Kinship Terms: Intercultural Linguistic Markers of Teknonymy

Esra Abdelzaher
Department of English Linguistics
University of Debrecen
esra.abdelzaher@gmail.com

Bacem Essam
Department of Computer Science
Cairo University
literaryartrans@gmail.com

Abstract

This study proposes that teknonymy (i.e., the act of referring to someone by relating them to a kinship, dominantly a father or a mother) is more than an anthropological practice that varies across cultures. We argue that teknonymy and teknonymy-like are well-structured relations that take different patronymic, matronymic and charactonymic patterns in Semitic, Germanic, Slavic and Romance languages. We analyze the semantics, polarity and sociolinguistic aspects of kinship terms in 3K constructions in Arabic to build an automatic classifier that separates teknonyms (e.g. Abu Ahmed/father of Ahmed) from near teknonyms (Abu Alfasad/father of corruption) and sub-classify the usage of the kinship as patronym, matronym or charatonym, among others. We also provide a user-friendly web-based version of the most frequent 1k kinship constructions in Arabic (https://arabic-studies.com/TI/index.html). Our results reveal both universal and cultural-specific patterns in teknonymic structure and significant cross-linguistic variations, providing insights into the interface between language, culture, and cognition and implications for including teknonymic structures in multilingual wordnets.

1 Introduction

Teknonyms, a type of naming convention that begins with a kinship term such as *father*, *mother* or *son*, are prevalent in many cultures but vary in form and function across languages. Classical Arabic, for instance, incorporated teknonyms into its metaphorical framework, allowing things to be referred to without using their original given names (Alenizi, 2019; Ebraheme, 2016; Salamh et al., 2022). These designations encompassed the anthroponymic (e.g. *Abu Lahab* for a man named *Abd Al-Ozza*), zoological (e.g. *Abu Al-Harith* [lions], *Umm Amer* [hyenas]), topographical (e.g. *Umm Al-Qura* [Mecca]), and conceptual (e.g. *Umm Qashaam*, epitomizing warfare) (Almuhanna, 2023) names. Contemporary Arabic lexicographers prioritize the consideration of teknonyms as a cornerstone of the Arabic naming system over consideration of their generic metaphorical usage. Arabic nomenclature incorporates proper nouns, teknonyms, and epithets (Ebraheme, 2016).

Old English also used forms of teknonyms in the naming system, such as "Fitzwilliam" or "Fitzwilliam" (son of William), to denote lineage, evolving in Middle English into the form of "Williamson" (Tait, 2006). In the Renaissance era (1500–1660), English borrowed teknonyms from Gaelic; for example, Macbeth, stemming from Mac Bethad (son of life) (Davis, 2012), which is similar to the contemporary usage of McDonald or McMahon in American English, and to O'Connor or O'Brien in Irish English (Tait, 2006; Tucker, 2006). In modern British and American English, informal parental sobriquets are employed to reference offspring (e.g. Johnny's mum, Susie's dad). However, they are not frequently used as addressing terms. The use of kinship terms in teknonym-like constructions, such as "father of modern Philosophy" is currently prevalent in English and other languages. A sample of Arabic and English teknonyms is available in Appendix 1.

We argue that teknonymy and near-teknonymy are integral to using kinship terms in different languages. The scope of the current study is limited to the Arabic language. Our study aims at:

(a) proposing an annotation schema to separate teknonyms and near-teknonyms from each other and from standard uses of kinship terms in Arabic

(b) exploring the possibility of automatically classifying the uses of kinship terms in

teknonyms, near-teknonyms and non-teknonyms

(c) introducing a browsable database of the most frequent uses of Arabic kinship terms

2 Teknonymy: A multilingual kinship-based relation

The use of kinship terms in addressing and naming constructions has been studied in different Semitic, Slavic and Romance languages. Teknonyms exhibit remarkable universality and diversity across languages and cultures, as a fundamental linguistic feature identifying individuals through their familial roles. This concept manifests in various forms, from the complex Korean system of address (incorporating kinship terms, honorifics, and pronouns, with teknonyms often using birth-order designations among parents) to the extensive use in Arabic naming conventions (employing "Abu" for father and "Umm" for mother, alongside patronymics and epithets). Other languages demonstrate related practices, such as Icelandic (using patronymic and sometimes matronymic naming), Spanish (incorporating both paternal and maternal surnames), and even constructed languages like Klingon (deliberately including patronymic elements). Semitic languages similar to Arabic also utilize teknonyms or related concepts, as seen in Hebrew (using "Avi" for father and "Em" for mother, often as given names), Aramaic (employing "Abba" for father and "Imma" for mother in teknonym formation), and Amharic (using a patronymic system where a person's name is followed by their father's name). Maltese, a Semitic language with significant Romance influence, offers an interesting example of teknonymic evolution. While it retains some Arabic-origin patronymic forms, these have primarily transitioned into surnames rather than active teknonyms. For instance, surnames like "Bencini" (from "Bin Cini," meaning "son of Cini") demonstrate the historical use of teknonyms in Maltese culture, even as their current function has shifted.

In the present study, we differentiate between teknonyms and near-teknonyms based on the original function of the relation in Arabic (i.e. further specification of the referent). Therefore, constructions like Ibn Misr/Son of Egypt which can be generically used to refer to anyone born in Egypt or of Egyptian origin will be considered teknonym-like, mainly if they have frequently used cross-lingual equivalents (e.g., son of Rome, Son of England).

3 Creating Arabic teknonymy dataset

3.1 Data compilation

We collected Arabic teknonyms from classical and contemporary resources, whereas other constructions embedding kinship terms were retrieved from corpora and databases using lists of Arabic kinship terms. The collection of classical Arabic teknonyms was the easiest because there are multiple classical biographical lexica, and dictionaries which are sorted according to the uses of patronyms and matronyms (e.g. the book of the proper names of people known by their teknonyms, the book of poets' matronym teknonyms). For Modern Standard Arabic (MSA), we bootstrapped constructions starting with kinship terms in contemporary Arabic dictionaries, newspapers, and literary works such as ArTenTen (a web-crawled Arabic corpus available through Sketch Engine). For contemporary Arabic dialects, sources include dialect-specific crowdsourced dictionaries (e.g. http://ar.mo3jam.com), social media corpora (e.g. Refaee & Rieser, 2014, Essam et al., 2019; Essam & Abdo, 2021), and databases (e.g. Bouamor, Habash, et al., 2019; Bouamor, Hassan, et al., 2019). The raw lists of teknonyms and other kinships, including constructions, included more than 7K constructions. After removing the duplicates and false positive constructions (e.g. words sharing the same form with one of the kinship terms), the list included 4K constructions, which were further classified manually during the annotation process. A sample of teknonymic patterns, their frequency and a cross-linguistic reference to English examples are presented in Appendix 2.

3.2 Linguistic annotation

Each construction was annotated by 3 native speakers of Egyptian, Tunisian and Saudi Arabic. The annotation schema included the number of words in the construction, the identification of the kinship terms and their number, the detection of their literal or metaphoric use, the identification of the proper and common nouns, the selection of a literal or metaphoric use of the common word, the identification of the singularity, plurality and gender of the referent, the documentation of the polarity of the construction, its frequent association

with a specific dialect, the patronymic, matronymic or charactonymic use of the kinship term and, finally, its label as a teknonym, near-teknonym or non-teknonym. A sample of the annotated dataset is viewable as supplementary material. Our annotation schema showed variation in the distribution of kinship patterns in the categories of teknonyms and near-teknonyms. Table 1 shows a sample of the patterns associated with teknonyms (Tek.) and near-teknonyms (N.Tek) in our dataset.

Table 1. Dichotomy of teknonymic constructions

Pattern	Tek.	N.Tek
Kinship term (literal) + proper noun	√	×
Kinship term (literal) + definite common noun	×	√
Kinship term (metaphoric) + proper noun	×	√
Kinship term(metaphoric) + indefinite common noun	√ less frequent	√
Kinship term + definite common noun + definite common noun	×	√
Kinship term (literal) + indefinite common noun	×	√
Kinship term (metaphoric) + definite common noun	√	√
Kinship term (literal) + kinship term (literal) + proper noun	√	×
Kinship term (literal) + kinship term (metaphoric) + definite common noun	√	√ less frequent

Abbreviations: **N.Tek**: Near-teknonym; **Tek**: Teknonym

3.3 Automatic classification of teknonyms

We trained a classifier to predict the type of construction and the use of the kinship term in the construction. We used cross-validation to train a classifier to predict the type of the construction as teknonym, near-teknonym or non-teknonym. The results of the classification task are reported in Table 2. Whereas the Area Under Curve (AUC) is comparable for the three classifiers, the Classification Accuracy (CA) was highest for the Random Forest (RF) algorithm, followed by Logistic Regression (LR) and Naïve Bayes (NB). The Precision (P) and Recall (R) were also the highest for the RF classifier, which was also associated with the highest Matthews Correlation Coefficient (MCC).

Table 2. Results of the classification task

	AUC	CA	F1	P	R	MCC
RF	0.99	0.97	0.97	0.97	0.97	0.94
LR	0.99	0.94	0.94	0.95	0.94	0.86
NB	0.98	0.94	0.94	0.94	0.94	0.86

Unlike the successful prediction of the teknonymy labels, the classifier was far less successful in predicting the use of the kinship term. Table 3 shows the average results of the automatic classification of the kinship term across 12 classes.

Table 3. Mean values for classifying kinship terms

	AUC	CA	F1	P	R	MCC
RF	0.94	0.86	0.83	0.81	0.86	0.80
LR	0.94	0.85	0.81	0.80	0.85	0.78
NB	0.92	0.74	0.76	0.80	0.74	0.65

The misclassified cases were mostly associated with the female gender of the referent (OffspringF) which ended up in male counterpart category in 99% of the cases. Also 48% of the matronymic uses appeared in the patronym category. Figure 1 shows the scatter plot of the misclassified cases.
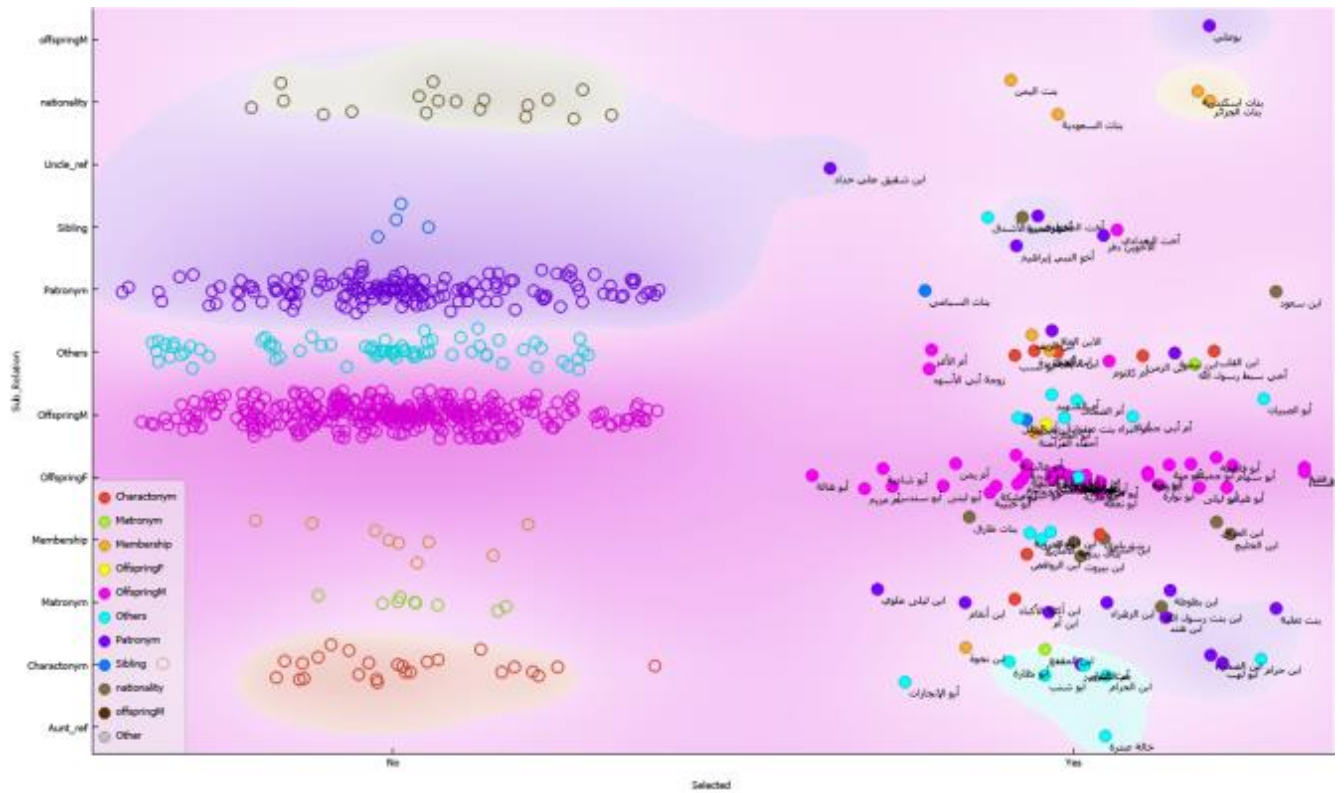
Figure 1. Scatter plot of the misclassified cases

4 Browsable database of teknonyms

We developed a web-based searchable version of the annotated dataset. The web version currently contains the most frequent 1K constructions embedding kinship terms.

The database can be browsed using any Arabic kinship term. The results retrieve the possible constructions in which this term can be used and recall the annotated semantic, pragmatic and sociolinguistic information recorded for the construction in the dataset. For each restored construction, the retrieved information includes the kinship term, which is the head of the construction and the total number of kinship terms in the construction for cases like *the son of the mother of*

x. Based on identifying the head kinship term, further information about the gender and countability of the referent are provided. Including a proper noun, such as the father of Ahmed or a common noun, such as the father of virtue, is also clarified in the recalled information. Additional information about the category of the construction as a teknonym, near-teknonym or non-teknonym is also provided in the user-friendly interface, as well as the relation holding between the kinship term and the rest of the words in the construction, e.g. patronymy, matronymy, offspring in cases of the literal use of the kinship term or nationality, membership or charactonym in cases of the metaphoric uses of the head kinship term. Figure 2 shows the interface of the web version of the database.

Figure 2. Searchable Web-based Interface

5 Conclusion

We proposed an annotation schema to separate teknonyms and near-teknonyms from each other and from standard uses of kinship terms in Arabic. Our schema depended on identifying the head kinship term and its literal or metaphoric use, classifying the following noun into proper or common, and deciding whether it was used literally or metaphorically, too. We included other corpus-based information about the association of the construction with a specific dialect or sentimentality. We also explored the possibility of the automatic classification of the uses of kinship terms into three broad categories (i.e. teknonyms, near-teknonyms and non-teknonyms) and more specific classes (e.g. matronym, charactonym, membership). Our results showed promising results for most classification algorithms at the broad level, but the classification accuracy significantly dropped for the sub-classification. The gender factor appeared to be the most influential in the misclassified cases as most of the matronym patterns were classified as patronym, and several offspring_female cases were misplaced in the offspring_male.

Finally, we proposed a browsable database of Arabic teknonyms, which is a valuable resource for both linguistic research and cultural studies, potentially inspiring similar projects for other languages to enable the analysis of comparable constructions in different languages. Teknonyms demonstrate linguistic universality while maintaining cultural specificity, suggesting a common human tendency to identify individuals through kinship terms that can adapt and persist even as their primary function changes. The successful automatic inter/intra-language classification of teknonyms, near-teknonyms, and non-teknonyms can advance syntagmatic relationships in NLP for identifying and categorizing kinship-based naming conventions.

Acknowledgements

References

Michael Adams. 2013. *From Elvish to Klingon: Exploring invented languages*. Oxford: Oxford University Press

Reima Al-Jarf. 2023. The Interchange of Personal Names in Muslim Communities: An Onomastic Study. *Journal of Gender, Culture and Society*, 3(1):42–56. https://doi.org/10.32996/jgcs.2023.3.1.5

Aied Alenizi. 2019. The Norms of Address Terms in Arabic: The Case of Saudi Speech Community.

International Journal of English Linguistics, 9(5). https://doi.org/10.5539/ijel.v9n5p227

Amin Almuhanna. 2023. On the Origin of Kuwaiti Surnames. *Al-'Arabiyya*, 55–56:175–221.

Houda Bouamor, Nizar Habash, Mohammad Salameh, Wajdi Zaghouani, Owen Rambow, Dana Abdulrahim, Ossama Obeid, Salam Khalifa, Fadhl Eryani, Alexander Erdmann, and Kemal Oflazer. 2019. The madar Arabic dialect corpus and lexicon. *In LREC 2018 - 11th International Conference on Language Resources and Evaluation*.

Lucien Brown and John Whitman. 2015. Honorifics and politeness in Korean. *Korean Linguistics*, 17(2). https://doi.org/10.1075/kl.17.2.001int

Frans Ciappara. 2010. Religion, kinship and godparenthood as elements of social cohesion in Qrendi, a late-eighteenth-century Maltese parish. *Continuity and Change*, 25(1). https://doi.org/10.1017/S0268416010000019

Madison Davis. 2012. *The Shakespeare Name and Place Dictionary*. Routledge

Kenneth Dimaggio. 2016. From Egypt to the Arizona desert to places still to come: The ongoing meta-literary journey of eliza's escape to freedom in uncle Tom's Cabin. *International Journal of Literary Humanities*, 14(3):41–49.

Barahmi Ebraheme. 2016. The semiotics of the proper nouns: The case of teknonyms (in Arabic). *Dyrassat*.

Miseon Lee, Sorin Huh, and William O'Grady. 2017. Korean subject honorifics: An experimental study. *Journal of Pragmatics*, 117. https://doi.org/10.1016/j.pragma.2017.06.001

Eshrag Ali Refaee. 2022. Detecting Hadith *Authenticity Using a Deep-learning Approach. Scientific Journal of King Faisal University Basic and Applied Sciences*, 23(1):80–84.

Eshrag Refaee and Verena Rieser. 2014. Evaluating Distant Supervision for Subjectivity and Sentiment Analysis on Arabic Twitter Feeds. *In ANLP 2014 - EMNLP 2014 Workshop on Arabic Natural Language Processing, Proceedings*, pages 174–179. https://doi.org/10.3115/v1/W14-3624

Sami Ben Salamh, Zouheir Maalej, and Mohammed Alghbban. 2022. To be or not to be your son's father/mother. *Pragmatics. Quarterly Publication of the International Pragmatics Association* (IPrA). https://doi.org/10.1075/prag.00001.ben

Clodagh Tait. 2006. Namesakes and nicknames: Naming practices in early modern Ireland, 1540-1700. *Continuity and Change*, 21(2):313–340. https://doi.org/10.1017/S0268416006005935

Kenneth Tucker. 2006. A comparison of Irish Surnames in the United States with those of Eire. *A Journal of Onomastics* (54)1: 55-75 https://doi.org/10.1179/nam.2006.54.1.55

A Appendices

Appendix 1. Sample of Arabic and English Teknonyms

Teknonym	Variety	Meaning	Source
Abu Abed	Syrian Arabic	Father of Abed	Dialect database
Abu Adnan	Gulf Arabic	Father of Adnan	Dialect database
Abu Alaa'	Egyptian Arabic	Father of Alaa'	Social media corpus
Abu Al-Duhhak	Classical Arabic	Father of the beaming person	Biographical lexica
Abu Al-Kasim	Modern Standard Arabic	Father of Kasim	Literary texts
Addison	Middle English	Adam's son	Biographical lexica
Son of King Charles III	Contemporary British English	Son of King Charles III	General reference corpora
The son of Marcus Cato	Early Modern English	The son of Marcus Cato (Cato the Elder)	Literary corpus

Appendix 2. A cross-linguistic sample of teknonymic patterns

Teknonym pattern	Example	Trend	
Patronymic teknonym	Ibn/ben Mousa Williamson	Maintained (Ar)	Diminished (Br En)
Matronymic teknonym	Ibn Angham Mollison	Increasing (Ar)	Diminished (Br En)
Father of x(daughter)	Father of Helen Keller Abu Fatima	Increasing (Ar)	Decreasing (Br En)
Father of x(son)	Abu Ahmed Father of Boris	Maintained (Ar)	Decreasing (Br En)
Three- to four-word teknonyms	Ibn Abu Al-No'man Mother of Queen Elizabeth	Decreasing (Ar)	Decreasing (Br En)
Charactonymic teknonym	Um Al-Noor Grayson	Maintained (Ar)	Diminished (Br En)

B Supplementary Material

The most frequent usages of kinship terms in Arabic, according to our dataset, are accessible at https://arabic-studies.com/TI/index.html .