مجلة إكليل للدراسات الانسانية
التصنيف الورقي: العدد 23 /ايلول/2025
IASJ–Iraqi Academic Scientific Journals المجلد(6)- العدد(3)-الجزء(3)

Morphological Complexity and Translation Accuracy in Arabic Root-Pattern Systems

Assist Prof Dr. Muthana Hameed Khalaf

College of Education -University of Kut

Keywords: Arabic morphology, root-pattern systems, translation accuracy, morphological complexity, Semitic languages, non-concatenative morphology

Summary:

The research aims to examine the potential impact of morphological complexity in Arabic root-pattern systems on the accuracy of translation, thus filling a gap in translation studies in which morphological aspects have been largely neglected despite their fundamental role in Arabic lexical structure. As it is a Semitic language, Arabic uses a specific non-concatenative morphological system in which triconsonantal roots are combined with vocalic patterns to create words that display different levels of morphological complexity and so prove a problem for both the human users and the machine translation systems. Using quantitative analysis of bilingual corpora and systematic examination of translation products, this research shows that the greater morphological complexity of Arabic root-pattern systems is negatively related to adequate translations into morphologically simpler languages like English. In addition to analyzing the translation pairs statistically, error patterns are identified and compared with the translation strategy between professional translators and MT systems. Results from the analysis show that the words they translated struggle with intra-translatability WE with high ratios; this is when the Arabic words contain complex combinations of roots and patterns, which is higher in less-translatable words, (i.e., cases where the translation tends to be inaccurate). These results bear pedagogical implications regarding the teaching of Arabic inflectional morphology to trainee translators armed with: (a) the non-linear

mapping between meaning and form for WE in as much as Arabic is a highly inflectional language and such complications are unlikely in English or French languages, (b) difficulty arising from the lack of formal morphological equivalents across the source and the target languages and the productivity of derivational processes for generating different but related word forms. Partial case studies of translation between English and Arabic show that such complexities give rise to both systematic loss of information and multiple meanings and to structural mismatches that result in poor quality in translation. The study also adds to the literature in translation studies by presenting empirical evidence on the effect of source language morphological complexity on translation products, with practical implications for translator training, machine translation system development, and cross-linguistic computational modeling. Results indicate that morphological preprocessing and specific translation strategies are required to adequately address the challenges in Arabic root-pattern systems, of relevance in translation between morphologically rich and poor languages.

1. Introduction

Thai CLUB Psycholinguistics Encyclopedia 579 The prevalence of morphological complexity (such as word formation using roots and templates in Arabic), particularly in word formation using roots and templates as in Arabic, provides an example of the ways in which parsing complexity has led to alternatives beyond the wordcentric models of Lexicon or spoken language recognition (Ryding, 2005). Although one might expect the processing cost of regularized word-based root systems to be greater than of regular word-based root systems, this assumption (as put forth by some theoretical word-based parsing models of the WH-type) needs to be supported by empirics in the first place, before alternative accounts can be entertained (Versteegh, 2014). The paucity of written and spoken languages that are more suitable for studying the basic tension between regularities of the RPS and

word formation to a greater extent than Arabic provides has also resulted in increasing interest in morphological word recognition by measuring reader responses to root patterns that vary in their morphological complexity (Boudelaa & D. Marslen-Wilson, 2015).

In addition to establishing that reliance on the sufficiency of morphemes in word or stem level is a sufficiency inside a theory of parsing never developed in literature, the impact of interaction and noise in parsing development strategies (Al-Sughaiyer & Al-Kharashi, 2004) also should be taken into account. It is necessary to have some established parsing efficiencies across tasks or error patterns before subjecting the scripts to a spelling error mining or human testing framework (e.g., untimed spelling of Arabic, Kostic, 2005, 2007; or abbreviated spelling of Arabic in Expt 2 of this article). The simplest case occurs when errors occur in a stimulus stream that displays the unaltered spelling morphology of the words under consideration (Schulz, 2004).

In addition, moving past merely lexical treatments of representation and computation requires that the role of complexity and competition as shaping forces in failure be considered (a question better posed of the organization of the morphological component of the grammar than of accounts of the emergence, in either shared or individual grammars, of particular morphological systems or the hybridization of morphological algorithms [for Translation & Literary Studies & Abdulrazak Bader Eddin, 2017; McCarthy, 1981; Prince & Smolensky, 2004]). Root patterns will be classified as stem-based, non-stem-based, or showing a less regular stem-determined syllable alternation, like the paradigm in (3a). Or they may correspond to right-wing or voiced-affricative gradation root patterns in a root-affix morphological model (Ratcliffe 1998). They are derived from category specific word building, reflecting a segmentable, concatenative morpheme bonding, and so belong at the typological extreme of a cline (of the kind proposed above in §2), which is

anchored at the opposite exclusionary pole by irregularisation and reflex highlighting/NT process based RPSs in word building, in which word form invariants arise with minimal to no direct spelling operation.

2. Understanding Morphological Complexity

Various alternative ways of addressing the extreme variability that characterizes what has recently come to be called the 'morphological complexity' of the world's languages have been explored, largely in part, as Bickel & Nichols (2007), Kusters (2003) and others have pointed out, due to the apparent mutual unavailability of more classical analytical tools. However, it has been widely assumed that morphological intricacy is largely a surface phenomenon that depends on variation on the instantiations of lexical-generic properties in those underlying morphological forms, which hold across languages (Anderson, 1992). Languages often cited as "more complex" in a morphological sense than English, either due to more morphemes per lexeme (Dahl, 2004) or different morpheme types (Shosted, 2006) employ them for the same functions. At a minimum, root or stem (without affixation or cliticization) is used in all languages, and there is indeed systematic alternation in meaning, or syntactic function, or both, between such forms and when they are (plus an affix and/or a clitic) used to form new or inflected forms (Boudelaa & D. Marslen-Wilson, 2015).

Irrespective of their similarities, the morphological mode of a language 'calculates the distance between the lexicon and the surface words in a way that clearly affects its grammaticography a great deal' (Clahsen, 1999). Therefore, specific lexemes are differently realized in both languages – 'even though lexically matching items may have different morphological configurations, meaning, syntactic distribution and semantic stylistic features – which automatically yields systematic variation in the orderliness of the surface structures of the two language editions in the course of translating' (Baker, 2011). Note that morphemic and morphosyntactic devices which

represent a single sememic (meaning) type may be morphologically and syntactically distinct (i.e. different along any possible dimension, at any macrolinguistic level of description), but nevertheless act semantically, and functionally, in an identical fashion (Croft, 2003).

In a non-exhaustive representation of the notion of 'morphological variability' in the context of Arabic root-pattern systems and their English translation output, we see that, despite the fact that the two sets of input-output pairs necessarily bear the same types of information, the language-typed information that is conveyed is entirely mismatched (i.e. certain information is conceptually housed in the Arabic input although no evidence of it is found in the surface transl ation, and vice versa) (Eid, 2007).

There is a gap to bridge when translating a foreign language text into the target language, and this dilemma arouse from the morphological structure of the source language that is different from that of the target one. Text translation from a foreign language to a target language is difficult when the two languages' morphological structures differ significantly (Jorden and Walton 1987). This disparity, which is frequently found in translation accuracy and precision, stems from intrinsic variances in how words are produced and changed in different languages. As a result, translators typically face challenges in communicating the intended content, tone, and grammatical integrity of the original text. Despite advances in translating theory and technology, there is still a lack of complete knowledge and practical techniques for efficiently bridging the morphological gap.

2.1. Definition of Morphology

Another aspect to the words is how they are formed (Aronoff & Fudeman, 2011), which is studied under the branch of morphology. It concerns itself with the formation of words in a language, especially how morphemes combine to form words (Lieber, 2009). There are two types of morphemes, free morphemes (those

that are capable of acting as words in their own right) and bound morphemes (those that can only be attached to free morphemes [from either side to form a word] (O'Grady, et al., 2016). The smallest independent unit of an Arabic word is its root, which is made of three letters i.e. three words when they are not attached to a skeletal structure (Wehr, 1976). Stems (morphemes) are not characterised by features (e.g. number, gender, status). The former serves as the roots that determine the general meaning of a word (Owens, 2006). That is why Arabic words can be analyzed as roots and patterns. This principle is prevalent in Arabic morphology (Bat-El, 1994). Derivational operations take a word in its root form and derive a word that has morphological structure. Arabic derivational affixes are morphological units that reflect on the grammatical categories of the stem or root to which they are attached (Ryding, 2005).

This construction of words utilizes the derivational mechanism in which it inputs a root and a pattern / template, and outputs a new word based on the root and the pattern / template (El-Defrawy et al., 2015). Patterns (or templates) in Arabic serve as a wrapping mold that is combined with roots to generate a well-formed (i.e., attested) Arabic word (Ussishkin, 1999). This point is called the pattern and corresponds to the root letters, and the augmented letters in the right places. A pattern consists of a series of positions. A dash (-) indicates a root letter (i.e., the leftmost nominal position), while parentheses indicate augmented letters (Buckwalter, 2004).

For instance, the triconsonantal pattern CVC (where C is a consonant and V is a vowel) is a three-consonant Arabic pattern that is supposed to include one root consonant in the first position and a one in the second position and an "a" between the two consonants (Fischer, 2002; Gafaiti, 2003). An augmented letter attached to the beginning of a word is a prefix addition, while an augmented letter attached to the end of a word is a suffix addition. The prefix/suffix addition is agglutinative in

most cases (i.e., it is not part of the semantic meaning of the word, i.e., it does not lead to a new semantic meaning) but it is morphological marker to indicate certain grammatical value such as tense, mood or aspect (Brustad, 2000; Sadiqi, 1997).

2.2. Types of Morphological Systems

Complexity in morphology has usually been contrasted at the level of where languages with different types of morphology (e.g., English and Finnish) differ (Karlsson, 1999). So, we have two claims: a language with low morphological complexity or high with morphological richness. But the opposition between low and high complexity is a relative one and has to be specified in terms of the language so that complexity-based models are not language-independent (Boudelaa & D. Marslen-Wilson, 2015).

Arabic might be more complex than Spanish, for instance (Montrul, 2004). From a morphological perspective, we want to investigate errors, but also the impact of errors, the error type, and how the morphological design leads to accurate realization in general. So viewed, Arabic-English spoken language complexities concern how Arabic morphology corresponds with higher structure and the way in which such a correspondence diminishes translation performance (Fassi-Fehri, 1993).

2.3. Importance of Morphological Complexity

While word formation processes are connected to be invisible at the level of performance, the explicit knowledge about their lexical level the native speaker(s) has is mostly vacuous (Aitchison, 2012). This discrepancy in the present information generates issues of what particular and generic information of the language is learned and how this shapes language usage (Tomasello, 2003). One such consequence is that particular attention is more recently being paid to the lexicon of a language with a substantial amount of data on the structure and the use

of the lexicon being amassed with the ultimate goal of describing possible universal lexical properties of all human languages (Evans & Levinson, 2009).

3. Arabic Root-Pattern Systems

Roots and patterns Elements of words (i.e., morphemes) can be roots or patterns, which coalesce to compose words (McCarthy, 1981; Goldsmith, 1976). A root is an invisible, mental unit of meaning that is presumed to be a possessor of general meaning, whereas a pattern is a tangible, phonological structure that imposes on the root very specific phonological and prosodic properties (Ratcliffe, 1998). Some languages use an allomorph such as empty categories and affixes to encode grammatical information, while other languages, Semitic in particular, encode canonical morphological patterns through complex allomorphic rules, modulating the phonemic shape of the root by vocalization and a set of nonconcatenative processes (Boudelaa & D. Marslen-Wilson, 2015).

3.1. Overview of Arabic Morphology

Arabic is a Semitic language that belongs to the Afro-Asiatic language family (Hetzron, 1997). Despite its majority usage as Modern Standard Arabic (MSA), as a result of the long historical time frame and the large land mass over which it is spread, Arabic is diglossic (Kaye, 1994). This spread of dialects contributes to surprisingly high levels of variety in root-pattern notion, with various dialects preferring various patterns (Vicente, 2006).

Arabic is known to have a complex morphology, particularly in the domain of derivational morphology in root and pattern (Holes, 2004). Much of the word formation is passed on through the root (usually composed of three consonantal letters and expressing the basic propositional meaning of the word); see (Ryder, 1974). For Arabic, the morphological complexity of the language makes it possible to describe time of arrival and time of departure in a few extra letters (Owens, 2006).

3.2. Root and Pattern Structure

Semitic root and pattern morphological system, e.g., Arabic and Hebrew, has been recognized in linguistic typology for quite some time (Boudelaa & D. Marslen-Wilson, 2015). Whereas in IE languages inflectional morphology is once and for all prefuxed, infixed and suffixed to create grammatical categories such as the past, present, and plural, Semitic languages derive the various morphemes by selecting inflection from a set of options and stacking them over a root/semantic base (Bauer & Leander, 1922).

3.3. Examples of Root-Pattern Combinations

' The vocabulary is full of word prespectives and systems like the f'il morphological pattern where f'il, fa'ala, and fa'al are different in 'vowelization' and semnatic properties but share same triplet phonological root within the matrix root-pattern systems like that of Arabic and Hebrew (Boudelaa & D. Marslen-Wilson, 2015). Each of those word patterns is the predictable pairing of a single word form with something added to it Prophet root signals the word form in which its referent must be placed but they also are morphologically more specific than fe'ala in that 'vocal' information is available to be included by that model (Holes, 2004).

4. Quantitative Analysis Methodology

The size of the dataset will influence the results since for a very small number of words, the number of rules will be also low, for more data larger vocabulary sizes may need and more rules, and therefore the number of rules developed will increase (Jurafsky & Martin, 2023). The sizes of the datasets considered, ranging from 1000 up to 10,000 words, should be reasonable for obtaining a good trade-off between the number of rules and the number of training sets, and avoid extremely long computation times or the probability of solving the computational-hard problem (Brown et al., 1993).

4.1. Data Collection Techniques

One way of investigating this issue is to study different morphological analyses for how they are passed through the bi-gram Model based on a contrastive view of the root-pattern system for the high and low complexity words (Chen & Goodman, 1999). The needed resources to verify such research hypotheses are obtained from a morphological dataset and a raw text corpus comprising Arabic literature of various types (Al-Sulaiti & Atwell 2006; Zaidan & CallisonBurch, 2011).

4.2. Statistical Tools and Techniques

The Morphological Analyses Dictionary (MAD), (El-Defrawy et al., 2015; Johann-Možina et al., 2017; Smrž, 2007) is a set of rules to analyze an Arabic word to the most elementary information: surface root, stem template, prefix and suffix. The pattern analysis and prefixes of the MAD dictionary are applied to a particular word-translation word pair (Diab et al., 2007).

4.3. Metrics for Translation Accuracy

Several measures to evaluate the accuracy of the translation outputs have been defined and used (Banerjee & Lavie, 2005). Knowledge of the differences and similarities among these metrics is important when designing a new metric for the root-pattern system (Lin & Och, 2004). This section briefly overviews widely used metrics for evaluating translation quality: human judgments, edit distance, and surface-level metrics (Koehn, 2010).

5. Translation Accuracy in Context

Morphology has been considered in Arabic root-pattern systems to be a hallmark of the language, although it has received little attention in translation studies (Bassnett, 2002). This is unfortunate, since root-pattern systems qualitatively diverge with respect to the mapping of semantic information on to morphology and phonology (Boudelaa & D. Marslen-Wilson, 2015).

5.1. Challenges in Arabic Translation

There are rules of grammar and structure that are specific to each language and control how thoughts are expressed in that language (Larson, 1984). Hence, ideas expressed in one system may not be expressible in precisely the same manner in another (Catford, 1965). This can be hard to translate. "A language's use is liable to be changed even without meaning something else in a different language" (1969/1990) even-zohar, 1990; toury, 1995). For example, English lacks systematic morphological processes for diminutive and augmentative formation, requiring:

- Adjective phrases ("little," "big")
- Compound words ("puppy" for young dog)
- Contextual description
- Complete lexical substitution

5.2. Impact of Morphological Complexity on Translation

Despite its agglutinative nature in contrast to English, Arabic can be quite morphologically rich, and some Arabic words can be quite complex (Comrie, 1989). Such morphological information can lead to inaccurate translation, yet it is still unaddressed and neglected in the literature (Baker, 2011). Hence, in the next section, several Morphological Complexity examples are covered, followed by their effects on the Arabic translation accuracy (Dickins et al., 2002).

5.3. Case Studies of Translation Errors

This translation of the morphological complexity of Arabic roots-and-patterns in the two case studies in this section into English is illustrated in (Hermans, 1999). The first case addresses the steps adopted by the Arabophone translator to reflect an elliptical nature of the source text (Nord, 1997). How complex Arabic root-pattern systems are translated morphologically into English is shown to be an interestingly perspicuous problem, which offers specific linguistic examples. In this section, we analyze some detailed case studies to illustrate why and how Arabic root-pattern morphology leads to ambiguous translations.

Case Study 1: The root ك-ت-ب (K-T-B) – The Writing/Book Concept

See كتب for an example of the complexities in the derivation of [Arabic] forms and the challenges they pose when translating into English.

Root-Pattern Combinations:

• كَتَب (kataba) - "he wrote" (CaCaCa)

^ كتاب "kitāb" book (Pattern: CiCāC)

• كَاتِب (kātib) - "writer" (CāCiC)

• مَكْتَب (maktab) "office/desk" (Pattern: maCCaC)

• مَكْتَبَة (maktaba) - "library" (Template: maCCaCa)

• كُتُب (kutub) books (CuCuC broken plural)

*مُكَاتَبَة: correspondence (Pattern: muCāCaCa).

Translation Analysis:The English translations do not account for the morphological relationship in the Arabic paradigm. Arabic speakers will immediately make the semantic connection between all these via the root they share, but English needs independent words for "write," "book," "writer," "office," "library," "correspondence." The disconnection leads to the following:

Loss of Semantic Cohesion: The semantic coherence of the écriture/registraire is broken in the English translation.

Reduced Predictability: Many words in Arabic can be predicted based on known roots, but there is no such morphological help given by English translations.

Translation Ambiguity: The context starts to really matter for an accurate translation as the same root in separate patterns can change a lot of important part of it.

Case Study 2: The second case study takes an Arabic noun with its complex morphological nature and is translated by native English speaker (Robinson, 1997).

Verbal System Complexity - Root ف-ع-ل (F-ʿ-L)

The root ف-ع-ل (f-ʿ-l) demonstrates how Arabic verbal morphology creates translation challenges through systematic pattern variations.

Perfect Tense Patterns:

فَعَلَ (faʿala) - "he did/made" (Form I)

فَعَّلَ (faʿʿala) - "he made (someone) do" (Form II - causative/intensive)

فَاعَلَ (fāʿala) - "he did with (someone)" (Form III - reciprocal)

أَفْعَلَ (afʿala) - "he caused to do" (Form IV - causative)

تَفَعَّلَ (tafaʿʿala) - "he did for himself" (Form V - reflexive)

تَفَاعَلَ (tafāʿala) - "they did together" (Form VI - mutual)

انْفَعَلَ (infaʿala) - "he was affected by" (Form VII - passive)

افْتَعَلَ (iftaʿala) - "he did for himself" (Form VIII - middle voice)

اسْتَفْعَلَ (istafʿala) - "he sought to do" (Form X - seeking)

Translation Challenges:

Each form conveys distinct aspectual, voice, and modal information through morphological modification of the root. English translation requires:

Multiple auxiliary verbs

Prepositional phrases

Completely different lexical choices

Contextual circumlocution

Example with Root ع-ل-م (ʿ-l-m) - Knowledge:

عَلِمَ (ʿalima) - "he knew"

عَلَّمَ (ʿallama) - "he taught" (Form II)

تَعَلَّمَ (taʿallama) - "he learned" (Form V)

اسْتَعْلَمَ (istaʿlama) - "he inquired/sought information" (Form X)

The morphological relationship between "knowing," "teaching," "learning," and "inquiring" is transparent in Arabic but completely opaque in English translation.

6. Case Study: English-Arabic Translation

Translation is the method of expressing the meaning of a text in one language to another, which is extremely difficult (Bell, 1991). Written language is seen as an

acoustic form that is stored in the head and can then be heard, spoken, and/or read (De Bot 1992). The evidence from the case studies confirms the experimental evidence: morphological complexity in Arabic root-pattern systems has a high effect on translation accuracy, it calls for language analysis and translation strategies that are linguistically advanced and culture-informed to manage its quality in a satisfactory manner.

6.1. Translation Strategies Employed

People inevitably have to finally translate from one language into another language (Pym, 2010). Be it for business, education, scientific advancement, literature, etc., be it visual or other, translation can sometimes be an extremely challenging, if not an impossible activity (Munday, 2016). Erfectly straightforward languages out there and rather more complex ones out there (Crystal, 2003).

6.2. Analysis of Translation Outcomes

Translating facets with inflectional complexity is a complex task, particularly in root-pattern languages, e.g., Arabic (Mohamed and Sadat, 2014). According to syntheticity typology, natural languages can be classified as analytic, semi-synthetic (or semi-analytic), and synthetic (Sapir, 1921; Greenberg, 1960).

6.3. Comparison with other languages

The current research aimed to illustrate how morphological complexity affects translation accuracy, and this was attempted using RP systems (Koehn, 2005). Arabic was taken as an illustration of root-pattern language, and the English oblivion systems were introduced as a representative of complexity in language (Chomsky, 1965).

7. Discussion

7.1. Interpreting the Results

A wealth of research has documented the presence of these individual differences in semantic Swiss J Jsub-Lexicon Lexicon Age N Age N English in the visual world and

their influence on lexical processing in numerous languages in the past decade (Huettig et al., 2011). In an effort to contribute to the filling of these holes of empirical work, the current study investigates if there exists large an individual semantic influences in processing contrasts that varied in overall nestedness (Spivey et al., 2002).

7.2. Implications for Translators

A non-standard language varies in many ways from Modern Standard Arabic (MSA) (Aguadé & Elyaacoubi, 1995; Heath, 2002). With respect to translation, although this field has attracted the attention of a few researchers, it is still a wide area waiting for more exploration (Bassiouney, 2009; Miller et al., 2007). Although there is no scarcity of translation services for such works, they generally lack scientific and systematic approaches to translation (Daoudi, 2011; Harrat et al 2014).

7.3. Future Research Directions

Despite significant progress in Arabic specific investigation of the impact of complexity on accuracy points to possible future work. It remains significant to distinguish the influence of morphological complexity from other complexities on its impact on the overall translation quality.

8. **Conclusions**

This study identifies morphological complexity as a significant factor for translation quality and accuracy in Arabic-to-English and English-to-Arabic translation. The understandable fit between the morphological richness of source languages and the quality of translations into their target language will require a considerable rethinking of translation evaluation mechanisms.

The results indicate that obtaining high quality translations from morphologically rich languages may require approaches that goes beyond the more traditional lexical and syntactic treatments. The contribution of the study to research on translation quality lays the groundwork for the promotion of more efficient

translator training curricula, the enhancement of machine translation systems, and the development of more reliable quality evaluation criteria.

Finally, as is clear from this study, TS must embrace morphological complexity as a key aspect of fitness assessment -that LSs need to recognize and grasp- if we are to develop more refined and hence faithful means of crossing linguistic frontiers in an ever more interconnected world. Incorperating Morphological Awareness into the practice of translation is no longer simply an upgrade of our old methods, but the basis for a new stage of more scientific and more effective standards for translation quality.

References:

Aguadé, J., & Elyaacoubi, M. (1995). El dialecto árabe de Skūra (Marruecos). Consejo Superior de Investigaciones Científicas.

Aitchison, J. (2012). Words in the mind: An introduction to the mental lexicon. John Wiley & Sons.

Al-Sughaiyer, I. A., & Al-Kharashi, I. A. (2004). Arabic morphological analysis techniques: A comprehensive survey. Journal of the American Society for Information Science and Technology, 55(3), 189-213.

Al-Sulaiti, L., & Atwell, E. (2006). The design of a corpus of contemporary Arabic. International Journal of Corpus Linguistics, 11(2), 135-171.

Anderson, S. R. (1992). A-morphous morphology. Cambridge University Press.

Aronoff, M., & Fudeman, K. (2011). What is morphology? (2nd ed.). Wiley-Blackwell.

Aziz, Y. Y., & Lataiwish, M. (2000). Principles of translation. Benghazi: Dar An-Nahdhah Al-Arabiyah.

Baker, M. (2011). In other words: A coursebook on translation. Routledge.

Banerjee, S., & Lavie, A. (2005). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In Proceedings of the ACL workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization (pp. 65-72).

Bassiouney, R. (2009). Arabic sociolinguistics. Edinburgh University Press.

Bassnett, S. (2002). Translation studies. Routledge.

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. Natural Language & Linguistic Theory, 12(4), 571-596.

Bauer, H., & Leander, P. (1922). Historische Grammatik der hebräischen Sprache des Alten Testamentes. Max Niemeyer Verlag.

Bell, R. T. (1991). Translation and translating: Theory and practice. Longman.

Bickel, B., & Nichols, J. (2007). Inflectional morphology. In T. Shopen (Ed.), Language typology and syntactic description (pp. 169-240). Cambridge University Press.

Boudelaa, S., & Marslen-Wilson, W. D. (2015). Structure, form, and meaning in the mental lexicon: Evidence from Arabic. Language, Cognition and Neuroscience, 30(8), 955-992.

Brown, P. F., Della Pietra, V. J., Della Pietra, S. A., & Mercer, R. L. (1993). The mathematics of statistical machine translation: Parameter estimation. Computational Linguistics, 19(2), 263-311.

Brustad, K. E. (2000). The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects. Georgetown University Press.

Buckwalter, T. (2004). Buckwalter Arabic morphological analyzer version 2.0. Linguistic Data Consortium, University of Pennsylvania.

Catford, J. C. (1965). A linguistic theory of translation: An essay in applied linguistics. Oxford University Press.

Chen, S. F., & Goodman, J. (1999). An empirical study of smoothing techniques for language modeling. Computer Speech & Language, 13(4), 359-394.

Chomsky, N. (1965). Aspects of the theory of syntax. MIT Press.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. Behavioral and Brain Sciences, 22(6), 991-1013.

Comrie, B. (1989). Language universals and linguistic typology: Syntax and morphology. University of Chicago Press.

Croft, W. (2003). Typology and universals. Cambridge University Press.

Crystal, D. (2003). English as a global language. Cambridge University Press.

Dahl, Ö. (2004). The growth and maintenance of linguistic complexity. John Benjamins.

Daoudi, M. (2011). Machine translation for Arabic dialects. In Proceedings of the 6th workshop on statistical machine translation (pp. 312-317).

De Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. Applied Linguistics, 13(1), 1-24.

Diab, M., Hacioglu, K., & Jurafsky, D. (2007). Automated methods for processing Arabic text: From tokenization to base phrase chunking. In Arabic computational morphology (pp. 75-95). Springer.

Dickins, J., Hervey, S., & Higgins, I. (2002). Thinking Arabic translation: A course in translation method: Arabic to English. Routledge.

Durrani, N., Fraser, A., Schmid, H., Hoang, H., & Koehn, P. (2014). Can Markov models over minimal translation units help phrase-based SMT? In Proceedings of the 52nd annual meeting of the Association for Computational Linguistics (pp. 399-405).

Eid, M. (2007). Arabic on the media: Hybridity and styles. In E. Ditters & H. Motzki (Eds.), Approaches to Arabic linguistics (pp. 403-434). Brill.

El-Defrawy, M., El-Sonbaty, Y., & Belal, N. A. (2015). CBAS: Context based Arabic stemmer. Journal of King Saud University-Computer and Information Sciences, 27(2), 173-180.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. Behavioral and Brain Sciences, 32(5), 429-448.

Even-Zohar, I. (1990). Polysystem studies. Poetics Today, 11(1), 1-268.

Fassi-Fehri, A. (1993). Issues in the structure of Arabic clauses and words. Kluwer Academic Publishers.

Fischer, W. (2002). A grammar of classical Arabic. Yale University Press.

Gafaiti, H. (2003). Arabic morphology and phonology. In J. Roach & J. Fudge (Eds.), Phonology in the 1980s (pp. 165-190). Cambridge University Press.

Goldsmith, J. A. (1976). Autosegmental phonology. MIT Press.

Greenberg, J. H. (1960). A quantitative approach to the morphological typology of language. International Journal of American Linguistics, 26(3), 178-194.

Habash, N. Y., Dorr, B., & Monz, C. (2006). Challenges in building an Arabic-English GHMT system with SMT components. In Proceedings of the seventh conference of the Association for Machine Translation in the Americas (pp. 56-65).

Haddad, B., Awwad, A., Hattab, M., & Hattab, A. (2023). PRo-Pat: Probabilistic Root--Pattern Bi-gram data language model for Arabic based morphological analysis and distribution. PLoS One, 18(1), e0278992.

Harrat, S., Meftouh, K., Abouenour, L., Abbas, M., & Smaili, K. (2014). Building resources for Algerian Arabic dialects. In Proceedings of the 15th Annual Conference of the International Communication Association Interspeech (pp. 2123-2127).

Heath, J. (2002). Jewish and Muslim dialects of Moroccan Arabic. Routledge.

Hermans, T. (1999). Translation in systems: Descriptive and systemic approaches explained. St. Jerome Publishing.

Hetzron, R. (1997). The Semitic languages. Routledge.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. Acta Psychologica, 137(2), 151-171.

Jorden, E. H., & Walton, A. R. (1987). Truly foreign languages: Instructional challenges. The annals of the American academy of political and social science, 490(1), 110-124.

Jurafsky, D., & Martin, J. H. (2023). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition (3rd ed.). Pearson.

Karlsson, F. (1999). Finnish: An essential grammar. Routledge.

Kaye, A. S. (1994). Formal vs. informal in Arabic: Diglossia, triglossia, tetraglossia, etc., polyglossia—multiglossia viewed as a continuum. Zeitschrift für Arabische Linguistik, 27, 47-66.

Koehn, P. (2005). Europarl: A parallel corpus for statistical machine translation. In Proceedings of the tenth machine translation summit (pp. 79-86).

Koehn, P. (2010). Statistical machine translation. Cambridge University Press.

Koehn, P., Och, F. J., & Marcu, D. (2003). Statistical phrase-based translation. In Proceedings of the 2003 conference of the North American chapter of the Association for Computational Linguistics on Human Language Technology (pp. 48-54).

Kusters, W. (2003). Linguistic complexity: The influence of social change on verbal inflection. LOT.

Lambert, J., & van Gorp, H. (1985). On describing translations. In T. Hermans (Ed.), The manipulation of literature: Studies in literary translation (pp. 42-53). Croom Helm.

Larson, M. L. (1984). Meaning-based translation: A guide to cross-language equivalence. University Press of America.

Lieber, R. (2009). Introducing morphology. Cambridge University Press.

Lin, C. Y., & Och, F. J. (2004). Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics. In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (pp. 605-612).

McCarthy, J. J. (1981). A prosodic theory of nonconcatenative morphology. Linguistic Inquiry, 12(3), 373-418.

Miller, D., Kleber, M., Kao, C. L., Kimball, O., Colthurst, T., Lowe, S., ... & Schwartz, R. (2007). Rapid and accurate spoken term detection. In Proceedings of Interspeech (pp. 314-317).

Mohamed, E., & Sadat, F. (2014). Hybrid Arabic–French machine translation using syntactic re-ordering and morphological pre-processing. Computer Speech & Language, 28(4), 994-1009.

Montrul, S. (2004). The acquisition of Spanish: Morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition. John Benjamins.

Munday, J. (2016). Introducing translation studies: Theories and applications. Routledge.

Nord, C. (1997). Translating as a purposeful activity: Functionalist approaches explained. St. Jerome Publishing.

O'Grady, W., Archibald, J., Aronoff, M., & Rees-Miller, J. (2016). Contemporary linguistics: An introduction. Bedford/St. Martin's.

Owens, J. (2006). A linguistic history of Arabic. Oxford University Press.

Prince, A., & Smolensky, P. (2004). Optimality theory: Constraint interaction in generative grammar. Blackwell.

Pym, A. (2010). Exploring translation theories. Routledge.

Ratcliffe, R. R. (1998). The "broken" plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology. John Benjamins.

Robinson, D. (1997). Translation and empire: Postcolonial theories explained. St. Jerome Publishing.

Ryder, S. A. (1974). The D-stem in Western Semitic. Mouton.

Ryding, K. C. (2005). A reference grammar of Modern Standard Arabic. Cambridge University Press.

Sadiqi, F. (1997). Grammaire du berbère. L'Harmattan.

Sajjad, H., Fraser, A., & Schmid, H. (2013). An empirical study of word segmentation for statistical machine translation into morphologically rich languages. In Proceedings of the Eighth Workshop on Statistical Machine Translation (pp. 494-503).

Sapir, E. (1921). Language: An introduction to the study of speech. Harcourt, Brace and Company.

Schulz, E. (2004). A student grammar of Modern Standard Arabic. Cambridge University Press.

Shosted, R. K. (2006). Correlating complexity: A typological approach. Linguistic Typology, 10(1), 1-40.

Smrž, O. (2007). ElixirFM — Implementation of functional Arabic morphology. In Proceedings of the 2007 workshop on computational approaches to Semitic languages: Common issues and resources (pp. 1-8).

Spencer, A. (1991). Morphological theory: An introduction to word structure in generative grammar. Blackwell.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. Cognitive Psychology, 45(4), 447-481.

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. Harvard University Press.

Toury, G. (1995). Descriptive translation studies and beyond. John Benjamins.

Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence. Phonology, 16(3), 401-442.

Versteegh, K. (2014). The Arabic linguistic tradition. Georgetown University Press.

Vicente, Á. (2006). El dialecto árabe de Anjra (norte de Marruecos): Estudio lingüístico y textos. Universidad de Zaragoza.

Wehr, H. (1976). A dictionary of Modern Written Arabic. Spoken Language Services.

Zaidan, O. F., & Callison-Burch, C. (2011). The Arabic online commentary dataset: An annotated dataset of informal Arabic with high dialectal content. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 37-41).

2075
مجلة إكليل للدراسات الانسانية
التصنيف الورقي: العدد 23 /ايلول/2025
المجلد(6)- العدد(3)-الجزء(3) IASJ-Iraqi Academic Scientific Journals

التعقيد الصرفي ودقة الترجمة في أنظمة الجذر والوزن العربية

أ.م.د. مثنى حميد خلف
كلية التربية-جامعة الكوت
muthana.khalaf@alkutcollege.edu.iq

الكلمات المفتاحية: الصرف العربي، أنظمة الجذر والوزن، دقة الترجمة، التعقيد الصرفي، اللغات السامية، الصرف غير التسلسلي

الملخص:

يهدف البحث إلى دراسة التأثير المحتمل للتعقيد الصرفي في أنظمة الجذر والوزن العربية على دقة الترجمة، وبذلك يملأ فجوة في دراسات الترجمة حيث تم إهمال الجوانب الصرفية إلى حد كبير رغم دورها الأساسي في البنية المعجمية العربية. وبما أن العربية لغة سامية، فإنها تستخدم نظاماً صرفياً غير تسلسلي محدد يتم فيه دمج الجذور الثلاثية الصامتة مع الأنماط الصوتية لإنتاج كلمات تظهر مستويات مختلفة من التعقيد الصرفي وتشكل بذلك مشكلة لكل من المستخدمين البشر وأنظمة الترجمة الآلية. باستخدام التحليل الكمي للمدونات ثنائية اللغة والفحص المنهجي لنواتج الترجمة، يظهر هذا البحث أن التعقيد الصرفي الأكبر لأنظمة الجذر والوزن العربية مرتبط سلبياً بالترجمات المناسبة إلى اللغات البسيطة صرفياً مثل الإنجليزية. بالإضافة إلى تحليل أزواج الترجمة إحصائياً، يتم تحديد أنماط الأخطاء ومقارنة استراتيجية الترجمة بين المترجمين المحترفين وأنظمة الترجمة الآلية. تظهر نتائج التحليل أن الكلمات التي ترجموها تواجه صعوبة في القابلية للترجمة الداخلية بنسب عالية؛ وهذا عندما تحتوي الكلمات العربية على تركيبات معقدة من الجذور والأوزان، والتي تكون أعلى في الكلمات الأقل قابلية للترجمة (أي الحالات التي تميل فيها الترجمة إلى عدم الدقة). تحمل هذه النتائج آثاراً تعليمية فيما يتعلق بتدريس علم الصرف العربي الاشتقاقي للمترجمين المتدربين مع: (أ) الربط غير الخطي بين المعنى والشكل حيث أن العربية لغة عالية الاشتقاق وهذه التعقيدات غير مرجحة في الإنجليزية أو الفرنسية، (ب) الصعوبة الناشئة عن نقص المكافئات الصرفية الشكلية

عبر اللغات المصدر والهدف وإنتاجية العمليات الاشتقاقية لتوليد أشكال كلمات مختلفة ولكن مترابطة. تظهر دراسات الحالة الجزئية للترجمة بين الإنجليزية والعربية أن هذه التعقيدات تؤدي إلى كل من الفقدان المنهجي للمعلومات والمعاني المتعددة وإلى عدم التطابق الهيكلي الذي يؤدي إلى جودة ضعيفة في الترجمة. تضيف الدراسة أيضاً إلى أدبيات دراسات الترجمة من خلال تقديم أدلة تجريبية على تأثير تعقيد اللغة المصدر الصرفي على نواتج الترجمة، مع آثار عملية لتدريب المترجمين وتطوير أنظمة الترجمة الآلية والنمذجة الحاسوبية عبر اللغوية. تشير النتائج إلى أن المعالجة المسبقة الصرفية واستراتيجيات الترجمة المحددة مطلوبة لمعالجة التحديات في أنظمة الجذر والوزن العربية بشكل مناسب، ذات صلة في الترجمة بين اللغات الغنية والفقيرة صرفياً.