaped political rhetoric in Turkey since the early 2010s, and foreign policy–making is no exception. Thus, a great portion of his foreign policy speeches are parochial in character, targeting opposition parties or electoral processes (Bayulgen et al. 2018). Working with the Turkish text corpus will help future scholars mitigate, if not eliminate, the corroding effects of rampant populism at the elite level.

Piggybacking on populism, the audience effect has become ever more pronounced in polarized polities, and Turkey is one of the primary cases of this phenomenon (Kesgin 2019a). Audience effect also manifests itself in Turkish leaders' foreign policy speeches, and the effect is more evident when leaders deliver speeches in their native Turkish language. For example, toward his own constituency, President Erdoğan uses inordinately humble language populated by self-effacing utterances such as "Bu fakir hiçbir zaman Sultan olma gayretinde olmadı." (This destitute person (I) never tried to become a Sultan).[10] Erdoğan's foreign policy speeches in Turkish have been the most vitriolic and belligerent during critical electoral cycles such as the 2015 and 2018 general elections and the 2017 constitutional referendum. In these periods, Erdoğan clung to hawkish

[9] Our data excludes acting and technocratic Turkish leaders whose term in office was shorter than two years. We decided to start our dataset from 1946, when Turkey transitioned into a multiparty system.

[10] English translation our own. For Turkish-language coverage of this speech, see: Hürriyet Daily News "President Erdoğan Spoke in Sincik, Adıyaman," available at http://www.hurriyet.com.tr/gundem/erdogan-sincikte-konustu-29067331 (last accessed: February 1, 2020).

Table 5. Davutoğlu and Erdoğan's master belief scores in English (E) and Turkish (T) materials compared to norming groups on state leaders*

	World leader average	Rogue leader average	Davutoğlu (E)	Davutoğlu (T)	Erdoğan (E)	Erdoğan (T)
Nature of political universe	0.30	0.15	0.23	0.30	0.32	0.21
Strategic direction	0.40	0.25	0.31	0.41	0.41	0.29
Perception of control	0.22	0.18	0.20	0.36	0.19	0.23
Speech N	164	52	25	25	30	30
Years			2014–2016	2014–2016	2014–2109	2014–2109

Source: Own depiction.
*The norming sample scores on rogue and average world leaders are courtesy of Stephen Benedict Dyson and Akan Malici.

foreign policy themes in his campaign speeches and threatened Syria with military interventions, which are more pronounced in his domestic speeches in Turkish targeting Western countries and Israel.[11]

A comparison between two Turkish Islamist leaders, Davutoğlu and Erdoğan, (table 5) shows that while Erdoğan employs harsher and more hawkish foreign policy rhetoric toward domestic audiences, he switches to a much softer tone when he addresses foreign audiences about the same topic. In contrast, Davutoğlu's speeches in Turkish are more modest and peaceable, while those in English have a more conflictual tone (unlike Erdoğan, Davutoğlu has a command of English and chose to speak in English when he was addressing foreign audiences). Further research can focus on such potentially statistically significant differences between English and Turkish text corpora and help disentangle the relationship between populism, audience effect, and foreign policy decision-making.

The Turkish OCA coding scheme may be the most accessible tool to address significant research questions including:

1. How do Turkish leaders' political beliefs affect their politics?
2. How do beliefs of Turkey's secular leaders differ from those of political Islamists?
3. How do beliefs of certain Turkish leaders influence their critical foreign policy decisions such as the Cyprus issue, the second Gulf war, Syrian civil war, and the Kurdish issue?

All the advantages of a Turkish operational code construct notwithstanding, there are at least three avenues for further research. First, over time, as we convert source documents from pdf to txt formats, we will create a norming group for Turkish-speaking leaders to establish a basis for comparison with future Turkish political leaders and with Turkic national leaders in different parts of Eurasia, including the Caucasus and Central Asia. This effort will help researchers draw a general profile, or a lack thereof, of Turkish decision-makers. In-group and cross-regional comparisons may also lead to new insights into the psychological idiosyncrasies of leaders by revealing a new set of intervening variables, such as the studied language, culture, history, and religion.

A second avenue for research is the development of a Turkish leadership trait analysis (LTA) coding scheme as a complementary tool for the Turkish OCA coding scheme. One of the limitations of the latter is that the action words and deeds are not always embedded in Turkish verbs because certain nouns and adverbs could

[11] Recep Tayyip Erdoğan, "Trump Is Right on Syria. Turkey Can Get the Job Done." New York Times, January 7, 2019, https://www.nytimes.com/2019/01/07/opinion/erdogan-turkey-syria.html (last accessed: February 1, 2020).

serve as main utterances of action and positive-negative sentiments. Erdoğan's following words back in 2012 as the Turkish premier are illustrative: "İnşallah biz en kısa zamanda Şam'a gidecek, Emevi Camisi'nde namaz kılacağız." (God willing, we will go to Damascus very soon, and will pray in the Umayyad mosque).[12] Our Turkish coding scheme scored this sentence as zero because the verb "kılmak" (perform) is a neutral transitive verb. Yet, a nuanced interpretation of Erdoğans sentence reveals Turkey's threat to soon intervene in Syria militarily. Idiosyncrasies of the Turkish language could require an analysis of certain nouns and adverbs, and the creation of a Turkish LTA might be another promising pathway for further research.

Last, future researchers could expand on the Turkish operational code construct to include distinct dialects of Turkish spoken in different parts of Turkey or in different Turkic nations and ethnic groups in the broader region. This could enable researchers to profile leaders from other Turkish language centers, such as Azerbaijan and Turkic regions of the Middle East. Such research could begin by beefing up the Turkish language dictionary and adjusting actor and self-reference tables.

Profiling Leaders in Arabic*

SERCAN CANBOLAT

University of Connecticut

Individual leaders play a crucial role in MENA politics and foreign policy. Particularly, Arab MENA is viewed as one of the few regions in the world characterized by predominant and charismatic political leaders. Although Hinnebusch (2015) argues that this is partially a consequence of the prevailing regime types, including hereditary monarchies and presidencies-for-life, and feeble democratic institutions in MENA, Dekmejian (1975) makes a case for the emergence of high-profile leadership due to rampant political turbulence and a cascade of wars and revolutions producing predominant and charismatic leaders in the region. From the early nation-builders like Gamal Abdel Nasser and Abdullah bin Hussein to late rogue leaders such as Saddam Hussein and Muammar Qaddafi, and from modern secular leaders to Islamist insurgents such as Osama bin Laden and Abu Bakr al-Baghdadi, MENA politics is associated with high-profile political personalities.

However, there has been a dearth of systematic approaches to MENA leadership analysis. Hinnebusch (2015, 84) stresses the dominance of a rational choice model (RCM) in FPA-style actor-specific studies at the expense of psychological approaches to leadership analysis in the region. The prevalence of RCM approach and limitations in accessing and coding Arabic speech material reduced leadership assessment to historical anecdotes, and data-based leadership profiling remained rather atypical of FPA scholarship in MENA (Malici and Buckner 2008; Duelfer and Dyson 2011; Özdamar and Canbolat 2018; Kesgin 2019b). By providing an Arabic coding scheme for operational code analysis (OCA) (Walker et al. 1998; Schafer and Walker 2006b), this project aims to open new horizons in the study of political leaderships in MENA.

Translation efforts from Arabic into English should factor in four main cultural and linguistic characteristics of the Arabic language to ascertain nonskewed results: internal consonants, compound words, omission of words, and nondeclarative word order. First, while English usually has consistent stem content forms called "lemmas" with endings attached for grammatical differences, Arabic words and verbs

[12] Translation our own. For English-language coverage of this speech, see: Hürriyet Daily News, "Premier vows to pray in Damascus mosque soon," available at http://www.hurriyetdailynews.com/premier-vows-to-pray-in-damascus-mosque-soon-29505 (last accessed: February 1, 2020).

* I wish to acknowledge the helpful comments of Doug Fuller, Daniel Weiner, Zaid Eyadat, Stephen G. Walker, Stephen Benedict Dyson, and Akan Malici and the journal's anonymous reviewers.

are predicated on three internal consonants (ك, ت, ب), which is also called "Semitic root." Vowel pattern distinctions constitute the heart of Arabic grammar, and internal consonants function as the determinative root of the word, producing most of the Arabic verbs. Moreover, there are many "irregular" words that do not follow the internal consonant rule, for instance "هجوم" (to attack). Such differences require nuanced content-marking rules for the Arabic coding scheme, in order to code values for each transitive verb correctly.

Second, compound words, an amalgamation of two or more nouns, are very common in the Arabic language. While grammatical function words—prepositions, pronouns, auxiliary verbs, etc.—are mostly separated from the verbs in English, some of them are always attached to either the beginnings or ends of Arabic verbs, such as ب (bi-), ل(li-), و(wa-), and كم(-kum). Compound verbs are also commonplace in Arabic language, and Arab leaders use them heavily in their speeches. I adjusted the coding scheme by adding more rules to find the core verb and match it with the correct value in the dictionary.

Next, some Arab decision-makers, and particularly the leaders of nonstate actors (NSA), shy away from using certain words, which have defeatist connotations in the context of history of MENA and the Arab-Israeli conflict. My text corpus on the Arabic-speaking leaders of violent nonstate actors (VNSA), which contains around eighty-five thousand transitive verbs, does not have the word "إسرائيل" (Israel) as a name/subject for other. Strikingly, furthermore, certain words such as *leader*, *leadership*, *military*, *treaty*, and *international organization* are almost nonexistent in my Arabic speech material. The root cause of the omission of such words is that the Arab VNSA leaders do not perceive the international and political environment in the same way as official state leaders, including Arab national leaders. For example, Islamist militant groups use the words *caliph* and *sharia*, respectively, instead of *leader(ship)* and *treaty*, due to their sheer religious connotations for their audience. The following part focuses on main data limitations associated with the Arabic source material.

The availability and accessibility of empirical data is always a prerequisite for any type of FPA-style research, and the data challenge is even more pronounced in non-North American contexts such as the Arab world and MENA. As noted by Hinnebusch (2015, 176), data challenges are particularly endemic in nondemocracies, as data availability and accessibility is more constrained in such regime types. Although Arabic speech material provides more data points for leadership analysis than those of English sources, future practitioners of an Arabic coding scheme should be aware of the following potential impediments concerning data quality: the availability of relevant data and database permanence.

The first challenge in creating an Arabic text corpus for VNSA is finding adequate and pertinent speech data to establish foreign policy profiles for the terrorist leaders. Particularly, ISIS leader Al-Baghdadi's and Nusra leader Al-Jolani's public speeches on politics and foreign policy have been elusive, and so locating and retrieving them was the most time-consuming part of the research.

Second, the platforms or databases providing Arabic source material for Islamist militant leaders are not permanent. Due to broad censorships on terrorist propaganda and communication material, the VNSA leaders' speeches are only temporarily available for the researchers (Jacoby 2019). For example, while I first had a cursory look on al-Baghdadi's political statements at online platforms such as *Dabiq* and *al-Amaq*, via Twitter I located seventeen speeches from 2014 to 2019, but my final attempt produced a collection of only nine speeches in total because of the censorship.[13] Additionally, unlike Arab or Western governments, VNSA lack resources and know-how to build and maintain their online platforms to make their institutional and propaganda material accessible to the public. I had this

[13] The SITE Intelligence Group Enterprise attempts to provide researchers with the translated speeches (from Arabic to English) of jihadist leaders: https://ent.siteintelgroup.com/Jihadist-News/Statements/ (last accessed: April 18, 2019).

Table 6. Badie and Mashal's master belief scores in English (E) and Arabic (A) materials compared to norming groups on state leaders*

	World leader average	Rogue leader average	Badie (E)	Badie (A)	Mashal (E)	Mashal (A)
Nature of political universe	0.30	0.15	0.18	−0.06	0.11	−0.25
Strategic direction	0.40	0.25	0.21	0.03	0.22	−0.33
Perception of control	0.22	0.18	0.19	0.29	0.16	0.39
Speech N	164	52	10	10	15	15
Years			2010–2013	2010–2013	2011–2016	2011–2016

Source: Own depiction.
*The norming sample scores on rogue and average world leaders are courtesy of Stephen Benedict Dyson and Akan Malici.

problem throughout the creation of Arabic text corpus for VNSA in MENA, except for Egypt's Muslim Brotherhood, whose speeches are always available on their official website.[14] However, as shown in table 6 below, speech unavailability and media censorship hinder the collection of English text corpora more than Arabic corpora.

The Arabic operational code coding scheme provides prospective researchers with three advantages concerning theoretical and empirical growth of FPA beyond the North-American context: (1) possibility of testing established insights including audience effect and propaganda rhetoric; (2) studying more Arab leaders whose English source material is limited; (3) extending OCA to study influential NSA and VNSA in MENA.

First, how do certain MENA militant leaders succeed in confounding scholarly and governmental expectations in the Western capitals regarding their political personality and leadership style? Such enigmatic leaders include Hamas' Mashal and Iraq's al-Sadr (Lazarevska et al. 2006). Constructing an Arabic scheme to analyze MENA's main militant leaders is a stepping stone for addressing this puzzle. One possible explanation could be leaders' use of Arabic to cajole domestic constituents and potential recruits and their utilization of English in making episodic charm offensives to the West (Özdamar and Canbolat 2018). To test such an explanation, I focus on their master political beliefs in two different sets of language materials covering the same time periods and foreign policy themes: (1) leaders' original Arabic speeches and (2) the translated transcripts of the former into English. Treating the audience (domestic or international) as a control variable, I measured the key political beliefs of Hamas' Mashal[15] and the Muslim Brotherhood's Badie in the post–Arab Spring era.

Table 6 below suggests that audience effect factors in Islamist leaders' strategies, whose dichotomous rhetoric lends support to the "two-level game" logic of diplomacy and domestic politics (Putnam 1988). In addition, Islamist leaders of VNSA in MENA exhibit notable differences in their political beliefs depending on the preferred language of their speeches and target audience. Lastly, the leaderships of Islamist VNSA may be similar to leaders of rogue states.

The Arabic coding scheme makes Arabic data sources and text corpus more useful increasing the volume of useable material beyond the English source material. For example, Özdamar and Canbolat's (2018, 23) research on Muslim Brotherhood leadership could only utilize twenty-six speeches and find 2,134 transitive verbs in total from which to draw foreign policy profiles of three Brotherhood leaders because the authors were restricted to solely using English source material for Arab leaders. Conversely, my preliminary research on two Brotherhood leaders, Mashal and

[14] Mohamed Badie's speeches can be accessed at the Muslim Brotherhood's official website: http://www.ikhwanonline.com/ (last accessed: April 18, 2019).

[15] Mashal's speeches in English are available at Al Jazeera (https://www.aljazeera.com/), and his Arabic speeches are available at Al Hadath, a local media platform in Gaza Strip (alhadath.ps).

Badie, uses twenty-five Arabic texts and analyzes 11,546 transitive verbs due to the extensiveness and greater accessibility of the Arabic source material (see table 6). Moreover, Hinnebusch (2015) suggests that translations of Arab leaders' speeches into other languages including English have only recently become common practice. An Arabic scheme helps scholars extend their empirical leadership analysis to the understudied political personalities in MENA.

The Arabic OCA coding scheme also broadens FPA's individual-level lenses beyond executive state leaders; NSA and particularly VNSA are significant political actors shaping the modern politics and foreign policy of the MENA region (Dalacoura 2001). VNSA leaders from organizations such as Hezbollah and the Islamic State control territory and exert sizeable influence in the region's politics.

Although the Arabic OCA coding scheme provides significant advantages, additional research is needed on three fronts: (1) factoring different dialects of Arabic in the analysis, (2) creating a regional norming group for Arab leaders, and (3) dealing with nonverbal communication and purely religious references.

First, Arabic has been a language of many cultures and religions and has been used in a gigantic swathe of territory in the MENA region, and therefore, there are several different dialects and language centers of Arabic. For example, it is almost impossible to view an Arabic dialect spoken in Morocco as formal standard Arabic (FSA) or Levantine Arabic (Shami). Future research might focus on the subgroups of the Arabic language and beef up Profiler Plus dictionary and rule tables to include political leaders from other Arabic language centers beyond the current FSA format.

Second, the creation of a norming group for Arabic-speaking MENA leaders will provide a basis for statistical comparison with both future and former Arab national and militant leaders. Future scholars can emulate the scholars of leadership trait analysis (Hermann 2005b), who created regional norming groups including the one for MENA leaders. The development of an Arabic operational code norming group could produce both state-level and regional-level comparisons and reveal insights into leadership styles and the effects of history and culture on leaders' psychologies.

Finally, an updated version of the Arabic coding scheme could find some solutions or proxies to decipher and quantify nonverbal communication, which is rather common among Islamist militant groups in MENA. Acts of nonverbal communication include propaganda videos containing religious songs, fatwas, and horrendous acts of violence. Additionally, Islamist VNSA heavily use Quranic texts and references in their speeches, rendering their statements exclusive and arcane, particularly for nonreligious researchers. Many Quranic references are politically charged symbols, and some function as dog-whistle politics considering the target audience. Native speakers play a significant role in determining and quantifying culture and religion-specific linguistic symbols by their content judgements. Future developments of the Arabic coding scheme could address this significant research avenue, which should help us to better comprehend the psyche of militant Islamist leaders and VNSA.

Profiling Leaders in Spanish

Consuelo Thiers

University of Edinburgh

In the Latin American context, the figure of the president has a gravitating role in policy decision-making processes. This stems mostly from the distinctive features of the presidential regime, characterized by the high concentration of power in the executive. Thus, the role of political leaders in foreign policy decision-making in

Latin American, particularly the role of presidents, has been the focus of much research. Scholars have addressed the decisive role of political leaders in contexts such as regional organizations (e.g., Chodor and McCarthy-Jones 2013; Jenne, Schenoni, and Urdinez 2017) and in the management of bilateral conflicts and agenda-setting (e.g., Cason and Power 2009; Wehner 2011). However, there are no in-depth analyses of leaders' specific beliefs that shed light on decision-making processes in foreign policy. Likewise, there are no studies focused on the systematic assessment of leaders' characteristics that allow for comparisons and help understand foreign policy decisions.

For scholars interested in Latin American politics, operational code analysis (OCA) provides a toolkit to study the beliefs of political leaders that can shed light on decision-making processes in foreign policy issues. The Spanish coding scheme for OCA will give researchers access to a specialized method and enable the development of more consistent and comparable research on political leaders in the region.

The process for developing the Spanish OCA coding scheme required two simultaneous tasks: the improvement of the Spanish Token Tagger (STT) that provides part-of-speech and lemma information[16] and the adaptation of the operational code rules to the Spanish language.

The improvement of the STT mostly entailed the addition of all verbs contained within the OCA coding scheme and their testing in all tenses and persons of the Spanish indicative mode. As a result, approximately seven hundred verbs were added and classified according to their conjugation pattern, some missing rules were added, and some others were modified to fit into the OCA requirements.

I based the development of the OCA rules on the English version of the coding scheme, using the tables of the original version as a starting point to guide the creation of the Spanish version. The content of the tables was modified for Spanish grammar and syntax by adding nouns and adjectives expressing conflict and cooperation, as well as transitive verbs and some new actors. The *Real Academia Española* (*RAE*) online dictionary provided the source for information on the transitivity of the verbs, types of words, and definitions.

The development of the Spanish scheme presented some difficulties related to both the context and structure of the sentences. The OCA coding scheme provides scores based on leaders' use of transitive verbs and their associated parts of speech. The scores take into account whether the action was performed by self or others. In the English language, the recognition of verbs and the self-other context does not present great difficulties; English verbs are generally accompanied by an easily identifiable pronoun. In the Spanish language, the use of the subject pronoun is not required and is usually omitted. Generally, it is the form of the verb that indicates who is performing the action. For instance in: *Firmé un tratado de paz* (I signed a peace treaty), the conjugation of the verb *firmar* (to sign) clearly indicates that the first person singular is the one performing the action, so there is no grammatical need for the subject pronoun (in Spanish "yo").

Therefore, developing the Spanish version of the OCA coding scheme required the creation of two main rules to assign a subject pronoun in the cases where it was omitted:

1. The first rule recognizes verbs that lack a subject and then adds a truth-value (subject), class (actor), and other-self, depending on the conjugation of the verb.
2. The second rule inserts a new token before the verb and copies its values, adding the corresponding tense and person.

[16] The STT was created by Doug Fuller in 2012 (Social Science Automation, Inc.).

Another characteristic of Spanish language texts is to use more words and longer sentences to express similar ideas than corresponding English language texts. This characteristic makes it more complicated for the software to recognize subjects and objects. To navigate this issue, I implemented three measures. The first measure consisted of the addition of several phrases and words to the first table of the scheme with the aim of reducing the noise. This table either shortened some sentences or deleted unnecessary words, reducing text length and complexity. The second measure addressed those sentences that contained several verbs and commas that obstructed the clarification of their object and subject and therefore were altering the final scores. In order to exclude those verbs whose subject was not omitted, the rule only covers verbs that did not have a subject word up to two tokens before them. This layout creates a problem for those verbs that do not have a subject and that are placed close to a verb that has one. For instance: *Yo capturé, encarcelé y luego liberé a los soldados* (I captured, imprisoned, and then freed the soldiers). Since this type of sentence formation usually contains commas, the issue was partially solved by adding a rule that is meant to recognize verbs without subjects following a comma. Although this solution proved to be quite useful, the problem remains in sentences where there is no punctuation. The third measure tackled the problem of lengthy sentences containing verbs used in the infinitive form. In this case, the subject assignment rule cannot provide a subject to the verb, as it does not have an identifiable person. Another rule assigns the person of any conjugated verb positioned within ten tokens before the infinitive form verb.

The following excerpt of a speech delivered by president Michelle Bachelet in 2015 illustrates the rules:

> Estamos trabajando a través de una gran reforma fiscal que nos proporcionará dinero permanente para reformas muy importantes que estamos llevando a cabo, como la Reforma Educativa, para garantizar educación para todos, con calidad, también gratuita.[17]

In this example, there are three verbs whose subject is omitted, namely: *estamos* (twice) (to be) and *garantizar* (to guarantee). In the first case, the subject assignation rule adds a token before the verb *estamos* providing the classification: pronoun, first person plural. In the case of the infinitive verb *garantizar*, the rules created take the person of the closest verb (*estamos*) and assigns it to the verb in the infinitive form.

A preliminary assessment of the Spanish language coding scheme indicates that the "translation" from English has been successful. A comparison of five speeches (∼7000 words) by Vicente Fox available in Spanish and English translation yields a similar number of coded verbs (302 vs. 314), with the same rank order in the coding categories, and very close I1 and P1 scores (0.67 vs. 0.70 and 0.51 vs. 0.58, respectively). However, finding the information and collecting the source texts necessary for this type of analysis is itself a labor-intensive endeavor with two main challenges: access to the verbal material and the variations in the information available.

In many Latin American countries, access to decision-makers' verbal material presents some difficulties due to the lack of archives or databases that compile this information. Source texts are usually disorganized and scattered around different governmental websites, presidential libraries, and ministerial archives. Likewise, the information found online mostly corresponds to verbal material that has been produced from the year 2000 onward. While the lack of organization makes the data collection process somewhat difficult, there are clear efforts from different presidents and governments to make this information available. For

[17] "We're working through a big tax reform that it'll give us permanent money for very important reforms that we're carrying on like Education Reform, to ensure education for all, with quality but also free of charges" (World Leaders Forum, Columbia University).

Table 7. Spanish norming group

	N	Mean	Standard Deviation
I1 (Strategic approach to goals)	15	0.66	0.08
I2 (Tactical pursuit of goals)	15	0.38	0.08
P1 (Nature of the political universe)	15	0.49	0.15
P2 (Realization of political values)	15	0.32	0.12

Source: Own depiction.

instance, former presidents have created personal websites and foundations where researchers can easily access source texts (e.g., Cristina Fernandez and Ricardo Lagos).

Another difficulty is the variation in the information found online. The data obtainable on governmental websites, and sometimes the URL, changes when a new administration takes office. While regaining access to the old data can be challenging, there are places such as the Wayback Machine website, presidential foundations, national libraries, and archives where some of this information can be retrieved.

The main empirical implication of the development of the Spanish OCA is the increase of the verbal material and leaders that can be analyzed using this framework. Considering that Latin American leaders are hardly ever required to speak in English, the current language barrier confines analyses to the very few translated speeches available. In cases where verbal material is not accessible in English, the application of OCA requires prior translation of leaders' utterances, which can be costly and time-consuming in a region where access to research funding is quite insufficient. The current use of OCA and other at-a-distance schemes by scholars interested in Latin America is quite limited compared to their application to leaders from other regions. The restricted use of this technique contrasts with widespread interest in, and relevance attributed to, political leaders in the region by researchers working in the field. Therefore, the Spanish version of the OCA coding scheme allows for the assessment of a myriad of political leaders that have not been analyzed before due to language and material restrictions.

Furthermore, an initial coding of material for fifteen presidents has generated a preliminary norming group for Latin American leaders (see table 7) that can be used to conduct comparative research on decision makers' beliefs.

Spanish is the official language in nineteen countries in Latin America, accounting for over 400 million people (Instituto Cervantes 2018); the Spanish OCA coding scheme allows for assessing leaders in their own language, furthering efforts to decenter FPA (Brummer and Hudson 2015). The at-a-distance assessment of political leaders using OCA or leadership trait analysis (LTA) has been predominantly in the United States and in Western countries. Hermann identifies this as a concern: "indeed, the U.S. bias in the decision-making literature has made it difficult to generalize to other countries and has given researchers blind spots regarding how decisions are made in governments and cultures, not like the American" (Hermann 2001, 49). Brummer and Hudson (2017) address this concern and discuss the boundedness of foreign policy analysis theory. They conclude that "mainstream FPA theories can be sharpened and further specified based on insights drawn from non-Western settings" (Brummer and Hudson 2017, 157).

The Latin American context has some distinctive characteristics that may shape leadership style, as well as the way leaders perceive foreign policy issues. Factors such as heavy foreign intervention in the region, which led to destabilization and a rise of dictatorships, as well as the large number of presidential regimes with concentrations of power in the executive, beg for a more region-specific analysis. Addressing

these issues is a step forward in theory-building in a region where the field of foreign policy analysis remains underdeveloped. Moreover, the characteristics of the Spanish language will expand the assessment of leaders' operational codes by incorporating the subtleties and different uses of Spanish verb modes and tenses. For instance, leaders' preferences for different types of past tenses as well as the use of the subjunctive mode are two characteristics worth analyzing further, for they may be associated with linguistic strategies utilized by Latin American leaders in their public addresses.

The Spanish OCA coding scheme still has room for improvements. A follow-up version of the scheme should include: (a) the addition of more verbs and conflict/cooperation words to improve the accuracy of the assessment; (b) additional rules to score verbs that lack a subject and have identical conjugations for the first, second (formal), and third person singular; and (c) additional rules to detect and interpret language subtleties and different social and cultural contexts.

Furthermore, the Latin American region is rather heterogeneous. There are social, historical, and cultural differences within its population. This diversity is also reflected in the use of the Spanish language, whose words and expressions vary throughout the region. These differences, and how they can influence the assessment of decision makers' operational codes, are elements that require further investigation.

In a nutshell, there is plenty to be done in terms of research utilizing the Spanish version of the scheme. Making the OCA coding scheme available in Spanish constitutes a first step that will allow researchers to conduct analyses that consider the specificities of Latin American leaders and the regional context, contributing to theory building.

Profiling Leaders in German

Christian Rabini

Katharina Dimmroth

Klaus Brummer

All Catholic University of Eichstätt-Ingolstadt

and

Mischa Hansel

Development and Peace Foundation

Chancellors matter in German politics. Their special position is due to both formal and informal prerogatives, which is why Germany is often called a "chancellor democracy" (Niclauß 2015). Apart from a chancellor's power to determine policy guidelines, noncodified powers further enhance their status. Since most chancellors have simultaneously been leaders of the main ruling party, they profoundly shape the policy-making process through party control and agenda-setting powers. The influence of chancellors is further enhanced by the media tendency to personalize politics by putting leaders center stage (Niclauß 2015). This personalization is exemplified by chancellors' nicknames, which refer to a perceived personality trait. Hence, Schröder is known as the "Basta-Kanzler," a term to describe his propensity to cut off cabinet discussions to make his own decision. Brandt was called "Willy Wolke" ("Willy Cloud") to describe his visionary thinking instead of focusing on bread-and-butter politics. Merkel even received her own verb "merkeln" which

denotes the process of waiting out a difficult situation until there is no other option but to react.

Since chancellors and foreign ministers matter in German foreign policy, a systematic approach for leadership analysis sheds further light on the personality of these leaders that goes beyond anecdotal evidence. By providing German coding schemes for leadership trait analysis (LTA) (Hermann 2005a), we broaden the scope of this method to all German foreign policy leaders (i.e., chancellors and foreign ministers from 1949 to 2017).

The creation of a German coding scheme for LTA goes beyond simply translating the rules from English into German. Cultural and linguistic idiosyncrasies need to be taken into account. Language is more than just a medium of communication: "The *natural* differentiation of languages has become a positive phenomenon underlying the allocation of peoples to their respective territories, the birth of nations, and the emergence of the sense of national identity" (Eco 1995, 339; emphasis in the original).

Every language reflects the cultural and historical identity of its speakers, which has serious repercussions for the creation of the coding scheme. There are three linguistic and cultural idiosyncrasies of the German language that might distort results when English translations of German leaders' utterances are used: compound words, omission of words, and separable verbs (Rabini et al. forthcoming). Compound words consisting of two or more nouns or adjectives and nouns such as "Arbeitsmarktreform" (labor market reform) are a common tool in German to create new words, especially in the political context (Girnth 2015, 67–68).

Our speech material has also proven that German leaders omit existing words they deem too tainted by National Socialism. Three examples from our speech material illustrate this point: leader(ship), honor, and hero. Whereas these terms are part of the standard rhetorical toolbox of current non-German political leaders, and had also been used in Germany before 1945 (Schmitz-Berning 2008,163; 240; 306–8), they have been virtually nonexistent in spontaneous speech material of German leaders since. Our text corpus contains 146,000 words; the word "leader" (*Führer*) or related words such as "leadership" (*Führung* or *Leitung*) were only used six times; the word "honor" (*Ehre*) or related adjectives like "honorous" (*ehrenhaft*) were used only once; and the word "hero" (Held) or related words such as "heroic" (*heldenhaft*) were not used at all. In the few instances in which these words were used by German leaders, they never referred to themselves or their in-group but to external actors.

Additionally, grammatical idiosyncrasies of the German language also demand new coding rules. Separable verbs consist of a core verb and a separable prefix and are very commonplace in German (e.g., "Wir haben Land X angegriffen" [We have attacked country X] or "Wir greifen Land X an" [We attack country X]). Since they can be separated but do not have to be, at least two rules are needed to identify them. In addition, the separable prefix may be situated several tokens[18] away from the core verb and can be mistaken with a preposition; rules ought to be sufficiently flexible to match the separable prefix with the core verb and sufficiently rigid not to count unrelated prepositions as part of the verb. For example, there are two rules for the separable form of *angreifen*. The first looks for a token with a self-ingroup marker, followed immediately by a token with lemma *greifen* and within nine tokens the lemma *an*. The second looks for a token with a self-ingroup marker immediately preceded by a token with lemma *greifen* and within nine tokens after a token with the lemma *an*. The nine-token distance captures most instances of the separable form without coding too many unrelated prepositions. Overall the process was a success. A comparison of results between the German and the English versions,

[18] A token is the smallest entity in a sentence, comprising words, numbers, punctuation and special characters (Schiller et al. 1995, 4).

which was based on some 150,000 words of source material per language, resulted in an average deviation below three percent for all seven categories.

These idiosyncrasies hold an important lesson for coding-scheme creation in general. They highlight the fact that not only should leaders be analyzed in their native language but language and cultural varieties also need to be considered. There are ten linguistic centers of the German language; three of these—Germany, Austria, and Switzerland—are also nation states (Ammon 2018, 71). The combination of nation state and language center has created a variety of words that are understandable to all German speakers but carry additional meaning in the respective centers.[19] The term "Wiedervereinigung" (reunification) is a common term in German. It was a core political concept for many German leaders and—depending on the context and the speaker—could either be a call for unity or a form of threat (for leaders of the GDR). Our LTA version is limited to German leaders and reflects their vocabulary. When analyzing leaders from other language centers, the self-reference and actor tables need to be adjusted. Furthermore, a thorough dictionary and corpora analysis is necessary to find relevant indicators from other language centers.

Despite the much greater availability of German speech material for German leaders, there are still three challenges pertaining to source material. They concern the availability, permanence, and pertinence of sources. Regarding *availability*, it is important to note that while LTA draws on interviews as its main object of analysis, not all interviews fulfill LTA's requirement of spontaneity to the same degree (Hermann 2005a, 179–81). For instance, interviews in newspapers are typically edited prior to publication, which limits the spontaneity of the response. To navigate this issue, we collected as many televised and radio emitted interviews as possible—our assumption being that in those settings, the interviewee is not in a position to demand an edit of the given answer and has to rely on their own ability to formulate a response. While this strategy was successful for the more recent leaders, we could not locate a sufficient amount of televised or radio emitted material for other leaders in our data set to hit our goal of fifty speech acts per leader. Given the nearly seventy-year span of our analysis, this is hardly surprising. When Adenauer took office in 1949, radio interviews were few and televised ones almost unheard of. In general, the further an analysis stretches into the past, the harder it might be to find speech acts beyond print.

Regarding *permanence*, we unexpectedly encountered a lack of online source material for the set of leaders whose office terms fell around the turn of the century. Although the Chancellery and the Federal Foreign Office (*Auswärtiges Amt*, AA) maintained their own online presences at the time, they along with their archived material have since been taken offline. Archival sites, such as the Wayback Machine, allowed us to recover some of the material, but the lack of stable, permanent hosting sites is definitely an issue to consider when gathering material for empirical analysis. Even with recently published material, we sometimes found that news outlets would pull their content after a certain amount of time. This problem can be combatted by downloading the material to ensure traceability.

Finally, concerning *pertinence*, the issue with making printed interviews part of the analysis is that the answers could have been edited. Since we perused archival material of the Federal Archive, the AA, and the Konrad-Adenauer Foundation, we found exchanges between leaders' offices and the news outlet containing editing requests. Thankfully, the archive material oftentimes contained the unedited version for transcription purposes, which we were able to use for our analysis. This also brought our attention to what can be called "letter interviews." Especially in the early days of the FRG, interviews were not always given in person. Instead, questions were sent by the news outlets to the office of the person one was seeking a

[19] See the *Variantenwörterbuch des Deutschen* (Ammon et al. 2016).

Table 8. German foreign policy leaders' norming group

Trait	Belief in ability to control events	Need for power	Conceptual complexity	Self-confidence	Task focus	Distrust of others	In-group bias
German leaders (n = 17)	Mn = 0.28 SD = 0.06	Mn = 0.26 SD = 0.06	Mn = 0.61 SD = 0.03	Mn = 0.37 SD = 0.06	Mn = 0.64 SD = 0.05	Mn = 0.14 SD = 0.06	Mn = 0.12 SD = 0.02

Source: Own depiction.

response from. The response oftentimes came from the staff, which obviously defies the "spontaneous" aspect that LTA requires. We only made an interview part of our corpus once we could ensure that it was the result of an actual meeting of the interviewee with a journalist.

Drawing together the newly developed German coding scheme and the compiled source material in German, the advantages of a German LTA version in empirical terms are threefold. First, it extends the analysis to leaders that could hitherto not be analyzed due to a lack of source material; second, it sheds new light on the role of contextual variables; and third, it creates a base for a German norming group.

The greater availability becomes evident when we look at the corresponding English source material. While we gathered fifty speech acts with at least one-hundred words each for all seventeen chancellors and foreign ministers from 1949 to 2017, this task proved much more difficult for German leaders' source texts in English. We met the threshold for just seven of the leaders and only because they included both spontaneous and prefabricated texts (Rabini et al. forthcoming). Moreover, six of these seven leaders were in office after reunification. This might suggest that English translations of German leaders' utterances have only recently become standard practice or that they simply had not been archived before. Either way, a German version broadens the scope for empirical analysis by including all foreign policy leaders of the Federal Republic.

The German LTA version also adds value on the impact of contextual variables on leadership traits and the extent to which the latter can be seen as stable over time more generally. For instance, comparing leaders who served in two different functions—such as Brandt as foreign minister as well as chancellor—renders it possible to examine the influence of bureaucratic roles on leaders. The question of the impact of contextual variables touches on the topic of the stability of traits over time. By enabling comparative case studies, we can now examine with a rigorous methodology the extent to which German leaders' traits have changed before and after major external shocks, such as the fall of the Berlin Wall or 9/11.

Finally, the German LTA version allows the creation of a German-speaking leaders' norming group (see table 8). This could serve as a basis for comparison with both future German foreign policy leaders and those from other world regions. Both intra- and cross-regional comparisons could shed new light on the particular leadership styles of German foreign policy leaders and might also provide new insights into the role of history, language, and external variables on the characteristics of leaders.

Although the German-language coding scheme for LTA contributes to the understanding of German foreign policy as well as foreign policy leadership more generally, there are several avenues for future research. First, in terms of probing the external validity of the coding scheme, the German foreign policy leaders' biographies can be studied to assess their leadership style and how well this corresponds with the profile LTA offers for each individual. Second, in order to

analyze leaders from other German language centers, the self-reference and actor tables need to be adjusted. A thorough dictionary and corpora analysis has to be conducted to find further relevant indicators from other language centers.

Finally, a follow-up "German LTA version 2.0" could add even more context to the indicators. The current version was created by using dictionaries to search for synonyms of indicators and corpora to find out more about the indicator's contexts. Thanks to corpora such as the *Digitales Wörterbuch der Deutschen Sprache* (DWDS 2018) or the Leipzig Corpora Collection (LCC 2018), rules were created that are more sensitive to context and contain numerous collocations, which are defined as "any holistic lexical, lexico-grammatical or semantic unit normally composed of two or more words which exhibits minimal recurrence within a particular discourse community" (Siepmann 2005, 438). This particular combination of words could serve as an example for a more context-sensitive LTA version of the future, since collocation profiles retrieved from corpora for every indicator provide an opportunity to refine rule creation and help to get a better picture of the context of an indicator.

Profiling Leaders in Persian

Ameneh Mehvar

University of St. Andrews

Foreign policy decision-making in Iran is a pluralistic process that involves various state, nonstate, and semistate entities. Understanding the role of individuals therefore requires more than the sole focus on the personality of the supreme leader—the highest authority in the country—and should also involve the study of several other influential actors, such as the president, who can impact the process directly or indirectly to varying degrees. Consequently, we need to go beyond single qualitative case studies and employ approaches, such as the LTA framework, that facilitate the systematic assessment and comparison of different stakeholders.

This contribution discusses the adaptation of the coding scheme for conceptual complexity (CC), as one of the seven traits covered by LTA, into Persian using PersianAnalyzer, a custom program developed by David Symons for the purpose of this project. This process involved resolving several issues related to particularities of the Persian language and its differences to English.

The first task was creating complexity lists in Persian. Certain English complexity words and their synonyms are not indicative of complexity in Persian. For example, the high complexity word "also" can be translated to several Persian terms, including "همچنین," "نیز," "بعلاوه," and "هم." However, the last of these is frequently used as a meaningless modifier. Conversely, several different English words map to a single Persian word. For instance, the complexity words "must" and "ought," alongside the noncomplexity words "should" and "shall," all translate to "باید" in Persian. The final complexity lists were compiled by identifying words that were indicative of complexity in several Persian political statements. Ambiguous cases that require human judgment had to be precluded.

Next, several computational challenges pertaining to automated coding in Persian had to be addressed. One issue concerns problems arising due to the "lack of standardization in Persian orthography in terms of different writing styles, spacing and font encoding" (Seraji et al. 2014). For example, words can be written in several different forms in Persian when they include imported "Tanwin" and "Hamza" sounds from Arabic. Moreover, numerous Persian webpages encode text using Arabic letters. While Persian and Arabic share the same character encoding for most letters, this is not always the case, such as with the letters "ی" and "ک." Other times, Western digits or punctuation are used instead of their Persian equivalents.

A further issue is that compound words or words that have affixes can be typed in three different ways in Persian: with no space, whitespace, or short-space between them. The short-space or the so-called zero-width non-joiner (ZWNJ) is a "non-printing character in computerized typesetting placed between two characters to keep the word forms intact and close together without being attached" (Mirzanezhad and Feizi-Derakhshi 2016, 1095). But the use of ZWNJs can lead to problems in detecting words. To remedy these issues, PersianAnalyzer standardizes all texts by substituting non-Persian sounds, letters, digits, and punctuation with Persian characters and deleting the nonprinting characters.

A further challenge concerns words in the complexity lists that, when used in certain combinations, are no longer indicative of complexity. For instance, while "جامع" (comprehensive) is a low complexity word, it should not be counted as such when referring to the "برنامه جامع اقدام مشترک" (Joint Comprehensive Plan for Action)—that is, the nuclear agreement reached between Iran and the P5 + 1. Another example is the low complexity word "اساسی" (fundamental). In Persian, the term for constitution is "اساسی قانون" (fundamental law), which is not a complexity word. Therefore, an exception list had to be created to instruct the program to skip over any such phrases.

Another issue is the recognition of complexity words that include an affix or are written joined up with the neighboring word. Persian is a language rich in affixes. Some affixes are always connected to the word stem, while others may be written connected or disconnected to the host word. For instance, the plural forms of nouns can be formed by adding the postfixes "ات," "ون," "ین," "ان," and "ها" to the word. The word "alternatives," for instance, can be written as "جایگزین ها" or "جایگزینها." In the latter case, the token needs to be dissected in search of a concealed stem. The language also allows words such as the preposition "به," the determiner "این," the postposition "را," or the relativizer "که" to be connected to the preceding word (Megerdoomian 2004). The stem of words in the current project is obtained through the affix stripping method—that is, by "removing some morphemes from one or both sides of the word" (Sharifloo and Shamsfard 2008, 583).

But this process is not always as straight forward as it might seem. Consider the following example: When the high complexity word "برخی" (some) is preceded by the preposition "به" (to), they can be written joined up as "ببرخی." Since PersianAnalyzer cannot immediately recognize "ببرخی" as a complexity word, it will proceed to obtain the stem by stripping pre- and/or postfixes. For a machine this means sequentially attempting all possible combinations. Our program starts by checking for a possible postfix and would recognize "ی" as a candidate. "ی" can be a postfix in several instances, including when it appears after words that end with vowels in ezafe constructions (Mirzanezhad and Feizi-Derakhshi 2016, 1096), or it can act as the indefinite article (Dolamic and Savoy 2009, 389). But in the case of "برخی," the "ی" is simply the last letter of the word stem and not a postfix. The remainder of the token—that is, the token without the candidate postfix—is then again checked against the complexity lists. Still no match would be found due to the extant prefix, which is detected in the next step. But even after its removal from the token, the erroneously stripped "ی" prevents recognition. For the complexity stem to finally be identified, the "ی" ending must be rejoined. In this way, the ambiguity is resolved, leading to a successful match.

A further necessary step for our program is to record all affixes, both connected and disconnected. These may become relevant for scoring complexity since some affixes in Persian, such as "بی," "غیر," "نا," "ن," and "ناپذیر," may flip the classification of the complexity word. An overview of how recognition is achieved is illustrated in figure 1.

Another case in which classification might flip is in the presence of negating verbs. Due to the difficulties of associating negating verbs with their corresponding

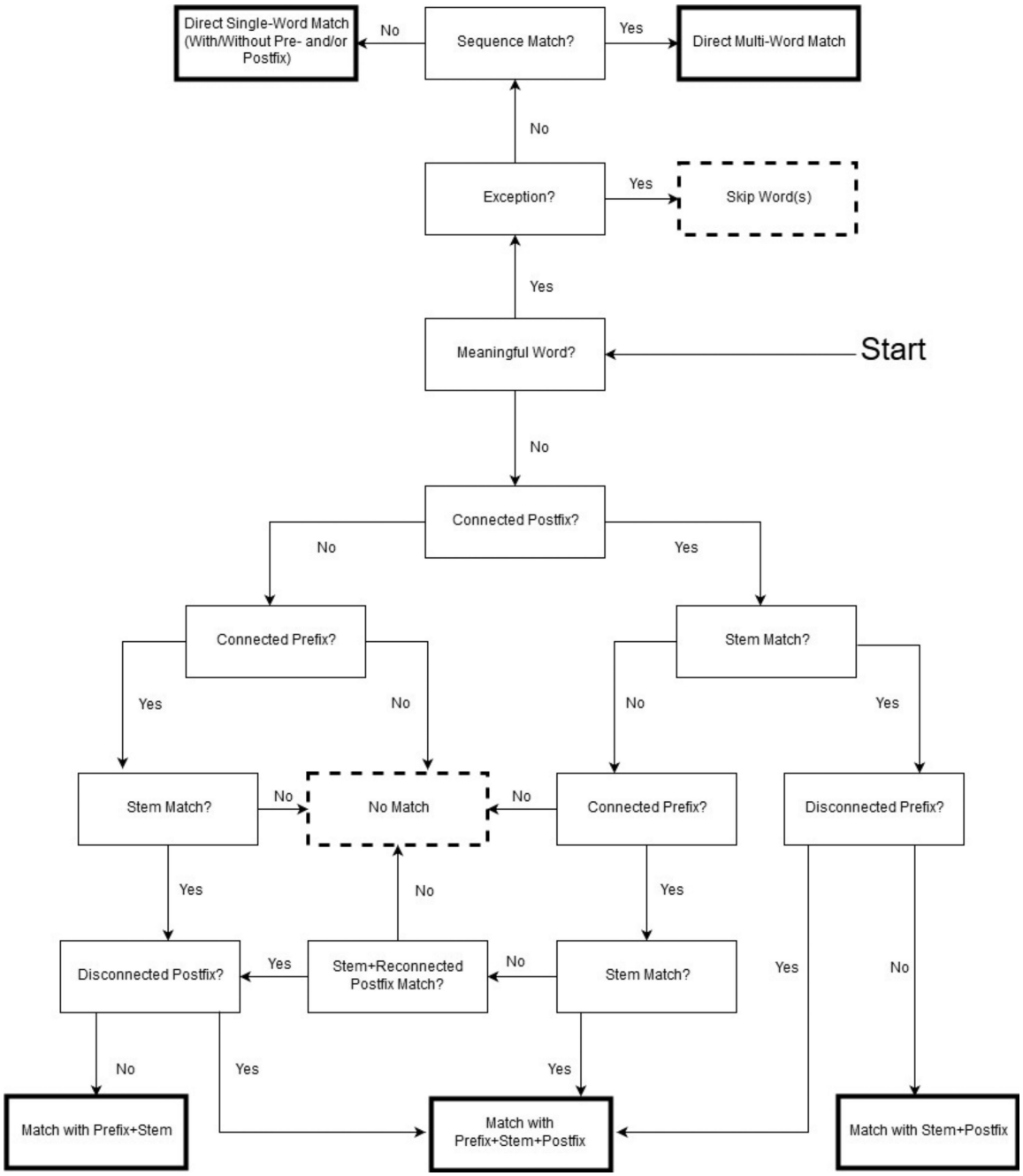

Figure 1. Complex word recognition process.
Source: Own depiction.

complexity words, PersianAnalyzer is currently limited to the recognition of trailing negations.

In order to assess the reliability of the developed program, its output had to be validated against some other measures. To this end, a test-test procedure (McKenny et al. 2018, 2914) was performed by coding a sample of fifteen randomly selected statements by three Iranian presidents by both hand and computer, which contained 1,133 coded complexity words. The two coding procedures achieved an interrater reliability of 86.3 percent and 0.728, using percent agreement and Cohen's Kappa methods, respectively. Therefore, a satisfactory level of agreement has been established.

The utility of the developed CC coding scheme is partially dependent on the availability of Persian texts, but obtaining these is not always an easy and straightforward task. The number of statements available in a digital format can vary significantly between different politicians. For example, the supreme leader, the sitting president, and some other politicians, such as former president Rafsanjani, have

Table 9. Ahmadinejad's CC scores in Persian and English

	Ahmadinejad interviewed by	Persian score	English score
1	Wallace-CBS (2006)	0.49	0.61
2	Aust-Spiegel (2006)	0.49	0.59
3	Columbia University (Q&A session 2007)	0.46	0.53
4	Snow-Channel 4 (2007)	0.45	0.60
5	Sawyer-ABC (2007)	0.36	0.52
6	Stephanopoulos-ABC (2009)	0.56	0.51
7	Curry-NBC (2009)	0.40	0.51
8	King-CNN (2009)	0.50	0.56
9	Zakaria-CNN (2011)	0.46	0.53
10	RT (2011)	0.40	0.42
	Mean:	0.46	0.54

Source: Own depiction.

websites that contain large volumes of their statements. But remarks by many other key politicians are much harder to come by. At times, remarks are taken off the internet after a while, and it becomes much more difficult to retrieve them. In other cases, audio or video materials are available on the web but require transcription. Furthermore, statements by politicians from the prerevolutionary era or the postrevolutionary period before the internet age have not been widely digitalized and need to be retrieved from physical archives. The digitalization of archival material in Iran can however be connected to some degree of risk, common in conducting field research in authoritarian states. Therefore, depending on the subject of study, the effort required for data collection can vary significantly. Despite this, there is still a much larger pool of spontaneous source material available in Persian than in English. Except at the most official occasions, such as presidential inaugurations or when addressing foreign dignitaries, the majority of Iranian politicians speak freely in public, and many of these statements are made available in a digital format.

The developed Persian coding scheme also offers several other benefits. Most importantly, our confidence in the validity of the CC scores is higher since adverse impacts of suboptimal translations leading to skewed or even invalid results can be ruled out. The following example demonstrates how the adapted coding schemes can produce superior results by making a direct comparison between the scores of ten translated and original spontaneous statements made by former president Ahmadinejad.

As seen in table 9, the analyses of original and translated texts lead to very different scores: a two-tailed paired sample t-test yields a t-value of 4.1201 and a p-value of 0.0026, suggesting a statistically significant difference at $p < 0.05$ across the scores of the two different groups. These results lead to different interpretations of Ahmadinejad's CC. The assessment based on translated texts indicates that the former Iranian president has a moderate CC,[20] while the score obtained from original documents suggests that he is low on this trait. In order to determine which of the two scores is more valid, one may check the validity of the analyses through assessing their predictive validity. A content analysis is predictively valid if it can "accurately anticipate events, identify properties, or describe states of affairs, knowledge of which is absent or did not enter that analysis" (Krippendorff 2004, 321). Research on conceptual complexity has linked this trait to an individual's need for information and advice (Preston 2001). Based on this, one would expect Ahmadinejad to avoid broad collection of information and to be closed to advice if the results of the

[20] Based on comparison with the norming group of 284 world leaders.

Table 10. Iranian leader's norming group

Iranian Politician	CC Score
Khamenei	0.50
Khomeini	0.51
Rafsanjani	0.54
Khatami	0.56
Ahmadinejad	0.44
Rouhani	0.55
Ali Larijani	0.56
Jannati	0.42
Sadeq Larijani	0.52
Jalili	0.53
N = 10	
Mean	0.51
Standard Deviation	0.04

Source: Own depiction.

original Persian material were valid. Going by the results of the translated analysis, Ahmadinejad would exhibit these behaviors in moderation.

A qualitative case study based on interviews given by several Iranian politicians who worked closely with Ahmadinejad confirms that his actual behavior fulfills the predictions suggested by the low CC score: they all indicate that he was averse to advice, consultation, feedback, and criticism. Therefore, it can be argued that the Persian content analysis has delivered the valid CC score for Ahmadinejad. In order to make a more general statement about the method, the validity of more cases needs to be assessed. A first step has been made in confirming the results for four other politicians: qualitative case studies suggest that Rafsanjani, Khatami, Rouhani, and Khamenei's actual behaviors concerning openness to advice and information are in line with the predictions of the Persian content analysis.

A further opportunity that the Persian coding scheme offers is the creation of a norming group for Iranian leaders. This facilitates comparison between politicians in Iran, enabling a more accurate interpretation of their CC scores that is sensitive to linguistic and cultural specificities of the Persian language. A small norming group of ten people can be seen in table 10, which can be expanded once remarks by more politicians have been collected. With a mean score of 0.51, the CC score for the Iranian leaders' norming group is lower than the mean score of 0.59 for the norming group of 284 world leaders. Depending on which norming group is used, the interpretation of a CC score can differ.

The benefits offered by the conceptual complexity coding scheme, including higher confidence in its validity and the opportunity to study more individuals, incentivize the adaptation of the other six LTA coding schemes into Persian. This task will entail dealing with further particularities of the Persian language, such as the free word order in conversational Persian; omission of subject; the creation of new words by combining affixes, stems, nouns, and adjectives; and compound verbs with long-distance dependencies. Addressing some of these points will require the extension of the program to include part-of-speech tagging capabilities. This would also improve the current CC coding scheme's ability to associate negating verbs with the appropriate complexity word. In order to allow other researchers to further validate, improve, and expand on the current project, and to mitigate the issue of "black box" measurement (Neuendorf 2002, 129), the PersianAnalyzer's source code and word lists have been made available via GitHub.[21]

[21] See https://github.com/DavidSymonz/PersianAnalyser.git.

Acknowledgements

The development of the German coding scheme for LTA was funded by a grant from the German Research Foundation (DFG) (project number 288437573). The development of the Arabic operational code scheme was funded by a research grant from the office of Global Affairs at University of Connecticut.

References

Ammon, Ulrich. 2018. "Deutsch Als Plurizentrische Sprache Und Der Intertourismus Zwischen Den Zentren." *Sociolinguistica* 32 (1): 69–78.

Ammon, Ulrich, Hans Bickel, and Alexandra Nicole Lenz, eds. 2016. *Variantenwörterbuch des Deutschen—Die Standardsprache in Österreich, der Schweiz, Deutschland, Liechtenstein, Luxemburg, Ostbelgien und Südtirol sowie Rumänien, Namibia und Mennonitensiedlungen*. Berlin: De Gruyter.

Bayulgen, Oksan, Ekim Arbatli, and Sercan Canbolat. 2018. "Elite Survival Strategies and Authoritarian Reversal in Turkey." *Polity* 50 (3): 333–65.

Breuning, Marijke. 2007. *Foreign Policy Analysis: A Comparative Introduction*. New York: Palgrave Macmillan.

Brummer, Klaus, and Valerie M. Hudson, eds. 2015. *Foreign Policy Analysis beyond North America*. Boulder, CO: Lynne Rienner.

Brummer, Klaus, and Valerie M. Hudson. 2017. "The Boundedness of Foreign Policy Analysis Theory?" *Global Society* 31 (2): 157–66.

Cagaptay, Soner. 2017. *The New Sultan*. London: IB Tauris.

Cason, Jeffrey, and Timothy Power. 2009. "Presidentialization, Pluralization, and the Rollback of Itamaraty: Explaining Change in Brazilian Foreign Policy Making in the Cardoso-Lula Era." *International Political Science Review* 30 (2): 117–40.

Chodor, Tom, and Anthea McCarthy-Jones. 2013. "Post-liberal Regionalism in Latin America and the Influence of Hugo Chávez." *Journal of Iberian and Latin American Research* 19 (2): 211–23.

Cuhadar, Esra, Juliet Kaarbo, Baris Kesgin, and Binnur Ozkececi-Taner. 2017. "Personality or Role? Comparisons of Turkish Leaders Across Different Institutional Positions." *Political Psychology* 38 (1): 39–54.

Dalacoura, Katerina. 2001. "Islamist Movements as Non-state Actors and Their Relevance to International Relations." In *The Non-State Actors in World Politics*, edited by Daphne Josselin and William Wallace, 235–50. New York: Palgrave.

Dekmejian, Richard H. 1975. *Patterns of Political Leadership: Egypt, Israel, Lebanon*. Albany: State University of New York Press.

Demir, Ali F. 2007. *Turk Dis Politikasinda Liderler*. Istanbul: Baglam Yayinlari.

Digitales Wörterbuch der deutschen Sprache [DWDS]. Accessed February 12, 2019. https://www.dwds.de.

Dolamic, Ljiljana, and Jacques Savoy. 2009. "Persian Language, Is Stemming Efficient?" *Paper presented at the 20th International Workshop on Database and Expert Systems Application*, Linz, March 2009.

Duelfer, A. Charles, and Stephen B. Dyson. 2011. "Chronic Misperception and International Conflict: The US-Iraq Experience." *International Security* 36 (1): 73–100.

Eco, Umberto. 1995. *The Search for the Perfect Language*. Oxford: Blackwell.

Ethnologue. 2019. "Language Size." Accessed May 17, 2019. https://www.ethnologue.com/statistics/size.

George, Alexander L. 1969. "The 'Operational Code': A Neglected Approach to the Study of Political Leaders and Decision-Making." *International Studies Quarterly* 13 (2): 190–222.

Girnth, Heiko. 2015. *Sprache und Sprachverwendung in der Politik: Eine Einführung in die linguistische Analyse öffentlich-politischer Kommunikation*. Berlin: De Gruyter.

Hart, Roderick. 1985. "Systematic Analysis of Political Discourse: The Development of DICTION." In *Political Communication Yearbook: 1984*, edited by Keith R. Sanders, Lynda Lee Kaid and Dan Nimmo, 97–134. Carbondale, IL: Southern Illinois University Press.

Heper, Metin, and Sabri Sayari, eds. 2002. *Political Leaders and Democracy in Turkey*. Lanham, MD: Lexington Books.

Hermann, Margaret G. 1980. "Explaining Foreign Policy Behavior Using the Personal Characteristics of Political Leaders." *International Studies Quarterly* 24 (1): 7–46.

_____. 2001. "How Decision Units Shape Foreign Policy: A Theoretical Framework." *International Studies Review* 3 (2): 47–81.

_____. 2005a. "Assessing Leadership Style: Trait Analysis." In *The Psychological Assessment of Political Leaders. With Profiles of Saddam Hussein and Bill Clinton*, edited by Jerrold M. Post, 178–212. Ann Arbor: University of Michigan Press.

_____. 2005b. "Saddam Hussein's Leadership Style." In *The Psychological Assessment of Political Leaders: With Profiles of Saddam Hussein and Bill Clinton*, edited by Jerrold Post, 375–86. Ann Arbor: The University of Michigan Press.

_____. 2008. "Content Analysis." In *Qualitative Methods in International Relations*, edited by Audie Klotz and Deepa Prakash, 151–67. London: Palgrave Macmillan.

Hinnebusch, Raymond. 2015. "Foreign Policy Analysis and the Arab World." In *The Foreign Policy Analysis beyond North America*, edited by Klaus Brummer and Valerie M. Hudson, 77–100. Boulder, CO: Lynne Rienner.

Hudson, Valerie M. 2014. *Foreign Policy Analysis. Classic and Contemporary Theory*, 2nd edition. Lanham, MD: Rowman & Littlefield.

Instituto Cervantes. 2018. "El Español: Una Lengua Viva. Informe 2018."

Jacoby, Tim. 2019. "Islam and the Islamic State's Magazine, Dabiq." *Politics and Religion* 12 (1): 32–54.

Jenne, Nicole, Luis Schenoni, and Francisco Urdinez. 2017. "Of Words and Deeds: Latin American Declaratory Regionalism, 1994–2014." *Cambridge Review of International Affairs* 30 (2–3): 195–215.

Kesgin, Baris. 2013. "Leadership Traits of Turkey's 'Islamist' and 'Secular' Prime Ministers." *Turkish Studies* 14 (1): 136–57.

_____. 2019a. "Turkey's Erdogan: Leadership Style and Foreign Policy Audiences." *Turkish Studies* , advance view. DOI:10.1080/14683849.2019.1575735.

_____. 2019b. "Features of Foreign Policy Birds: Israeli Prime Ministers as Hawks and Doves." *Cooperation and Conflict*, advance view. DOI:10.1177/0010836719850208.

Krippendorff, Klaus H. 2004. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage.

Lazarevska, Elena, Jayne M. Sholl, and Michael D. Young. 2006. "Links among Beliefs and Personality Traits: The Distinctive Language of Terrorists." In *The Beliefs and Leadership in World Politics: Methods and Applications of Operational Code Analysis*, edited by Mark Schafer and Stephen G. Walker, 171–86. New York: Palgrave Macmillan.

"Leipzig Corpora Collection" [LCC]. Accessed February 12, 2019. http ://corpora.uni-leipzig.de/en.

Levine, Nick, and Michael D. Young. 2014. "Leadership Trait Analysis and Threat Assessment with Profiler Plus." In *Proceedings of ILC 2014 on 8th International Lisp Conference*, 50–59. Association for Computing Machinery Digital Library.

Malici, Akan, and Allison, L. Buckner. 2008. "Empathizing with Rogue Leaders: Mahmoud Ahmadinejad and Bashar al-Asad." *Journal of Peace Research* 45 (6): 783–800.

McKenny, Aaron, F. Herman Anguinis, Jeremy C. Short, and Aaron H. Anglin. 2018. "What Doesn't Get Measured Does Exist: Improving the Accuracy of Computer-Aided Text Analysis." *Journal of Management* 44 (7): 2909–33.

Megerdoomian, Karine. "Developing a Persian Part of Speech Tagger." *Paper presented at the 1st Workshop on Persian Language and Computer*, Tehran, May 2004.

Mirzanezhad, Zeinab, and Mohammad-Reza Feizi-Derakhshi. 2016. "Using Morphological Analyzer to Statistical POS Tagging on Persian Text." *International Journal of Computer Science and Information Security* 14 (8): 1093–103.

Neuendorf, Kimberly A. 2002. *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage.

Niclauß, Karlheinz. 2015. *Kanzlerdemokratie: Regierungsführung von Konrad Adenauer bis Angela Merkel*. Wiesbaden, Germany: Springer.

Norris, Clayton. 2016. Petrarch 2: Petrarcher. Posted July 13, 2017. http://dx.doi.org/10.2139/ssrn.2733766.

Özdamar, Özgür. 2017. "Leadership Analysis at a 'Great Distance': Using Operational Code Construct to Analyze Islamist Leaders." *Global Society* 31 (2): 167–98.

Özdamar, Özgür, and Sercan Canbolat. 2018. "Understanding New Middle Eastern Leadership: An Operational Code Approach." *Political Research Quarterly* 71 (1): 19–31.

Pennebaker, James W., Martha E. Francis, and Roger J. Booth. 2001. *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Erlbaum Publishers.

Preston, Thomas. 2001. *The President and His Inner Circle*. New York: Columbia University Press.

Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42 (3): 427–60.

Rabini, Christian, Klaus Brummer, Katharina Dimmroth, and Mischa Hansel. Forthcoming. "Profiling Foreign Policy Leaders in Their Own Language: New Insights into the Stability and Formation of Leadership Traits." *British Journal of Politics and International Relations*.

Schafer, Mark. 2000. "Issues in Assessing Psychological Characteristics at a Distance: An Introduction to the Symposium." *Political Psychology* 21 (3): 511–27.

Schafer, Mark, and Stephen G. Walker. 2006a. "Operational Code Analysis at a Distance: The Verbs in Context System of Content Analysis." In *Beliefs and Leadership in World Politics. Methods and Applications*

of Operational Code Analysis, edited by Mark Schafer and Stephen G. Walker, 25–51. Basingstoke, UK: Palgrave Macmillan.

SCHAFER, MARK, AND STEPHEN G. WALKER, eds. 2006b. *Beliefs and Leadership in World Politics: Methods and Applications of Operational Code Analysis*. New York: Palgrave Macmillan.

SCHILLER, ANNE, SIMONE TEUFEL, CHRISTINE THIELEN, AND CHRISTINE STÖCKERT. 1995. *Vorläufige Guidelines für das Tagging deutscher Textcorpora mit STTS*. Technischer Bericht, Universitäten Stuttgart und Tübingen.

SCHMITZ-BERNING, CORNELIA. 2008. *Vokabular des Nationalsozialismus*. Berlin: De Gruyter.

SERAJI, M., CARINA JAHANI, BE'ATA MEGYESI, AND JOAKIM NIVRE. 2014. "A Persian Treebank with Stanford Typed Dependencies Mojgan." *Paper presented at the 9th International Conference on Language Resources and Evaluation*, Reykjavik, May 2014.

SHARIFLOO, AMIR, AND MEHRNOUSH SHAMSFARD. 2008. "A Bottom Up Approach to Persian Stemming." *Paper presented at the Third International Joint Conference on Natural Language Processing*, Hyderabad, January 2008.

SIEPMANN, DIRK. 2005. "Collocations, Colligation and Encoding Dictionaries. Part I: Lexicological Aspects." *International Journal of Lexicograph* 18 (4): 409–43.

VAN ESCH, FEMKE, AND MARIJ SWINKELS. 2015. "How Europe's Political Leaders Made Sense of the Euro Crisis: The Influence of Pressure and Personality." *West European Politics* 38 (6): 1203–25.

WALKER, STEPHEN G. 2003. "Operational Code Analysis as a Scientific Research Program. A Cautionary Tale." In *Progress in International Relations Theory. Appraising the Field*, edited by Colin Elman and Miriam F. Elman, 245–76. Cambridge, MA: MIT Press.

WALKER, STEPHEN G., AND MARK SCHAFER. 2006. "Belief Systems as Causal Mechanisms in World Politics: An Overview of Operational Code Analysis." In *Beliefs and Leadership in World Politics. Methods and Applications of Operational Code Analysis*, edited by Mark Schafer and Stephen G. Walker, 3–22. Basingstoke and New York: Palgrave Macmillan.

WALKER, STEPHEN G., MARK SCHAFER, AND MICHAEL D. YOUNG. 1998. "Systematic Procedures for Operational Code Analysis: Measuring and Modeling Jimmy Carter's Operational Code." *International Studies Quarterly* 42 (1): 175–90.

_____. 2005. "Profiling the Operational Codes of Political Leaders." In *The Psychological Assessment of Political Leaders: With Profiles of Saddam Hussein and Bill Clinton*, edited by Jerrold M. Post, 215–45. Ann Arbor: University of Michigan Press.

WEHNER, LESLIE. 2011. "Developing Mutual Trust: The Othering Process between Bolivia and Chile." *Canadian Journal of Latin American and Caribbean Studies* 36 (71): 109–38.

WOHLSTETTER, ROBERTA. 1962. *Pearl Harbor: Warning and Decision*. Stanford, CA.: Stanford University Press.

YANG, YI EDWARD. 2010. "Leaders' Conceptual Complexity and Foreign Policy Change: Comparing the Bill Clinton and George W. Bush Foreign Policies toward China." *Chinese Journal of International Politics* 3 (4): 415–46. ğ ü ç ş