

The efforts of the Algerian linguist Abdel Rahman Al-Haj Saleh in the computational processing of the Arabic language. The Linguistic Corpus Project as a Model.

Dr. Boughar Mohamed¹

¹: University Center Morsli Abdellah, Tipaza, (Algeria), boughar.mohamed@cu-tipaza.dz

Received: 11/10/2024, Published: 13/11/2024

ABSTRACT:

The call for computational processing of the Arabic language in today's world, especially with the technological boom and the rise of studies that digitize languages, is of great significance. This call plays a crucial role in preserving the Arabic language and facilitating easier access to information for those seeking it. It also represents a vital necessity for Arabic speakers to keep up with their time and maintain their linguistic security, which is inherently linked to intellectual security, as both are inseparable. One of the most prominent Arab scholars who contributed to this field was the Algerian linguist Abdel Rahman Al-Haj Saleh (may he rest in peace). Through his project, the Linguistic Corpus, he sought to computationally process the Arabic language in an intelligent way, enabling easier access to information and its meanings, especially for Arabic language learners. Based on this introduction, this research paper is titled: "The Efforts of Abdel Rahman Al-Haj Saleh in the Computational Processing of the Arabic Language: The Linguistic Corpus Project as a Model." The main questions it addresses are: What were the contributions of Abdel Rahman Al-Haj Saleh to the computational processing of the Arabic language? And what were his key efforts in developing computational terminology for this process?

Keywords: Abdel Rahman Al-Haj Saleh, Arabic language, The Linguistic Corpus, computational processing.

Introduction:

The call for a scientific approach to the computational processing of the linguistic heritage produced by the Arabic language, as well as by Arab scholars and writers in both spoken and written forms, is of utmost importance in this era. It is fair to say that this call represents a vital existential necessity for the survival of the Arabic language and the legacy of its scholars and writers. Abdel Rahman Al-Haj Saleh, a prominent Algerian scholar (may Allah have mercy on him), recognized the need for a scientific and technological program that could enable the computational processing of Arabic in a structured, technological format. This program would help pave the way for a modern Arab renaissance, particularly in linguistic and linguistic research.

The project in question is the Thesaurus of the Arabic Language, proposed by Abdel Rahman Al-Haj Saleh to the Arab Organization for Education, Culture, and Science in 1987. The primary aim of this ambitious project was to computationally process the entire linguistic heritage of the Arabic language using modern information and communication technologies and engineering. Haj Saleh designed a unique scientific approach to achieve this. The ultimate goal of the project was to create an automated database, allowing researchers from any part of the world to search for any term in the Arabic language and retrieve all relevant information in a short amount of time. There is no doubt that if this process were realized for the Arabic language, it would bring about a significant scientific leap forward, as Haj Saleh firmly believed that language is the bridge and soul of all sciences.

Based on the above, this article aims to explore the computational processing of the Arabic language through the lens of Abdel Rahman Al-Haj Saleh's Thesaurus project and investigate the innovative scientific aspects of this project. The key questions we seek to answer are: What were the efforts of Abdel Rahman Al-Haj Saleh in the computational processing of the Arabic language? And what were his most significant contributions to establishing computational terminology for Arabic language processing? To address these questions, we have divided the study into two sections. The first examines the Thesaurus project, while the

second discusses the key scientific and technological terms coined by Haj Saleh in the field of computational intelligence. We conclude the paper with a summary of the main findings and recommendations.

The Linguistic Corpus: Thesaurus of the Arabic Language

The Linguistic Corpus Project, also referred to as the Arabic Linguistic Internet, is a project proposed by the scholar Abdel Rahman Al-Haj Saleh to the Arab Organization for Education, Culture, and Science. Its general goal is to computationally process the Arabic language through the use of modern technologies. Regarding this, Al-Haj Saleh stated: "This led us to propose a major and highly significant project, the Arabic Linguistic Corpus or the Arabic Linguistic Internet. We had proposed it to the Arab Organization for Education, Culture, and Science in December 1987. The executive council adopted it at the time and requested that the costs for overseeing it be determined, but this has yet to be achieved" (Saleh, 2012, pp. Vol. 2, p. 112)

Al-Haj Saleh further defines the project, stating: "The Linguistic Corpus is a project linked to the automated processing of Arabic texts. Its aim is to establish a database of linguistic information by gathering as many texts as possible, including what Arab thought has produced in literature, science, technology, and other fields. This database will be automated, allowing it to be accessed via terminals in any Arab country" (Saleh, 2012, pp. Vol. 1, p. 95).

He elaborates on the functionality of the project, stating: "This project arose from the idea of utilizing computers and exploiting their tremendous speed in data processing and their remarkable ability to store billions of pieces of information in memory, to create an automated database that contains:

- The most significant works written in Arabic, whether of scientific, literary, or historical value, and more.
- The most valuable contributions of Arab thought, both ancient and modern, and what will be produced in the future.
- The primary goal of this project, or this automated database, is to be accessible to any researcher anywhere in the world, allowing them to query the computer for any information at the speed of light.
- This database will include ancient and modern Arabic texts from the pre-Islamic era to the present day. The most notable feature of this database, according to Al-Haj Saleh, is the ease with which researchers can quickly obtain the desired information, alongside the comprehensiveness of the available data.
- One of the key advantages of this automated database is its inclusion of data that enables researchers to understand the authentic usage of the Arabic language across different eras and Arab countries"* (Saleh, 2012, pp. Vol. 2, p. 112).

The Importance and Benefits of the Linguistic Corpus Project:

Al-Haj Saleh believes that the Linguistic Corpus Project has several benefits, including:

- Relying on this automated database to create terminology in various linguistic research fields and to search for all linguistic data within a specific area based on the usage of the Arabic language, both in the past and present. This is done by compiling the corpus of terms and words from which the researcher can derive or use the term they want within seconds. This task would be difficult, if not impossible, through traditional manual work, as the corpus provides all the words used throughout the history of Arabic, allowing the researcher to base their study or derive the terms they need.
- Relying on the automated database to select common words with the intended meaning. The corpus allows the researcher to retrieve all the common words they wish to express or write about in seconds, with precision in identifying the meaning they want and in the contexts they are searching for. In this way, the corpus enables the researcher to choose the most accurate word, with the precise meaning, in the appropriate

context.

- Relying on the automated database to track the semantic evolution of the Arabic language. This allows researchers to create a historical dictionary of the Arabic language that enables Arabic language students and others to monitor the evolution of any given word's meaning.

- The automated database allows researchers to create an index of all Arabic texts, both those published and those that will be published, at the level of the Arab world. This includes words specific to Arab culture and civilization, scientific terms, figures, tribes, and more.

- One of the significant benefits of this automated database is that it enables scholars of Arabic linguistics to compile a comprehensive dictionary of Arabic words, allowing researchers to conduct precise linguistic studies of any word or dictionary entry.

The Function of the Corpus According to Abdel Rahman Al-Haj Saleh (Saleh, 2012, pp. Vol. 1, p. 402):

Al-Haj Saleh explained that the corpus serves several important functions and tasks for the Arab researcher and others, including:

1. Obtaining information about Arabic words, whether ordinary or technical terms:

The corpus provides the ability to know if a word has been used in the Arabic language, whether in written or spoken texts, or both. It also specifies the exact context in which it was used and its meaning in each context. Some questions a researcher might ask include:

- Does the word (X) appear in written or spoken usage, or both? Where did it appear, and what meaning did it hold in each of its sources? What contexts did it appear in?

- What conceptual field does (X) belong to? Does it have synonyms, and what are they?

2. Obtaining information about the roots and forms of Arabic words:

The corpus allows the researcher to find out whether a certain root or form has been used in Arabic, both in ancient and modern times. Some questions a researcher might ask include:

- Were the original materials A..B..C.. used by a writer or speaker? What words were derived from them, and how did the author use them?

- List all the words derived from the forms A, B, C, indicating the meaning of each word.

3. Obtaining information about the types of Arabic words:

The database allows the identification of proper nouns, verb sources, and other word forms in Arabic. The corpus also enables researchers to find the specific instances where a proper noun, verb source, or adjective was mentioned, along with the period in which it was used. Some questions a researcher might ask include:

- What are the proper nouns, verb sources, trilateral or quadrilateral verbs (both simple and augmented), etc.?

- What is the frequency of each in a single text or multiple texts?

4. Obtaining information about particles and meanings, borrowed words, sentence structures, and poetic meters:

The corpus provides information about particles, foreign words used in Arabic, sentence structures (both active and frozen), poetic meters, elisions, rhymes, and other related aspects.

The Most Important Terms Coined by Al-Haj Saleh in the Field of Computer Science for the Computational Processing of Arabic Linguistic Heritage:

Before delving into the important terms that the scholar Abdel Rahman Al-Haj Saleh (may he rest in peace) coined in the field of computer science and artificial intelligence, it is necessary to highlight the

significance of terminology. The importance of a term stems from the essential role scientific knowledge plays in human existence in general. A term is the key to sciences and knowledge, serving as the true intention of anyone seeking expertise in any field. Science itself is nothing more than a system of specific concepts represented by specific terms. When we master a term and its concept, we master the very essence of that science. A term in any discipline represents the boundaries of that science. It is imperative for a student of a particular field to understand the terms and concepts used by experts in order to grasp the system of knowledge that defines the boundaries and deeper meanings of that discipline.

More precisely, a term reflects a particular mental structure and a specific mode of thought within a given culture. In other words, terminology represents a unique worldview and understanding of existence, which differs between peoples. Every group perceives existence through its own unique system of terms. Thus, we can say that terminology is the hidden aspect of a nation's consciousness and is closely tied to its progress, prosperity, decline, or extinction. When we speak of a term, we are not merely referring to a word; rather, we are talking about a broader and deeper concept—the thought structure of a civilization and culture with all its components.

Among the most important terms coined by Abdel Rahman Al-Haj Saleh in the field of computational processing are:

Al-Rattab (Computer):

Al-Haj Saleh uses the term Al-Rattabas a synonym for "computer." He mentions terms such as "Computational Linguistics" (Saleh, 2012, pp. Vol. 1, p. 86), "Computational Engineers" (Saleh, 2012, pp. Vol. 1, p. 87), and "Computations" (Saleh, 2012, pp. Vol. 1, p. 85). Al-Haj Saleh defines the Al-Rattab or "computer" as "the most appropriate machine for studying these matters (referring to forms and structures and processing them in an informational way), as it performs computational operations of the type known as transformational operations: a known arrangement that leads to an unknown result. The computer (Al-Rattab) receives a set of data and a set of rules to perform the task, and the researcher instructs it to perform a certain task to reach a specific result by giving it a series of instructions in a language it understands and in the required order. This set of instructions is called a program. The computer (Al-Rattab) can, using a special program, automatically extract all the words in a given text and, at the same time, the original root forms from which these words are derived" (Saleh, 2012, pp. Vol. 1, p. 279).

The linguistic root of the term "Al-Rattab" is derived from the verb "rattaba" (to arrange) in the form of an instrumental noun (fa'aal). According to Lisan al-Arab, the verb means: "rattaba something, yartubu, rattaban, and tartaba, meaning it stabilized and did not move; it is said, 'it stood firm like a heel'; and 'rattaba it' means to stabilize or fix something. Rattaba also implies maintaining a steady and fixed state" (Manzūr, 2003, pp. Vol. 1, p. 477-479).

Al-Firuzabadi also defines it in Al-Qamus as: "Rattab, rattaban: to stabilize without moving, similar to tartaba; I stabilized it, and tartab: something stable and fixed, and Rattabwith movement means firmness and intensity" (Al-Firuzabadi, 2003, p. 666).

From these definitions of the verb "rattaba," we understand notions of stability, consistency, and firmness. These three elements are integral to the function of Al-Rattab(the computer). Due to the vast amount of data and forms it processes, the computer requires constant stability and consistency, which can only be achieved through the necessary engineering strength and intensity to maintain this stability at every moment a user requests information from it. This is likely what Al-Haj Saleh had in mind when he coined the term "Al-Rattab" using an intensive form to reflect the large volume of computational operations and the rapid and precise responses the computer provides when handling requests.

Al-Ala and Al-Ayyala (Robot):

Abdel Rahman Al-Haj Saleh uses the term "Al-Ayyala" as a synonym for "Al-Ala" (machine) when discussing artificial speech synthesis. He states: "Scientists have achieved remarkable results in artificial synthesis, where it has become possible for both machines and robots (Al-Ala and Al-Ayyala) to speak" (Saleh, 2012, pp. Vol. 1, p. 92-93).

By "Al-Ala" or "Al-Ayyala," Al-Haj Saleh refers to artificial devices that rely on artificial dictionaries to recognize words and express the intended meaning.

In *Lisan al-Arab*, the term "Al-Ayyala" is derived from: "Al (verb): to manage or guide; it also refers to curdled milk, and 'Al-Mal' means to manage property or wealth well" (Manzūr, 2003, pp. Vol. 6, p. 41). It also means: "To return or to become; and 'Al-Rajul' refers to the man's family, followers, or supporters" (Omar, 2008, pp. Vol. 1, p. 141). Al-Jurjani defines it as: "It is the intermediary between the agent and the affected in conveying its effect, such as the saw for the carpenter" (Al-Jurjani, 2004, p. 31).

The terms "Al-Ala" and "Al-Ayyala" here are used to describe the function of transformation and return. This is because the machine (robot) transforms speech from one form to another or from one usage to another. Originally, this function is a human task, and the idea of "return" in this context refers to the machine's return to imitating human speech. It also returns to performing the function of voicing and pronouncing in the way the human programs it to do.

Al-Mammawaj (Kymograph):

Al-Haj Saleh uses the term "Al-Mammawaj" when discussing the occurrence of artificial speech in its acoustic aspect. He states: "It is well known that sound is a vibrational disturbance of the air, or as our ancient scholars said, it is the undulation of air or the way air is affected by undulation. These undulations or vibrations can be recorded in the sequence of speech by various devices, one of which is the previously mentioned Mammawaj (kymograph), which consists of a rotating cylinder and several sensors that capture sound and transform it into brushes that draw those vibrations onto the cylinder" (Saleh, 2012, pp. Vol. 1, p. 270). He also says regarding the Mammawaj: "It is what we call linguistic technology, a field that began weak and very limited in the late 19th century with the use of the Mammawaj (kymograph) to study linguistic sounds in Europe (Saleh, 2012, pp. Vol. 1, pp. 265-266).

The term "wave" appears in Arabic dictionaries, including *Al-Mu'jam Al-Wasit*, where it is defined as: "The sea surged (maja) with waves: its water rose and became turbulent. It is said, 'the people surged' when their affairs became chaotic, and 'people surged in turmoil,' as did the turmoil itself. Al-Mammawaj [Kymograph]: a device for recording movements of internal organs and others" (Al-asatitha, 2004, p. 891).

From this, we understand that the term "Mammawaj" is derived from the motion of sea waves and their turbulence. The term is based on the pattern of *mif'al* (instrument noun), referring to a device used to measure the wave-like movements of sound produced by humans as they spread through the air, causing a disturbance between low and high frequencies due to their interaction with the air.

Al-Mihazaz (Oscillograph):

The "Al-Mihazaz" refers to a machine that enables the analysis of speech into specific sound elements. Al-Haj Saleh explains: "Both the kymograph and the oscillograph allow for the breakdown of speech into its sound components and track their transformations as they occur simultaneously. This means that we can observe the changes in the glottal sound along with the transformations in the nasal sounds, the sounds produced from the mouth, and air pressure, all at the same time. These analyzing machines specialize in displaying tonal changes on the screen during speech, sometimes along with intensity changes" (Saleh, 2012, pp. Vol. 1, p. 271).

Here, it may seem that the functions of the kymograph and the oscillograph are the same or similar, but each has a distinct role. The kymograph measures the

movements of sound waves after they spread through the air, while the oscillograph is used to analyze sounds produced by the mouth. Al-Haj Saleh adds: "They have now replaced the brushes and cylinders with cathode rays in a machine we call the oscillograph, and its benefit is enormous because there is no friction in drawing the vibrations, making the drawings much more precise" (Saleh, 2012, pp. Vol. 1, p. 271).

In the language, "mihazaz" (the shaker) refers to: "Shaking something, as one shakes a spear causing it to vibrate and move. 'He shook it' means to move it vigorously, as in the verse from the Holy Quran: ﴿وَهَزِي﴾ (إليك بجذع النخلة), meaning 'move it.' The Arabs say: 'He shook it' when referring to moving something. In the Hadith of the Prophet (peace be upon him), 'The Throne of the Most Merciful shook at the death of Mu'adh.' In the Hadith of Umar (may Allah be pleased with him), 'We set off hastily, shaking the two sacks,' meaning they hurried their pace with them. 'The earth shook,' meaning it moved and sprouted. 'Al-Hazza' refers to a camel's gait that causes the caravan to sway, and a 'hazzaz camel' refers to one with a strong sound" (Manzūr, 2003, pp. Vol. 5, pp. 491-492).

According to Al-Mu'jam Al-Wasit, "Hazz" refers to a man's enthusiasm or joy, the bubbling sound of a pot, or the sound of a meteor as it descends with a roar. It also describes the movement of wind shaking trees or thunder resounding in the sky. "Hazz" the tree, or something similar, means to shake it intensely, and oscillation refers to the state of a body moving in a vibratory motion" (Al-asatitha, 2004, p. 974).

From this, we understand that the term "Al-Mihazaz" is derived from the concept of sound vibrations and the movement of objects, particularly in the air. "Al-Mihazaz" is an instrumental noun based on the mif'al form, and its function is to monitor the occurrence of sounds, their vibrations, and to analyze these sound waves.

Al-Mutayaf or Al-Mushbah (Sonograph):

Abdel Rahman Al-Haj Saleh uses the term "Al-Mushbah" or "Al-Mutayaf" as a synonym for "sonograph," referring to the device responsible for analyzing speech. He explains: "There is another analytical device invented by the engineers of the American Bell Company during World War II, called the sonograph or mushbah. It is one of the most remarkable inventions in this field, greatly advancing acoustic sciences, to the point that it enabled the creation of artificial speech. The sonograph consists of a microphone that captures sound and records it on a magnetic cylinder. The recorded sound passes through a filter with a frequency range between 80 and 8000 Hertz, breaking it down into its frequency components. These components are then drawn on a closed tube covered with special paper that rotates, revealing the frequency changes, as well as the intensity changes. Together, they form a sound spectrum of the spoken speech" (Saleh, 2012, pp. Vol. 1, p. 272).

The term mutayaf (derived from the Arabic root "T-W-F") is related to the concept of circling or encompassing. According to Lisan al-Arab: "Tafa bih al-khayal: sleep touched him. Tafa bilqawm (he circled around the people); al-tufan refers to anything that is overwhelming and encompassing, like a flood that covers entire cities, or a devastating death. A person described as 'ta'if' is one who circles frequently" (Manzūr, 2003, pp. Vol. 5, p. 278).

From this, it becomes clear that the term "mutayaf" derives from the idea of circling and encompassing something quickly and lightly. The sonograph, in turn, reflects these characteristics as it rapidly captures and records sounds in a circular motion, highlighting its speed and agility in processing sound waves.

Al-Mushbah and Al-Zumra (Group):

Al-Haj Saleh used the term "Al-Mushbah" as a synonym for "Al-Mutayaf" (sonograph). It is an instrumental noun derived from "al-shabah," a word related to the concept of something elusive or hidden, much like a ghost. He used the term in a metaphorical sense, drawing on the similarity between the

sonagraph's function and the elusive nature of a shadow or phantom.

Al-Zumra (Group):

Al-Haj Saleh also introduced the term "Al-Zumra" (group) when discussing mathematical patterns and computational tools in the context of processing the Arabic language. He states: "We draw attention to the importance of the concept of 'group' (groupe) as a mathematical concept that can accommodate complex linguistic patterns" (Saleh, 2012, pp. Vol. 1, p. 102).

Modern scholars agree that the concept of the group might be the structure upon which many refined actions are built. It consists of a set of operations arranged in a mathematical pattern leading to a specific and intended result in a particular field. For instance, in education, the teacher aims to develop skills in a specific subject through a set of principles and patterns that include organized and systematic procedures to facilitate learning for students. In the field of automated information processing, these are procedural mathematical operations that aid in building software. This software relies on what is now called a "monoid," which is a group missing the property of returning something to its original state. Since it lacks this, the monoid is incomplete and not fully effective (Saleh, 2012, pp. Vol. 1, p. 79).

Thus, the term "Zumra" is a mathematical term, and it corresponds to the modern concept of the "monoid," which lacks the ability to restore something to its original form. Al-Haj Saleh suggests that the concept of the group is broader and more precise than the monoid. He encouraged analytical linguists to recognize that the software developed since the inception of computational linguistics does not fully meet the needs of modern analytical theory. Therefore, he advised modern scientists to rely on the group in procedural operations rather than being content with the monoid (Saleh, 2012, p. Vol. 1).

Al-Haj Saleh's use of the term "Zumra" was not arbitrary, but rather rooted in scientific tradition, specifically referencing the work of the classical Arab scholar Al-Khalil ibn Ahmad Al-Farahidi. Al-Haj Saleh noted: "Al-Khalil's analysis is renowned for his invention of prosody, but many may not realize that he also introduced other significant scientific ideas, such as the mathematical concepts known today as the factorial and the cyclic group. These two ideas underpinned the creation of the first Arabic dictionary, 'Kitab Al-Ain,' which scholars agree was based on the mathematical principles laid out by Al-Khalil" (Saleh, 2012, p. Vol. 1).

Al-Muthul (Generator Pattern or SchémeGénérateur):

•Al-Haj Saleh introduced the concept of Al-Muthul in the context of Arabic linguistic analysis. He explained: "As for Arabic analysis, it operates both horizontally and vertically, functioning on two axes simultaneously. It considers verbal clues that reveal the unit and all the units that can occupy its place. This leads to a very useful concept, which is the idea of 'the pattern' or 'Al-Hadd' (the boundary). It is a structure or form in which several units are incorporated. Interestingly, this concept does not exist in Western linguistics. This formal procedural limit that defines the operations forming the units, and therefore determining them from a syntactical perspective, produces a recursive branching pattern that starts from an origin and extends to infinite branches. Our early grammarians referred to this as 'Al-Mithal' (the pattern), and we have named it the 'generator pattern' or schémegénérateur in English and French so that it can have its place in general linguistics" (Saleh, 2012, pp. Vol. 1, p. 240).

•He also stated: "We conclude that the 'Mithal' according to Arab grammarians is a set of conceptual positions arranged in a particular order. Some of these positions may lack the original elements, while others may contain additional ones. The Mithal is not limited to the level of words (i.e., morphological patterns), but exists at all linguistic levels, including syntactic structures and beyond" (Saleh, 2012, pp. Vol. 1, p. 240).

•In this explanation, Al-Haj Saleh highlights how the Arab grammarians' notion of "Mithal" extends

beyond mere word patterns (such as morphological forms) and applies to various levels of language structure, including sentence formation and even higher levels of linguistic organization.

- **Morphological Level (Al-Mustawa Al-Sarfi):** This is the lowest level of meaning, referring to the word itself. There are many patterns (muthul) at this level, but they are limited in number. In Arabic, the pattern of the word is referred to as its wazn (weight) or bina'(structure).

- **Verbal Level (Al-Mustawa Al-Lafthi):** This level pertains to the utterance. It has its own patterns, one for nouns and three for verbs. It represents the level of independent units.

- **Syntactic Level (Al-Mustawa Al-Tarkibi):** This level is not a result of word combinations but can involve other elements. The syntactic units at this level include the amil (governing element), the first ma'mul (governed element), the second ma'mul, and the mukhasis (specifier).

This clearly demonstrates, without a doubt, that Arab scholars were ahead of many modern linguistic theories in describing the linguistic system by analyzing speech into its original material, along with all its branches and forms. Al-Haj Saleh (may Allah have mercy on him) drew the attention of contemporary linguistics scholars to the possibility of leveraging this Arab precedence in the field of computational linguistics and processing the Arabic language corpus using automated computational methods. He believed that such analyses could assist linguistic engineers in developing software for computer-based word processing by compiling all forms with their meanings from actual language use and classifying speech into organized semantic networks. This compilation would later facilitate the creation of scientific and technical terminology.

Al-Haj Saleh also emphasized the importance of the concept of al-mithal (the pattern) in the computational processing of the Arabic language. He said: "Since the machine will search within this database for the necessary tools to identify units and their meanings, it is essential to store all individual, verbal, and syntactic patterns, including their origins and variations, such as morphological patterns, to make it easier for the machine to find what it needs" (Saleh, 2012, pp. Vol. 1, p. 93-94).

Conclusion:

In conclusion, we can summarize some key findings and recommendations as follows:

- The need for the computational processing of the Arabic linguistic heritage is, today, an existential necessity for the preservation of the Arabic language and the intellectual preservation of its recipients and learners, especially in light of the advancements in language technology.

- The Linguistic Corpus Project, proposed by Abdel Rahman Al-Haj Saleh, represents the true and scientifically effective approach for integrating artificial intelligence and computational science into the study and access of the Arabic language.

- In our view, the Linguistic Corpus Project or the "Arabic Internet" is the most deserving linguistic project in the field of Arabic linguistic studies. It should capture the attention of Arab scholars in linguistics to ensure its implementation and reap its benefits. This project is a scientific renaissance for Arabs and the Arabic language, on par with the Arab Renaissance of the past. Through this project, Arabs can achieve significant scientific results in various fields, facilitating a knowledge, methodological, and historical breakthrough in both the humanities and the exact sciences.

Recommendations:

- All relevant institutions and Arabic linguistics scholars must rally around the Linguistic Corpus Project

proposed by Abdel Rahman Al-Haj Saleh. It has been unequivocally proven that this project is the true and effective pathway to a modern Arab renaissance, aligned with the sciences of the current era. This project enables Arabs to achieve scientific breakthroughs in various fields, comparable to the achievements of the West at the dawn of the 19th century. We emphasize this because we firmly believe that language is the bridge of sciences, its vital spirit, and the law by which both language and neighboring sciences evolve.

References:

1. Al-asatitha, A. G. (2004). *Al-Mu'jam Al-Wasīṭ*. Egypt: Al-Shorouk International Bookstore.
2. Al-Firuzabadi. (2003). *Al-Qāmūs Al-Muḥīṭ*. Cairo: Dar Al-Hadith.
3. Al-Jurjani, A.-S. (2004). *Al-Taarifāt*. Cairo: Dar Al-Fadilah.
4. Manzūr, I. (2003). *Lisān al-‘Arab* (Vol. 1). Beirut: Dar Al-Kutub Al-Ilmiyah.
5. Omar, A. M. (2008). *Dictionary of the Arabic Language*. Cairo: Alam Al-Kutub.
6. Saleh, A. A.-H. (2012). *Research and Studies in Arabic Linguistics*. Alger: MOFEM.