

Exploring Proportionality between Translation
Universals & Generic Properties of Scientific Writing
in Historical Authored and Contemporary Translated
/Non-translated Arabic Corpora

استقراء العلاقات النسبية بين عموميات الترجمة وخصائص الأسلوب
العلمي: دراسة مقارنة لثلاث ذخائر عربية مؤلفة و مترجمة، تاريخية
ومعاصرة"

Dr. Amal Mohammad Abdel Maqsoud
Senior Lecturer – English department
Faculty of Al-Asun - Ain Shams University

د. أمل محمد عبد المقصود
مدرس بقسم اللغة الإنجليزية
كلية الألسن - جامعة عين شمس

Exploring Proportionality between Translation Universals & Generic Properties of Scientific Writing in Historical Authored and Contemporary Translated /Non-translated Arabic Corpora

1. Introduction

Universal properties of brevity, informativity, objectivity, logicity, precision and impersonal style are given norms of scientific writing, particularly books and treatises. Linguistic realizations in languages eligible for scientific writing are sought through features of in/transitivity, voice systems, SVO and/or VSO sentence order, nominalizations, pluralities, standard terminology. In original and, presumably, translated scientific genre, the prime concern is to present pure facts, phenomena, hypotheses, diagnoses, medicinal prescriptions, chemical reactions and formulas. Neutral, direct, emotionless elaborations are parameters of a scientific textual profile. Pertaining to three selected linguistic features; and by means of a comparable corpus-based analysis of three selected Original and Translated Arabic Corpora and a direct eye-search of random samples, the present study seeks to explore the positive/ negative relations between these constant scientific properties and translators' norms and preferences of target text production, generally known as translation universals.

1.1 Lingua francas of Science

Historical background of the status of Arabic and English in the world of science helps provide grounds for the criterion of the small-scale corpora selection in this paper. The rise of Arabic as a lingua franca of science started during the Umayyad Caliphate (661-750), continued into the Abbasid rule (750- 1258), and was increasingly asserted, reaching top cultural peaks during the Abbasid Golden Age of science and civilization due to factors of power, far-reaching Islamic Conquests and translation activity. Recognition of Arabic as the standard official language of administration in the entire Islamic State in the 8th century during the reign of the Umayyad Caliph, Marwaan (r. 685-705) helped consolidate the status of Arabic as a language of science for the years to come. Major translation movements were initiated and sponsored since the early 8th c. and took two opposite directions: first came a flow of translations into Arabic which was later turned from Arabic into Latin in the late 10th c.

(Kunitzcshe, 2007, pp. 2-12). “The Arabs were credited with initiating the first organized translation activity in history which started during the Umayyad reign and reached its zenith under the Abbasids, particularly during al Maʿmuun (813-33), known as the Golden Era of translation”. A translation chamber was set up by al Mansuur (r.753-75) and was considerably expanded by ar-Rasheed (r. 786-809); Baytu l- hikmah, House of Wisdom, the most important institute of higher learning in Islam was founded by al Maʿmuun in 830 (Baker,1998, p. 318). In the meantime, during the same epochs, in the western part of Europe, Old and early Middle English continued to be repressed by the French-Latin dominance of the Norman Conquest (1066-1200) and forced into the minor status of a language spoken by the masses (Baugh, 1957, pp.127-149). A main consequence, historically regretted about the Norman Conquest, is “the total eclipse of the English vernacular as the language of literature, law, and administration. Superseded in official documents and other records by Latin and then increasingly in all areas by Anglo-Norman, written English hardly reappeared until the 13th century” (<https://www.britannica.com/event/Norman-Conquest>)

Source languages in the first wave of translations into Arabic were Indian, Persian, Aramaic, Syriac and particularly Greek; the Arabs were “eager learners” of all sources of knowledge in the civilizations and nations they came to be in contact with. All fields of science were translated, giving the way to a rich period of Original Arabic writing in the fields of medicine, alchemy, pharmacy, theology and philosophy. This was produced by “ethnic and non-ethnic” Arab scientists and scholars who, more than often, had no knowledge of Greek and relied mainly on existing Arabic translations as Ibn Sina(Avicenna), ar-Raazi (Rhazes), Ibn Hayaan (Geber)...(Baker, 1998, pp.317- 23). **Accordingly, the Originality of authored Arabic scientific works of the time was ascertained and Arabic was consolidated as a major language of science.** ¹

The second opposite wave of translation from Arabic into Latin and Old Spanish since the late 10th c contributed to further assert this fact right until the mid-13th c. European scholars in the 12th c. are ascertained to have come into Spain to collect the scientific wealth of the Arabs;

scholarly monks of Barcelona wrote their first Latin treatises based on Arabic writings (Kunitzsch, 2007, p.5). The newly translated works soon spread in France, Germany and England establishing the fame of Arabic Sciences and paving the way for the European Renaissance.

The case is reversed in modern times. The scientific status of Arabic went on the wane during the long dormant four centuries of Ottoman rule (since 1517). Only when refreshed in the early 19th.c under the ruler of Egypt, M.Ali (r.1805- 1849), Arabic turned into a target language of western science and literature; the main source languages at al-Tahtawee's al-*Alsun*, school of translation and languages, founded in 1835 on his return from a mission in France, used to be French, Turkish and occasionally English. Conversely, English continued to consolidate itself as a standard and official language of state, administration and literature since the Rise of the Standard London Dialect in the early 15th c; its use became 'a matter of precept and practice in the 16thc' (Baugh, 1957, pp. 231-237). **It was not until the end of WWI that English was gradually introduced as the one and only global language of science.** Following WWI, German and Austrian scientists were boycotted and banned from publishing in western European journals. Newly established organizations of science, medicine and applied chemistry used to function only in English and French; an anti-German war swept the US with a group of new laws criminalizing the use of German; there followed a public American tendency of paying no attention to German or French. During and ever since the 1930's- 40's the American scientific establishment came to dominate the world (Gordin, 2015, pp.175-184).

Thus, the grounds for a comparable translation corpus study are provided: linguistic features of contemporary translated scientific Arabic works are compared to Arabic as an Original historical source language of science; considerations of the time dialect variety are observed; that is why further evidence is sought by incorporating a third small-scale corpus of contemporary authored, i.e., non-translated scientific Arabic. This is carried out against the backdrop of the acknowledged properties and features of Original scientific English genre as established by translation theorists (Pinchuk 1977; Dickens, Hervey, & Higgins 2002; Baker 2011; Olohan 2016) and in view of its being the contemporary global language of science. As far as the strict constraints of copyrights allow, care has

been taken to choose similar domains of the selected corpora, focusing mainly on medicine, pharmacy and physics.

1.2 Rationale of Study

Trying to find smoother ways to render the required messages from a given SL into a TL, translators develop habits of translation. Positively, these might become norms of translation, but, negatively, it may lead to deformities or undesirable features. Norms and deformities stand out only when translations are tested against their originals and/or source texts. No primacy is held any more in recent Descriptive Translation Studies (DTS) to the source text properties or linguistic features. Long observations drawn from translation quality assessment and contrastive analysis lead translation scholars to make predictions about particular lexical and syntactic features that carry simplification, explicitness, normalization, and conventionalization _in short, features that realize translation universals. Large and small-scale electronic corpora are in use since the early 1990's to prove the high and low frequencies of translation universals, whereupon new directions and recommendations can be made about translating particular genres, and observations are registered to describe a given language- pair corpus or a multilingual set of source languages and a target.

Variability of translation universals is acknowledged in DTS. Frequencies differ according to the TL in question, genre or register chosen for investigation and time span. Here lies the point of departure in the present study: properties of scientific genre particularly when they belong to the lingua franca of science at a given age are invariable. Science is science across ages. Based on (1) a monolingual comparable corpus study of Arabic scientific translations, Historical and contemporary non-translations, (2) selected syntactic features of the Arabic emphatic particle/?inna,?anna, /, intransitivity and passive verb forms, predicted to carry specific translation universals, the paper investigates proportionality between the variability of translation universals and the invariable properties of scientific genres. The aim is to obtain observations about contemporary Arabic scientific translations and reach possible redirections of relevance.

1.3 Objectives of Study

Carried out for the purposes of E<>A scientific translation teaching and training, and adopting a non-pejorative approach to translation assessment, the present corpus-based descriptive translation study **attempts to:** - Redirect translation universals of over/under representation of emphatic particles, intransitivity, and conventionalized passive voice verb forms in contemporary translated scientific Arabic, - Prove, as far as the selected linguistic features realize universal scientific properties, the eligibility of authored historical and modern non-translated Arabic for the description of science.

1.4 Corpus Design

According to the agreed classification of corpora in CTS (Baker 1995, pp.223-243; Olohan 2004, pp.35-44; Laviosa 2003 pp.105-115; Fernandes 2006), the present is a small-scale monolingual comparable corpus set up by the researcher; it consists of three collections of texts. The first and second 'mono-SL' sub-corpora comprise original historical (HASC) and contemporary Scientific Arabic texts (NTASC); the third comprises Arabic scientific translations from English (TASC), as listed below. Similarities of text genre, domain and an attempted relatively comparable length have been observed as far as access to permissions of copyright allowed.

Setting up an individual researcher's own corpus, the case of this paper, is a recognized difficulty of CTS; twenty years ago Baker(1995) predicted the difficulty of affording institutional funding for the monolingual comparable corpora, and Olohan ascertained the predicted case in 2004 (p.49); apprehensions shown by translation publishing institutions and houses against seekers for electronic forms of published translations, my case, is still maintained, too:“ People tend to assume that you want to get hold of their translations to criticize them. This is natural reaction to the kind of discourse that has so far dominated the field” (Baker,1995, 234). Searching for electronic forms of typed and edited versions of Abbasid treatises readable to Sketch Engine has been a lengthy and tiring process. Other established parameters of setting up monolingual corpora in CTS include time span (Baker, 1995, pp.223-243; Olohan, 2004, pp.35-44; Laviosa, 2003 pp.105-115; Fernandes, 2006);

reasons for testing a contemporary language variety against an older one are to provide a balanced time span of selected varieties, henceforth a compilation of a historical original and contemporary non-translated and translated Arabic scientific writing is in question.

Small and large scale corpora are commonly used in CTS. Following Sinclair (1992), Baker (1995, p.226), Olohan (2004), the term is nearly always associated with huge amounts of data extracted from +million-word corpora; however Bowker (1996), Bowker, Cronin, & Pearson (1998) and Olohan (2004) acknowledge the need for small scale corpora. Sometimes, qualitative aspects of analysis may be as or more relevant than the quantitative; therefore “corpus size in CTS is a relative issue” (Fernandes, 2006, p.225). Accordingly, three small scale scientific corpora are compiled:

1- Historical Arabic Scientific Corpus (HASC), 1,127,020 w.

includes three Arabic scientific classics: - ar-Razii's *al-haawee fi t^ʿ-t^ʿib*, (2000 edition; uploaded 2012; [https://www.kutub-pdf.com/book/الحاوي في الطب.html](https://www.kutub-pdf.com/book/الحاوي%20فى%20الطب.html)) - Ibn Sina's *al-Qanuun fi t^ʿ-t^ʿib*, known as *The Canon Of Medicine, I, II, IV, V* (1999 edition; uploaded 2012; <http://waqfeya.com/book.php?bid=7482>) - Ibn al-Haitham's *al-Manaad^ʿer*. (uploaded 2016; [http://www.mediafire.com/file/كتاب+المناظر.rar](http://www.mediafire.com/file/كتاب%20المناظر.rar))

2- Non-translated Arabic Scientific Corpus (NTASC), 156,258w.

is made up of two contemporary authored Arabic books:- M. Mushrafaa's (2013) *al-Qanaabel aḏ-ḏarriyyah, Atomic Bombs* (<https://www.kutub-pdf.com/book/>), - M. Hegazee's (2001) *al-amraad^ʿ al-geldiyyah ?inda al-At'faal, Principles of Pediatric Dermatology* (<http://www.gulfkids.com/pdf/geldea.pdf>)

3- Translated Arabic Scientific Corpus (TASC), 375,669w.

contains three books published by the Cairo-based National Center for Translation: - *at-t'ab^ʿ ṣabr at-t'abu^ʿ*, (2013) ([https://www.kutubpdfcafe.info/kutubpdf/الطبع والتطبع.html](https://www.kutubpdfcafe.info/kutubpdf/الطبع%20والتطبع.html)) Arabic translation of Matt Ridley's (2003) *Nature via Nurture: Genes, Experience, & What Makes Us Human*. - *Mufkilat al-fizyaa?*, (2016) (<https://www.kutubpdfcafe.info/kutubpdf/مشكلة-الفيزياء.htm>), Arabic

translation of Lee Smolin's (2006) *The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next*.

- *Istifraaf al-mustaqbal: thawrat at-tiknolojia an-Nanoniyya*, (2016)

(<http://www.mediafire.com/file/-+استشراف+المستقبل+ثورة+التكنولوجيا+النانوي>

.pdf) Arabic translation of K. Eric Drexler, Chris Peterson, and Gayle Pergamit's (1991) *Unbounding the Future: the Nanotechnology Revolution*.

Random Samples have also been extracted and compiled from the files of each corpus with roughly **equal word counts: Historical RS Corpus, 24538w; Non-translated RS Corpus, 19966w; Translated RS Corpus, 21218w**. These are subjected to a direct manual analysis and a digital one via Sketch Engine, as well. Care has been taken to cut out the samples at appropriate breakpoints and with word counts roughly proportional to the overall word count of each corpus. By random sampling standards as the rule goes in natural sciences, a minimal level of representativeness varies between 2 to 4% and 5 to 10% (Wilson, 2005). Historical RS constitutes 2.2 % of HASC, Non-translated RS 12% of NTASC, and Translated RS 5.7 % of TASC.

1.5 Method and Procedure of data extraction and analysis

Held and compiled in an electronic form readable to Sketch Engine corpus processor, selected Corpora and Random Samples are subjected to a quantitative analysis via steps of search types. The concordance, a most common corpus tool, is applied to show occurrences of a key word, phrase or a collocation cluster. These are entered with co-text selections: right/left tokens and an assigning of the part of speech (PoS) of summoned tokens. This concordance display is also called KWIC. A search by character+ left token is resorted to in the search for /?inna/. Search results are downloaded in txt. files, then converted into docx. files for easier tracing of the sorted concordances, and finally uploaded as Appendices for future verification on Outlook drives via

<https://alsunasuedu->

[my.sharepoint.com/:f:/g/personal/amalabelmaqsoud_alsun_asu_edu_eg/EI2Pn](https://my.sharepoint.com/:f:/g/personal/amalabelmaqsoud_alsun_asu_edu_eg/EI2PnXfiLHhCgo-PR5vHug4BIJVyocyBoR7grehXvNh7hQ?e=JXm0JN)

[XfiLHhCgo-PR5vHug4BIJVyocyBoR7grehXvNh7hQ?e=JXm0JN](https://my.sharepoint.com/:f:/g/personal/amalabelmaqsoud_alsun_asu_edu_eg/EI2PnXfiLHhCgo-PR5vHug4BIJVyocyBoR7grehXvNh7hQ?e=JXm0JN). Upon the search results, frequency lists and tables of particles and verb forms corresponding to selected features are drawn in number and percentage.

Where Sketch Engine Stanford Arabic parser fails to provide a code for a search by tag, a manual quantitative eye-search of the representative random samples is carried out.

2. Theoretical framework

2.1 Translation Universals (TU)

First sprouts for the emergence of TU were laid by Zohar's(1979) polysystem theory and Toury's tripartite model of norms (1978).Both views were a fertile soil for a growing dissatisfaction on the part of translation scholars and theorists towards traditional theories of equivalence and an expressed reluctance to continue the old doctrine of the primacy of the source text to the target, known in DTS as the "pejorative approach" to translation assessment. DTS was then onset with the interface between TU and corpus analysis as a central field of study since the 1990's. TU are "features which typically occur in translated text rather than original utterances and which are not the result of interference" (Baker, 1993, p.243). The term 'descriptive universals' is used by Chesterman (2001): "I call this approach descriptive because, relying on electronic corpora, it doesn't incorporate ideas about what translation should or should not be, but about what translations typically are" (p.7). Olohan (2004) adopts the same definition and taxonomy as set by Baker (1993, 1995); both scholars have collaborated to investigate lexico-syntactic features of comparable English corpora extracted from BNC and TEC (2000).

Baker's simple, clear and applicable taxonomy of TU was set as early as 1993 in her seminal article *Corpus linguistics and translation studies*; since then onwards, DTS and CTS scholars have made considerable contributions; just to name a few; Tirkkonen-Condit (2004) uses Baker's universal of normalization to develop her Unique Items Hypothesis . Chesterman (2001) prefers setting up a model where he distinguishes between S-universals and T-universals; in fact, a considerable overlap between both resists smooth application; Mauranen (2004) introduces interference as a positive and negative impact of source texts.

The probabilistic and hypothetical nature of TU is a common acknowledgment in DTS. Chesterman (2001) maintains that proposals about universals are descriptive hypotheses about (a) the relation between

translation and source texts, (b) the relation between translations and comparable non-translations in the TL. Mauranen and Kujamaki (2004) attribute the controversy of TU to its presumed probabilistic nature. TU, in their view, are a debate between two parties of translation scholars: the first maintains clear-cut findings which support presumed hypotheses about general linguistic properties of translated language; the other believes that reaching conclusions about translation universals is inconceivable since no way is provided for capturing evidence from all translations in all languages (pp.4-9). Toury (2004) writes a full chapter on *Probabilistic explanations in translation studies: Welcome as they are, would they qualify as universals?* Hatim and Mason (2004) prefer a disclaimer of “*it is hypothesized*” in defining the term, TU: “specific characteristics that, *it is hypothesized*, are typical of translated language as distinct from non-translated language” (p.5). **Thus, the variability of TU is evidenced;** they differ in terms of TL, genre comprising both properties and linguistic features, and time span. For instance, conventionalization can by no means be monitored with a constant frequency in all genres of translated Arabic. **Invariability tends to be associated to the properties of a given genre in the source language, target translations and non-translations.** Henceforth, a direct or inverse proportional relation can be held between both poles. For this purpose, the present paper relies for its theoretical framework on: (1) Baker’s universals of conventionalization and normalization, (2) Tirkkonen-Condit Unique Items Hypothesis, (3) Mauranen’s Interference.

Conventionalization is “the strong preference to use what is conventionally grammatical”; this is manifested in interpreters’ tendency to round off unfinished sentences and grammaticize ungrammatical utterances” (Baker, 1993, p.244). This TU applies to the frequency tests of Arabic passive verb forms. **Normalization or exaggeration** is “a general tendency to exaggerate features of the target translated text” or more explicitly to “conform to patterns and practices that are typical of the target language, even to the point of exaggerating them” (Baker 1993, p.245; 1996 quoted in Olohan, 2004, p.91). This has been further developed by Tirkkonen-Condit’s (2004) **hypothesis of over /under represented Unique Items** in the translated language, which is based on

the assumption that “every language has linguistic elements unique in the sense that they lack straightforward counterparts in other languages”. These can be sought on lexical, phrasal, or textual levels and “need not to be untranslatable; they are not similarly manifested in other languages”. However, Condit tends to assume the opposite end of exaggeration as identified by Baker: while exaggeration means an over-representation by translators in the translated texts of elements that are not found in the originals, the Unique Items hypothesis suggests that there is a possible universal tendency in translations to manifest smaller proportions which do not have similarly manifested counterparts in the source language. It is used in Condit’s case to test the frequencies of some verbs and particles using the Translated Finnish Corpus (pp. 177-183). Similarly, the present paper draws on an integrated view of Baker’s and Condit’s over/under representation of linguistic items to test the frequencies of the emphatic particle/?inna, ?inn,?anna/ and the multiple Arabic intransitive verb forms in the compiled corpora.

Interference provides support to a dimension of study observed through corpus-based research (Mauranen, 2004). The impact of SL cannot be entirely excluded from TU. It is assumed in second language learning that an individual’s first language (L1) necessarily influences his/her second language (L2). In translation, the case is reversed: “it is the source language ... that influences the target, usually translator’s L1”. Language levels that are most likely to be affected by Interference are; lexical, syntactic as well as extra-linguistic pragmatic and text-contextual levels. In Mauranen’s view, the difference between Interference and Transfer is that transfer carries the positive face of the commonly assumed negative impact of interference. Comparable corpus analyses are a main tool for her empirical research about interference where the best predictor is contrastive analysis (pp.65-83).

As the quan-qual analysis below shows, representation of passive verb forms in TASC cannot be examined away from translators’ assumed preconception of a consolidated representation of passive structures in Scientific English, or from their presumed arbitrary lower representation of the passive voice in Original Arabic. Also, interference of English conjuncts as *moreover*, *furthermore*, and a translators’ preference to insert an explicit equivalent to the English coordinator ‘*and*’ cannot stand

excluded from TASC monitored occurrences of the passive verb /yudʿaaf/ يُضَاف, *added*, when stuck in almost all hits in the verb phrase of إلى يضاف ذلك, *added to this*; the direct deductive technical meaning of *combining* the verb readily indicates is absented from TASC. Neither can the hands of interference be stripped off the emergence of so-called Arabic auxiliaries as/Qaama bi, tamma/ قام ب, *carried out something, conducted something* in correspondence to a natural high frequency of the English passive voice.

2.2 Comparable Corpora

The need for comparable corpora based on systemic comparisons and contrastive analysis between translations and non-translations was first foreseen by Baker (1993, 1995) as commonly agreed (Olohan, 2004, pp.35-37; Fauntinuoli & Zanettin, 2015, p.3). Her definition of comparable corpora is quoted *as is* in Olohan (2004); used by Laviosa (2003) and considered as a key for the concept by Mason as he terms original texts “spontaneously sourced texts” (2001, p.68) and in Zanettin’s definition where non-translated texts are termed “spontaneously produced texts”(2000 p.106); however, the term ‘non-translated’ texts or ‘non-translations’, as Olohan (2004) asserts, is by now in common use for original texts:

Comparable corpora consist of two separate collections of texts in the same language: one corpus consists of original texts in the language in question and the other consists of translations in that language from a given source language or languages. The corpus of original texts is therefore a monolingual corpus of the type linguists have been using for several decades. Any existing monolingual corpus can be used, provided it is similar in design to the translation corpus. Both corpora should cover a similar domain, variety of language and time span, and be of comparable length. (Baker, 1995, p.234)

HASC, NTASC, and TASC, as abbreviated here, are designed following Baker (1995) and belong to the second type of comparable corpora as listed by Olohan (2004, p.35) following Stig Johanson (2003, p.136): - **Corpora of comparable original texts in two or more languages, - Corpora of original texts and their translations in two or more other languages.**

Inevitable hybridity of non-translations is a major problem in comparable corpus compilation; Fauntinuoli and Zanettin (2015) alert

scholars to the non-originality of some texts thought on compilation to be original; e.g., the Wikipedia in multiple languages which are produced by trans-editing and are partly original and partly translations (p.4). This is part of the reason why the present corpora are compiled by the researcher; firstly, because a corpus of Historical Arabic Scientific works does not yet exist; secondly, there is no guarantee that the proportion dedicated on the International Corpus of Arabic (ICA) for science and technology, 2% of the corpus volume, comprises purely non-translated published works and articles .

Isolationist blind scanning of corpus analysis, a claim much similar to the older rejected doctrine of de-contextualized translation analysis, is now raised as a shortcoming of corpus study; i.e., translated corpora are scanned, frequency lists are drawn, lemmatization, tagging and alignments are conducted away from the extra-linguistic considerations of the translated text-production as who the translator is, his/her motives, sociocultural circumstances of translated text production, constraints of censorship...etc. Therefore, large scale comparable corpora are progressively equipped, as Olohan explains (2004, p.40), with new browsers showing co-texts of search tokens and author backgrounds as in TEC. Likewise, software are updated for these purposes, e.g., Sketch Engine browser shows the co-paragraph of searched tokens and supports full word sketches of a given item. In a quest for neutrally published scientific translations, as presently compiled in TASC, publications of the Cairo-based National Center for Translation are sought. With the historical background of HASC provided and the digital/manual RS analysis incorporated as a procedure secondary to main corpus analysis, claims of a blind decontextualized search types are avoided in this paper. HASC, TASC and NTASC can be stored and shared via Sketch Engine later on as reusable corpora for future research. The aim of this present monolingual comparable corpus study is, partly, as commonly recognized by CTS scholars, analysts and software developers: **to provide evidence-supported clues for the development and refinement of translations.**

2.3 Generic Properties of Scientific Writing

Scientific writing under this study means the written language used in authoring and composing pure scientific books, treatises, factsheets, and

papers. These are different from press scientific and technical articles, manuals, and commercial brochures in terms of usability, orientation, preferred sets of standard and/or corporate-based terminology, and political backdrops of power and hegemony. So, properties of a scientific genre are agreed to be objectivity, informativity, colorlessness, timelessness, logicity; tendency to keep human doers at the back, which generates a prevailing impersonal style; brevity or an economized use of language. Extensive studies have been conducted on English as the language of science to explore linguistic realizations (Dickens *et al.*, 2002; Krein-Kuhle, 2003; Hassnawi, 2003). Objectivity and timelessness linguistically generate a preferable use of the present simple tense, pluralities and nominalization (Pinchuk, 1977). Linguistic features that tend to realize the impersonal style are passive voice structures, intransitive verbs termed as “passive-like verb forms”, or “notional passive verbs”, and abstract subjects (Dickens *et al.*, 2002, pp.191-192); multiplicity of conjugated Arabic intransitives is a best case in Arabic. In SVO /VSO languages as Arabic, logicity and informativity of scientific writing is influenced by the sentence order, e.g., modulations of cause/effect structures. Brevity means a lower clear, concise and precise word count used for the statement of facts; flexible compounding is a best case in English (Dickens *et al.*, 2002, p192), to which I add flexible itemization. To state that features as the prevalence of abstract subjects, present simple tense, passive voice, intransitivity, nominalization, pluralities, flexibly conjugated forms, concise itemization are also common in Arabic non-translated scientific writing needs extensive large- and small-scale corpus studies. This is where this paper belongs.

3. Quantitative and Qualitative Analysis

3.1 Exaggeration, Over-representation of Unique Items &Neutrality and Economy of Scientific Genre: Search for /ʔinna/

Arabic particle /ʔinna/ and, or /ʔanna/ is an accusative, emphatic particle whenever it precedes noun clauses,(subject-predicate); it is accusative, but non-emphatic when followed by the Arabic second simple form/ siiyat al- mudʔaaraʔah/. Before Verb clauses, (VSO), /ʔin/ is used as a conditional particle relating subordinate and main clauses. The semantic function assumed by the initial emphatic /ʔi/ʔnna/when preceding noun

clauses is to foreground an asserted association of the predicate to the subject, namely, /ʔinna/'s clause, e.g., إن الشجاعة في قول الحق , *It is brave to tell the truth*. Emphasis is fully conveyed by occurrences of /ʔinna/ or /ʔanna/ following حيث، حتى، لو، كما، in a left to right order, *in addition to, if, even, given*. Depending on the immediately preceding content word /ʔamelu ʔinna/, the subordinate infinitive /ʔinna/ or /ʔanna/ confers emphasis to its clause, known as /al-masʔdaru l- muʔawwal/, with variable degrees. So, if preceded by verbs of assumption as حسب، ظن، *think or guess*, or subjunctive verbs as، يرجى، *wish, hope*, the emphatic function assumed by default by the infinitive particle contradicts with the preceding ʔamel, hence a less or a zero degree of emphasis. /ʔan/'s infinitive verbal clauses carry no emphasis at all, e.g., أرجو أن تحسن إلى الضعيف , *I hope that you be good to the weak*. /ʔinna,ʔanna/'s infinitive clause can replace any of the main clause elements: subject, predicate, object, genitive head words, noun phrases lying after prepositions (Hasan,1999, V.1, pp.631- 652; Omar & Zahraan, 1994, pp.371-374, Wright,1996,V.2, pp.78-79)The subject and predicate in infinitive clauses formed by /ʔinna/, /ʔanna/or/ʔann/(/al-masʔdaru l- muʔawwal/), are longer in terms of word count than direct infinitive phrases,(/al-masʔdar asʔ- sʔareeh/): e.g.,

لا يعني ذلك حاجة كل الحالات الجدية إلى 2-) قائمة (9 words	لا يعني ذلك ان كل الحالات الجدية تحتاج إلى 1-) قائمة (11words
--	---

A further emphasis and a higher word count are marked in the object noun clause in (1) compared to the object noun phrase in (2).

Attempting to assert an original quality of their target text, contemporary into-Arabic translators and some writers tend to densely use this emphatic particle as an item that peculiarly marks Arabic noun clauses (Subject-predicate), even if no real emphasis is intended in the source text. The TU of exaggeration (Baker, 1995) and over-representation of unique items (Tirrkonen, 2004) can be applied to test the frequency of using emphatic and infinitive /ʔi,ʔanna / in TASC compared to HASC and NTASC. Given the neutrality, objectivity and the economy of language entailed by the language of science, a high frequency of over asserted facts, particularly when already established as basic facts of science, is excluded, and a reduced word count is necessitated. Carried on free and restricted scales, covering the random samples and the entire

three corpora, respectively, the following search attempts to test proportionality between the invariable scientific properties of neutrality and language economy, and the variable TU of exaggerating unique items peculiar to Original Arabic as represented in the use of emphatic and infinitive/?i,?anna/. The interrelated free quan-qual and restricted quantitative scales of the search can be elaborated as follows:

3.1.1 Restricted search for initial /?inna/ positioned at the beginning of sentences in the entire 3 corpora (a quantitative corpus-analysis)

For the purpose of showing the query types utilized to reach this search, it has been called the full stop+/?inna/ search. Within Sketch Engine corpus-analyzer, a full stop is placed in the character search; /?inna, ?anna/is placed in the word form filter; hence, no left or right tokens are specified; it is worth noting that Sketch Engine’s parser doesn’t read the difference between /?anna/, /?inna/ and/?in/, which means all possible meanings of the particle can be reached through this full stop+?inna search. The results came as follows:

Corpus	Word Count	Total Hits of % /?i,?anna,?inn/	Approx. %
HASC	1,127,020	14	0.001242214%
NTASC	156,258	46	0.0294384927%
TASC	375,669	284	0.0755984657%

Table 1: Sketch Engine fullstop+?inna frequency search result.

Discussion: Linguistic analysis of the HASC **14** hits of initial/?inna, ?anna, ?in/ monitored by Sketch Engine’s concordance are found to be conditional, only two of them are emphatic, i.e., a **minor occurrence of initial emphatic /?inna /** is registered as Appendix1-1 shows. The **46** occurrences in NTASC are categorized after scanning as initial emphatic, with a single hit in Mushraffa’s *al-Qanaabel*(Appendix 1-2). Scanning the **284** hits of TASC, **all are found out to be initial emphatic /?inna/ lying at the beginning of sentences** (Appendix 1-3;) as cross-sections show in Appendix 1-7, Figure1. This relatively high frequency compared to the Historical and Non-translated corpora is a strong indication of translators’ tendency to use initial emphatic /?i,?anna/. Worth to note is the parallel alignment with one of the source books doesn’t show a parallel use of conjuncts of assertion and /or English marked themes that

might have entailed equivalent markers of assertion and emphasis; a random selection of a full text in Figure 2, Appendix 1-7, shows 6 occurrences of initial emphatic /?inna/ to which unmarked theme structures are plainly observed in the source.

3.1.2 Free search for all possible occurrences of /?inna, ?anna/, emphatic, infinitive and conditional, in the RS's. With the RS's, a simple query search for the particle is used, a single right token is selected with an unspecified part of speech, PoS, to the effect that parts of speech of the word(token) immediately occurring after /?inna, ?anna/ incorporate nouns and verbs, in addition to all possible meanings of the particle. This is followed by an eye scrutinized analysis of all occurrences. Search results are classified as follows:

RS	Word Count	Total no. of hits	Infinitive /?inna,/?anna/ hits			Emphatic hits (Initial/inter-sentential)		Conditional /?in/
			Infinitive emphatic	Infinitive non-emphatic (followed by verbs)	% of infinitive occurrences to total hits	Hits	% of emphatic occurrences to total hits	
Historical	245 38w	18 9	58	3	32.27%	67	35.44%	15
Non-translated	199 66w	53 8	188	115	56.31%	20 0	37.1%	6
Translated	212 18w	23 2	148	2	64.6%	46	19.82%	0

Table 2: Frequency of Infinitive, initial emphatic/?inna,?anna/ and conditional /?inn / in the Random Samples

Discussion: Quantitatively, infinitive hits in the Translated RS, registering **64.6%** of the total hits, are double their Historical RS peers, **32.27%** of the total hits, and surpass the Non-translated hits by **8.3%**. There is a relative balance of using the infinitive particle and the emphatic (initial or inter sentential) in the Historical RS, **32.27%**, and **35.44%**, but the same doesn't hold true to neither of the Translated nor the Non-translated RS. The case is more acutely observed in the Translated RS,

with a **19.83%** for the emphatic particle and a **64.6%** for the infinitive. Qualitatively, upon these frequency hits and calculated percentages, some findings can be drawn. The use of emphatic /ʔinna,ʔanna/ in the Historical RS whether occurring at the beginning of and between full main clauses following the Arabic coordinator /fa/, or lying as an emphatic particle preceded by the consequential /fa/ in the main clause of a conditional structure /jawaab af-fard^s/is confined to underscore asserted observations and conclusions as in:

file5660566 اذا استغرق معظم الشمس ب القياس الى موضع من الارض ف > ان < مركز الشمس في ذلك الوقت مستتر . عن ذلك الموضع

The full stop+ʔinna Sketch Engine query results (Appendix 1-1) confirm a rare resort to the emphatic particle after full stops in HASC, i.e., at the beginning of main sentences. A cross-section extracted from the Historical RS (Appendix 1-7, Figure 3) provides a piece of evidence.

Worthy to note is that the hits where the initial emphatic particle is preceded by the Arabic consequential coordinator /fa/ are **63** out of the total **67** registered across the Historical RS; itself is a good evidence that Ancient Arab Scientists tended to use emphatic /ʔi,ʔanna/ for an assertion of conclusions and an underscore of observations as an auxiliary deductive, conclusive and emphatic particle, **but rarely for emphasizing already established facts**. A similar phenomenon has been observed throughout both files of the Non-translated RS; out of the total 200 initial emphatic hits, 53 were semantically figured out as deductive and conclusive, positioned in almost all cases in the conditional main clauses, or as a conclusion to an introduced premise, and preceded by consequential /fa/ as in:

file5730627 ان ه اذا تحرك جسيم مكهرب في مجال مغناطيسي ف > ان < ه يتحرك في دائرة , و يتوقف قطر الدائرة على سرعة

file5730627 ف اذا بدا جسيم في الحركة ثم ازدادت سرعة ه ف > ان < الدائرة التي يتحرك في ها , يكبر قطرها ; و ب ذلك يتحرك

29 of them are spotted in *al-Qanaabel* (Musharafa, 2013) alone. Figure 4, Appendix 1-7, incorporates 6 examples (cross-sectioned from Appendix 1-6_RS_Nontranslated, pp.16-18). The same conclusive and deductive function is a rare observation through the frequency of the initial emphatic /ʔi,ʔanna/ within the Translated RS; in the majority of cases the particle is arbitrarily positioned as emphatic due to its occurrence following to ،غير ،

كما ، بيد، حيث، إذ، كما or appearing after full stops, at the beginning of main sentences where its use is almost dispensable, e.g.:

بين هم , و كتب نتائج هذا البحث ب فرحة الانتصار . < ان > التوائم المتشابهة منذ الولادة تتشابه في : file5660502 جميع اماكن كل من هما

Only 10 hits are preceded by the consequential /fa/and used to highlight a fact, a prediction, or a contrast: Figure 5, Appendix 1-7, displays a compiled representative snapshot; grey highlight marks futurity or a repeated idea (cross-sectioned from Appendix 1-5_RS_ Translated, pp.2, 7, 8). However, a similar trait of a dispensable use of the arbitrarily initial emphatic particle is traced in the Non-translated RS first file, al ?amraad, *Pediatric Dermatology*, too; Figure 6, Appendix 1-7, shows a predominant use of ان كما (cross-sectioned from Appendix1-6_RS_Non-translated, pp.1,2);

e.g.: file5660598 لا تتعدى منطقة اتصال البشرة مع الادمة . كما < ان > الادمة تحوي شبكة من الالياف العصبية file5660598 ينتهي في جراب الشعر و الاخر ل ه نهايات منتفخة . كما < ان > هناك نهايات عصبية متخصصة في الادمة

The additive /kamaa/ plus the initial emphatic /?inna/ can be replaced by the coordinator/wa/, e.g., كما ان التغييرات نتيجة الاختلاف في نسبة مساحة الجلد تؤدي الى , literally, *this is in addition to the fact that changes occurring due to a variable dermal space cause...* , can be replaced with كما تؤدي التغييرات , or , وتؤدي التغييرات , literally, *Changes also lead to...*

Obvious dispensability of initial emphatic /?i/?anna/across the entire TASC and Translated RS, and some selections from the Non-translated RS as traced through the free query and restricted searches provides a strong sign on the overrepresentation of the particle as a unique item in TASC. Furthermore, the full stop +?i/?anna restricted character search in TASC mirrors a prevailing and constant use of the emphatic particle to assert facts that are already established, and a rare resort to the particle to highlight a conclusion or deduce a result. **High frequency of infinitive verb- and noun- clause hits in the Translated RS (64.6%) and the Non-translated RS (56.31%) compared to a much lower frequency monitored in the Historical peer hits (32.27%) asserts:** 1- an over-representation of /al-mas{daru l- mu?awwal/ in the Translated and Non-translated Corpora, alike, 2- a possible dispensability of the infinitive particle through the Translated and the Non-translated corpora; this is in view of the fact that infinitive /?i,?anna/ preceding noun clauses, i.e., subject predicate, carries a variable emphatic function and before verbs carries a zero degree of emphasis. Given the higher word count involved in the use of /?i,?anna/

and /ʔan/ infinitive clauses/al-masʕdar al-muʔawwal/, than direct infinitive phrases,/al-masʕdar asʕ- sʕareeh/, excessive resort to the infinitive particle runs counter to the dictates of the scientific property of language economy. **Thus, TU of exaggeration of unique items as represented in the findings of the Search for/ʔi/ʔanna,ʔan/ is inversely proportional to the generic scientific properties of neutrality and economy of language.** Henceforth, into Arabic translators are strongly urged to forsake a repetitive, void use of the emphatic particle /ʔi,ʔanna/ on producing target scientific Arabic .

Ancient Arab scientists, who originally set through their treatises the rules for science and the language appropriate for the description of science, medicine and pharmaceutical formulas, had no intention to prescribe for the scientific language. It is only through linguistic analysis that the universal features of Scientific Arabic can be figured out and established.

3.2 (Under)-representation of Unique items & Impersonal style: Intransitivity in Scientific Arabic

Intransitivity is recognized as a main feature realizing the impersonal style of scientific language. Description of diseases, prescription of medication, examination of cosmic manifestations and physical phenomena involve subjects who are not the true doers of actions, e.g., *Outbreaks of the disease occur in mental hospitals* (Dickens *et al.*, 2002, pp.191- 192). The real doer of the verb ‘*occur*’ is the eye that monitors the given occurrence, or the scientists, health care institutions who have detected the disease outbreak. Scientific objectivity entails an impersonal style to express facts whereby humans as scientists, physicians, chemists are kept at the background; phenomena, their direct causes, concrete observations and direct findings are left to speak for themselves. English, the lingua franca of modern scientific writing, is recognized to incorporate a frequent occurrence of intransitive ‘passive-like’ verbs (Dickens *et al.* 2002, p.192), the same syntactic category is recognized as a general feature of English, which is known as ‘notional passive’ (Jespersen 1961, p.365), or pseudo intransitive constructions (Bakir 1994, p.40-42)

With the lack of a code for Arabic intransitive verbs on Sketch Engine’s CQL tagger, a quantitative search of the full three corpora grew

impossible; instead, a direct eye search of the RS's is conducted. Through the quantitative search of the RS's and a quan/qual analysis of cross-sections extracted from each, the paper attempts to: 1- test the (under)-representation of the Arabic intransitive verb (IV) forms in TASC compared to HASC and NTASC; 2- subsequently, explore proportionality between Tirrkonen's variable TU of under-representation of unique items and the invariable generic scientific property of impersonal style. A question remains: why to consider Arabic intransitive verbs a unique feature of Original Arabic? Intransitivity in Arabic per se may not be unique, but the diversity of the consonant-vowel root clusters of Arabic intransitive verb forms, the morphological richness of the Arabic verb system, alongside the semantic functions each form assumes and a translator's/writer's competence for a multiple use in scientific writing are what is unique about it.

To convey a full meaning, an intransitive Arabic verb suffices with its subject (doer of the action) and does not need an object unless at some cases through a preposition: *مشى، ظهر*, *to appear, to walk*, are intransitive verbs while *قعد في الحجرة ، انتهى من الدرس*, *sit down in the room, get done with one's lesson*, are examples of verbs which are intransitive in form and transitive in content. In this paper this type of formal intransitive verbs are counted as intransitives. Grammarians of Arabic have relatively identified a number of consonant-vowel root cluster forms for coining intransitive Arabic verbs (Hasan, 1999, 2, pp.150-157; Shalash, 1971, pp.191-196), these are known to European scholars as basic forms of the Arabic trilateral verb form and its conjugations in addition to the quadrilateral verb forms (Wright, 1996, 1, pp.30-35). Intransitivity forms and the semantic functions each assumes are clearly identified by Arab grammarians; the most common of which are listed as follows (Hasan, 1999, 2, pp.150-157; Shalash, 1971):

Intransitive verb forms	Semantic function	Example	Literal meaning in English
1-/ faʕala/	The most common intransitive Arabic verb which indicates a happening of an action provided that it is not followed by an object	حدث- قام- ظهر	to happen, to stand up, to emerge

/faʕala, faʕila/	expresses a change from a state to another	ذل الضعيف - قوى الرجل	The weak got humiliated, the man grew more powerful
2-/faʕula/	indicates a permanent attribute associated to the subject	قَصُر - نَحَف	to grow shorter, to grow thinner
3-/faʕila, faʕula/	conveys a temporary state the doer is subjected to; it soon comes to an end when the factor causing it disappears	مرض المتعرض للعدوى	The infected got sick.
4- /ifʕalla/	indicates a change of color	احمرّ	It Changed into red
5- /ifʕalalla/	expresses a sudden state	إقشعر	- to feel an inner chill
6-/infaʕala, istaʕala, ʔaʕala/	a change from a state into another	انطلق - استأسد القط - أمكن	- to go ahead - the cat came to act like a lion - to become possible
7- Modulation formats, sʕiiyat al-mutʕaawafah: /infaʕala/ /tafaʕʕala/, /i ftaʕala, tafaaʕala/	An action the subject is caused to do, or a state in which it is driven into	دفعت الباب -فانفتح - وفرت المال - فتوفّر - مددت الحديد - فامتد	- I pushed the door, so it opened. - I provided the money and it became available
8- /tafaʕlala/		تدحرج	to roll down

Table 3: Consonant-vowel root clusters of Arabic IV forms

Obviously, semantic functions assumed by each form contribute to the impersonal style. The following table shows the number of RS's intransitive Arabic verbs classified according to the above-listed root clusters. The form of /ʔifʕalalla/ registered nil results in all RS's; so it is excluded from Table 4.

RS	W.	Total Intransitive Hits	% of total word count	/faʕala/	/faʕula/	/faʕila/	/ifʕalla/	/ʔafʕala infaʕala, istafʕala/	Modulation formats, sʕiiyat al-muʕaawafah: /tafaʕʕala, iftaʕʕala, tafaaʕʕala/	/tafaʕʕala/
Historical	24 53 8	579	2.3 5 %							
al-haawee		127		5 8	7	12	1	23	19	7
al-Qanuun I		129		3 8	1 6	5	0	27	38	5
al-Qanuun II		65		3 8	2	0	0	3	17	0
al-Qanuun IV		131		4 7	1 2	16	0	13	29	0
al-Qanuun V		5		1	0	1	0	1	2	0
al- manaaðʕer		122		2 9	2 3	11	0	18	41	0
Non-translated	19 96 6	396	1.9 9 %							
alʔamraadʕ		143		6 4	1	0	0	22	56	0
al-qanabel		253		1 0 7	9	4	0	36	97	0
Translated	21 21 8	303	1.4 %							
at-ʕʕabʕʕar at- tʕʕabuʕ		71		3 7	1	1	0	3	29	
Muʕkilat al- fizyaaʔ		158		6 3	0	1	0	10	83	1
Istifraaf al-		74		3	0	0	0	11	27	0

mustaqbal				6					
-----------	--	--	--	---	--	--	--	--	--

Table 4: Frequency of Arabic Intransitive Verb-forms in the RS's

Discussion: Frequency of Translated RS intransitivity, **1.4%**, as the figures above show, registers roughly half the rate compared to a peer tendency of intransitive verb use in the Historical, **2.35%**, and stands lower than the Non-translated RS's, **1.99%**. This provides a strong sign on the low tendency of contemporary Arabic translators to use IV forms in scientific contexts. Furthermore, categorized scanning of the diverse intransitive root clusters proves a poor multiplicity of IV forms in the translated RS compared to a diversified and spontaneous norm of IV use in the Historical and Non-translated RS's alike. A more balanced distribution of the multiple intransitive root clusters is noticed in the Historical RS. Verb forms as /faʕala/, /faʕila/, /ifʕalla/ show few or rare occurrences in the Non-translated and Translated RS's, respectively. Three main forms are predominant in the Non-translated RS: /faʕala/, /ʔafʕala, infaʕala, istafʕala/ and modulation forms. Translated RS intransitive hits are concentrated in the two forms of /faʕala/ and modulation forms. It might be useful to seek more evidence by referring to relevant cross-sections (Figure 1, Appendix2-4; Appendix 2-1_RS_Historical, pp.59-60).

In his description of chickenpox, Ibn Sina shows a high awareness of the impersonal style exemplified in his frequent resort to diverse root clusters of IV forms (Table 5). 7 hits are observed within the first line of Figure1, Appendix 2-4, to describe the appearance of the disease. Relying on 24 hits (grey highlights) of IV use (10% of the total 262 paragraph words, Figure 1, Appendix 2-4) alongside the passive voice (underlined verb forms) and a didactic imperative mood(italicized verb forms), the ancient physician details the later advanced stages of the disease and interrelated symptoms shared by chickenpox and measles.

/faʕala/	/faʕula/	/ifʕall a/	/infaʕal a/	Modulation formats,	/tafaʕlala /
يسوء- يظهر- يموت- يجري- يطراً- يغوريلج	يعسر- يبطن - يبرد- يجدر	تخضر	تندفع	يختلف- يتتابع- يشدد- اجتمع- تزايد	تتأذى

Table 5

A similar awareness of the impersonal style is observed in *al-Qanaabel* . The 170 w paragraph in Figure 2, Appendix 2-4 (cross-sectioned from Appendix 2-2_RS_Non-translated, p.51) manifestly shows a multiple use of Arabic IV root clusters through 15 hits classified in Table 6:

/faʕala/	/infaʕala/	Modulation formats, sʕiiyat al-mutʕaawaʕah
فتخرج- يحدث- ينشأ دار	ينحصر	تكتسب- يزداد- فيقترب- ازدادت - يتفق

Table 6

Constellations of a dense use of Arabic IV forms are not recurrent in the Translated Corpus. The rule throughout the Translated RS is that two or four occurrences are figured out each paragraph. Sometimes, all of them might follow a single root cluster within a paragraph of 131 words, e.g.

; extract in Figure 3, Appendix 2-4 is a case in point

(Appendix 2-3_RS_Translated, p.37-38). اكتمل-امتد-اتفق-انضح

The maximum number of Translated RS hits per paragraph might reach 10. More than often a random parallel alignment shows equal intransitive hits in the target Arabic compared to the English source cross-sections; a poorer multiplicity of IV root clusters is now an ascertained observation in the translated RS. Considerations of translation equivalence, translation shifts or translators' resort to explicitation may cause higher occurrences in the target compared to the source as exemplified in Figure

4, Appendix 2-4, where three IV forms appear:

/faʕala/	/ʔinfaʕala/	Modulation formats, sʕiiyat al-mutʕaawaʕah
نشأ- تعود- تعمل	انعكس- انخفضت	ترجع- توقف

Table7

Two intransitive hits are explicit occurrences that are not found in the source English extract, نشأ- تعمل على. Verb *to be* is a most common intransitive verb in English to which a variety of subject- predicate structures and passive voice forms in Arabic are used away from intransitive verb forms. The verb/kaana/, its mate-verbs/ʔsʕbaħa, ʔdʕħaa, ʔamsaa, baata/and their derivative forms are excluded from the Arabic intransitive verb search in RS's; by Arabic grammar rules, these are deficient verbs not intransitives.

Search for /qaama bi/: Verb /qaama bi/ قام ب is counted within the frequency of intransitive hits in the translated RS. It is intransitive in form, yet transitive in content. It has become a norm by contemporary Arabic translators to modulate a number of English transitive verbs into the use of قام ب+ المصدر. Frequency of verb/qaama/, *do something*, registers **166 hits**, approximately, **0.044%** of TASC (375,669 w). In **99.8%** of the hits (Appendix3-3_Search for Qaama _Translated Corpus) the verb, "قام " /qaama/ is followed by " ب /bi/ then an infinitive derived from the main verb form that could have otherwise been used: e.g. قام باكتشاف instead of *to pay a visit* instead of *to visit* and *to carry out a discovery* instead of *to discover*. Henceforth, /qaama/ is used by translators as an auxiliary form that may facilitate more flexible use of a grammatical pattern. The result is an intended or unintended resort to explicitation; itself runs counter to the economy of language as one of the prime properties of a scientific language. Some of the registered hits could have otherwise been reduced into one word instead of two or three, were the different patterns of Arabic transitive and/or intransitive verb forms used: e.g. قام علماء السلوك بتطبيق , *behaviorists carried an application on...* could have been rephrased as طبق علماء السلوك , *behaviorists applied*. Out of the **14 hits** of /qaama/ in NTASC, 13 are accompanied by /bi/ (Appendix 3-2_Search for Qaama_Non-translated Corpus). In HASC, (1,127,020 w), there are **only 11 hits of /qaama/**, with no single occurrence of /qaama bi /+ noun, pronoun or nominalization (Appendix3-1_Search for Qaama _Historical Corpus).

Assuming that the intransitive hits in the Translated RS are roughly equal, if not higher due to explicitation, translation shifts and the use of /qaama bi/, to their original English peers and given the ascertained quantitative ratio as shown by the RS's above, it can be safely assumed that, as far as the present Corpus-based study is concerned, Original Historical and Contemporary Scientific Arabic writing surpasses Original English in terms of intransitivity by a ratio of 2:1; itself is a proof that intransitivity is densely used in Scientific Original Arabic writing, let alone the multiplicity of root clusters. By so assuming, Translated Arabic cannot be held responsible for a lower representation of intransitivity, which is shown by now to be a direct interference from English.

In short, the comparable corpus analysis of intransitivity in the RS's shows a rate of **seven** IV multiple root clusters in the Historical RS compared to **two** most common forms in the translated RS. Quantitative analysis reflects a ratio of **2:1** between the Historical and the translated

RS and 2:1.4 between the Non-translated and Translated RS's. Henceforth, a multiple and a more use of IV root clusters can safely be recommended for an aware use in translated Scientific Arabic. Inverse proportionality is shown to stand between the constant generic property of impersonal style and variable TU of under-representation of:1-intransitivity forms in the translated Arabic RS as a direct reflection of Original English, 2- multiple Arabic IV forms (Translated RS).

3.3 Conventionalization and the Impersonal Style: Passive Verb forms in Scientific Arabic

Linguists and translation theorists agree on the common no-agent nature of using the voice grammatical subcategory of the passive in most languages; even when agentive structures occur within the passive constructions, they are less foregrounded or usually play a role secondary to the effected entity. Baker (2011) asserts that the main function of the passive in English and in a number of other languages is “to avoid specifying the agent and to give an impression of objectivity” (p.116). In Arabic, attention is directed more to the person affected by the act, ‘the patient’ than to the doer of it ‘the agent’ (Wright 1996, pp. 49-50). Out of the three possibilities Lyons (1968) lists for the passive two are passive with a non-specific agent, *Bill was killed by someone*, and passive without mention of the agent, *Bill was killed* (p.376). The reader- hearer’s attention is focused on the goal rather than the agent in the receptive process- oriented clauses of the passive structures in Halliday’s (1967) theorizing (p.41). Predominant use of the passive voice in English, particularly technical and scientific writing, is ascertained to the extent of becoming a norm or a convention: “Scientific and technical writing in English relies heavily on passive structures. This is done to give the impression of objectivity and to distance the writer from the statements made in the text. It has come to represent the norm in technical writing...” (Baker, 2011, p.193). Accordingly, a direct proportional relation between using the passive voice grammatical subcategory which realizes the constant scientific properties of objectivity and impersonal style is well established in English scientific and technical writing (Dickens *et al.*, 2002, pp. 191-192).

The question is whether the same convention is maintained in translated scientific Arabic corpora by way of a norm (as the case is in HASC) or by way of interference from English due to E/A translations. The first part of the question entails a brief overview of the norm of predominant or less common use of passive voice in modern Arabic writing; while the second part involves a talk on conventionalization and Arabic translators' tendency to or escape from the passive voice use presuming its oppositeness to the norm in Arabic.

Arabic passive is formed to fulfill a set of purposes and situations entailing an omission of the doer or actor only to be replaced by the effected entity, patient or object of the act as above explained (Hasan 1999, 2, p.97). Purposes include verbal functions as intended brevity, i.e., economy of language, e.g., *لما فاز أحمد في السباق كوفيء*, *when Ahmad won the race, he was rewarded*, with no exact naming or identification of the rewarder. Rhymic appropriation is the second verbal purpose for passive voice formation, e.g., *من حسن عمله عُرف فضله*, *the one with good deeds is recognized indeed*, (p.97).

Semantic functions assumed by the Arabic passive voice include:

- indication of Allah Almighty or some higher entity as the doer of the act (Wright 1996, 1, p. 50); e.g., *جُبِلَتِ النُّفُوسُ عَلَى حُبِّ مَنْ يَحْسُنُ إِلَيْهَا*, *souls are made to love those who do good to them*. It is already understood that it is Allah Almighty who has created souls with such innate liking (Hasan 1999, 2, p.97).

- drawing the hearer/reader's attention to the happening of the act or the fact that the process is already done particularly when the speaker doesn't wish to name the doer or the doer is unknown; e.g., *ضُرِبَ زَيْدٌ*, *Zayd was hit*, and the speaker is the one who has hit him (97-98). This particular semantic purpose proves that the Arabic passive verb form itself contributes to achieve a function away from its agent omission. The same stylistic and rhetorical function is assumed when passive verb forms are used to describe the grave events of dooms day in the Holy Quran (Mohammad 2006, p.48):

إِذَا الشَّمْسُ كُوِّرَتْ وَإِذَا النُّجُومُ انْكَدَرَتْ وَإِذَا الْجِبَالُ سُيِّرَتْ وَإِذَا الْعِشَارُ عُطِّلَتْ (التكوير ١-٤)

When the sun shall be darkened, when the stars shall be thrown down, when the mountains shall be set moving, when the pregnant camels shall

be neglected (Surah of at-Takweer, *Darkening*, verses1-4; trans, A.J. Arberry1982, p.637). This is of observable relevance to scientific writing where objectivity is a common prerequisite as the act or process is what matters not observers, physicians, or scientists.

- foregrounding of the effected entity whether with omission of the agent or with a complete agentive prepositional phrase; e.g., (أمن الرسول بما أنزل) إليه من ربه (البقرة ٢٨٥), *the Messenger has believed in what is revealed to him from his Lord* (Surah of al-Baqara, *The Cow*, verse 285). This is termed by modern Arab grammarians a full passive construction (Mohammad, 2006, p.46). Thus, an agentive passive is an original feature of the Arabic passive voice, not a recent development as claimed by some western writers as Cantarino (1975).

- implying a diversity and plurality of the omitted actor and universality of the reasons for the act performed; e.g., (إتق الله أخذته العزة بالإثم (البقرة ٢٠٦)) وإذا قيل له *whenever he is told to fear Allah, he gets overwhelmed with a sinful pride* (Surah of al-Baqara, *The Cow*, verse 206). A plurality of advisers or tellers is well indicated by the passive voice. Properties of brevity, plurality and informativity of scientific writing entail a need for this purpose.

- a set of other purposes are recognized by ancient and modern Arab grammarians as fear of or glorifying, and /or an intended undermining of the omitted actor (Hasan, 1999,2, p.97-108; Mohammad, 2006, p.56).

Some Arab grammarians have attributed the less common use of the passive voice in modern Arabic writing to a number of reasons. First, heavy pronunciation of the Arabic trilateral verb when conjugated as a passive verb form with an /u/ placed after the first radical consonant /f/ and an/i/ placed immediately after the second radical; e.g., /akala/, /ukila/(Mohammad 2006, pp.21-22). Second, more than often, the passive voice has been tackled as a sub-category to the active voice, based on a claim that each of the rhetorical and semantic functions assumed is only due to an omission of the agent and an utter overlooking of the functions attributed to the passive verb form itself (40). This, in my view, has partly led modern language users to consider the Arabic active voice as a norm or a conventional rule and to rely much less on the passive voice. Third, Arabic draws on the rich morphological resources of the verb system

producing the multiple varieties of the derived forms of the strong verb as termed by Wright (1996), notional passive structures as opposite to true formal passive structures (Khafaji, 2000, pp. 32) or the multiple root clusters of the intransitive Arabic verbs as enumerated above. Fourth, the passive verb forms have been widely replaced in colloquial Arabic varieties with the derived forms of the strong verb (Mohammad, 2006, pp.26-27); e.g., / qutila/, *murdered* is replaced in some Arabic colloquial varieties with /ʔitqatal, or, ʔinqatala/. Fifth, to avoid the passive verb forms there appeared a tendency by modern users and translators by default to develop new smoother grammatical patterns of active forms comprising auxiliaries as /tamma/, *done* and /qaama bi /, *conducted* + the Arabic infinitive; *تم انجاز مجموعة من المشروعات*, *a number of projects have been achieved*. The same view is asserted by Rosenhouse (1988) and Khafaji (2000).

This phenomenon of smoother active voice constructions is proved through a corpus-based scanning of the entire three corpora at hand. Throughout the **26** hits of /tamma+infinitive/ in HASC the verb /tamma/ means *to complete* or *finish* as *تم الانكسار*, *digested*, *said*, *refracted*; while TASC registers **562** occurrences of /tamma+infinitive/ where the verb / tamma/ is used with the meaning of the auxiliary, *do*; e.g., *تم قطع*, *cut off*, *made*, *replaced*; only **14** hits are monitored through the simple query search on Sketch Engine's processor where both usages of the verb are observed (Appendix 4-1, 4-3_Search for tamma_HASC, TASC).

Obviously, due to these reasons, a less common use of the Arabic passive came to be the norm in modern Arabic writing. In their ever-maintained quest for originality of the target product, translators always mimic original writing and consider some of its norms wherever, and whenever applicable as conventions. Scanning the RS's for the Arabic formal passive verb forms proves a virtual tendency of translators to avoid a passive verb direct use due to the presumed conventionalized use of the active verb forms as a rule of norm. Again, the RS's are eye-searched as Sketch Engine's tagger does not support a code for the Arabic passive verb forms. A higher frequency is observed in the Historical RS with **274** hits **1.11%** of the total 24604 w, this grows a bit lower in the

Non-translated RS with **184** hits, **0.92 %** of the 19966 w, and finally reaches an all-time low in the translated RS, in spite of the recognized prevalence of the passive voice in the source English as above evidenced, with **42** hits, **0.19%** of the total 21218 w.

RS's	Passive verb form hits
Historical (24538 w)	
1- al-ḥaawee 4474 w	45
2- al-Qanuun I 4369 w	22
al-Qanuun II 3856 w	109
al-Qanuun VI 4271 w	87
al-Qanuun V 3243 w	8
3-al-Manaaḏ'er 4325 w	6
	Total 274, 1.11%
Non-translated (19966 w)	
1-al-qanaabel aḏ-ḏarriyyah 9736 w	87
2- alʔamraad' al-geldiyyah 10230 w	97
	Total 184, 0.92 %
Translated (21218 W)	
1- at't'ab' ṣar at-t'abu' 4821 w	22
2-mushkilat al- fizyaa' (in 3 parts) 11675 w	16
3- Istifraaf al-mustaqbal 4200 w	4
	Total 42, 0.19%

(Table 8: Arabic passive verb form occurrences in RS's)

Cross sections help provide qualitative elaborations. Passive verb form frequencies are densely monitored in Ibn Sina's *al-Qanuun II* with an obvious integration with intransitive verb forms (*italicized and underlined in selected cross-sections*); ^{٨٢} w extract in Figure 5-a, Appendix 2-4 (cross-sectioned from Appendix 2-1_RS_Historical pp.50-51) describes medicinal benefits of the pumpkin, where 8 passive verb hits (*highlighted in grey*) appear. 24 passive verb occurrences and 2 intransitives are used in a 123w extract from *al-Qanuun IV* to detail a pharmaceutical prescription for preparing an anti-blood fever analgesic medication at cases of measles and chickenpox (Figure 5-b, Appendix 2-4, cross-sectioned from Appendix 2-

1_RS_Historical, pp.61- 62). Similarly, an integration of 6 passive verb forms and 8 intransitive verbs are incorporated in a 237 w Non-translated cross section from *al-Qanaabel* (Figure 6, Appendix 2-4, cross-sectioned from Appendix 2-2_RS_Non-translated, p.53)

A comparable search for some passive verbs found in the Non-translated RS has been conducted by means of Sketch Engine's simple and phrase concordance across the entire TASC and NTASC. Interesting findings are reached: 1- /Yuqaas bi, tuqaas bi/ verb forms, *measured by*, with the feminine marker in the second, are used 6 times in NTASC and only 3 in TASC; whereas / tamma qiyaas / shows nil results in NTASC and 5 hits in TASC (Appendix 5-2_Search for yuqaas_ TASC&NTASC; Appendix 5-3_Search for tamma qiyaas) 2- / d^f uriba/، ضُرب، *multiplied by*, as in إذا ضُربت السرعة في ثلاثة ضُرب الارتفاع ,shows empty results in TASC with only two noun occurrences of the same sense, *multiplication* , 3- Musharrafa uses some everyday simple passive structures as: تدار الآلات, يُضاف إلى المادة to express some mathematical or chemical operations. Looking for the passive verb /yud^faaf/, يُضاف in TASC, KWIC concordance hits 37 occurrences; in 36 hits it appears in the phrase يضاف إلى ذلك ,*added to this*; resorted to by translators to provide an Arabic explicitation equivalent to the coordinator *and*, and conjuncts as *moreover* and *furthermore*, *in addition to*, or to express intra/ inter-sentential connectedness in Arabic where necessary or not.. The 7 hits of /yud^faaf/, *added*, in NTASC are to denote a chemical or medicinal compound as, يضاف إلى تلك العلاجات, يضاف إليها مضادات الاكسدة, يضاف إلى الماء عند, *added to water on...*, *added to anti-oxidants*, *added to medications*. (Appendix5-1_ Search for yudaaf _TASC&NTASC) 4- Passive verbs as أُستحدث, *introduced*, appear 4 times in *al-Qanaabel*, and shows empty results on the KWIC concordance via Sketch Engine when searched in TASC. Instead, alternative active voice structures as, وذلك تم ادخال, مع ادخال are recurrently used in TASC (Appendix 5-4_Search for ustuhditha _TASC&NTASC.)

These results indicate that ancient Arab Scientists were more aware of the functions of the Arabic passive voice and tended to integrate it with intransitivity to serve the impersonal and objective purposes of science and medicine; their use came true, natural and expressive of an Arabic

norm where both active and passive are two reliable voice categories in a once lingua franca of science. Similarly, if not as perfectly, modern Arab scientists carried on with the same practice. In view of the descriptive quality of the present study, contemporary translators of scientific Arabic are urged to reconsider a more representative use of the passive voice. The aim is not to take them away from using the active voice or to replace each active verb form with a modulated passive; rather, it is always “the semantic function of a category rather than the form it takes that is of paramount importance in translation” (Baker 2011, p.119). According to the findings monitored in the RS’s and subsequent observations drawn from the entire three Corpora, a conventionalized avoidance of the Arabic passive verbs in the Translated RS is inversely proportional to the impersonal style.

Conclusion

Objectives of the present study are set to provide evidenced clues for reconsidering translation universals of over/under representation of emphatic particles, intransitivity, and conventionalized passive voice verb forms in contemporary translated scientific Arabic, and to explore the authentic eligibility of Original Arabic for scientific writing. Both objectives are supported by the findings of the quan-qual corpus analysis of corpora and RS’s, and manual quan-qual search of RS’s, as follows:

- High frequency of using initial emphatic /ʔi,ʔanna/ in TASC (284 hits,0,08%) compared to minor occurrences in HASC (14hits, 0,001%) and NTASC (46 hits, 0,03%) – entire corpora are subjected to a query search restricted by character,
- High frequency of infinitive/ʔi,ʔanna/ in Translated RS (64.6% of total hits) compared to (32.7% of total hits) in HASC and (56.31 % of total hits) in NTASC. Henceforth, a higher frequency of /al-msʔdar al muʔwwal/ is well evidenced in TASC; hence, a higher word count which runs counter to the dictates of brevity and an unneeded overrepresented emphasis counter to required scientific objectivity. RS’s are subjected to a corpus-based free search for the particle followed by a quan-qual manual analysis,

- Use of initial emphatic /ʔi,ʔanna/ in Historical RS is mainly to underscore asserted observations and conclusions; but rarely for emphasizing established facts,

- Possible dispensability of initial /ʔi,ʔanna/ across the entire TASC and Translated RS in addition to some selections from Non-translated RS is evidenced through cross-sections.

Over-representation of the emphatic and infinitive particle is thus proved in TASC and Translated RS. **TU of exaggeration, as far as the search for /ʔinna/ shows, is evidenced to be inversely proportional to generic scientific properties of neutrality and economy of language.** A reconsideration of the repetitive void use of the particle is recommended for users and translators on producing Arabic scientific writing.

- Frequency of multiple intransitivity forms registers 2.35% in the Historical RS, 1.99% in the Non-translated RS, and a 1.4% in the Translated RS. Random parallel alignments of Arabic Translated RS extracts compared to their English source excerpts prove an equal, if not a higher frequency of intransitive verbs. Thus, translated scientific Arabic compared to the Historical and non-translated as exemplified in the representative RS's cannot be held responsible for a lower frequency of intransitivity which is shown to be a direct interference from the source English. **TU of underrepresentation of unique items as multiple intransitive verb forms in TASC runs inversely proportional to the scientific universality of the impersonal style.**

- Low frequency of passive voice occurrences is monitored in the Translated RS, in spite of a recognized norm of the passive voice structures in scientific English. Presumed conventionalized avoidance of the passive voice by modern Arabic writers and translators is shown to be the reason for this phenomenon. Conversely, Original Arabic scientific writing is evidenced to incorporate high densities of the passive voice running as a norm in integration with intransitivity. Direct eye search of RS's shows a high frequency of the passive verb forms in the Historical RS with 284 hits, 1.11% of the total word count, which becomes a bit lower in the Non-translated RS with 184 hits, 0.92%, only to reach its lowest level in the Translated RS with 42 hits, 0.19%. **Inverse proportionality stands between a lower conventionalized representation of Arabic passive voice in TASC and generic dictates of the impersonal style.**

- Eligibility of Original Arabic for scientific elaborations is ascertained; ancient Arab scientists are proved to have been more aware of the linguistic features required for the realization of the language of science. Arabic translators are recommended to redirect their preferences of linguistic representations of emphatic and infinitive/?i,?anna,?ann/, multiple Arabic intransitive and passive verb forms in a way directly proportional to generic scientific properties.

References

- Arberry, A. (1982) *The Koran Interpreted*. Oxford: Oxford University Press.
- Baker, M. (1993). Corpus Linguistics and Translation Studies: implications and applications. In Mona Baker, Gill Francis & Ellena Toginini- Bonelli (Eds.) *Text and Technology: in honor of John Sinclair* (233-250). Amestrdam: John Benjamins.
- _____. (1995). Corpora in Translation Studies: An overview and some suggestions for future research. *Target*, 7 (223-243). Amestrdam: John Benjamins.
- _____. (1998). Arabic Tradition. In Mona Baker & Kirsten Maelmakjaer (Eds.) *Routledge Encyclopedia of Translation Studies* (316-324). London: Routledge.
- _____. (2011). *In Other Words: a Coursebook on Translation*. London: Routledge.
- Bakir, M. (1994). Notes on Passive and Pseudo Intransitive Constructions in English and Arabic. Basra: Mu?tah University. Retrieved from <http://ifa.amu.edu.pl/psicl/files/31/03Bakir.pdf>.
- Baugh, A. (1957). *A History of the English Language*. London: Routledge & KeganPaul
- Bowker, L. (1996). Towards a Corpus-based Approach to Terminography. *Terminology*, 3(1).1996. (27-52).
- _____, Cronin, M., Kenny, D., & Pearson, J. (Eds.) (1998). *Unity in Diversity: Current Trends in Translation Studies*. London: Routledge.
- Cantarino, V. (1975). *Syntax of Modern Arabic Prose*. (Vols1-3).Bloomington: Indiana University Press.
- Chesterman, A. (2001). Hypotheses about Translation Universals. In Gayde Hansen, Kirsten Malmkjaer & Daniel Gile (Eds.) *Claims, Changes and Challenges in Translation Studies* (1-13). Amsterdam: Benjamins.
- _____. (2004). Beyond the Particular. In Anna Mauranen & Pekka Kujamaki (Eds.) *Translation Universals: Do They Exist* (33- 49).Amsterdam: John Benjamins.
- Dickens, J., Hervey, S., & Higgins, I. (2002). *Thinking Arabic Translation*. London: Routledge.
- Fauntinuoli, C.,& Zanettin, F.(Eds.) (2015). *New Directions in Corpus-based Translation Studies.Translation and Multilingual Natural Language Processing 1*. Berlin: Language Science Press.

- Fernandes, L. (2006). Corpora in Translation Studies: revisiting Baker's typology. Retrieved from <https://periodicos.ufsc.br/index.php/fragmentos/article/viewFile/8217/7690>
- Gordin, M. (2015). *Scientific Babel: How Science Was Done Before and After Global English*. Chicago: University of Chicago Press.
- Halliday, M. (1968). Notes on Transitivity and Theme in English. *Journal of Linguistics* 3 (47). (37-81).
- Hasan, A. (1999). *An-naħw al-Waafii*. (Vols. 1-4). Cairo: Daar al Ma'aarif.
- Hassnawi, A. (2003). Aspects of Scientific Translation. English into Arabic Translation as a Case Study. *Journal Ibri College of Education*, The Sultanate of Oman. Retrieved from: <http://www.translationdirectory.com/article10.Htm>
- Jespersen, O.(1961). *A Modern English Grammar On Historical Principles*. Copenhagen: Einar Munksgaard.
- Johansson, S. (2003). Reflections on corpora and their uses in cross-linguistic research. In F.Zanettin, S.Bernandi,& D. Stewart (Eds.) *Corpora In Translator Education*. (133-44) Manchester: St Jerome.
- Khafaji, R. (2000). *Arabic Translation Alternatives for the Passive in English*. Retrieved from <http://ifa.amu.edu.pl/psicl/files/31/02Khafaji.pdf>.
- Krein-Kuhle, M. (2003). *Equivalence in Scientific and Technical Translation: A Text-in-Context based Study*. Ph.D thesis. Salford: University of Salford.
- Kunitzsche, P. (2007). Scientific Contacts between the Islamic World and Europe: the Case of Astronomy. In Mohamad el-Gomati & Salim Iyduz (Eds.) *Foundation for Science Technology and Civilization* (2- 12). UK: FSTC Ltd.
- Laviosa, S. (2003). Corpora and the Translator. In Harold Somers (Ed.) *Computers and Translation* 35(105-119). Amsterdam: John Benjamins.
- Lyons, J. (1968). *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Mason, I. (2001). Translator Behavior and Language Usage: Some Constraints on Contrastive Studies. *Hermes* 26 (65-80).
- Mauranen, A. (2004). Corpora, Universals and Interference. In Anna Mauranen & Pekka Kujamaki (Eds.) *Translation Universals: Do They Exist* (65-82).Amsterdam: John Benjamins.
- _____ & Kujamaki, P. (Eds.) (2004) *Translation Universals: Do They Exist*. Amsterdam: John Benjamins
- Mohammad, A. (2006). Alfiḡl al-mabnee lilmaghuul fi l-luayah al-arabiyyah ?γ raad^uuh, mus^talahaatuh, ?ahmayatuh. *Journal of Damascus University* 22 (2+1). (17-65). Retrieved from <http://www.damascusuniversity.edu.sy/mag/human/old/human/2006/22-2006/mohammad.pdf>
- Olohan, M. (2004). *Introducing Corpora in Translation Studies*. London: Routledge.
- _____. (2016). *Scientific and Technical Translation*. London: Routledge.
- Pearson, J. (1998). *Terms in Context*. Amsterdam/Philadelphia: John Benjamins
- Pinchuk, I. (1977). *Scientific and Technical Translation*. London: Andre Deutsch.

Rosenhouse, J. (1988) Occurrence of the passive in different types of text in English, Hebrew and Arabic. *Babel* 34(2). (90-103).

Shalash, T. (1971). *Awzaan al-Fi'l wa ma'aneehaa*. Iraq, Najaf: al-Adaab Publishing House.

Sinclair, J. (1992). The Automatic Analysis of Corpora. In J. Svartvik (Ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm 4-8 August 1991*. (379-97). Berlin & New York: Mouton de Gruyter

Sketch Engine. <https://the.sketchengine.co.uk/login/?next=/auth/corpora/>

Tirkkonen-Condit, S. (2004). Unique items – over- or under-represented in translated language? In Anna Mauranen & Pekka Kujamaki (Eds.) *Translation Universals: Do They Exist*. (187-184). Amsterdam: Benjamins.

Toury, G. (1978). The Nature and Role of Norms in Literary Translation. In James S Holms, José Lambert & Raymond van den Broek (Eds.) *New Perspectives in Literary Studies*. (83-100). Leuven: ACCO.

_____. (2004). Probabilistic explanations in translation studies: Welcome as they are, would they qualify as universals?. In Anna Mauranen & Pekka Kujamaki (Eds.) *Translation Universals: Do They Exist*. (15-35). Amsterdam: John Benjamins.

Wright, W. (1996). *A Grammar of the Arabic Language*. (Vols. 1-2). Beirut: Librare du Liban.

Wilson, M. (2005). *Field Methods in Vegetation Science: Online Textbook*. Mark V. Wilson and Oregon State University. Retrieved from

<http://oregonstate.edu/instruct/bot440/wilsomar/Content/Index.htm>

Zanettin, F. (2000). Parallel Corpora In translation Studies: Issues in Corpus Design. In Maeve Olohan (Ed.) *Intercultural Faultlines: Research Methods in Translation Studies I: Textual and Cognitive Aspects*. (105-118) Manchester: St Jerome.

Zohar, I. (1979). Polysystem Theory. *Poetics Today*, 1: 1-2 (287-310). Retrieved from http://www.tau.ac.il/~itamarez/works/papers/papers/EvenZohar_1979--Polysystem%20Theory%20Poetics%20Today%20I.pdf

Compiled Corpora

The researcher has compiled three electronic small-scale corpora and uploaded them via Sketch Engine Corpus processor: <https://the.sketchengine.co.uk/login/?next=/auth/corpora/>

Following is a list of the three compiled corpora and the electronic primary sources each comprises:

1. Historical Arabic Scientific Corpus (HASC):

1.1 Ar-Razii, *al-ḥaawee fi t^l-t^{ib}*, 2000 edition; uploaded 2012;

[https://www.kutub-pdf.com/book/الحواري في الطب.html](https://www.kutub-pdf.com/book/الحواري%20في%20الطب.html)

1.2 Ibn Sina, *al-Qanuun fi t^l-t^{ib}*, (Vols. I, II, IV, V), 1999 edition; uploaded 2012;

<http://waqfeya.com/book.php?bid=7482>

1.3 Ibn al-Haitham, *al-Manaaḏ^{er}*. uploaded 2016; <http://www.mediafire.com/file/>

كتاب+المنظر.rar

2. Non-translated Arabic Scientific Corpus (NTASC),

2.1 Mushrafaa, M. (2013) *al-Qanaabel aḏ-ḏarriyyah*.

<https://www.kutub-pdf.com/book>,-

2.2 Hegazee, M. (2001) *al'amraad^s al-geldiyyah ?inda al-At'faal*,

(<http://www.gulfkids.com/pdf/geldea.pdf>)

3. Translated Arabic Scientific Corpus (TASC),

3.1 *at-t'ab^s ṣabr at-t'ab^sabu^s*, (2013)) <https://www.kutubpdfcafe.info/kutubpdf>

الطبع والتطبع .html,. Arabic translation of Matt Ridley (2003) *Nature via Nurture: Genes, Experience, & What Makes Us Human*. Cairo: National Center for Translation.

3.2 *Mujkilat al-fizyaa?*, (2016)

<https://www.kutubpdfcafe.info/kutubpdf>مشكلة-الفيزياء .htm, Arabic translation of Lee Smolin (2006) *The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next*. Cairo: National Center for Translation.

3.3 *Istifraaf al-mustaqbal: ḥawrat at-tiknolojia an-Nanoniyya*, (2016)

<http://www.mediafire.com/file/>استشراف+المستقبل+ثورة+التكنولوجيا+النانوي.pdf. Arabic translation of K. Eric Drexler, Chris Peterson, and Gayle Pergamit's (1991) *Unbounding the Future: the Nanotechnology Revolution*, Cairo: National Center for Translation.

Appendices

Appendices are uploaded on Outlook drive via

https://alsunasuedu-my.sharepoint.com/:f:/g/personal/amalabdelmaqsoud_alsun_asu_edu_eg/EI2PnXfiLHhCgo-R5vHug4BIJVyocyBoR7grehXvNh7hQ?e=JXm0JN

Appendix 1_Search for Inna_pdf

Appendix 1-1 Full stop+Inna in the entire Historical Corpus

Appendix 1-2 Fullstop+Inna in the entire non- translated corpus-

Appendix 1-3 Fullstop+Inna in the entire Translated Corpus

Appendix 1-4 RS_ Historical_ Free Search_ Inna

Appendix 1-5 RS_ Translated_ Free Search_ Inna

Appendix 1-6 RS_ Nontranslated_ free Search for Inna

Appendix 1-7_ Cross sections _ Parallel alignment _ Search for Inna

Appendix 2_Random Sample_pdf

Appendix 2-1_RS_ Historical Arabic_ Compiled

Appendix 2-2_RS_ Non-translated Arabic_ Compiled

Appendix 2-3_RS_ Translated Arabic_ Compiled Appendix 2-

4_RS_ Crossections_ Parallel alignment

Appendix 2-4_RS_ Crossections_ Parallel alignment

Appendix 3_ Search for Qaama bi_pdf

Appendix 3-1 _ Search for Qaama_ Historical Corpus

Appendix 3-2_ Search for Qaama _ Non-translated

Appendix 3-3 _ Search for Qaama_ Translated Corpus

Appendix 4_ Search for Tamma_pdf

Appendix 4-1_ Search for Tamma_ HASC

Appendix 4-2 Search for Tamma_ NTASC

Appendix 4-3_ Search for Tamma_ TASC

Appendix 5_ Search for yudaaf, yuqaas, tamma qeyaas_pdf

Appendix 5-1 search for yudaaf_ TASC & NTASC

Appendix 5-2_ Search for yuqaas_ TASC & NTASC

Appendix 5-3_ Search for tamma qiyaas_ TASC & NTASC

Appendix 5-4_ Search for ustuhditha_ TASC&NTASC

¹ - A bold font is sometimes used by the researcher to mark an emphasis and to highlight observations and conclusions based on the given premises.