An interdisciplinary corpus-based analysis of the translation of كرامة (karāma, 'dignity') and its collocates in Arabic-English constitutions

Claire Brierley, University of Leeds
Hanem El-Farahaty, University of Leeds/University of Mansoura, Egypt

ABSTRACT

We examine whether occurrences of the human rights term كرامة (*karāma*, 'dignity') in the Leeds Parallel Corpus of Arabic-English Constitutions imply a shared understanding of this term from source to target text. Our approach combines quantitative and qualitative techniques from corpus linguistics and Arabic legal translation and contributes to theory and practice in computer-assisted legal linguistics and translation. Our methodology includes: specification of morphological variants of كرامة; scrutiny of parallel concordance lines; and analysis of the semantic prosody of target terms via their collocations. We identify 65 instances (or variants) of كرامة in the Arabic data: its raw frequency is highest in the constitutions of Egypt and Sudan but missing in that of Palestine. We find that while the indefinite noun كرامة is always translated as 'dignity', the definite form (الكرامة, *al-karāma*) is often rendered as 'treatment' plus a qualifying adjective. The combination of كرامة and negation results in qualification of 'treatment' with notions of humiliation and cruelty, as evidenced via collocation discovery over both sub-corpora of 19 constitutions. This suggests a common understanding of كرامة and dignity as an inviolable human right across these different languages and cultures, fostered perhaps by the theological significance of these terms.

KEYWORDS

Arabic translation, legal translation, human rights, 'كرامة' (*karāma*, 'dignity'), constitution, corpus linguistics, collocation.

1.   Introduction

Since the *Universal Declaration of Human Rights* in 1948, the concept of 'dignity' (Arabic counterpart كرامة, *karāma*) has become "the chosen underlying principle and source of rights in international human rights law and many domestic constitutions" (O'Mahoney 2012: 551-552; cf. Neuman 2003). More recently, it is identified and further defined as the primary fundamental right protected by the *European Charter of Fundamental Rights* in 2000. In both the aforementioned documents of the United Nations and the European Union, 'dignity' is construed as an inherent right of every human being: we are all "born free and equal in dignity and rights" and hence "inviolable" (see Article 1 of the former, and Chapter 1, Article 1 of the latter, respectively).

The dignity of humankind is also a longstanding theological concept in the Qurʾān as well as the Bible. Such shared religious connotations may have facilitated what Biel (2015: 140) refers to as the "cross-linguistic negotiation of [the] meaning" of 'dignity' in international human rights. Evidence of the association of universally-held modern values surrounding human rights and religious ideas has been noted by Edzard (1996: 54),

who observes that the stylistics of modern diplomatic Arabic persistently reflects a historical religious conscience.

In this paper, we present an interdisciplinary study combining tools and techniques from corpus linguistics and Arabic legal translation to ascertain whether there is a shared understanding of the concept of 'dignity' in translated legal language from Arabic as source language to English as target language. To address this research question, we examine how the deeply embedded, lexicalised (and Quranic) concept of كرامة (*karāma*, 'dignity') in Arab constitutions is translated into English and understood in the recipient vocabulary of human rights. To achieve this aim, our study investigates instances, variants, and translations of كرامة (*karāma*) in the *Leeds Parallel Corpus of Arabic-English Constitutions*. Our approach contributes to theory and practice in the emerging field of computer-assisted legal linguistics and translation. Our research methodology includes: specification of morphological variants of كرامة in the Arabic corpus, close scrutiny of parallel concordance lines, and collocation discovery.

The dataset used for the study is the *Leeds Parallel Corpus of Arabic-English Constitutions* from nineteen states across the Arab world, namely: Algeria, Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, Palestine, Qatar, Saudi Arabia, Sudan, Syria, Tunisia, United Arab Emirates, and Yemen. The most recent, ratified version of the constitution is used in each case, the earliest dating back to 1992 (Kuwait), and the most recent from 2016 (Algeria). The corpus (comprising 169,861 Arabic words and 205,893 English words) has been uploaded into Sketch Engine (Kilgarriff *et al*. 2014) and this is the toolkit used in our empirical study.

The paper is structured as follows. In Section 2 we review research in corpus-based legal linguistics and translation and discuss compilation of the *Leeds Parallel Corpus of Arabic-English Constitutions*. The denotation of 'dignity' is then discussed in relation to international human rights law in Section 3, and the Bible and the Qurʾān in Section 4. Section 5 focuses on linguistic specification of all derived forms of the Arabic root ك ر م (*k-r-m*) as top-level node for the semantic field of 'dignity' prior to corpus exploration. Our presentation and discussion of results in Section 6 includes: (i) counting raw frequencies, (ii) inspecting Arabic and English concordance lines for terms and their translations as they appear in context in the corpus, and (iii) significance testing and analysis of collocations pertaining to key terms via metrics implemented in Sketch Engine. We present our conclusions in Section 7.

2.   Corpus-based legal linguistics and legal translation

The analysis of legal language using techniques from corpus linguistics over electronic corpora is an important recent development in the field of

law. In legal proceedings, a corpus-based approach to disambiguating meaning has been recommended over use of the dictionary as an interpretative tool (Mouritsen 2011; 2010); notwithstanding, corpus-based lexicography is now standard practice (Buckwalter and Parkinson 2013)[1].

Also, since legal data is largely text-based, legal scholarship in the United States is now attuned to the potential for big data Text Analytics over vast quantities of legal documentation, following the successful deployment of Text Analytics within the business and medical industries (Fagan 2016). This trend is reflected in Europe. For example, projects associated with the interdisciplinary research group in Computer Assisted Legal Linguistics (CAL2) include: machine learning for argumentation mining, and automated detection of stylistic violations in legislative drafting in German legal texts. Researchers are also keen to analyse: legal interpretations, legal discourse, the comprehensibility of legal texts, conflict resolution, and linguistic human rights (Vogel et al. 2017; Biel 2015; Williams and Milizia 2008).

Research on legal translation has focused more on equivalence, accuracy (Alcaraz Varó and Hughes 2002), genre analysis (Bhatia 1993) and on "terminological incongruity" (Biel and Engberg, 2013: 3). Researchers have been applying corpus-based and corpus-driven methods to translation studies since the 1990s (Biel 2010) and corpus-based research on legal translation between different European languages (e.g. Dutch, Spanish, Polish, Italian) has recently been undertaken (Biel and Engberg, 2013; Biel *et al*. (eds) 2019).

Within Arabic-English legal translation, not a lot of research has been undertaken with almost no research on corpus-based Arabic legal translation. El-Farahaty (2015; 2016) discusses translation accuracy and the asymmetry between Arabic and English legal discourse in terms of linguistic features (e.g. modal auxiliaries, passive structures and gender-specific terms) plus culture-specific and system-based terms (e.g. Islamic law terms, abstract terms). More specifically, El-Farahaty (2015) scrutinises the translation of modal auxiliaries and passive structures in a small corpus of key UN and other international documents as well as legislative and official documents. One of the major findings in this study is that translating modal auxiliaries does not follow well-defined semantic and syntactic rules. Translation of Islamic legal texts has been the focus of Alwazna's research over the past five years (2013a; 2013b; 2016; 2017). He has discussed the problems of translating Islamic legal texts into English, the translation of Islamic legal terminology and the impact of Islamic legal culture on the comprehension of Islamic statements.

While accuracy is still an important aspect of legal translation, researchers such as Šarčević (1997) and Engberg (2002) have, according to Biel (2010: 7), started to focus more on target texts (TTs) as "independent

texts on their own, emphasizing the importance of translated texts in receiving cultures." Likewise, researchers in legal translation stress the need for more interdisciplinary research and less fragmented research efforts across languages and cultures "to obtain a multidimensional view of legal translation" (Biel and Engberg 2013: 2; Engberg 2013). The application of corpora in legal translation and the importance of parallel corpora in shifting from "prescription to description" (Baker 1995: 231) are much needed to obtain informed observations.

2.1 Translation of transnational legal terms

In this paper, we focus on the fundamental human rights concept of كرامة (*karāma*, 'dignity') as realised in legislative language in a parallel corpus of Arabic-English constitutions. There is widespread consensus that dignity and human rights principally emerged as legislative concepts in the *Charter of the United Nations* of 1945 (see Section 4 for a more detailed discussion). As such, the meaning of these transnational legal terms is a construct of cross-linguistic negotiation and "translator-mediated communicative events" (Biel 2015: 159; Baker 1993: 243). This is further complicated by the proliferation of dignity discourse and 'loose usage' of the term which has tended to blur the boundaries between this and other human rights concepts (O'Mahoney 2012: 551). The inherent difficulties of legal translation from English into Arabic (and vice-versa) arising from systemic and cultural legal diversity, including incongruence of legal terminology, are exacerbated by linguistic asymmetry between these languages (El-Farahaty 2016: 474).

2.2   Plain meaning and legal meaning

The sense of a word is context-sensitive, and meanings consistent with legislative goals differ from everyday usage. Nevertheless, the words of a statute serve a dual function: "as guides in the attribution of general purpose" and "as factors limiting particular meanings" (Hart and Sacks 1994 in Mouritsen 2011: 169)[2]. Elsewhere, the judiciary has introduced the concept of "objectified intent," where the meaning (intent) of a statute is to be found "in the understanding of the objectively reasonable person" (Hart and Sacks 1994 in Mouritsen 2011: 174). When discussing the readability of EU law, Williams and Milizia (2008: 2226-2227) go further, citing Hassen Ebrahim (Executive Director of the South African Constitutional Assembly): "constitutions are about basic values affecting society and should be understood by even the least educated." Evidence that 'dignity' (and its absence) means the same the world over, in *deed* as well as word, is brought home in a recent broadcast about the White Helmets in Syria: "Where is the world, man? Where are the Arabs? Nobody cares about anybody anymore. All dignity is dead." (BBC4 2017).

Abstract terms are crucial in international law, so in addition to distinguishing between legal and non-legal or general meaning (Garre

1999: 116), it is important to consider the implications of these differences on translation (El-Farahaty 2016: 481-482). In spite of the indeterminacy of some of these abstract terms such as 'sufficient cause' (Engberg and Heller 2008: 146), and the vagueness of abstract English legal terms such as 'fair and reasonable' and 'justice' (Cao 2007: 19), and despite the many interpretations abstract terms will have in the legal arena, abstract international human rights terms are directly translated into the target language with no attempt to disambiguate them (El-Farahaty 2016: 491).

2.3 A parallel corpus of Arabic-English constitutions

The dataset used in this study is the *Leeds Parallel Corpus of Arabic-English* Constitutions[3]. This is currently a raw text corpus of 169,861 Arabic words and 205,893 English words compiled from reputable websites such as the World Intellectual Property Organisation and CONSTITUTE. The corpus comprises the most recent versions of the constitutions of 19 Arab states, and includes a sub-corpus of Preambles (6059 Arabic words and 7396 English words). The 19 states are: Algeria, Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, Palestine, Qatar, Saudi Arabia, Sudan, Syria, Tunisia, United Arab Emirates, and Yemen. The constitutions of Somalia and Djibouti, and also the constitution of Comoros, were not included since there is no Arabic version of the former, and no English version of the latter.

The corpus has been uploaded into the Sketch Engine corpus management and query toolkit (Kilgarriff *et al.* 2014) but is not yet open-source. Metadata for each corpus file includes attributes such as country, date, and source URL, and their corresponding values (e.g. Algeria; 2016; www.joradp.dz/har/consti.htm). The most challenging task prior to upload was sentence-level alignment of Arabic-English data. This alignment was directional from the original Arabic source text to the English translation, and was first attempted automatically via the SDL Trados WinAlign translation tool (2014 version). However, it was found that manual intervention was necessary within the WinAlign Workbench interface due to the mismatch in sentence length between Arabic and English. The result is correspondence on a one-to-many basis, since Arabic sentences differ from English in length and punctuation. The sub-corpus of Preambles is only aligned at paragraph level.

3.   Human   dignity   and   human   rights: the   international perspective

In this section, we examine the concept of 'dignity' in its secular international context to provide important background knowledge of the term and its transnational meanings which will then inform the investigation of this concept in our parallel Arabic-English corpus. The

term 'human dignity' denotes a central concept in "the transnational vocabulary of constitutionalism and human rights" (Jackson 2004: 15)[4]. Its emergence in the *Charter of the United Nations* in 1945, less than two months after the end of World War II, and in the *Universal Declaration of Human Rights* in 1948 represented the international community's determination to safeguard succeeding generations from the "barbarous acts" of war crimes committed in "disregard and contempt for human rights" that "outraged the conscience of mankind (*Preamble* to the *Declaration* 1948). It also represented a re-affirmation of faith in humanity and in "the dignity and worth of the human person" (*UN Charter* 1945). Article 1 of the *Declaration* provides that "all human beings are born free and equal in dignity and rights," thus setting up associations between several concepts: freedom, equality, dignity, and rights as the "birthright" of every human being. The logic of argument put forward in the *Charter* and *Declaration* whereby the "inherent dignity of the human person" guarantees the "common people" (*cf. Preamble* to the *Declaration* 1945) "equal and inalienable rights" constituting "the foundation of freedom, justice and peace in the world," is repeated in successive international conventions from the General Assembly on: *Elimination of All Forms of Racial Discrimination* (1966a); *Civil and Political Rights* and then *Economic, Social and Cultural Rights* (1966b, 1966c); *Elimination of All Forms of Discrimination against Women* (1979); *Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment* (1984); *Rights of the Child* (1989). Finally, the *Vienna Declaration* (1993) re-affirms the principle of intrinsic human dignity and the need for protection and fuller observance of this principle in the international human rights system. Afshari (1994: 248) makes the point that the *Declaration* of 1948 and successive covenants "define what is needed to protect a life of dignity and equality in a modern state.

While the unifying concept/value of 'human dignity' is fundamental to international human rights law, it is also open to a variety of interpretations. Critics such as O'Mahoney (2012: 565) argue that "loose usage" of the term has resulted in inconsistent application of dignity safeguards in domestic constitutional law. However, the European Union (EU) has sought a normative legal definition. The *European Charter of Fundamental Rights* (2000) identifies six overarching principles protected by the EU, where 'dignity' appears first and foremost. These are: dignity, freedoms, equality, solidarity, citizens' rights, and justice. It situates dignity (and each of the other rights) within constitutional traditions, international obligations, and importantly, European case-law. The EU definition of dignity as a generic term subsumes human dignity as an "inviolable" human right (Article 1), and also specifies "positive" rights (i.e. the right to life; and the right to the integrity of the person); and "negative" rights (i.e. the prohibition of torture and inhuman or degrading treatment or punishment; and the prohibition of slavery and forced labour). This definition differentiates (and could be used to differentiate) the jurisdiction of dignity from the jurisdiction of other generic

principles/values, namely: freedoms (e.g. freedom of thought, conscience and religion); equality (e.g. non-discrimination); solidarity (e.g. fair and just working conditions); citizens' rights (e.g. freedom of movement and residence); and justice (e.g. presumption of innocence and right of defence).

4.   Human dignity as a theological concept

It has been argued that human dignity is an entirely secular, "judge-made" concept (Dupré 2011) and as such very modern, originating in the *Charter of the United Nations* (1945) and the *Universal Declaration of Human Rights* (1948). However, we concur with Hallaq (2014: 28) and Schmitt (1985: 36) that human dignity: "[like] all significant concepts of the modern theory of the state [is a] secularised *theological* concept [our italics]," and posit that its deeply-embedded theological significance has facilitated the cross-linguistic understanding and cross-cultural acceptance of dignity as an inviolable human right. The association of religious and secular values in a Christian and European context is evidenced in the EU *Charter* itself: "conscious of its *spiritual* and moral heritage, the Union is founded on the indivisible, universal values of human dignity, freedom, equality and solidarity" [our emphasis]. Similarly, the equivalent term for human dignity in modern diplomatic Arabic (كرامة, *karāma*) is a deep-rooted Islamic concept and part of the Quranic vocabulary. In the rest of this section, we consider the significance of dignity in religious texts.

4.1 Human dignity and the Bible

A search of the *King James Bible Online* (conducted on 07.01.19 at: https://www.kingjamesbibleonline.org/search.php?q=dignity&bsec=Z&order=0) uncovers 10 specific instances of the word 'dignity', all appearing in the Hebrew Scriptures; these include celebrations of the bond between father and son (Genesis 49: 3; Ecclesiasticus 10: 28). However, the idea of dignity as intrinsic to human nature and the human person, compelling ethical treatment of fellow human beings, is a common theme throughout, from Genesis to the Pauline epistles. Examples include: "what does the Lord require of you but to do justice, and to love kindness" (Micah 6: 8); "whatever you wish that others would do to you, do also to them" (Matthew 7: 12); "love one another" (John 15: 17); "There is neither Jew nor Greek, there is neither slave nor free, there is no male and female, for you are all one in Christ Jesus" (Galatians 3: 28). Pin (2017: 57-61) traces the transformation of 'dignity' from its association with social status to the prevailing concept of inherent human worth, attributing this in part to the prominent role of Catholicism in the development of modern Christian thought. In the *Catechism of the Catholic Church* (Libreria Editrice Vaticana 1993), one manifestation and implication (responsibility) of belief in God is: "knowing the unity and true dignity of all men" (Part I, Section II: *The Creeds*, Article 1.225).

4.2 Human dignity and Islam

Searching the Qurʾān online via the dictionary tool in the Quranic Arabic Corpus (Dukes 2014), we find 47 occurrences of the trilateral root ك ر م (k-r-m) in 8 derived forms. One of the most pertinent to our discussion of human dignity is the Form II verb كَرَّمَ (*karrama*) meaning 'to honour' or 'to dignify'. As with all Form II verbs, this intensifies the basic, infinitival form (Form I) by doubling the middle radical (i.e. the middle letter in the triliteral/triconsonantal Arabic root). An example from the text of the Qurʾān is: وَلَقَدْ كَرَّمْنَا بَنِي آدَمَ (*wa laqad karramnā banī ādama*, 'we have honoured the children of Adam' Q.17.70). We also note the passive participle مُكرَمونَ (*mukramūn, 'honoured'*) associated with the Form IV verb (i.e. a causative or transitive form of the verb requiring an object) أَكْرَمَ (*akrama*), as in: أُولَئِكَ فِي جَنَّات مُكْرَمُونَ (*ʾulāʾika fī jannātin mukramūn*, 'They [will be] honoured in the gardens [of Paradise]' Q.70.35), where to be honoured thus is the reward for keeping one's word, standing firm in testimony, and sincere and regular prayer. These behaviours may be said to exemplify the exercise of human dignity in the Qurʾān. They are also reminiscent of the second part of Article 1 in the *Universal Declaration of Human Rights*, namely: that all human beings 'are endowed with reason and conscience and should act towards one another in a spirit of brotherhood'. This statement is imbued with religious as well as humanitarian values, and has universal appeal for all faiths. For Islam, the appeal to human reason or intellect (العقل البشري, *al-ʿaql al-basharī*) is particularly influential: the root ع ق ل (ʿ-q-l) itself represents a major Quranic concept with the overarching meaning of being endowed with the faculty of reason, and occurs 49 times in the *Qurʾān* as the Form I verb عَقَلَ (*ʿaqala*) 'to reason' or 'to understand'.

We observe that in formulating their constitutions, Arab states map the concept of human dignity to Islamic principles. These are upheld in two related documents: the *Universal Islamic Declaration of Human Rights* (1981), and the *Cairo Declaration of Human Rights in Islam* (1990). The former is very explicit about the divine origin and purpose of human rights as "decreed by Divine Law" and "designed to eliminate oppression and injustice." Written in formulaic language, it borrows from the Qurʾān, Hadith Qudsi (Divine Sayings), and Hadith throughout. The latter introduces the notion of a 'dignified life' in its preamble as well as affirming 'basic human dignity' and serving as a set of guidelines for member states in the Islamic Conference. It is written in a more secular style. The parity between 'human dignity' and كرامة الإنسان (*karamat al-ʾinsān*) as theological concepts may have posited an 'assumed equality' between Western and Islamic perspectives on human rights (cf. Edzard 1996: 54) and may yet help to reconcile them. Pin (2017: 61) maintains that the trajectories of *karāma* and 'human dignity' towards their current signification of inherent human worth in the context of human rights are strikingly similar. She attributes this to the adaptability of religious cultures to read the signs, and respond to the needs, of the times. In the

next section, we turn our attention to *karāma* as a linguistic phenomenon in Arabic, and consider its variety of derived forms from the abstract root entity ك ر م (k-r-m) as these are pertinent to our corpus-based study.

5.   Establishing search terms for analysis: derived forms of the Arabic root ك ر م (k-r-m)

The primary concept in the Arabic morphological system (as with any other Semitic language) is the 'root', typically represented by a set of three consonants in a certain order (Ryding 2005: 47); our target triliteral root is ك ر م *k-r-m*. This root entity constitutes a "nucleus or core around which are constelled a wide array of potential meanings, depending on which pattern is keyed in" (Ryding 2005: 47). The pattern, in turn, is defined as a "discontinuous morpheme [...] of one or more vowels and slots for root phonemes (radicals)" (Ryding 2005: 47-48). This combines with the root on a stand-alone basis or with up to three derivational affixes to form the morphologically complex Arabic 'word' (Ryding 2005: 48).

The intrinsic meaning of the root k-r-m is: to be noble; to honour or revere or treat with deference; to call someone noble and high minded (Wehr 1994: 692). The main part-of-speech signifying this intrinsic meaning is the verbal noun or *maṣdar* كرامة (*karāma*, 'dignity'). The Form I verb كَرُمَ (*karuma*, to be noble or generous) in turn enables morphological derivation of other words such as the masculine and feminine adjectives كريم (*karīm*, 'noble') and كريمة (*karīma*, 'valuable'). We are primarily interested in the maṣdar form كرامة *karāma*, where some of the associated meanings are: nobility, high-mindedness, noble-heartedness, generosity, magnanimity, liberality, munificence, honour, dignity, respect, esteem, standing, prestige, mark of honour, token of esteem, and favour (Wehr 1994: 693).

Derivation (إشتقاق, *ishtiqāq*) is the main method of word formation in Arabic (Al-Jurjānī, 1983: 44-5; Stetkevych 1970: 7; Elmgrab 2016: 77). As we have seen, this includes formation of verbs which represent a combination of root meaning plus a particular pattern that adds a functional meaning. The second basic verb paradigm of interest is the enhanced Form II verb which expresses intensity by doubling the middle radical (Al-Jurjānī, 1983: 299). Thus, from *k-r-m* we get the verb كَرَّمَ (*karrama*, 'to honour') which emphasises the notion of honouring someone. This verb gives rise to the verbal noun/ maṣdar تكريم (*takrīm,* 'honouring'), and the past participle form مكرّم (*mukarram*, 'honoured'). Our set of search terms derived from the root *k-r-m*[5] appears in Table A.

6.   Quantitative and qualitative analysis of search results

The *Leeds Parallel Corpus of Arabic-English Constitutions* is a raw text corpus of Modern Standard Arabic and therefore our search began with

identifying target derivational forms of the root ك ر م (*k-r-m*) associated with the concept of dignity. These have already been presented and discussed in Section 5. We then inspected KWIC (key word in context) concordance lines for each term in the entire corpus of Arabic constitutions in Sketch Engine, and then conducted a parallel search for each term to identify its equivalent translation(s) in the English data. For example, we retrieved three instances of the prepositional phrase بكرامة (*bi-karāma,* 'with dignity') via the concordance function (Figure 1). The token بكرامة is an Arabic word (defined by whitespace) where the indefinite form of the noun كرامة is affixed by the preposition and morpheme بـ (*bi-*) as proclitic.

Figure 1. Three concordance lines for بكرامة (with dignity) in the Sketch Engine interface, with running text from the constitution of Kuwait at bottom of screen

By clicking on the node word بكرامة in each concordance line, users can retrieve more running text in the vicinity of that word; and by clicking on the file ID, users can locate the source text (i.e. constitution) for a given concordance. In this case, بكرامة (with dignity) appears once in each of the constitutions of Algeria, Kuwait and Libya. Starting a new search with the same Arabic word (بكرامة) but also selecting the parallel English corpus retrieves the English translation for each occurrence of the search term. Corresponding English translations for the prepositional phrase بكرامة in our data are as follows: 'infringe *upon the dignity*' (Algeria); 'pride *in the dignity*' (Kuwait); 'living *with dignity*' (Libya). Results of our comprehensive search over the Arabic data are tabulated in Table 1.

Arabic Word	Romanised Form	Syntax	Count
كرامة	karāma	noun: fem. indef.	24
بكرامة	bi-karāma	as above, prefixed with preposition (bi)	3

الكرامة	al-karāma	noun: fem. def.	23
الكريم	al-karīm	adjective: masc. def.	3
الكريمة	al-karīma	adjective: fem. def.	4
كريم	karīm	adjective: masc. indef.	1
كريمة	karīma	adjective: fem. indef.	4
تكريم	takrīm	verbal noun (maṣdar): indef	1
تكرم	tukarim	verb: present tense	1
مكرماً	mukaraman	past participle: masc. indef.	1
TOTAL COUNT			65

Table 1. Raw counts for all forms derived from the root ك ر م (*k-r-m*) in the Arabic corpus

One interesting finding is that the total count of 65 represents mentions of 'dignity' in every Arabic constitution bar one: neither the term كرامة, or any variant form, occurs even once in the current constitution of Palestine, although it was mentioned three times in *The Palestinian National Charter: Resolutions of the Palestine National Council* (July 1-17, 1968) in Articles 17 and 24. Constitutions with the most mentions of كرامة (*karāma*, 'dignity') are Egypt (12 instances) and Sudan (11 instances). Another interesting finding is the frequency of the indefinite form كرامة (*karāma*, 'dignity'). This is the most frequent variant with a total count of 24. The majority of concordance lines for كرامة in the Arabic corpus are shown in Figure 2. The full breakdown of counts for كرامة (*karāma*, 'dignity') per Arab constitution is given in Appendix 1.

Figure 2. The first 20 concordance lines for the search term كرامة (*karāma*, 'dignity') from the Arabic corpus

6.1 English translation data

We now discuss the full set of parallel English translations. The feminine noun كرامة (*karāma*), in its indefinite form, is always translated into English with the word 'dignity' in the corpus. This is also the case when it occurs within the prepositional construction بكرامة (*bi-karāma*, 'with dignity'). When this noun is made definite, as in الكرامة (*al-karāma*), it is translated as 'dignity' in most cases (i.e. 18 out of 23), but whenever it occurs within a negative construction, it is translated as: 'degrading treatment' (2), or 'demeaning treatment' (1), or 'ignominious treatment' (1), or 'undignified treatment' (1). In the single instance of 'undignified treatment', for example, which appears in the Constitution of Bahrain (2002), the original Arabic sentence begins with negation: 'لا يعرض اي الإنسان' (*no person shall be subjected to*), and then negates each item in a list of forbidden behaviours ending with: 'أو للمعاملة الحاطة بالكرامة' (*or undignified treatment*). We comment further on connotations of the word 'treatment' in Section 6.2.

English translations for the adjectival forms: الكريم (*al-karīm* [masc. def]); الكريمة (al-karīma [fem. def]); كريم (*karīm* [masc. indef]); and كريمة (*karīma* [fem. indef]) display another interesting pattern. These Arabic forms are mostly translated as 'decent' or 'dignified' in the context of reasonable expectations for a person's life or standards of living. For example, in the Constitution of Iraq (2005), there is an undertaking that the State 'shall guarantee to the individual and the family — especially children and women – social and health security, the basic requirements for living *a free and decent life*' (our italics). Furthermore, in the Constitution of Sudan (2005), we find: 'redressing imbalances of income and achieving *a decent standard of life for all citizens*' identified as a major Millennium Development Goal (our italics and bold).

The remaining items for comment are translations for: the Arabic verbal noun or gerund تكريم (*takrīm*); a present tense verb form تكرم (*tukarim*); and the past participle مكرماً (*mukaraman* [masc. indef]). These are translated respectively as: 'honor' (2) and 'dignified' (1). We note the American spelling in the English translation here which suggests that American rather than British English is the dominant norm-providing variety used in that country/region, in this case Egypt and Libya. This contrasts with the spelling 'honour' which appears in the constitutions of Algeria, Tunisia, Sudan, Oman, Bahrain and Qatar, and is suggestive of a British English influence. We also note that in spite of the asymmetry of the Arabic and English linguistic systems, the concept 'dignity' and its variants discussed in this section translate the equivalent meaning to the TT but it should be noted that "it is not possible to recover this exact equivalent meaning in any translation, into any language, at any given time, without some kind of temporal, cultural or political interference" (Vidal Claramonte 2013: 187).

6.2 Semantic prosody of 'treatment' in the English corpus

The notion of semantic prosody refers to the accumulation of positive or negative semantics for an ostensibly neutral word via its collocational patterning (Stubbs 1995; Sinclair 2000; Hunston and Francis 2000). Instances of the word 'treatment' preceded by a negative adjective in the English data are interesting because they suggest that the word 'treatment' in English carries a negative semantic prosody in this language domain (human rights law), namely: it habitually collocates with a particular semantic set where the attitudinal or pragmatic meaning is negative. An example that springs to mind from general English usage would be: 'ill treatment.' Inspecting concordance lines for the word 'treatment' in the English corpus, we find that it occurs 28 times in total; that it is neutral in meaning when associated with the medical domain (11 instances); that it is twice qualified positively in the phrase 'humane treatment' of prisoners and animals (2); but that elsewhere it has an extremely negative human rights connotation with *humiliation* in English (*degrading* (6); *demeaning* (2); *ignominious* (1); *undignified* (1)), and also with *cruelty*: 'cruel treatment' (3); 'inhumane treatment' (1); 'torture, inducement, or such treatment' (1). Thus, the Arabic noun الكرامة within negative constructions invokes, or is understood to invoke, protection against human rights violations as a constitutional and legal responsibility. This is apparent from its English translation in the following constitutions: 'degrading treatment' (Algeria, Mauritania, Morocco, Qatar, Sudan and UAE); 'demeaning treatment' (Oman); 'ignominious treatment' (Kuwait); 'undignified treatment' (Bahrian); 'cruel treatment' (Palestine and Mauritania), and 'inhumane treatment' (Iraq).

6.3 Collocates of 'dignity' in the English corpus

Another function available in Sketch Engine is the word sketch, which generates a one-page summary of a word's grammatical and collocational behaviour. The definition of collocation adopted in this paper is standard in corpus linguistics and refers to 'a sequence of words or terms that co-occur more often than would be expected by chance within the context of a specific word' (Gómez 2009: 149; cf. Lehecka 2015: 2). We are interested in statistically significant collocates of 'dignity' and 'كرامة' in our data. In corpus linguistics, these are defined via an association score which draws on raw frequencies within a given corpus to determine the degree of statistical association between two words. The score is computed for all possible word pairs and the highest-scoring pairs are then presented as collocation candidates. The preferred association measure for the word sketch function in Sketch Engine is logDice since it scales well on different corpus sizes (Rychlý 2008). However, other measures are also made available for researchers under the *collocations* tab in the toolkit, one of them being log likelihood (LL), where statistical significance at a confidence level of 99% is generally reflected in LL scores of 6.63 and over (Rayson 2009). The collocations tab appears on

the user menu in Sketch Engine when conducting a word search and enables users to select one or more association measures and to specify range (i.e. number of words removed from the node word in right/left positions).

We find that the most frequent modifier and highest-scoring collocate of 'dignity' in the English data is 'human' in the grammatical construction ADJ+NOUN: the phrase 'human dignity' occurs 15 times in total. Comparative statistics for the top ten (content word) collocates of 'dignity' (in various grammatical relations ranging from -5 to +5 either side of the node) generated from our data are given in Table 2.

Collocation Candidate	Co-occurrence Count	Candidate Count	LL Score	logDice Score
Human	15	133	167.411	11.415
Preserves	3	9	39.879	10.777
Justice	6	141	53.499	10.030
Individual	3	54	28.180	9.926
Equality	3	70	26.590	9.714
Prohibited	3	73	26.334	9.678
Freedom	5	168	40.845	9.573
Respect	5	187	39.765	9.451
Integrity	3	103	24.242	9.356
Social	4	205	29.193	9.022

Table 2. Raw counts and comparative association scores for the top ten most significant collocates of 'dignity' in our English data

6.4 Collocates of 'كرامة' and 'الكرامة' in the Arabic data

For the original Arabic data, كرامة الإنسان (*karamat al-ʾinsān,* 'human dignity') occurs 6 times in total and is variously translated as 'human dignity' (4), 'the dignity of Man' (1), and 'man's dignity' (1). This phrase (كرامة الإنسان) is an example of the possessive *iḍāfa* (إضافة) construction which relates two nouns in Arabic grammar; the head word (in this case كرامة) does not carry the definite article marker and the dependent noun is in the genitive case (الانسان). A further co-occurrence of 'كرامة' with 'الإنسان' which is included in the calculation for this particular collocation is: 'حرية الإنسان و كرامته مصونة' (*human freedom and dignity are safeguarded*) in the Constitution of Iraq. In Table 3 we present the top 5 collocates for the Arabic noun forms 'كرامة' (*karāma*) and 'الكرامة' (*al-karāma*) for further discussion, including linkage with some of the concepts that emerge as collocation candidates of 'dignity' in Table 2. These collocates are all content words apart from the

possessive pronoun هم (their). They are sorted by the logDice score and the LL score is also given.

Variant	Collocation Candidate		Co-occurrence Count	Candidate Count	LL Score	logDice Score
كرامة	الإنسان	(human)	7	111	76.345	10.730
	امن	(security)	3	48	32.082	10.415
	حرية	(freedom)	3	191	23.682	8.836
	حقوق	(rights)	3	217	22.916	8.672
	هم	(their)	6	995	36.778	7.592
الكرامة	الحاطة	(degrading)	5	5	91.809	12.514
	المعاملة	(treatment)	6	16	89.315	12.299
	الإنسانية	(human)	9	58	117.185	11.830
	التعذيب	(torture)	3	20	37.885	11.158
	الحرية	(freedom)	4	71	42.478	10.445

Table 3. Raw counts and comparative association scores for the top 5 most significant collocates of 'كرامة' and 'الكرامة' in our corpus of Arabic constitutions

6.4.1 Collocates of the indefinite form كرامة (*karāma*, 'dignity')

The main comment to make on the top-scoring (content word) collocates of كرامة (*karāma*) in Table C is that they have positive connotations. Furthermore, these positive connotations are reflected in the collocates of 'dignity' in the English translation (Table B). However, we note that in the Constitution of Saudi Arabia, the safeguarding of 'كرامة الإنسان و حقوقه' ('man's dignity and rights') is embedded within a list of 'actions' prohibited by statute, namely: 'sedition or division or harm[ing] the state's security and its public relations.' We also note that all top-scoring collocates of كرامة coalesce in a single sentence in the constitution of Syria:

الحرية حق مقدس وتكفل الدولة للمواطنين حريتهم الشخصية وتحافظ على كرامتهم وأمنهم

This is translated as: 'freedom shall be a sacred right and the state shall guarantee the personal freedom of citizens and preserve their dignity and security'. Co-occurrence of the possessive pronoun هم (*hum*, 'them') with كرامة (*karāma*, 'dignity') identifies the following referents: the martyrs and their dependents (Algeria); people with special needs (Sudan); the elderly (Sudan); and all citizens of the state (Sudan, Syria, Yemen)[6].

6.4.2 Collocates of the definite form الكرامة (*al-karāma*, 'dignity')

The two top-scoring collocates of 'الكرامة' (الحاطة and المعاملة; *al-ḥāṭah* and *al-muʿamālah*) tend to appear together in the same phrase, and are translated as a unit, consisting of a qualifying adjective plus the word

'treatment.' This supports our finding in Section 6.2, namely: the word 'treatment' in English appears to have a negative connotation in international human rights as evidenced in our corpus, and is used as an antonym to the concept of human dignity to define human rights abuses. Two examples from the corpus are:

أو للمعاملة الحاطة بالكرامة ('or demeaning treatment') [Oman 2011];

للتعذيب أو للمعاملة الحاطة بالكرامة ('torture or [any] degrading treatment') [Kuwait 1992; Qatar 2004; UAE 2011].

We note that in the second example, they also appear in context with another significant collocate of 'الكرامة', namely: 'التعذيب' (*al-taʾdhīb*, 'torture'). This explicates the principle that violations of 'human dignity' (كرامة الإنسان, *karamat al-ʾinsān*) involve physical as well as moral violence:

و يحظر اي عنف بدني او معنوي او اي مساس بالكرامة ('any form of physical or moral violence or any infringement of dignity shall be prohibited') [Algeria 2016].

The collocation of the definite form 'الكرامة' with 'التعذيب' in the Arabic corpus is also supported by the English data, where the top 3 collocates of the word 'treatment' are: 'degrading' (log likelihood: 100.884; logDice: 12.415); 'cruel' (log likelihood: 82.986; logDice: 12.192); and 'torture (log likelihood: 86.968; logDice: 12.061).

We have already commented on the combination: 'كرامة الإنسان' (*karamat al-ʾinsān*). The collocation 'الكرامة الإنسانية' occurs 9 times and is composed of head word and *nisba* (اَلنِّسْبَة) adjective, where both constituents agree in gender, number, definiteness and case. An example is the prepositional phrase: 'في الكرامة الإنسانية' ('in human dignity'), where the genitive case ending may be apparent in formal spoken Arabic: '*fi al-karama al-ʾinsāniyya*'. This collocation is mostly translated as 'human dignity' in our English data and in two cases is explicitly associated with treatment of prisoners (Constitution of Sudan 2005; Constitution of Morocco 2011).

7.   Conclusions

In this paper, we focus attention on the primary human rights concept of كرامة (*karāma*, 'dignity') in the language of Arabic constitutional law and its English translation. The main methodological contribution is our interdisciplinary approach combining corpus linguistics and Arabic legal translation to discover meaning, and parity of meaning, through statistical profiling and qualitative analysis of this term and its collocates in contemporary, naturally-occurring legal language; this advances theory and practice in the newly-emerging field of corpus-based legal linguistics and legal translation. Our corpus-based study is conducted in the Sketch Engine toolkit over our novel, parallel corpus of Arabic-English

constitutions. This contributes to much-needed Arabic language resources.

Our search methodology over the Arabic corpus entails specification of all derived forms of the Arabic root ك ر م (k-r-m) as hypernym for the semantic field of 'dignity'. In total, we identify 65 instances of the term كرامة (or one of its variants) spread throughout the Arabic corpus but notably missing in the constitution of Palestine.

Our methodology also involves close scrutiny of parallel concordance lines for each occurrence of the target term. We find that while the indefinite noun كرامة is always translated into English as 'dignity', the definite form الكرامة is often rendered via a qualifying adjective plus 'treatment' in English, especially in negative constructions. Moreover, the combination of الكرامة and negation in Arabic maps to qualification of the English word 'treatment' with notions of humiliation and cruelty (e.g. 'degrading,' 'demeaning,' 'ignominious,' 'undignified,' 'cruel') such that this word sense acquires a negative connotation or semantic prosody.

The positive connotations of كرامة versus the negative connotations of الكرامة (in sentential negation) are ascertained through their statistically significant collocates verified by stringent metrics in Sketch Engine (Table C). The most significant collocate of كرامة is الإنسان ('human'); and the most significant collocate of الكرامة is الحاطة ('degrading'). We also find this same collocational patterning reflected in the English translation data. While the collocates of 'dignity' (كرامة) have positive connotations, the most significant collocates of 'treatment', used with a qualifying adjective to translate the definite form 'الكرامة' in negative constructions, are: 'degrading', 'cruel' and 'torture'. Thus the concept of 'dignity' as signified by كرامة and its variants in Arabic constitutions presupposes negative as well as positive human rights in accordance with international law (e.g. the EU *Charter of Fundamental Rights*).

Our interdisciplinary study is one of the few studies in corpus-based analysis of Arabic legal language and Arabic-English legal translation, and thus contributes to these emerging fields. It suggests a successful 'cross-linguistic negotiation of meaning' (Biel 2015) between كرامة and 'dignity', and الكرامة (in negation) and 'treatment' (with qualifying adjective) in the specification of positive and negative human rights as evidenced in our parallel corpus of constitutions. This conclusion is supported and strengthened via corpus statistics over patterns of collocation: there is parity of meaning in the collocates of كرامة/dignity (e.g. الإنسان/human) and الكرامة/treatment in negative contexts (e.g. الحاطة/degrading). This may partly be due to the resonance of 'human dignity' and كرامة الإنسان (*karāmat al-ʾinsān*) as analogous theological concepts.

Finally, this study paves the way for further research in Arabic legal translation and highlights the role of corpus-based analysis in this under-

researched area. A future useful endeavour will be to compare the translations of the term (*karāma*, 'dignity') in the previous versions of the Arabic constitutions to check if these have undergone any change across time and whether this change (if any) has impacted on the negotiation of meaning. Another possible future endeavour will be a corpus-based investigation of Arabic-English translation of deontic modals, one of the most relevant linguistic areas that conveys obligation, hence mistranslation would have an impact on the binding nature of a legal document.

8.   Acknowledgements

9.   References

- **Afshari, Reza** (1994). "An essay on Islamic cultural relativism in the discourse of human rights." *Human Rights Quarterly* 16(2), 235-276.

- **Alcaraz Varó, Enrique and Brian Hughes** (2002). *Translation Practices Explained: Legal Translation Explained*. Manchester: St. Jerome.

- **Al-Jurjānī, ᶜAli ibn Muhammad** (1983). *kitāb al-taᶜrīfāt*. With an Introduction by Ibrahim El-Ibyāri. Cairo: dār al-Rayān li-turāth.

- **Alwazna, Rafat Yahya** (2013a). *Translating Ḥanbalī Sharīʾa Code from Arabic into English*. Saarbrücken: Lap Lambert Academic Publishing.

- **Alwazna, Rafat Yahya** (2013b). "Testing the precision of legal translation: The Case of testing Islamic legal terms into English." *International Journal for the Semiotics of Law* 26(4), 897-907.

- **Alwazna, Rafat Yahya** (2016). "Islamic Law: Its sources, interpretation and the translation of it into laws written in English." *International Journal for the Semiotics of Law* 29(2), 251–260.

- **Alwazna, Rafat Yahya** (2017). "Culture and law: The cultural impact on Islamic legal statements and its implications for translation." *International Journal of Legal Discourse* 2(2), 307–323.

- **BBC Four (2017).** *Storyville: Last Men in Aleppo*. Broadcast 27.11.17, 22:00.

- **Baker, Mona** (1993). "Corpus linguistics and translation studies: implications and applications." Mona Baker, Gill Francis and Elena Tognini-Bonelli (eds). (1993). *Text and Technology: In honour of John Sinclair.* Amsterdam/Philadelphia: John Benjamins, 233-250.

- **Baker, Mona** (1995). "Corpora in translation studies: an overview and some suggestions for future research." *Target* 7(2), 223-243.

- **Bhatia, Vijay Kumar** (1993). *Analyzing Genre: Language Use in Professional Settings*. London: Longman.

- **Biel, Łucja** (2010) "Corpus-based studies of legal language for translation purposes: Methodological and practical potential." Carmen Heine and Jan Engberg (eds) (2010)*. Reconceptualizing LSP. Proceedings of the XVII European LSP Symposium 2009.* Aarhus. http://asb.dk/fileadmin/www.asb.dk/isek/biel.pdf (consulted 12 April 2019).

- **Biel, Łucja** (2015). "Phraseological profiles of legislative genres: complex prepositions as a special case of legal phrasemes in EU law and national law." *Fachsprache* 3-4, 139-160.

- **Biel, Łucja and Jan Engberg** (2013). "Research models and methods in legal translation." *Linguistica Antverpiensia NS-TTS* 12, 1-11.

- **Biel**, **Łucja, Jan Engberg, Rosario Martín Ruano and Vilelmini Sosoni (eds)** (2019). *Research Methods in Legal Translation and Interpreting: Crossing Methodological Boundaries*. London/New York: Routledge.

- **Buckwalter, Tim and Dilworth B. Parkinson** (2013). "Modern Lexicography." Jonathan Owens (ed.) (2013). *The Oxford Handbook of Arabic Linguistics*. Oxford: Oxford University Press.

- **Cao, Deborah** (2007). *Translating Law*. Clevedon: Multilingual Matters.

- **Central Intelligence Agency** (CIA) (2019). *The World Factbook*. https://www.cia.gov/library/publications/resources/the-world-factbook/geos/le.html (consulted 15.05.19).

- **Dukes, Kais** (2014). *Statistical Parsing by Machine Learning from a Classical Arabic Treebank*. PhD Thesis. School of Computing, University of Leeds.

- **Dupré, Catherine** (2011). "What does dignity mean in a legal context?" *The Guardian* Mar. 24. https://www.theguardian.com/commentisfree/libertycentral/2011/mar/24/dignity-uk-europe-human-rights (consulted 29.12.17)

- **Edzard, Lutz** (1996). "Stylistic elements in the use of Arabic as language of diplomacy: Recent developments in the United Nations context." *Die Welt des Islams* 36(1), 25-58.

- **El-Farahaty, Hanem** (2015). *Arabic-English-Arabic Legal Translation.* London/New York: Routledge.

- **El-Farahaty, Hanem** (2016). "Translating lexical legal terms between English and Arabic." *International Journal for the Semiotics of Law* 29(2), 473-493.

- **Elmgrab, Ramadan Ahmed** (2016). "The creation of terminology in Arabic." *American International Journal of Contemporary Research* 6(2): 75-85.

https://www.researchgate.net/publication/303843062_The_Creation_of_Terminology_in_Arabic (consulted 15.5.19).

- **Engberg, Jan** (2002). "Legal meaning assumptions – what are the consequences for legal interpretation and legal translation?" *International Journal for the Semiotics of Law* 15(4), 375-388.

- **Engberg, Jan** (2013). "Comparative law for Translation: The key to successful mediation between legal systems." Anabel Borja Albi and Fernando Prieto Ramos (eds) (2013). *Legal Translation in Context: Professional Issues and Prospects*. Oxford: Peter Lang, 9-25.

- **Engberg, Jan and Dorothee Heller** (2008). "Vagueness and indeterminacy in law." Vijay Kumar Bhatia, Christopher N. Candalin and Jan Engberg (eds) (2008). *Legal Discourse Across Cultures and Systems*. Hong Kong: Hong Kong University Press, 145–162.

- **European Union** (2000). *Charter of Fundamental Rights of the European Union*. http://www.europarl.europa.eu/charter/pdf/text_en.pdf (consulted 08.01.19)

- **Fagan, Frank** (2016). "Big data legal scholarship: Toward a research program and practitioners' guide." *Virginia Journal of Law and Technology* 20(1), 1-81.

- **Garre, Marianne** (1999). *Human Rights in Translation: Legal Concepts in Different Languages*. Denmark: Copenhagen Business School Press.

- **Gómez, Pascual C.** (2009). "Attempting to model sense division for word sense disambiguation." Maria Manuela Cruz-Cunha *et al*. (eds). *Handbook of Research on Social Dimensions of Semantic Technologies and Web Services* (2 Volumes). Hershey: IGI Global, 126-157.

- **Hallaq, Wael B.** (2014). *The Impossible State: Islam, Politics, and Modernity's Moral Predicament*. New York: Columbia University Press.

- **Hart Jr., M. Henry and Albert M. Sacks** (1994).*The Legal Process: Basic Problems in the Making and Application of the Law*. William M. Eskridge Jr and Philip P. Frickey (eds). Cambridge, MA: The Harvard Law Review Association.

- **Hunston, Susan and Gill Francis** (2000). *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: John Benjamins.

- **Islamic Conference of Foreign Ministers** (1990). *Cairo Declaration on Human Rights in Islam*. http://hrlibrary.umn.edu/instree/cairodeclaration.html (consulted 08.01.19)

- **Islamic Council** (1981). *Universal Islamic Declaration of Human Rights*. http://www.alhewar.com/ISLAMDECL.html (consulted 08.01.19).

- **Jackson, Vicki C.** (2004). "Constitutional dialogue and human dignity: States and transnational constitutional discourse." *Montana Law Review* 65, 15-40.

- **Kilgarriff, Adam *et al*.** (2014). "The Sketch Engine: ten years on." *Lexicography* 1, 7-36.

- **Lehecka, Tomas** (2015). "Collocation and colligation." Jef Verschueren, Jan-Ola Östman, Jan Blommaert, Chris Bulcaen (eds). *Handbook of Pragmatics Online*.

Amsterdam/Philadelphia: John   Benjamins,   1-20.
https://benjamins.com/online/hop/articles/col2 (consulted 08.01.19)

- **Libreria Editrice Vaticana** (1993). *Catechism of the Catholic Church*.
  http://www.vatican.va/archive/ENG0015/__P16.HTM (consulted 08.01.19)

- **Mellinkoff, David** (1963). *The Language of the Law*. Boston: Little/Brown.

- **Mouritsen, Stephen C**. (2011). "Hard cases and hard Data: Assessing corpus
  linguistics as an empirical path to plain meaning." *Columbia Science and Technology
  Law Review* 13, 156-205.

- **Mouritsen, Stephen C.** (2010). "The dictionary is not a fortress: Definitional fallacies
  and a corpus-based approach to plain meaning." *Brigham Young University Law
  Review* 2010(5), 1915-1980.

- **Neuman, Gerald L.** (2003). "Human rights and constitutional rights: Harmony and
  dissonance." *Stanford Law Review* 55, 1863-1900.

- **O'Mahoney, Conor** (2012). "There is no such thing as a right to dignity."
  International Journal of Constitutional Law 10(2), 551–574.

- **Pin, Andrea** (2017). "Arab Constitutionalism and Human Dignity." *The George
  Washington International Law Review* 50(1), 1-67.

- **Rayson, Paul** (2009). *Wmatrix: a Web-based Corpus Processing Environment*.
  Computing Department, Lancaster University. http://ucrel.lancs.ac.uk/wmatrix/
  (consulted 08.01.19)

- **Rychlý, Pavel** (2008). "A lexicographer-friendly association score." Petr Sojka and
  Ales Horák (eds). *Proceedings of Recent Advances in Slavonic Natural Language
  Processing* (RASLAN). Brno: Masaryk University, 6-9.

- **Ryding, C. Karin** (2005). *A Reference Grammar of Modern Standard Arabic*.
  Cambridge: Cambridge University Press.

- **Šarčević, Susan** (1997). *New Approach to Legal Translation*. The Hague: Kluwer Law
  International.

- **Schmitt, Carl** (1985). *Political Theology: Four Chapters on the Concept of
  Sovereignty*. Translated by George Schwab. Chicago: University of Chicago Press.

- **Shulztiner, Doron and Guy E. Carmi** (2014). "Human dignity in national
  constitutions: Functions, promises and dangers." *The American Journal of
  Comparative Law* 62(2), 461–490.

- **Sinclair, John** (2000). "Lexical Grammar." *Naujoji Metodologija* 24, 191-203.
  http://donelaitis.vdu.lt/publikacijos/sinclair.pdf (consulted 13.11.17).

- **Stetkevych, Jaroslav** (1970). *The Modern Arabic Literary Language, Lexical and
  Stylistic Development.* Chicago: Chicago University Press.

- **Stubbs, Michael** (1995). "Collocations and semantic profiles: On the cause of the
  trouble with quantitative studies." *Functions of Language* 2(1), 23-55.

- **United Nations** (1945). *Charter of the United Nations*. http://www.un.org/en/charter-united-nations/ (consulted 08.01.19).

- **United Nations** (1948). *Universal Declaration of Human Rights*. http://www.un.org/en/udhrbook/pdf/udhr_booklet_en_web.pdf (consulted 08.01.19).

- **United Nations** (1966a). *International Convention on the Elimination of All Forms of Racial Discrimination*. https://treaties.un.org/Pages/ViewDetails.aspx?src=IND&mtdsg_no=IV-2&chapter=4&clang=_en (consulted 08.01.19).

- **United Nations** (1966b). *International Covenant on Civil and Political Rights*. https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx (consulted 08.01.19).

- **United Nations** (1966c). *International Covenant on Economic, Social and Cultural Rights*. https://www.ohchr.org/en/professionalinterest/pages/cescr.aspx (consulted 08.01.19).

- **United Nations** (1979). *Convention on the Elimination of All Forms of Discrimination against Women*. http://www.un.org/womenwatch/daw/cedaw/text/econvention.htm (consulted 08.01.19).

- **United Nations** (1984). *Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*. https://www.ohchr.org/en/professionalinterest/pages/cat.aspx (consulted 08.01.19).

- **United Nations** (1989). *Convention on the Rights of the Child*. https://www.ohchr.org/en/professionalinterest/pages/crc.aspx (consulted 08.01.19).

- **United Nations** (1993). *Vienna Declaration and Programme of Action*. https://www.ohchr.org/en/professionalinterest/pages/vienna.aspx (consulted 08.01.19).

- **Vidal Claramonte, Maria Carmen África** (2013). "Towards a new research model in legal translation: future perspectives in the era of asymmetry." *Linguistica Antverpiensia NS-TTS* 12 , 182-196.

- **Vogel, Freidemann, Hanjo Hamman and Isabelle Gauer** (2017). "Computer-assisted legal linguistics: Corpus analysis as a new tool for legal studies." *Law and Social Inquiry* 42, 1-24.

- **Wehr, Hans** (1994). *Dictionary of Modern Written Arabic: Arabic-English,* 4th Ed. Weisbaden: Harrassowitz.

- **Williams, Christopher and Denise Milizia** (2008). "How (un)readable is the European Constitution? A comparison of the English version and the Italian version." Andrea Cannone *et al.* (eds) (2008). *Studi in Onore di Vincenzo Storace.* Naples: Editoriale Scientifica, 2209-2227.

- **Zemánek, Petr, Andrea Moustafa, Naděžda Obadalová and František Ondráš** (2006). *Arabsko-česky slovník* [Arabic-Czech Dictionary]. Set Out: Praha, 2006.

Biographies

Claire Brierley is a computational and corpus linguist affiliated to the University of Leeds as a Visiting Researcher in Computing and also in Arabic, Islamic and Middle Eastern Studies. Her research interests include natural language processing for Quranic Arabic, plus automatic extraction/classification of prosodic features in Arabic and English text corpora, as evidenced in her most recent journal articles for *Semitic Studies* (Brierley *et al*. 2018; 2016) and her PhD in Computing from Leeds (2011). She has spent over 26 years in higher education as a senior lecturer and senior research fellow.

E-mail: C.Brierley@leeds.ac.uk

Hanem El-Farahaty is a Lecturer of Arabic/English Translation and Interpreting and the Programme Manager for the MA in Arabic/English Translation at the University of Leeds. In 2011, Hanem was awarded a PhD for her research into Arabic Translation Studies with a particular focus on legal translation. She has taught Arabic and translation in a number of UK universities. She also taught English language and translation at the University of Mansoura, Egypt. Hanem is the author of *Arabic-English-Arabic Legal Translation* (2015) and she is a co-editor of the *Routledge Handbook of Arabic Translation* (2019).

E-mail: H.El-Farahaty@leeds.ac.uk

Appendix 1. Breakdown of counts for morphological variants of كرامة (*karāma*, 'dignity') in the Arabic corpus

Country	Word count	كرامة	الكرامة	بكرامة	الكريم	الكريمة
		dignity	the dignity	with dignity	the dignified	the dignified
		noun: fem. indef.	noun: fem. def.	noun: fem. indef. with preposition (bi)	adjective: masc. def.	adjective: fem. def.
Algeria	11,051	1	2	1	0	0
Iraq	9,089	1	0	0	0	0
Jordan	8,122	1	0	0	0	0
Lebanon	5,260	1	0	0	0	0
Mauritania	5,351	1	0	0	0	0
Morocco	12,085	1	2	0	1	0
Oman	9,046	1	1	0	0	0
Saudi Arabia	2,354	1	0	0	0	0
Sudan	22,669	6	4	0	0	0
Syria	6,329	2	1	0	0	0
Tunisia	9,385	3	2	0	1	0
Egypt	16,968	3	3	0	0	3
Bahrain	8,878	0	3	0	1	0
Kuwait	6,558	0	3	1	0	0
Libya	2,485	0	0	1	0	0
Palestine	6,866	0	0	0	0	0
Qatar	5,692	0	1	0	0	0
Emirates	8,406	0	1	0	0	1
Yemen	7,208	2	0	0	0	0
Total*	**163,802***	**24**	**23**	**3**	**3**	**4**

Country	كريم	كريمة	تكريم	تُكرم	مكرماً	Total count for variants
	dignified	dignified	dignifying (honouring)	dignifies	dignified	
	adjective: masc. indef.	adjective: fem. indef.	verbal noun (maṣdar): indef.	verb: present tense	past participle: masc. indef.	
Algeria	0	0	0	0	0	4
Iraq	0	2	0	0	0	3
Jordan	0	0	0	0	0	1
Lebanon	0	0	0	0	0	1
Mauritania	0	0	0	0	0	1
Morocco	0	0	0	0	0	4
Oman	0	0	0	0	0	2
Saudi Arabia	0	0	0	0	0	1
Sudan	1	0	0	0	0	**11**
Syria	0	0	0	0	0	3
Tunisia	0	0	0	0	1	7
Egypt	0	2	1	0	0	**12**

Bahrain	0	0	0	0	0	4
Kuwait	0	0	0	0	0	4
Libya	0	0	0	1	0	2
Palestine	0	0	0	0	0	**0**
Qatar	0	0	0	0	0	1
Emirates	0	0	0	0	0	2
Yemen	0	0	0	0	0	2
Total*	**1**	**4**	**1**	**1**	**1**	**65**

*This count in the above tables is for the Arabic corpus minus the sub-corpus of Preambles which was included for the purposes of future research. Introductory preamble is already embedded in each constitution.

Notes

[1] One example cited by Buckwalter and Parkinson (2013) is the Arabic-Czech dictionary by Zemánek *et al.* (2006) drawn from a balanced corpus of some 50 million words.

[2] Another classification of legal terms or "terms of art," as Mellinkoff (1963: 16) calls them, is offered by Alcaraz Varó and Hughes (2002: 16-18) and introduces a three-way categorisation of: technical, semi-technical and everyday language.

[3] We would like to thank Athil Khaleel Farhan for her summary of the compilation of the corpus and its alignment that she produced during her time on the project, and which was consulted during the composition of Section 2.3 above.

[4] Human dignity was mentioned in the constitutions of Mexico, Germany, Finland, Ireland and Cuba in the period between 1900 and 1944 (Shulztiner and Carmi 2014: 464).

[5] For full information about the meanings and translation of the term and its derivations, please refer to Hans Wehr Dictionary of Modern Arabic: Arabic-English (1994) and the Almaany Dictionary online: https://www.almaany.com/ar/dict/ar-ar/%D9%83%D8%B1%D8%A7%D9%85%D8%A9/.

[6] In relation to 'dignity' and underprivileged groups, 'dignity' is specifically accorded to women in the Sudanese constitution. Lebanon, as the first Arab constitution to invoke كرامة (*karāma*, 'dignity') in 1926, associates the concept with freedom of religions (Pin 2017:15); and it is known for its diversity of religious sects. According to the CIA *Factbook* (2019), these are: "Muslim 57.7% (28.7% Sunni, 28.4% Shia, smaller percentages of Alawites and Ismailis), Christian 36.2% (Maronite Catholics are the largest Christian group), Druze 5.2%, very small numbers of Jews, Baha'is, Buddhists and Hindus."