




*The Extent of Artificial Intelligence Accuracy within Translating Arabic  
Metathesized Forms into English*

*Alaa Hasan Abdullah  
Email: alaa88hasan@gmail.com  
9647508121327*

*Abu Baker Khalid Mustafa  
Email: Abubker.k.mustafa@nahrainuniv.edu.iq  
Department of Mechanical Engineering, College of Engineering, Al-  
Nahrain University, Jadriya– Baghdad*



مدى دقة الذكاء الاصطناعي في ترجمة تراكيب القلب المكاني من العربية الى  
الانكليزية

علاء حسن عبدالله

أبو بكر خالد مصطفى  
قسم الهندسة الميكانيكية، كلية الهندسة، جامعة النهرين، الجادرية – بغداد



## Abstract

Artificial intelligence (AI) has made significant progress in natural language processing and machine translation, but it still faces challenges when dealing with linguistic phenomena like metathesis. Metathesis is a common phonological process where sounds or letters are swapped within a word, resulting in considerable changes in word structure. This makes it a complex challenge for AI systems that rely on fixed patterns and rigid rules for translation. One of the primary obstacles AI encounters with metathesis translation is the abundance of linguistic variation across languages and dialects, Arabic in particular. Metathesis can also lead to phonetic discrepancies between Arabic and English, as Arabic relies heavily on guttural sounds, which may not have an equivalent counterpart in English. AI systems often struggle to apply context-sensitive rules to translate metathesized words accurately. Metathesis can occur in various languages and is often caused by phonological, morphological, or historical factors in addition to linguistic factors. The context in which metathesis occurs can greatly influence translation accuracy. Context-aware AI algorithms that take into account the broader linguistic context can bridge the gap caused by metathesis and deliver more accurate translations. In conclusion, AI systems must employ sophisticated algorithms and language models that can accurately identify compound words and comprehend their complex structures to ensure accurate translations.

Keywords: Artificial intelligence, metathesis, translation, accuracy, meaning, ambiguity

### المستخلص

لقد تطور الذكاء الاصطناعي (AI) بشكل كبير في معالجة اللغة الطبيعية والترجمة الآلية، لكنه لا يزال يواجه تحديات عند التعامل مع الظواهر اللغوية كالقلب المكاني. إن عملية القلب المكاني هي عملية صوتية شائعة حيث يتم تبديل الأصوات أو الحروف داخل الكلمة، مما يؤدي إلى تغييرات كبيرة في بنية الكلمة. وهذا ما يسبب تعقيداً لأنظمة الذكاء الاصطناعي التي تعتمد على أنماط ثابتة وقواعد صارمة للترجمة. إحدى العوائق الرئيسية التي يواجهها الذكاء الاصطناعي في الترجمة التبادلية هي وفرة التنوع اللغوي بين اللغات واللهجات، وخاصة اللغة العربية. من الممكن أن يؤدي القلب المكاني أيضاً إلى تناقضات صوتية بين اللغتين العربية والإنجليزية، حيث تعتمد اللغة العربية بشكل كبير على الأصوات الحلقية، والتي قد لا يكون لها نظير مكافئ في اللغة الإنجليزية. غالباً ما تواجه أنظمة الذكاء الاصطناعي مشاكل في تطبيق قواعد حساسة للسياق لترجمة الكلمات المترجمة بدقة. يمكن أن يحدث القلب المكاني في لغات مختلفة وغالباً ما يكون سببه عوامل صوتية أو مورفولوجية أو تاريخية بالإضافة إلى العوامل اللغوية. يمكن للسياق الذي يحدث فيه القلب المكاني أن يؤثر بشكل كبير على دقة الترجمة. يمكن لخوارزميات الذكاء الاصطناعي المدركة للسياق والتي تأخذ في الاعتبار السياق اللغوي الأوسع أن تسد الفجوة الناجمة عن القلب المكاني وتقدم ترجمات أكثر دقة. وفي الختام، يجب أن تستخدم أنظمة الذكاء الاصطناعي خوارزميات ونماذج لغوية متطورة يمكنها تحديد الكلمات المركبة بدقة وفهم بنيتها المعقدة لضمان ترجمة دقيقة.

الكلمات المفتاحية: الذكاء الاصطناعي، التحويل، الترجمة، الدقة، المعنى، الغموض

## **1.1 An Introduction to the Development of Metathesis**

Dwight (1975:318) states that metathesis is found occasionally in English as compared to ancient languages, which refers to a shift in sound order that sporadically occurs in the relative arrangement of word letters. Occasionally, it is attested as a formal metathesis rather than an analytical one with words from other modern languages. Bloomfield (1983:216) confirms that metathesis can be regarded as a main part of language history because it was originally a gradual process. Joseph and Hock (1991: 140) state that metathesis is frequent in old English dialects and even in the prehistory of old English, such as *brestan* > *berstan* in old English and *bersten* > *bresten* in Middle English. It is the process of analyzing a segment in a non-historical position, caused by phonetic effects. This phenomenon is common in current English, especially in dialects. As far as historical point of view is concerned, Hogg (1992:110) states that metathesis of adjacent consonants was commonly present in Old English, especially when one of the consonants was (s) like *wasp* – *waps*. Such a change is not structurally valuable, but it should be mentioned since metathesis persists throughout the history of language.

According to Katamba (1993: 346), metathesis is a difficult concept to avoid since it is linked to the historical evolution of language worldwide. In the relatedness of inversion of a word order, just like in English, the structure “*he can, can he?* or the stress elements inversion, as in, /i'mport/, /im'port/, . These could be a kind of metathesis. According to Fukazawa and Miglio (2008: 12), metathesis is thought of as an erratic, surface-level impact associated with speech faults. It would only be historical or developmental in the context of learning a first language. According to Moskal (2009: 8), the metathesis process is the outcome of diachronic phonetic alterations, which frequently

convert a non-metathesized form into a metathesized form that is now seen over time.

Metathesis, according to Bower and Crowely (2010: 32), is not a typical type of alteration that just involves changing the order of sounds rather than deleting or adding any. A famous example of a metathesis in English words that became normal is "bird," which was originally pronounced "brid" before changing to "bird" in the spelling system. The sound "ir" went through many changes in some dialects to become /ɜ:/ at present. Not all sounds are affected by metathesis, but some of them can be commonly affected, such as liquids and the /l/ and /r/ sounds.

Hume (2004: 57) indicates that the phenomenon of metathesis is striking in phonology as compared to syntax, which is typical. Metathesis traditionally resists a unified phonological theory, which leads to arbitrariness. The nature of metathesis results from the perpetuation of the metathesis view, which insists that metathesis nature is sporadic and irregular and is limited to performance errors, child language, and language change. Historical linguistics indicates that metathesis arises in nearly all languages as a result of child phonology and performance faults. Since metathesis is not governed by rules and is considered a little acoustic alteration, it has been historically examined and is marginal in nature.

Joseph and Hock (1991: 141) state that metathesis occurs in a way beyond individual words and affects the utterance as a whole; in this case, it is called spoonerism, which indicates the speech errors in the faulty phonetic production of metathesis and dissimilation. Metatheses and dissimilation occur when the person feels tired or drunk; his or her speech would be full of many contractions as well. They were historically attested as temporary mistakes and became a permuting feature in language.

Paunović and Čubrović (١٣٩ :٢٠٠٩) say that metathesis is the process of reordering segments; it is frequently treated as a historical change more than as a synchronic process. This interchange in Old English has an effect on many consonant clusters, whereas in Middle English, its effect was represented by the /r/ sound. This interchange is going on so far, and many dialects are affected. Hume (2004: 45) shows that the patterns of native language have a potential effect on metathesis in some varieties of English language, especially the pronunciation of /t-l/, such as in the word "chipotle," a kind of pepper used in chains of Mexican restaurants, which can be pronounced within the same individual with two alternated orders. As a result of two views: the unfamiliarity of the borrowed words and the indeterminacy of sound, which is based on experience of mastering.

In Arabic, Adhima (1988: 44) points out that Kufi scholars see that metathesis is a transposition of letters of the word with each other, resulting in two words without changing the meaning, on the condition that the basic root of both words is found, e.g., *lafata* > *falata*) *فلت لفت*. attracted). Ibn Cinni, one of the Basra scholars, on the other hand, maintains that whenever there is transpositioning or prepositioning resulting in two pronunciations for the word, they are not metathesized, but each has its own conjugation.

According to Abd Altawab (1988:49), it has been observed that a large number of Arabic terms have often been metathesized to the point where it appears that they have rudimentary linguistic foundations. Scholars link the uniqueness of metatheses to fundamental foundations since they didn't see that. Alnahaas and Alsaqawy assert that if words were metathesized, their metathesis would be limited to the origin of the unmetathesized term and would not have a fundamental root to avoid confusion with its

origin. Some Arabic grammarians believe that if each metathesized form has an origin, the word then is not metathesis, and this will be asserted as a dialect. Arabic linguists, on the other hand, say that metathesis is not a dialect.

Maamouri (2005:2) clarifies that each Arab country has its own use of spoken language, which is called Ammiyyah "vernacular." These vernaculars may make use of some metathesized forms, which differ considerably among speakers from different parts of the Arabian Peninsula.

According to Hallo (1996: 256), a few factors that are thought to be important causes of metathesis lead to the phenomenon. This is a common occurrence in language and can be produced by a variety of factors, such as speaker knowledge or dialect variations, which can result in a word being pronounced metathetically in two dialects. The main cause of some words becoming metathesized may be word mishearing, and Arab Bedouins and those living in rural regions are particularly prone to this problem. The underpinnings of metathesis are addressed, including language friction within a community, phonological development synchronically, ease of articulation, geographic aspects, and child learning.

## 1.2 Roots of Arabic Metathesis

Adhima (1988: 44) points out that Arabic has certain hypothetical rules for metathesis according to the verb root (فعل) [fa'ala]:

1-Prepositioning Lam (ل) to Ayeen (ع) in the form of فلع such as the word " لعلق lick "علق" lick.

2-Prepositioning the first Lam (ل) to Ayeen (ع) in the form of فعلول such as Ghadruf (cartilage) غضروف interchanged into "Ghardouf غرضوف".

3-Prepositioning the second Lam (ل) to Faa (ف) in the form of [تفعللت *taffa* 'alatu] such as budgetت زحزت is interchanged into 'budgetت زحزت' which became as a form ofتفعللت. *tala* 'faltu]

As to the form فعلاء *fa* 'ala] the Lam ل is transposed to ف to become لفعاء *laffa* ' ] such as the word اشياء things, which has a dipotot form فعلاء originated as شيئا, but it was metathesized according to transpositioning the ل to ف

4- Prepositioning Ayeen (ع) to Faa (ف) in the form ofتفعللت. *taffa* 'alatu]. Prepositioning Lam (ل) to Faa (ف) in broken plural can cause metathesized forms in Arabic, such as (صاع , sa'a , asa' اصع), (راس, ra's, ara's ارس), (بئر, bi'r, abaar ابار)

5-Postpositioning Faa (ف) to Lam (ل) in the form of (فاعل) *fa* 'il] which is interchanged into (عالف) [*3alif*] then (و)'waw] is interchanged into [yaa(ي) like in the Arabic poetic verse,

وما اعتاد حب سليمي حين معتاد ولا تقضي بواقي دينها الطادي

The word الطادي is the metathesis of the word واطد or وطف

Wajeeh (2010: 9) believes that according to AlKhaleel's opinion, metathesis occurs in case the agent is derived from a hollow-triple verb with glottalized Laam (ل) where it is interchanged into فاعل not فاعل. The verb "جاء came" is originated as "جاء" but the letter (ي) became dynamic and the letter before it has undergone to Fatha, so that it was interchanged into (ا). It is an agent derived from a triple verb on the basis of the verb form (فاعل) *fa* 'il] then the verb (جاء, جاء) became (جاء) according to the morphological rule argued that if (و) or (ي) were the basic letter of the agent and preceded by (ا), then they are changed into Hamza, because Arabic does not accept subsequent Hamzas.

Plural of فعيلة *fa* 'i'lah, glottalized laam ل which is interchanged into the verb form فعاللي *fa* 'ali] not فعائل such as the word خطيبة based on the verb form فعيله and it has the plural خطايب according to the verb form فعائل *fa* 'ali. In order to avoid doubled

Hamzas and to attenuate speech, the word خطيئة has undergone transposition, where Hamza is interchanged into (ي) to avoid triple (ل) in one word, e.g. "خطاء عاي" so, it becomes خطايا to ease pronunciation. Sibawayh confirmed that if there are two Hamzas at the end of a word, one of them should be changed into (ي).

As for Basra scholars, metathesis can be found when the stem verb فعل by fronting laam ل to Ayeen ع such as the case found in the verse of Katheer Ibn Ghzaah (saw) (رَأَى) in (رَأَى)

وَكُلُّ خَلِيلٍ رَاعِنِي فَهُوَ قَائِلٌ  
مِنْ أَجْلِكَ هَذَا هَامَةٌ الْيَوْمِ أَوْ غَدٍ

The origin of راعني in the verse above is (رَأَى).

Metathesis is also found when fronting Ayeen ع to Faa ف in the stem verb فعل *fa l/* such as, (أَيْسَ) (despaired) the metathesis of (يَيْسَ) and (هَفَا فُوَادِي) for (هَفَا فُوَادِي) and يَأْسَل for the word (يسال ask). This case is found in Tamam's verse.

إِذَا قَامَ قَوْمٌ يَأْسَلُونَ مَلِيكَهُمْ  
عَطَاءً فَدَهْمَاءُ الَّذِي أَنَا سَائِلُهُ

Metathesis also takes place when Faa ف is postponed to Laam ل such as واحد حادي. Additionally, when the Laam ل is fronted to Faa ف such as أَشْيَاءٌ أَشْيَاءٌ. Some Arab morphologists confirm that when metathesis is recognized, there would be two ways to figure it out: back to the base form and to the word derivations.

### 1.3 Metathesis as a Physical Error.

The distinctive features of metathesis led to the creation of the "metathesis myth," which was commonly believed to be chaotic and irregular and restricted to linguistic errors made by children, errors in performance, or fluctuations in sound. One prevalent aspect of child phonology is metathesis, which often occurs in all languages as a type of speech error. It sometimes appears as performance faults, such as tongue slips, or as the phenomenon called "spoonerisms," which is also historically thought to be a slight irregular and random modification in sound.



(Hume, 2001:32; Spencer, 1996:68; Crystal, 1997:240; Powell, 1985:106).

According to Hume (2001: 34), metathesis was considered sporadic and irregular, restricted to performance errors, child language, or sound change. Metathesis occurs frequently in all languages as a type of speech error and is a common feature of child phonology. The most well-known type of these categories is the error in sound that is usually called "malapropisms." Aitchson (1994: 44) states that this type of error suggests meaning and sound to be partially linked in the mind, and also that the linking up involves the activation of a number of words that are similar to the target.

### **1.3.2 Spoonerism as a Metathetic Factor**

Normal speech contains a fairly large number of slips of the tongue, which mostly pass unnoticed. These slips of the tongue are defined as "an unintended mistake made in speaking." (Trask, 1995:38).

These metathesized errors indicate that chunks of speech are pre-prepared for utterance, possibly in a tone group (a group of words spoken within the same intonation pattern), and that the activated words are organized in accordance with a rhythmic principle. (Ladefoged 2001:63).

### **1.4. Processing Arabic Text**

In communication, language is a sophisticated evolutionary construct that is a fundamental manifestation of human intellect. There are several obstacles in the way of a computer being able to "understand and respond" to such distinctly human concepts. The "portability" of English language approaches to other languages is also restricted to the great diversity of languages and their unique phonetic, grammatical, and cultural characteristics. Furthermore,

computers processing natural language is becoming increasingly necessary for many languages worldwide. (Bird, 2019: 13)

In classical Arabic, diacritical marks are often used in poetry, legal texts, and educational publications. They explain how a word should be pronounced and signified. The primary issue with modern Arabic is that, in an effort to save space, diacritical marks are omitted. As a result, a word in contemporary Arabic might have several meanings based on the various diacritical forms it can take. For instance, "كتب qtb" might indicate "books" (كُتِبَ qutub, a noun), "he wrote" (qataba كَتَبَ a verb), or "it has been written" (qutiba كُتِبَ, a past participle). (Saal, 2018, 3)

For processing Arabic text, tokenization, stemming, lemmatization, part-of-speech tagging, and named entity identification are a few of the processes involved in natural language processing, or NLP. Arabic has a complicated morphological structure with a rich system of prefixes, suffixes, and infixes that may alter a word's shape and meaning, making these tasks difficult. Furthermore, Arabic is a highly inflected language, which means that a word can take on several forms based on how it functions syntactically in a phrase. (Deep, 2023, 6) (Atwan et al. 2021, 23)

When dealing with Arabic natural language processing (NLP) activities, there are a few issues to resolve:

- Orthographic variations: Arabic script has many orthographic norms to express short vowels, long vowels, and various vowel sounds. Words are written without gaps between them. Because of this, correctly tokenizing Arabic text may be difficult.
- Morphological complexity: Words in Arabic can include a large number of inflexions and affixes due to the language's highly inflected and agglutinative morphology. Because of

this, it may be challenging to determine a word's basic form and part-of-speech tag with accuracy. (Isaa 2022: 10)

- Syntactic ambiguity: Arabic has a complex system of nominal and verbal suffixes, which can make it challenging to understand how a phrase is put together syntactically. A single verb, for instance, can convey several voices, moods, and verb tenses.
- Variations in speech: There are about 400 million Arabic speakers in the Middle East and North Africa, and these speakers speak the language in a variety of dialects. The vocabulary, syntax, and pronunciation of these dialects might differ greatly, making it challenging to develop NLP systems that are compatible with them.

In comparison to English or French, there are very few annotated Arabic language corpora and NLP tools accessible. This leads to the fact that creating excellent NLP systems for Arabic difficult. (Deep, 2023, 9)

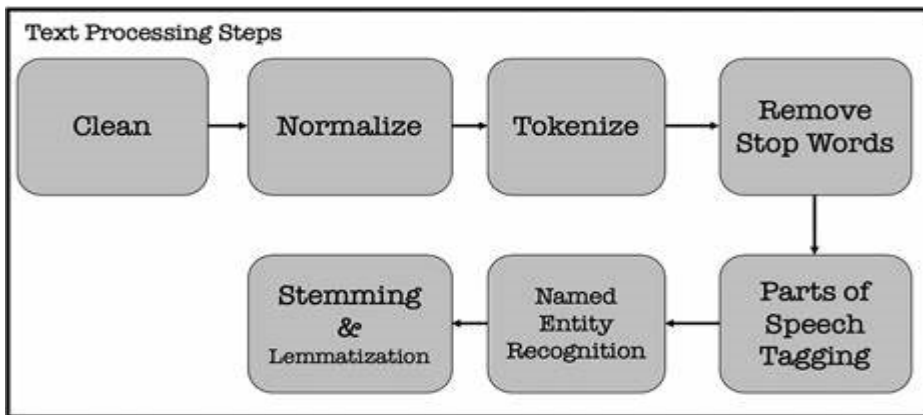
Arabic text is processed within artificial intelligence using natural language processing (NLP) techniques. These techniques involve tokenization, stemming, and lemmatization to break down the text into its basic components and understand the meaning of the words. Isaa (2022: 13) maintains that there are some of the NLP resources and tools that AI technologies offer:

- Tools for tokenization are those that divide Arabic text into discrete words or tokens.
- Stemming: Arabic stemmers specifically made to handle the language's intricate morphological structure.
- Lemmatization is the process of taking a word and breaking it down to its most basic form, or lemma, while taking the context and part of speech of the word into consideration.

- Tools for labeling words with their part of speech, such as nouns, verbs, adjectives, etc., are known as part-of-speech tagging tools.
- Named entity recognition is the process of locating and labeling named entities, such as proper nouns, businesses, locations, etc., using tools.

Additionally, machine learning algorithms are trained on large datasets of Arabic text to improve their understanding and processing of the language. This allows AI systems to perform tasks such as language translation, sentiment analysis, and text generation in Arabic.

Arabic metathesized form is processed within artificial intelligence by using various NLP techniques. These techniques involve analyzing and understanding the structure of the Arabic language, including its phonetics, morphology, and syntax. AI systems use algorithms to detect and handle metathesized forms, such as swapping letters or sounds within words, in order to accurately interpret and process Arabic text. Additionally, machine learning models are trained on large datasets of Arabic to improve their ability to recognize and process metathesized forms.



Source: Kasaraneni (2021) *Understanding NLP Pipeline: An introduction to phases of NLP pipeline*. Egen DatTeam Publication.

## **1.5 Artificial Intelligence (AI)**

Garnham (2017: 22) confirms that Artificial intelligence (AI) is a behavioral analysis technique that operates under the premise that replicating intelligence is the most effective way to study it. Reproduction, in practical terms, refers to computer simulation. AI falls under the category of computer science as a result. Its history is brief; as an autonomous subject of study, it began in the middle of the 1950s. For a long time, one of the main concerns in psychology has been intelligence. This is not only intelligence as it is measured in IQ tests, but intelligence in a broader sense that is necessary for thinking, reasoning, and learning, as well as for their prerequisites: high-level perceptual skills, the ability to represent information in the mind, and language proficiency.

According to Klepeis (2019:4), AI enables robots to carry out jobs that ordinarily call for human intellect and solve challenges. Does this imply that AI-equipped machines are capable of independent thought and decision-making? Not quite. The first step in programming a computer is for people to use software that gives them clear instructions. The term "algorithm" is often used in relation to AI, which is an algorithm is a thorough set of guidelines for completing a task or resolving a problem. So, the set of instructions required to build a chair is called an algorithm.

Mitchell (2019: 17) asserts that most people would agree that humans are intelligent and specks of dust are not. A single scale is used to test IQ, although there are other aspects of intelligence as well, such as emotional, linguistic, spatial, logical, artistic, social, and so on. For this reason, intelligence can be multidimensional (i.e., someone might have high verbal intelligence but poor emotional intelligence), binary (i.e.,

something is or is not intelligent), or on a continuum. In fact, intellect is like an overstuffed suitcase with a zip that's about to burst. For better or worse, these different characteristics have not received much attention in the area of AI, but rather, the attention has been on two initiatives: a practical one and a scientific one. Scientists studying artificial intelligence are attempting to incorporate biological intelligence into computers in order to better understand its workings. Practically speaking, advocates of artificial intelligence (AI) only seek to write computer programs that can carry out jobs just as well as or better than people, without considering whether or not these systems are truly capable of thinking like humans. Many in the field of artificial intelligence quip that it depends on where their current financing comes from when asked if their motives are scientific or practical.

### **1.5.1 AI & Meaning**

Fitch (2010: 41) states that meaning is the message that words, phrases, and symbols transmit in a given context; this process is known as semantic or lexical meaning. He continues, "There are two kinds of meaning: pragmatic and semantic." It has long been believed that in order to comprehend the meaning a speaker intends to convey in a phrase, we must grasp two different kinds of meaning. Extra-pragmatic meaning derives from the specific context in which a phrase is stated. Semantic meaning, which makes up the majority of a sentence's content, is expressed in a sentence.

Syrett and Arunachalam (2018: 19) say that knowing a language allows you to understand things like if a word has two meanings, whether they are meaningful or meaningless, and what terms actually mean and what they don't imply in everyday speech.

Changder (2023: 55) clarifies that words can have two meanings: denotative and connotative. Dictionary meanings of

words or definitions that are agreed upon by the linguistic community at large are referred to as denotations. For instance, a man who tends to livestock is what the name "cowboy" signifies. A person who is autonomous and/or carefree is another meaning. Because it has so many meanings, a more abstract term like "change" would be more challenging to comprehend. Change and cowboy are both regarded as polysemic terms since they have several meanings. Definitions based on people's emotional or experiential associations with a term are referred to as connotations. To return to our earlier discussion, a person's experiences can determine whether change is seen positively or negatively.

### **1.5.2 AI: Accuracy vs Ambiguity**

Wilks (2008:12) argues that machine translation was not only practically but theoretically impassible, where "impossible" meant just that, and not merely difficult. Proficient human translators employ their prior knowledge, primarily unconsciously, to settle syntactic and semantic problems that computers will either have to ignore or resolve using a "mechanical" approach that frequently produces an incorrect translation.

Carl and Way (2012:7) believe that an exponential amount of material has to be translated, not just much more swiftly than in the past but also considerably more affordably, as competition drives innovation and technology grows more complicated. Therefore, it is now essential to comprehend how to expedite, accurately, and effectively translate text using both technological programs and manual methods. The use of machine translation and post-editing MT output has become more crucial in this situation. A related but equally important development is that translation scholars have shifted their attention from conceptual and pedagogical difficulties to systematic empirical data about real-

world translation tasks, specifically data concerning industrial and cognitive translation processes.

Soudi et al (2012:5) comment that Machine Translation (MT) is primarily concerned with the mechanical replacement of words from one language with another. However, because it requires the recognition of whole phrases and their closest counterparts in the target language, this function alone virtually never yields a correct translation.

Kenny (2022:35) admits that the largest circle in the category hierarchy is artificial intelligence (AI), which is the most generic. It is commonly described as the area of computer science that seeks to build devices, or more precisely, computer programs, that are capable of solving issues that would typically call for human intellect. It is reasonable to argue that the type of intelligence that strong AI strives for is necessary for translation as performed by qualified human translators, but that machine translation systems are presently unable to provide such intelligence. AI is crucial to the translation process. By giving them tools to automate some parts of the translation process, it may make their job more precise and efficient. For example, AI translation tools can help translators identify and correct errors in translations, suggest alternative translations, and even provide translations in real-time. Also, AI can enable people who do not speak the same language to communicate with each other. AI translation tools can provide instant translations of spoken or written content, which can be useful in situations where a human translator is not available.

In conclusion, since machine translation relies heavily on artificial intelligence and user translations, one of the biggest issues you may encounter while utilizing services like Google Translate is inaccuracy. For instance, machine translation reveals that the Arabic term علم has several English counterparts, including



flag, teach, knowledge, science, and more. As a result, selecting one of these equivalents will depend on the translator's experience and the ST's context. Because machine translation is not very good at understanding the meanings and cultures of the context, it might make a lot of mistakes. Additionally, it ignores linguistic norms since language rules differ from one another. As a result, machine translation can only be used to provide background information and context. The accuracy and consistency of translations may be enhanced in the future by AI translation technologies' potential to learn from and adjust to the writing styles of specific authors. Regarding artificial intelligence, the preceding elucidation has a close relationship to word translation through the utilization of AI techniques. owing to the information that many scholars have supplied on denotative and connotative meaning and how they affect meaning. For example, the word home refers to the place where you live; it could be a house, an apartment, etc. This is the word's denotation. For many people, the word home has a positive connotation—it's associated with safety, comfort, and a sense of belonging. By using one of the AI tools (Google Translate), the same word “home” gives more than 10 equivalents in Arabic which can be a noun like in “وطن”, “منزل” or an adverb like in في البيت etc.

AI is crucial to the translation process. By giving them tools to automate some parts of the translation process, it may make their job more precise and efficient. Artificial intelligence (AI) translation systems, for instance, may assist translators in locating and fixing translation mistakes, offering substitute translations, and even delivering translations instantly. AI can also facilitate communication between speakers of different languages. When a human translator is not available, artificial intelligence (AI) translation systems may translate spoken or written text instantly.

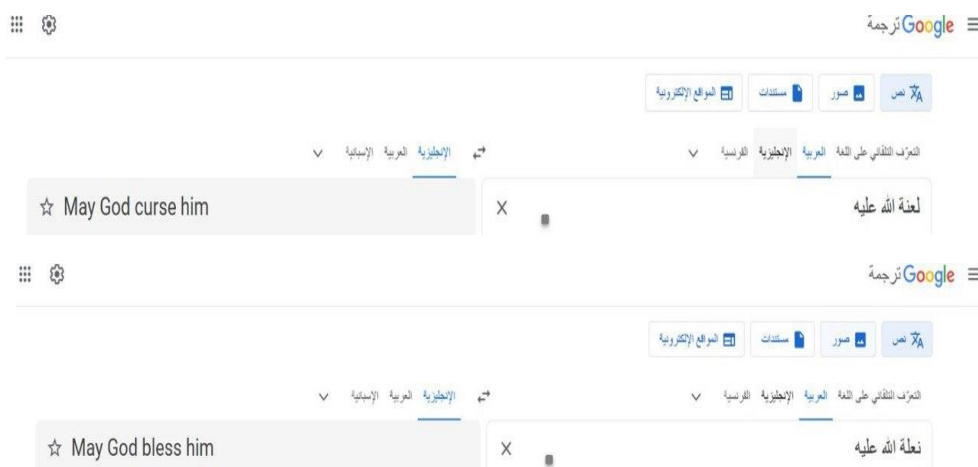
## **1.6. Discussion and Data Analysis**

This research will discuss the problems faced by AI in translating metathesis from Arabic to English, including linguistic differences, contextual ambiguity, and the limitations of training data. Therefore, Artificial Intelligence (AI) has made great strides in natural language processing and machine translation, aiding communication between people of different languages. However, it still faces significant challenges when dealing with linguistic phenomena like metathesis (the rearrangement of sounds or letters within a word). Metathesis is a common phonological process present in many languages, where sounds or letters are swapped within a word. For example, in English, the word "ask" has undergone metathesis to evolve into "aks" in certain dialects. Metathesis can result in considerable changes in word structure, rendering it a complex challenge for AI systems that rely on fixed patterns and rigid rules for translation. The first problem lies in capturing the true essence of the word and its intended meaning. Metathesis often alters the root of a word, which can completely change its semantic implications. Consequently, when translating metathesis, it becomes crucial to comprehend the underlying concept that the word seeks to convey. It necessitates more than a direct substitution of letters from Arabic to English. Instead, the translator must possess a deep understanding of the cultural, historical, and linguistic contexts to ensure an accurate representation of the original meaning.

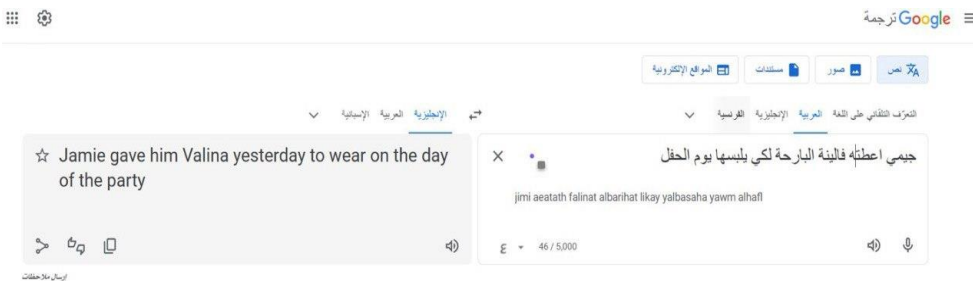
Linguistic variation is one of the primary obstacles AI encounters with metathesis translation is the abundance of linguistic variation. As metathesis varies across languages and dialects, capturing these nuances and creating accurate translation models becomes a difficult task. Lack of sufficient data could be another problem for metathesis translations in AI, because AI

algorithms heavily rely on data to learn and improve their translation capabilities. However, when it comes to metathesis, the availability of sufficient annotated data remains limited. The scarcity of comprehensive datasets exhibiting varied instances of metathesis inhibits the AI systems' ability to learn and generalize the rules accurately. Consequently, AI algorithms struggle to recognize, analyze, and translate metathesized words without a robust foundation of training data.

Some other problems regarding inaccuracies of metathesized forms in AI is the irregularities and ambiguities which relates to the fact that metathesis introduces irregularities and ambiguities in word structures, making translation even more challenging. Nonetheless, the rearrangement of sounds or letters often leads to morphological and phonological changes that alter the meaning of words. This inherent unpredictability in metathesis exacerbates the difficulties faced by AI models as they attempt to decipher the correct translation. Ambiguities arising from metathesis can lead to multiple plausible translations, increasing the chances of errors and inaccuracies. For instance,



The word “لعنة” curse” in Arabic which undergoes methathesis as “نعلة” curse” was translated completely opposite into English as ‘bless’ due to changing the order of the letters. Thus, changing a sound or a letter in Arabic led AI fail to fetch the meaning of the metathesized word. Moreover, metathesis can lead to phonetic discrepancies between Arabic and English because Arabic relies heavily on guttural sounds, which may not have an equivalent counterpart in English. So, when translating metathesized words, maintaining the integrity of these unique Arabic sounds becomes challenging. English, being a language with different phonetic rules, often cannot reproduce the original Arabic pronunciation accurately. As a result, the translated word may lose its intended phonetic impact and cultural significance. Contextual awareness translation becomes particularly problematic when metathesis occurs alongside other linguistic phenomena or in specific contexts, where AI systems often struggle to apply context-sensitive rules to translate metathesized words accurately.



AI in this sentence could not recognize the word “فألينة” as “فأنيلة” even in the context of the sentence. So, ambiguities can emerge if AI algorithms fail to recognize the underlying linguistic and cultural contexts, leading to inaccurate translations that can be confusing for the end-users.

Although the field of artificial intelligence (AI) has seen significant advancements, especially in the area of Natural Language Processing (NLP), which focuses on enabling computers to understand and generate human language, it still has an obvious challenge when translating Arabic metathesis into English. One of the major difficulties in translating metathesis from Arabic to English arises from the significant linguistic differences between the two languages. Arabic is a Semitic language that follows different grammar rules and has a complex morphology compared to English, which is an Indo-European language. Consequently, when encountering words with metathesis, AI often struggles to interpret them accurately due to its limited understanding of Arabic morphology. For example, in Arabic, the verb “قتل qatala” meaning "he killed" can undergo metathesis to become “قتلة qatla”, which may be erroneously translated as "cut" instead of "killed" by AI systems that lack the contextual knowledge of metathesis in Arabic. Moreover, contextual ambiguity poses another substantial challenge in translating metathesis. Arabic is known for its rich context-dependent language, where the meaning of words can change based on the surrounding context. This becomes particularly problematic when translating metathesized words since the proper interpretation of the word often relies on the context in which it appears. For instance, the word “ملس malsa” in Arabic can mean both "nightstand" and "sentenced." Without enough contextual information, AI may struggle to accurately determine the intended

meaning of metathesized words and could potentially produce mistranslations. Furthermore, the limitations of training data pose a significant obstacle in translating metathesis from Arabic to English using AI. Another example is the Arabic word “مكتوب” *maktub* which can mean “written” or “destiny” depending on the context. If metathesis occurs during translation, the word may inadvertently be switched, leading to a different translation. This issue can confuse both the translator and the recipient of the translated text, potentially causing misunderstandings and misinterpretations.

Metathesis can occur in various languages and is often caused by phonological, morphological, or historical factors. For instance, in English, metathesis can be observed in words like “comfortable” pronounced as “comfterble” or “nuclear” pronounced as “nucular”. However, these examples may seem insignificant, the proper understanding and interpretation of the intended meaning are crucial, especially in translation tasks. Artificial intelligence has made remarkable strides in natural language processing (NLP) and machine translation, although challenges persist when it comes to accurately translating languages that exhibit metathesis. In addition to linguistic factors, the context in which metathesis occurs can greatly influence translation accuracy. Metathesis can sometimes lead to a change in meaning or create ambiguity. Therefore, context-aware AI algorithms that take into account the broader linguistic context can bridge the gap caused by metathesis and deliver more accurate translations. Moreover, idiomatic and culturally specific Arabic expressions often use metathesis for rhetorical or poetic purposes. Translating these expressions into English requires a deep understanding of both languages and their linguistic features. Accurate translation using AI systems requires sophisticated

algorithms and language models to identify and understand complex structures.

## **1.7 Conclusions**

The field of artificial intelligence has seen great progress in recent years, with machine translation systems playing a key role in cross-language communication. However, translating metathesis from Arabic to English remains a challenging task for these systems. Arabic's complexity poses a significant obstacle to accurate translation. Arabic words often undergo metathesis due to root-and-pattern morphology. Translating metathesis in artificial intelligence from Arabic into English poses various challenges and complexities. This process involves the conversion of text, primarily linguistic or textual data, from one language to another, thereby facilitating effective communication between different cultures and societies. In the case of Arabic to English translation, metathesis can introduce ambiguities and inaccuracies, requiring translators and researchers in artificial intelligence to carefully analyze and address these issues. Metathesis can alter the intended meaning of a word or sentence and subsequently impede the translation process. Additionally, metathesis in Arabic can introduce difficulties in transliteration, a process that involves converting the Arabic text into its phonetic equivalent in English. Metathesis in Arabic words can cause pronunciation inconsistencies and phonetic irregularities when transcribed into English. This can make it difficult for AI systems to accurately process and reproduce affected words. The differences in structure and syntax between Arabic and English add to the complexity of translating metathesis in AI. This can impact not only individual words but also the overall sentence structure, affecting the meaning and context of the text. Arabic sentences tend to follow a different order, with the verb frequently placed at the beginning,

whereas English sentences commonly place the verb in the middle or towards the end. When metathesis alters the order of words, translating the sentence can become challenging, as the intended meaning may be obscured.

## **1.8 Recommendations**

The current research recommends potential solutions to address the challenges of translating metathesis in AI systems, several potential solutions can be explored:

1. **Enhanced Data Collection:** Collecting and annotating more diverse linguistic data, specifically focusing on metathesis, would provide AI algorithms with a broader understanding of this phenomenon. This increased dataset would boost the learning capabilities of AI models and enable more accurate translation.
2. **Linguistic and Cultural Context Recognition:** Developing AI systems that have a deeper understanding of linguistic and cultural contexts can significantly enhance translation accuracy. By recognizing specific linguistic patterns and contextual clues, AI algorithms can make informed decisions when translating metathesized words.
3. **Rule Adaptability:** Building rule-based AI models that are adaptable to different metathesis patterns across languages and dialects can improve translation accuracy.

Such models should incorporate context-sensitive rules to accurately analyze and translate metathesized words by considering their morphological and phonological implications. One approach to tackling the challenges of metathesis translation in AI is by incorporating robust language models that are trained explicitly on languages with metathesis patterns. By exposing AI algorithms to large datasets of languages that exhibit metathesis,



the systems can learn the underlying patterns and rules, thus allowing for more accurate translations. Furthermore, leveraging phonetic and phonological features of languages can aid AI in handling metathesis translations. By analyzing the phonetic structures and recognizing consistent patterns of metathesis, AI models can make more informed predictions during the translation process. Techniques such as phonetic transcription or phonological preprocessing can be employed to convert the input text into a phonetically aware representation, assisting the translation model in recognizing and reconstructing words with metathesis.

## References

- Abdul- Tawab, Ramadhan (1988). *Buhooth wa Maqalat fi Allugha* (Researche and Articles in Language). 2<sup>nd</sup> edn. Cairo: Maktabat Al-Khanchi
- Adhima, (1988) *The dictionary of verb conjugations [Almughni fi tasrif alafaal]*. Egypt. Al-Sa'ada Press.
- Aitchison, Jean (1994). *Words in the Mind: An introduction to the mental lexicon*. Blackwell:New York
- Atwan, J, Mohammad Wedyan, Qusay Bsoul, Ahmad Hammadeen, and Ryan Alturki (2021). *The Use of Stemming in the Arabic Text and Its Impact on the Accuracy of Classification*.
- Bird, S. (2019) *Language Archives*. Language community press.
- Bloomfield, L. (1983) *An Introduction to Study of Language*. Oxford university press.
- Bolinger, Dwight. *Aspects of language*. Yale University Press, 1975.
- Carl M. and Way A. (2012) *Recent Advances in Example-Based Machine Translation*. CUP.
- Changder N. (2023) *Connotation and Denotation*. Blackwell
- Crowley, T, & Bower, C. (2010) *An Introduction to Historical Linguistics* Oxford University Press.
- Crystal, D. (1997) *A Dictionary of linguistics and phonetics*. United Kingdom: Black Well publication LT 1<sup>st</sup> Edn.
- ————. (1997) *A Dictionary of linguistics and phonetics*. United Kingdom: Black Well publication LT 4<sup>th</sup> edn.
- Deep, (2023) *A Complete Guide to Natural Language Processing*.
- Dewight (1975). *Aspects of Language*, 2nd edn. New York
- Fitch W. Tecumseh (2010). *The Evolution of Language*. John Benjamine Publishing
- Fukazawa, Haruka and Viola Miglio. (2008) *What Metathesis Reveals about L1 and L2 Acquisition*. In *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*, ed. Roumyana Slabakova et al., 30-36. Somerville, MA: Cascadilla Proceedings Project
- Garnham A. (2017). *Artificial Intelligence: An Introduction*. Revivals
- Hallo, W, William (1996) *Origins: The Ancient Near Eastern Background Some Modren Western Institutions*. Eje.Jpril.

- Hogg R, M. (1992) the cambridge history of English.
- Hume, E. (2001) Metathesis: formal and functional considerations. Blackwell Press.
- ———. (2004) Indeterminacy of metathesis, ohio state university
- Isaa, I. (2022). Morphological Complexity in Arabic Spelling and its Implications for Cognitive Processing. Journal of psycholinguistics.
- Joseph and Hock (1991) Principles of Historical Linguistics. Oxford University press.
- Kasaraneni, K. (2021) Understanding NLP Pipeline: An introduction to phases of NLP pipeline. Egen DatTeam Publication.
- Katamba, F. (1993). Morphology. Great Britain: Mackays of Chatham PLC
- Kenny D. (2022) Machine Translation for Everyone. Language Science Press
- Klepeis A. (2019). Artificial Intelligence and Work. Raintree
- Ladefoged, P. (2001). Vowels and consonants: An introduction to the sounds of languages. Malden,MA: Blackwell.
- Maamouri, M. (2005) Variants of Arabic. Online article can be retrieved from ([www.variantsofarabicpdf.com](http://www.variantsofarabicpdf.com)).
- Mitchell M. (2019) Artificial Intelligence: A Guide for Thinking Humans. Springer.
- Moskal, B. (2009) The illusions of metathesis. Lieden University Press.
- Paunović and Čubrović (2009) Ta(l)king English Phonetics Across Frontiers. Cambridge school publishing.
- Powell, J. V. (1985) An occurrence of metathesis in Chimakuan. In V. Acson and R. Leed (eds.), For Gordon H. Fairbanks. Honolulu: University of Hawaii Press.
- Saal (2018) An Insight to Arabic Natural Language Processing (NLP) in AI Today. No publication.
- Soudi A., Farghaly A., Neumann G. and Zbib R. (2012) Challenges for Arabic Machine Translation. John Benjamine Publishing
- Spencer, Andrew. 1996. Phonology. Oxford: Blackwell.
- Syrett K. and Arunachalam S. (2018) Semantics in Language Acquisition. Routledge.
- Trask, R.L. (1995) Language: The Basics. New York :Routledge

- Wajeeh, M. (2010) Metathesis in Arabic. University of sciences. Egypt Press.
- Wilks Y. (2008) Machine Translation. Springer.