**Arabic Tweets in Turkish Twittersphere: Historical Repertoire and Symbolic Boundaries**

This study investigates the historical repertoires and symbolic boundaries employed by Arabs to include or exclude Turks within the Turkish Twittersphere. By leveraging generative AI-aided stance analysis and topic modeling on Arabic tweets mentioning Turkish President Recep Tayyip Erdoğan, the research identifies the symbolic boundaries drawn using historical narratives, primarily centered on the Ottoman Empire. Positive stance tweets often highlight Erdoğan's Muslim identity and political strength, portraying him as a competent in-group member. In contrast, negative stance tweets depict Erdoğan through the lens of Ottoman colonialism, framing him as an out-group member with varying levels of competence. The analysis utilized a dataset from the Politus project, filtering Arabic tweets, employing a large language model (ChatGPT 4) for stance annotation and BERT for topic modeling. The study finds that both positive and negative narratives draw from the Ottoman historical repertoire, but with differing emphases reflecting nationalist and religious perspectives.

**Keywords:** Symbolic Boundaries, Historical Repertoires, Stance Analysis, Topic Modeling

**Introduction**

**Turks, Arabs and Symbolic Boundaries**

Groups utilize symbolic boundaries to include or exclude any person/group from the in-group by making use of the existing historical repertoires at their disposal (Lamont, Pendergrass, Pachucki, 2015). The historical repertoires used vary depending on where the relevant symbolic boundary is drawn. This study investigates the historical repertoires used by Arabs to include and exclude Turks by conducting topic modeling on Twitter data labeled by large language models for stance and asks whether the texts with positive and negative stances use same historical repertoire.

The relationships between Turks and Arabs have a history rich enough to allow for extensive analysis. During the rule of the Ottoman Empire, the majority of Turks and Arabs lived under the same state; the relations between them became complex with the decline and eventual collapse of the Ottoman state in the face of Western powers, and the rise of nationalist movements and nation-states. In Turkish official historiography, Arabs are portrayed as representing the primitive life from which a modernizing Turkey was trying to escape or distance itself (Aktürk, 2010). On the other hand, in Arab nationalist discourse, the Ottomans are depicted as a colonial power pursuing Turkification policies. However, there have also been movements that viewed the Ottomans positively (Nafi, 2009).

Using the term "the Arab world" might lead to a superficial and generalized analysis. The relationships of each country and its people with Turkey and Turks are not based on the same dynamics, and the policies of countries can change over time. Although there have been developments in relations between Turkey and Arab countries since the 2000s, the fundamental determining factor for relations after 2010 has been the Arab Spring (Özcan, Köse & Karakoç, 2015). In this process, Turkey's stance towards the Arab Spring movements opened the door even to military interventions. While the attitude of the Arabic-speaking public towards Turkey's policies during this period varied, sectarian and national identities have emerged as significant factors in shaping these attitudes (Özcan, Köse & Karakoç, 2015). The overthrow of the government that came to power in Egypt at the end of the Arab Spring protests, the opposition's failure to achieve success against the current regime in Syria, Turkey's military presence in northern Syria and Libya after Gaddafi's overthrown where Turkey supports the Government of National Unity against General Haftar, who is supported by Egypt and the UAE, have defined the inter-country relationships. Additionally, as a result of this process, Turkey has opened its doors to a large number of Syrian refugees, leading to a rise in anti-refugee and anti-Arab sentiment in Turkish public opinion over the years. Siviş's (2022) study with Turkish citizens of non-Turkish ethnic origin shows that Turkish citizens exclude Syrians, especially the young men, by utilizing a militaristic historical repertoire. This negative sentiment towards Arabs, whether they are Syrian or not, has led to a counter-nationalist trend (Erdoğan, 2022).

This study examines the historical repertoires utilized by Arabic speakers in inclusion and exclusion processes through generative AI-aided stance analysis and topic modeling. Pro and anti stances users might use same or different events when doing boundary work; thus, this study investigates whether they use the same or different historical repertoire when including or excluding Erdoğan.

**Stance Analysis via Generative AI**

In this study, tweets from accounts associated with the Turkish Twittersphere were filtered to include those related to Turkish President Erdoğan, assuming they would reveal attitudes towards Turkey's policies and stance, given Erdoğan's prominence as a political figure. These tweets were then subjected to stance analysis. Stance analysis examines the author's attitude towards a specific target in a given text (Küçük & Can, 2020). In this study, tweets related to Erdoğan were classified into three categories as positive, negative, and neutral. Because the search was based on keywords, tweets containing the word "Erdoğan" were assumed to be about President Erdoğan, and irrelevant tweets were not inspected as a fourth category.

Labeling necessary for stance analysis in this study was performed using a large language model (ChatGPT 4). Large language models, especially ChatGPT, have been the subject of various studies regarding their performance in stance analysis (Zhang, Ding & Jing, 2022; Cruickshank & Ng, 2024). Considering human-created gold-standard datasets, they can make the data annotation task more economical. Of course, any labeling done with an LLM cannot surpass expert-adjudicated annotated data conceptually, as this data is ultimately used as the fundamental measure of reality. Nevertheless, the performance of LLMs can surpass annotations obtained through crowdsourcing or by a single annotator.

Despite promising results, using LLMs for any task without evaluation data is risky. Firstly, the performance of LLMs and the biases they exhibit in specific areas can vary greatly over time and depending on the task. Therefore, measuring the model's success is crucial to demonstrate the quality of the analysis at hand. Moreover, if the aim is to train a classifier with LLM-annotated data, the shortcomings of LLM annotation can exponentially increase during model training. Indeed, in the experimental phase of this study, a BERT-based classifier developed using data annotated by an LLM yielded results below 0.5, so it was not used.

**Method**

**Data**

In these analyses, Twitter data from the Politus project was used. While creating this dataset, first, 100 prominent accounts in the Turkish twittersphere were selected, and 55 million follower IDs were obtained from the followers of these accounts. Among these, those who did not have a Turkish name or did not have a Turkish place name in the location/description fields were filtered out. As a result of this process, 3.5 million users remained. With subsequent additions to the dataset, there are currently 6,681,771 users and 718,867,236 tweets.

Since Arabic tweets extracted from this dataset, these users can be defined as Arab users who follow at least one person from Turkey. The number of Arabic tweets in this dataset is 1,688,253. Among these, Arabic texts shared by Turkish users, especially Quranic verses and prayers, were included. To filter these out and access only tweets posted by Arabs, diacritical texts used by Turks but not by Arabs were used as criteria. If the rate of using diacritical marks in the letters of a text was more than 25%, these texts were filtered out. As a result, 53,942 tweets were removed, leaving 1,634,311 tweets. To further remove tweets posted by media channels, the users who tweeted the most were manually filtered, leaving 1,589,802 tweets. In this dataset, tweets containing the word "Erdoğan" were searched to access tweets related to Recep Tayyip Erdoğan (n=11,426).

**Stance Annotation**

To be able to evaluate the quality of the prompts and the final output, I created a hand-coded dataset. To address the possibility of class imbalance and avoid spending too much time finding examples of underrepresented classes, I first annotated a sample of tweets using Chatgpt 4, continuing until I obtained 100 examples from each class. Then, I checked the annotations made by the LLM and hand-coded them. These tweets were used for evaluation purposes.

Zero-shot, one-shot, few-shot, and chain-of-thought prompting techniques were tested using ChatGPT 3.5 and 4 models. The most promising technique, that is few-shot chain-of-thought, was refined on the ChatGPT 4 model. Using the refined prompt, all tweets were annotated on a dashboard.

As a result of the stance annotation, 5784 tweets were classified as positive, 2015 as negative, and 3626 as neutral. Since the neutral tweets mostly consisted of news tweets, the tweets with positive and negative stances were separated for topic modeling.

**Topic Modeling**

During topic modeling, tweets with positive and negative stances were first subjected to preprocessing, where stopwords were removed. In this step, a comprehensive list of Arabic stopwords found on GitHub[[1]](#footnote-1) was used, along with a list of common words and prepositions that were considered likely to disrupt the analysis. Additionally, URLs, mentions, hashtags, and RT expressions were removed.

The resulting texts were then processed using the bert-base-arabertv02 model developed by the MIND Lab at the American University of Beirut (Antoun, Baly & Hajj, 2020). The number of topics was not predetermined, and the model was allowed to determine the optimal number of topics on its own. The resulting topics were named manually with the help of ChatGPT-4 by selecting 15 random tweets from each topic.

**Results**

After trials and refinements, the final prompt achieved an F1 Macro score of 0.88 on manually coded evaluation data (see Table 1). It is observed that the model and prompt have high recall in identifying negative and positive stances. However, it is noted that they have relatively low recall (.65) for neutral tweets. The model shows very high precision in positive tweets, with precision scores over .9 for both positive and neutral tweets, but the precision drops to .73 for tweets with a negative stance.

**Table 1: Performance Scores of the LLM Stance Annotations**

|  |  |  |  |
| --- | --- | --- | --- |
| **Class** | **Precision** | **Recall** | **F1-Score** |
| Positive | 0.99 | 0.97 | 0.98 |
| Negative | 0.73 | 0.96 | 0.83 |
| Neutral | 0.93 | 0.65 | 0.76 |
| Average | 0.88 | 0.86 | 0.86 |

When we examine the confusion matrix to take a closer look at the model's performance in all three classes, it is seen that the model is successful in distinguishing between positive and negative tweets, but it tends to label neutral tweets as negative. This reduces the precision score of the model in the context of negative stance tweets (see Table 2).

**Table 2: Confusion Matrix of the LLM Stance Annotations**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Positive** | **Negative** | **Neutral** |
| **Positive** | 97 | 1 | 2 |
| **Negative** | 1 | 96 | 3 |
| **Neutral** | 0 | 35 | 65 |

The topics obtained from topic modeling on tweets with a positive and negative stance towards Erdoğan were selected if they had more than 20 tweets. From these, 15 random tweets were chosen and the topics named with the help of ChatGPT. Just these topics were included in the analysis (see tables 3 and 4). In tweets with a positive stance, topic 0 covers a larger portion of the tweets compared to other topics, while tweets with a negative stance showed a more balanced distribution.

**Table 3: Pozitif tweet topics**

|  |  |  |
| --- | --- | --- |
| **Topic No** | **Tweet Count** | **Topic Name** |
| **0** | 2491 | Support from Citizens for Erdoğan |
| **1** | 83 | Erdoğan's Strength and Power |
| **2** | 58 | Erdoğan's Visits to Arab Nations |
| **3** | 47 | Prayers for Erdoğan |
| **4** | 30 | Election Support for Erdoğan |
| **5** | 25 | Affection for Erdoğan |
| **6** | 24 | Outcomes of Elections |
| **7** | 20 | Discovery of Natural Resources |

Among the tweets with a negative stance, those mistakenly marked as negative due to relatively low precision for neutral stance tweets are particularly concentrated in topics 8 and 9. When examining the tweets gathered in these two topics and mistakenly labeled as negative stance, it is observed that they are mostly news texts. Although there are also neutral tweets among the negative stance tweets in other topics, they are quite few in these topics.

**Table 4: Negative tweet topics**

|  |  |  |
| --- | --- | --- |
| **Topic No** | **Tweet Count** | **Topic Name** |
| **0** | 638 | Critics of Erdoğan's Economic and Political Policies within Turkey |
| **1** | 222 | Anti-Democratic Practices during Elections |
| **2** | 116 | The Real Face of Erdoğan and Criticism of Supporting Arabs |
| **3** | 70 | Syria Policies |
| **4** | 58 | Military Operations at the Syria Border |
| **5** | 49 | Tensions with Egypt over Libya |
| **6** | 44 | Geopolitical Dynamics and Clashes |
| **7** | 40 | Public and Political Reactions to Erdoğan's Policies |
| **8** | 36 | Criticism of Incorrect Comments about Erdoğan |
| **9** | 33 | Turks and Their Relations with Other Nations |
| **10** | 30 | Turkish Activities in Libya |
| **11** | 22 | Erdoğan's Failed Policies |

**Discussion**

When analyzing the resulting topics from the perspective of "exclusion" and "symbolic borders," the first difficulty encountered is the possibility that the texts in question were posted by Turks who know Arabic, Arab citizens of Turkish origin, or propaganda accounts. Indeed, when examining the users, it is seen that there are accounts that fit these descriptions. But eventually, these accounts are addressing the Arab public, aiming to influence the "symbolic boundaries" and to shift the relevant boundary to desired direction, therefore operate within the Arab public opinion.

When tweets with a positive stance towards Erdoğan are examined on a topic basis, it is seen that the emphasis almost always revolves around Erdoğan's Muslim identity. Therefore, whenever Erdoğan is "included," the symbolic boundary is drawn through Islam, thus treating Erdoğan as an in-group member. Emphasizing Erdoğan's power and the strengthening of the country's economy through mentioning the discovery of natural resources also highlights Erdoğan not only as an in-group member but also as a competent in-group member, according to the stereotype content model (Cuddy, Fiske & Glick, 2008). Here, the historical repertoire emphasized is the Ottoman Empire as a strong Muslim political entity.

When it comes to tweets with a negative stance towards Erdoğan, it is seen that while the "inclusion" strategies are uniform, the "exclusion" strategies vary. Firstly, the emphasis on "power" that makes Erdoğan a strong in-group member here turns into the aggressiveness of an out-group member. The historical repertoire is again drawn from the Ottoman Empire, but it is described as an occupying, colonial state instead of a Muslim state. Erdoğan, as someone trying to revive the Ottoman Empire, appears as a competent out-group member.

Texts that mention Erdoğan as a competent out-group member, along with comments that belittle him, depict the image of an incompetent out-group member. These highlight Erdoğan’s failures in domestic and foreign policies. One of the main emphases in tweets with a negative stance is that Erdoğan’s real face is different from what people perceive. These serve as a critique of the "symbolic boundary" that makes Erdoğan an in-group member or the evaluations that portray him as competent. Examples include texts suggesting that Erdoğan is not a democrat but an oppressor. Another theme in tweets portraying Erdoğan as a competent, ill-intentioned out-group member is the supporters of Erdoğan in Arab countries. These individuals, often criticized as being members of the Muslim Brotherhood, are depicted as serving the out-group’s ambitions pointing to the memory of Ottoman Empire. The historical repertoire used here includes both the Ottoman Empire as a foreign, occupying force and the conflict between the Muslim Brotherhood and nation-states. One of the main topics addressed by referring to the Ottoman historical repertoire when discussing Erdoğan as a competent out-group member is Turkey's operations abroad, especially military ones. The conflicts in Syria and Libya and relations with Egypt and Saudi Arabia stand out in this context.

The Islamic identity used to draw the symbolic boundary to include Erdoğan in the in-group can also work against him. Indeed, among tweets on the "real face of Erdoğan," there are those that criticize Erdoğan and Turkey with Islamic arguments, mentioning secularism, LGBT, unreligious elements in Turkey, relations with Israel, and so on.

Overall, it is seen that the historical repertoire of the Ottoman Empire is referenced by both positive and negative stance tweet users in their efforts to draw a "symbolic boundary" around Erdoğan. Comments that draw the boundary to include Erdoğan portray him as a competent in-group member through his Muslim identity, while comments that draw the boundary to exclude him depict him as either a competent or incompetent ill-intentioned out-group member. Among negative stance comments, nationalist and religious tendencies are observed, while positive stance comments almost always have religious tendencies.

**Conclusion**

This study addresses the topics and historical references used in tweets about Erdoğan, who represents Turkish politicies as prominent figure of Turkish politics for the past 20 years, by Arabs in and around the Turkish Twittersphere during symbolic boundary work. While the Ottoman Empire serves as the primary historical reference for both positive and negative comments, Arab nationalist and religious interpretations of the Ottoman Empire differ from each other and within themselves. This study provides a successful example of using large language models for Arabic stance analysis and the BERT topic model for content analysis.

**Limitations and Future Work**

The first limitation is the issue of identifying neutral tweets. It has manifested itself through neutral tweets that are concentrated in specific topics. Since neutral tweets are generally composed of news tweets, distinguishing news texts first could make the study more successful. In this context, a model that differentiates news texts can be utilized, or this task can be handled through LLMs (Large Language Models).

In this study, only the results obtained by testing a single prompt on a single model are discussed. Using other models, especially open-source ones, apart from ChatGPT in this process can increase the reproducibility of the study. Moreover, using the ChatGPT dashboard instead of the API during the study can harm reproducibility.

**References**

Akturk, A. S. (2010). Arabs in Kemalist Turkish Historiography. *Middle Eastern Studies*, *46*(5), 633-653.

Antoun, W., Baly, F., & Hajj, H. (2020). Arabert: Transformer-based model for arabic language understanding. *arXiv preprint arXiv:2003.00104*.

Cruickshank, I. J., & Ng, L. H. X. (2024). Use of large language models for stance classification. *arXiv preprint arXiv:2309.13734*.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology*, 40, 61-149.

Erdoğan, M. (2022) *Syrians Barometer 2021. A Framework for Achieving Social Cohesion with Syrians in Turkiye*. Ankara: Eğiten Yayıncılık

Lamont, M., Pendergrass, S., & Pachucki, M. (2015). Symbolic boundaries. *International encyclopedia of social and behavioral sciences*, *2*, 850-855.

Nafi, B. M. (2009). The Arabs and modern Turkey: a century of changing perceptions. *Insight Turkey*, 63-82.

Özcan, M., Köse, T., & Karakoç, E. (2015). Assessments of Turkish foreign policy in the Middle East during the Arab Uprisings. *Turkish Studies*, *16*(2), 195-218.

Siviş, S. (2023). The use of cultural repertoires of everyday nationhood and citizenship in national identity boundary‐drawing: The case of Syrian refugees in Turkey. *The British Journal of Sociology*, *74*(4), 657-672.

Zhang, B., Ding, D., & Jing, L. (2022). How would stance detection techniques evolve after the launch of chatgpt?. *arXiv preprint arXiv:2212.14548*.

1. https://github.com/mohataher/arabic-stop-words/ [↑](#footnote-ref-1)