Received: 22 August 2025 | Revised: 20 April 2026 | Accepted: 12 October 2026 | Published online: 5 November 2026

RESEARCH ARTICLE

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–14
DOI: 10.47852/bonviewAIA62027336

BON VIEW PUBLISHING

Designing a Small Language Model Lexicon in Natural Language for Robotic Applications

Ioannis Giachos[1], Nikoletta Samaridi[2], Evangelos Papakitsos[1,*], Michail Papoutsidakis[1] and Nikolaos Laskaris[1]

[1] *Department of Industrial Design & Production Engineering, University of West Attica, Greece*

[2] *Department of Informatics and Computer Engineering, University of West Attica, Greece*

Abstract: In this paper, the aim is to present a project report of a method, based on which a small natural language lexicon is designed, or otherwise a small language model, for robotic applications. In this direction, a reference is made to Controlled Natural Languages, and their characteristics are determined for application in operational robotic vehicles. In addition, criteria are established for the selection of the nationality of the language, which constitutes the basis of the lexicon. A small introspection is made in the natural language unit of a human–robot interface, and a method is presented that examines the adequacy of the selected language in relation to the purpose of use of the system and the control criteria. The use of the lexicon's method is presented, and an example of the analysis and inclusion of selected Greek language words is given. The reproduction of words from the lexicon's stored data is also explained, as the reverse process that functions as a confirmation of the method.

Keywords: small language model, lexicon, intelligent human–robot interface, NLP

1. Introduction

For many years, up to recently, communication through written or spoken language between humans and robotic systems has been considered an interesting possibility that tends to become a necessary condition for their existence [1, 2]. Nowadays, large language models (LLMs) that can respond perfectly to a high-level dialogue are a fact, while at the same time they are constantly being improved. However, small language models (SLMs) are needed for small-sized systems, which cannot have the necessary power that the corresponding LLMs require. In case that is taken for granted that an LLM can simply exist in cyberspace and be used via a network connection, let's think of cases where this is not possible at all. Such cases may be some deep underground points, electromagnetically isolated areas, war zones, moving systems with quite limited energy resources, etc. In any case, the existence, use, and development of SLMs is also a fact. At least 15 SLMs are currently evaluated as the top of their kind [3]. The authors of the current work, during the process of developing an intelligent human–robot interface (iHRI) based on speech for the creation of a robotic system that could be used in several applications, were faced with a series of questions on the subject of the language model. In an earlier work, the artificial language SostiMatico [4] was used. This artificial language is based on Greek word roots and responded extremely well. In combination with the Organizational Method for Analyzing Systems (OMAS)-III computational model [4], the iHRI acquired exceptional capabilities in perception, understanding, and learning. A challenge arose when the artificial language had to be replaced by the natural language, which in this case is Greek.

All natural languages possess unique grammatical and syntactical idiosyncrasies. Specific rules apply to different word groups, while the meaning of many words is highly context-dependent within a sentence. This complexity is a decisive factor that precludes the establishment of universal rules for dictionary construction, both within a single language and across different ones. Consequently, the primary step in this process is determining the specific linguistic orientation of the dictionary. This selection is governed by predefined criteria related to the system's requirements, the target users, and the robot's operational environment.

In this paper, an attempt is made to identify the correct answers to all the questions raised for the choice of the language model and the final creation of the appropriate SLM lexicon so that the developing iHRI can maintain all the capabilities that were provided by the artificial language.

The paper is structured in six sections. Section 1 is the present introduction. Section 2 refers to Controlled Natural Languages (CNLs). Section 3 refers to the reasons that lead to the choice of the nationality of the language and the criteria for this choice. In Section 4, the architectural baseline of the natural language processing (NLP) unit within human–robot interaction is detailed, followed by the formal presentation of the two proposed methodology algorithms. Section 5 presents the operational implementation, application, and reconstruction results derived from executing both methods. Section 6 provides the discussion of the findings and, finally, concludes the paper.

***Corresponding author:** Evangelos Papakitsos, Department of Industrial Design & Production Engineering, University of West Attica, Greece. Email: papakitsev@uniwa.gr

© The Author(s) 2026. Published by BON VIEW PUBLISHING PTE. LTD. This is an open access article under the CC BY License (https://creativecommons.org/licenses/by/4.0/).

2. Controlled Natural Language (CNL)

We could say that a CNL nowadays is almost exclusively the language of robotic systems for operational use or small everyday applications. That is, the language for applications that are discussed in this paper. It was not always like this. In the first attempts to develop systems that could understand speech or systems that could translate, and for many years later, CNLs were the only tool that existed. However, as mentioned in the introduction, there are computational language models that are similar to natural language in full development. Consequently, many modern systems have now transitioned to using LLMs. Remaining on the issue of SLMs, CNL emerged from this idea (of limiting natural language) in order to make it more accurate and machine-readable, especially for tasks such as machine translation and knowledge representation. Early examples such as Ogden's Basic English [5] in the 1930s paved the way, followed by various industrial and research efforts [6] to create CNLs tailored to specific needs, such as technical documentation or formal reasoning. But what is a CNL? It is a subset of natural language, resulting from restricting grammar and vocabulary [7, 8]. The aim is to eliminate, or at least limit as much as possible, both the complexity of the language itself and the ambiguity that may arise when it is used by non-human brains, such as computers, robotic machines, etc. Within the framework of CNLs, a series of technical languages, such as Simplified Technical English (ASD) [9], Caterpillar Technical English [10], or IBM's Easy English [11], were developed and used to increase the quality of technical documentation and simplify the semi-automatic translation of documentation. These languages impose strong constraints on writers, with rules that require short sentences, strict vocabulary from an approved lexicon, only active voice, etc. In another version, such as the one we read in the paper by Caruso et al. [12], the authors focus mainly on a class of CNLs, called machine-oriented CNLs, which have a well-defined semantics that can be translated deterministically into formal languages for logical reasoning. This describes another type of CNLs that have been used in many application domains for solving problems and answering questions. These languages can be characterized as knowledge representation CNLs.

2.1. CNL characteristics for operational robotic vehicles

As follows, some basic characteristics are examined that a CNL should have. To make this analysis a little more targeted and within the framework of developing a lexicon for a specific use by the authors, it is noted that we are talking about operational robotic vehicles—that is, vehicles that participate in research missions, rescue, military, industrial, or space missions, etc., where it is necessary to ensure the accuracy, reliability, and functionality of communication between humans and robots. By examining indicative works [13–20], the following list of characteristics is identified and presented:

Clarity: In no case is a command or sentence allowed that introduces ambiguity. This means that the lexicon should include terms with unambiguous meaning. Thus, when formulating commands, ambiguities, ambiguous concepts, and anything else that may cause confusion in the robot's interpretation system will be limited or eliminated.

Limitation of grammatical complexity: In natural languages, there can be many syntactic structures. In the case of CNLs, a predefined one must be strictly adhered to. Usually, the subject–verb–object structure is chosen. Based on this predefined structure, the robot's analysis and interpretation system will be able to understand the command. Passive voice, indirect questions, syntactic embedding, etc., are also avoided.

Focus on the command (action-oriented vocabulary): The language lexicon should include verbs and nouns that accurately describe actions, locations, and objects, for example, "approach," "scan," "pick," "drop," "obstacle," "target," "left," "right," etc. That is, the combination of words should be explicit instructions for grounding the action of the verb.

Correspondence with navigation and execution models: Each word in the lexicon should directly correspond to functions supported by the robotic subsystems (sensors, navigation, arm, etc.). Obviously, the lexicon chosen should be designed for the particular robotic system so that there are no words that do not lead to a function.

Support for extensibility (modularity): The lexicon should be scalable for new missions or environments, without requiring a complete redesign. For example, changing the lexicon could add or remove features. This, of course, requires that the language mechanisms can support these changes. The iHRI that the authors have developed, as described in previous works by Giachos et al. [21], enables the system to enrich its lexicon, according to the needs of the mission.

Resistance to errors (error-tolerance): The words should be chosen in such a way that the system can handle possible voice recognition or typing errors. Such errors can arise from words that are very close in sound, for example, the words "STOP" and "STEP." Due to pronunciation or a noisier environment, the Automatic Speech Recognition (ASR) of the robotic system can confuse such words.

Multilingual and cross-cultural neutrality: If the mission is international or multinational, the lexicon should be based on international technical terms. In this case, we are talking about a language that is spoken internationally at a technical and scientific level. The most popular language is English, which is spoken by everyone in the academic community, in aviation, navigation, aerospace, etc.

Verification and reuse (verifiability and reusability): The words to be used are selected based on compatibility with formal methods, such as Planning Domain Definition Language (PDDL) [22] or Linear-Time Temporal Logic (LTL) [23], or verifiable execution models such as ROSPlan specifications [24].

Compatibility with Human Interface (human–robot interaction): The lexicon should facilitate user interaction with the system, either through speech or through an interface such as a Graphical User Interface (GUI) or a Command Line Interface (CLI). It is one of the very basic features of a CNL nowadays; moreover, this is the goal from the beginning: human-machine interaction through speech. Interaction through GUI or CLI further enhances this feature.

Support for semantic analysis (semantic mappability): Words and phrases can be mapped to conditions and goals within known semantic representations (e.g., ontology-based planning). Also, semantic analysis can be supported by Hole Semantics Theory in a different way, without this type of mapping. It is the method used by the authors in their own iHRI [21].

3. Language's Nationality

The use of a good artificial language, such as SostiMatico, has many advantages over natural language but one very important disadvantage. The advantages are the ease of handling words and the creation of new ones with simple additions. There are the

basic roots, and by putting specific letters before and after them, we have the verbs, nouns, adverbs, prepositions, numbers, and everything else that a complete language includes. The formalism is strict and specific, and thus, a sentence can be very easily deconstructed and analyzed. It should be noted that this finding is made by the authors herein, who studied the artificial language in combination with the OMAS-III computational model in the role of grammatical formalism. On the other hand, the great disadvantage of the artificial language is that it is required of the robot operator and interlocutor to know the artificial language well, which requires special training.

Conversely, while a natural-language-based SLM eliminates the need for specialized operator training, it introduces inherent structural ambiguity and a higher computational overhead for syntax parsing. In safety-critical or real-time robotic applications, this ambiguity presents a critical trade-off: a single misinterpreted natural language command could lead to operational failure, whereas artificial languages offer deterministic execution at the expense of general usability.

The nationality of the language for robotic systems seems to be a very popular issue. Although it sounds like a hypothesis that shows two paths, depending on where it is intended, that is, domestic or international use, it is not so simple. In a series of studies that have been done on speaking robotic applications, it seems that the nationality of the language plays an important role in more complex factors. Thus, even its involuntary choice is indirectly determined by certain criteria. Therefore, either indirectly or directly, the nationality of the language seems to be based on specific criteria related to the needs of the system, the users, and the operating environment of the robot. To systematically identify these criteria, an extensive review of the relevant literature was conducted, highlighting the multifaceted nature of language selection.

First, regarding the user-centric factors, in a paper by Godwin-Jones [25], speaking about virtual and augmented reality in the context of personalized language learning, the author identifies the choice of language as a requirement of the end user. In Reference [26], the author refers to the issue of minimal or no development of ASR systems or other language technologies in most minority languages of the world. The authors, referring to the Ngiemboon language, a Grassfield Bantu language spoken in the West Region of Cameroon (Africa) by about 400,000 people, highlight the need to adapt the systems to the language of the user community. In Reference [27], in the context of the evaluation of the language learning method with the help of social robots, rejection occurs when the language does not match the needs of the user, from the perspective of students, teachers, and second language technology developers. From the above, we can determine the language of these end users as the first selection criterion.

Following the previous findings and more specifically in work by Yemmene and Besacier [26], technical limitations must also be considered, giving a basis for the available ASRs. If there is no suitable ASR for recognizing a language, then an Human–Robot Interface (HRI) cannot be set up in this language. This can also be a problem in the use of artificial languages. This is because artificial languages, as well as less popular languages, require years of targeted study for the development of resources, as has been the case for so many years in other natural languages. In work by Besacier et al. [28], we read that speech processing for languages with insufficient resources has been an active field of research for several years. In work by Toshniwal et al. [29], a single sequence-to-sequence ASR model is presented, trained on nine different Indian languages, which have very little overlap in their alphabets. In the paper by Meelen et al. [30], three experiments were presented to test the most effective and efficient ASR to facilitate the documentation and preservation of endangered languages, which often have extremely low resources. From the above indicative works, it seems that for several years until today, research on ASR has been occupying the research community. And the reason, besides the effort for continuous improvement, is that it is still impossible to cover all cases, especially the cases of languages with insufficient resources. As a criterion for selecting the language of robotic systems, we can add the availability of resources for voice recognition.

Beyond technical resources, the deployment environment plays an equally vital role. In the paper by Mavridis [31], the author lists a series of robots that have been built from time to time to serve specific applications—that is, tour guide robots, surgical robots, industrial robots, robots to support autistic children, etc. Many of them are talking robots. Social robots can be used in education as teachers or as students by peers, as we read in the paper by Belpaeme et al. [32]. In multilingual schools, it is suggested that the robot uses the language of instruction or the predominant language of the children during recess (the one they hear most often). In the paper by Giachos et al. [33], the use of the iHRI they have developed was proposed to help people with brain damage in their daily lives. From the above indicative papers, we can define the next criterion for selecting the nationality of the language as that of the Target Market or Scope of Application of the robotic system. Thus, the selected language will cover both the needs of the application and the users, regardless of the country.

As the authors report in Reference [34], social assistive robots are designed to provide social and cognitive assistance, where they will seek to motivate and engage people in therapeutic activities. The work examines how cultural norms for eye contact and vocal tone affect the user–robot relationship. It is found that users attribute greater credibility and sympathy to robots that "mimic" the cultural expression of their own group. In Reference [35], it is found that users have a more positive view of robots that speak in a voice similar to their own or their cultural environment. Preferences are related to gender, accent, and style. In Reference [36], the applications of social robots for educational purposes were examined from the perspective of cultural awareness and responsiveness. The study concludes that cultural backgrounds can influence learning conditions and the overall climate in robot-based educational environments. Also, in Reference [1], among other things, the authors suggest that for robots to be effective, actions in the linguistic channel must be coordinated with other channels, such as natural gestures and gaze. This includes not only the selection of appropriate word sequences but also the prosody of the expression, etc. However, elements may differ, depending on the culture of the country. According to the above indications, among a multitude of studies over the years, we can identify another criterion, which is the cultural and social characteristics displayed by the robot, and one of them is language.

Furthermore, modern environments often require adaptability to diverse linguistic landscapes. In References [37, 38], the authors study language models for environments with mixed languages, which is a frequent phenomenon in multilingual populations. More specifically, they investigate language modeling for speech recognition through code switching between Mandarin and English. In a recent literature review by Agro et al. [39], the growing and undiminished interest of researchers in recent years in the multilingualism of robotic systems was demonstrated. The above could lead to a selection criterion of multilingualism when the application field requires it.

Finally, internal language variations must be accounted for. In Reference [40], in a review of error detection and correction methods in ASR units, Errattahi et al. emphasize that transcription errors are very important not only for improving the accuracy of speech recognition but also for avoiding the propagation of errors in subsequent language processing units, such as machine translation. Through this logic, several studies are dealing with and investigating methods for the correct recognition of various dialects of a native language. For example, the work by Hirayama et al. [41] proposes a system that statistically trains transformation rules between a common language and dialects and simulates a corpus of dialects for ASR based on a machine translation technique. Also, the work by Li et al. [42] in China, which is a multinational country with a large population and particularly endowed with diverse dialects, triggers significant challenges in speech recognition work. From the above, we identify a final criterion, that of dialect recognition and understanding. Perhaps for the time being, the answer to this criterion is to choose a language with a more "stable" or "neutral" vocabulary.

In conclusion, having answered the initial question: "what are the language selection criteria for a robot," we simultaneously determine the answer to the question: "what are the criteria that should be studied for an SLM lexicon development for a speaking robot application." The criteria are the same as those we identified above and are summarized in the following list:

1) End users' language;
2) Availability of resources for voice recognition (ASR);
3) Target market or scope of application;
4) Cultural and social characteristics;
5) Multilingualism;
6) Recognition and understanding of dialects.

Please note that the above are not prerequisites, but criteria that are adapted to each system. That is, given the specifications of a system that must be given a language, values are entered into them. For example, a system is being developed for language learning for students. In the criterion of "multilingualism," not only the "required" is given but also which languages these are. The criterion of the end users' language simply adds to the set of languages required: the national language of the students or the base language of the school. This will be the return language of the translation. As long as all the criteria receive the required values, a comparison is immediately made with a ready-made SLM at the level of identification with these values. If the language is not considered sufficient, then another one must be created, or an existing one must be modified, so that its lexicon has what is needed.

4. Proposed Methodology

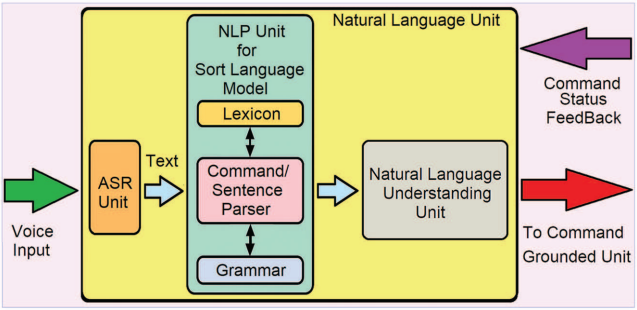
This section outlines the comprehensive methodological framework developed for the proposed SLM tailored for robotic applications. First, the primary theoretical and practical contributions of this research are delineated. Subsequently, the foundational architectural components of the NLP unit and its structural subcomponents are detailed. Finally, the two core mathematical and algorithmic methods—governing language adequacy verification and multi-functional lexicon development—are formally presented to demonstrate how natural language parsing is optimized for resource-constrained human–robot interaction.

The primary contribution of this study is the development and integration of a localized, highly compact SLM architecture specifically designed to support the authors' proprietary iHRI. Rather than relying on computationally heavy LLMs or vast, unstructured dictionary datasets, the proposed framework introduces a specialized methodology tailored for low-resource and operationally isolated environments. By utilizing dynamic word roots and systematic morphosyntactic encoding, this model directly expands the linguistic capabilities of the authors' existing iHRI. Consequently, it minimizes computational overhead while retaining the complex grammatical inflections and semantic adaptability required for precise, real-time human–robot interaction.

4.1. The typical NLP architecture

A human–robot interface necessarily contains a natural language unit. The natural language unit is not only the language processing unit. As shown in Figure 1, the NLP is included in it.

Figure 1
A typical natural language processing unit



Since we are talking about a speaking HRI, the natural language unit includes both the ASR and the natural language understanding (NLU) unit. NLP receives text that was said audibly and forwards the result to the NLU unit. The sound and the directions for the actions to be taken are the input and output of the unit, respectively. As feedback for the representation of the result of the command given, the status of the requested action is returned within the unit. As highlighted in the above sections, ASR and NLU are important, both for a CNL and for the nationality of its lexicon. In this section, we will talk about the NLP unit, since it contains the lexicon and grammar of the language that is the subject of this work. More specifically, in order to synthesize several of the characteristics of a CNL, the design concerns the lexicon, the grammar, and the parser that is the analyzer of the sentence-command. Therefore, when reference is made to an SLM, reference is made to the above three parts of NLP.

To establish the foundational rules of this model, the grammar is the set of rules set in the language of the system. Although we are talking about CNL, because of the term "Natural," it becomes clear that this grammar will include several elements from the corresponding one of the mother tongue. Of course, on the other hand, several rules are omitted in order to maintain some basic characteristics of CNL. The authors, through their own work and with the tools they worked with, which are OMAS-III and Hole Semantic Theory, keep more rules than usual. The sentence may state both the present, the past and the future; the first, the second and the third person can be used; there can be a definite, subjunctive, and imperative declension, in order to declare certainty, desire, and command [33]. There can certainly be prepositions, etc.

In close cooperation with these rules, the lexicon must, in cooperation with the grammar, provide correct inputs to the parser in order to correctly analyze the sentence-command. In the

lexicon compiled by the authors and the method of its compilation presented below, there are no words complete in various grammatical versions. There are word roots that can compose more than one word, while at the same time presenting the interpretation of each one [21]. Thus, a natural language root can, with the corresponding additions, become a verb, adverb, noun, or adjective, in singular or plural, in various tenses and with various prepositions. All of this arises from the semantics, both of the root of a word and of all its additives in a final connection. In short, the lexicon here is not a list of words, nor a complex ontology. It is a very simple database where each word root shows additional important syntactic and semantic information. The parser has access to this information during analysis and collects it both to check the correctness of the sentence that reaches it and to give the final meaning that will go to the NLU. The learning of a new word [21] is interesting, where the system, in order to register a word, requests all its structural elements. That is, its root, the type of verb it can create, the type of ending it can receive, and, of course, its interpretation and connection with a sensor of the system (walking, collecting, etc.), or even some synonym from the existing ones. And all this is done in the form of a dialogue. This method allows the lexicon to grow and the system to learn. The system can be easily adapted to various missions, even to a home assistant.

Finally, the primary operational tool that synthesizes these components is the parser. The parser, with one eye firmly on grammar, checks the words and structure of an incoming sentence one by one. It should be noted here that the parser in a CNL is a tool that is built on the capabilities of grammar and lexicon. That is, there are parsing methods, but these methods must be applied exclusively to the features of the current grammar and lexicon [43–45]. That is, as mentioned in the lexicon section above, the parser created by the authors of this work must collect information for each word from a database. Also, when a sentence has an elliptical structure, this particular parser does not reject it. It tries to collect information from its recent knowledge, from its database, and if it fails, it asks its interlocutor for further information like person, place, unknown words, etc. [4]. Also, this parser receives from the sentence the relevant information of place and time, which can be utilized in "Sequencing" issues, both in missions and in virtual assistants for people with brain damages [33].

4.2. The proposed model's architecture

As the present work is an extension of the intelligent human–robot interface (iHRI) framework that the authors are developing based on the Greek language, the native language of the lexicon is chosen to be Natural Modern Greek. There are two distinct methodologies proposed for the creation of this special-usage lexicon: "Checking the adequacy of the lexicon's selected language" and the "Lexicon development method." These methods are formally analyzed in the relevant subsections that follow.

4.2.1. Language adequacy checking method

The following methodology evaluates whether the language chosen for the development or application of a dictionary is the correct choice. This is achieved with the help of the criteria described in Section 3, more specifically in the summary subsection. All criteria are entered as inputs into an algorithm that yields a tri-state output. Although the term "tri-state" traditionally originates from digital electronic systems referring to "0," "1," and "high impedance (Hi-Z)" as state levels, in this algorithmic context, it maps to "No" (unsuitable), "Yes" (suitable), and "Doubtful/Undetermined," respectively, for additional clarity of the state level.

The schematic workflow of the proposed algorithm is illustrated in Figure 2, while its corresponding pseudocode representation is detailed in Figure 3. Based on this algorithm, a criterion-by-criterion test is performed to determine whether the selected language satisfies the application requirements. The diagram describes the inputs classified into two distinct groups:

1) "The Language Requirements Data of the Application," which defines the strict criterion according to the target application.
2) "Selected Natural Language's Data," which is the representative linguistic features and profile of the chosen language with respect to each criterion.

Thus, for each criterion, the system maps the requested data against the offered data. If these two profiles match, then the criterion is accepted; otherwise, it is rejected. All intermediate evaluation metrics reach a decision unit, where the ultimate verdict on the adequacy of the language is determined, based on the cumulative configuration of acceptances and rejections. According to the algorithm, absolute adequacy requires all criteria to be strictly met, whereas absolute insufficiency is triggered when all criteria are rejected. In the intermediate stages, the adequacy of the selection is flagged as undetermined. Indeterminacy is crucial because, depending on the application domain, certain criteria may be secondary, while others (e.g., ASR) are mission-critical. Thus, it is suggested that special attention be paid to the formulation of the algorithm's requirements (The Language Requirements Data of the Application) so that the proposed data (Selected Natural Language's Data) always seek the satisfaction of all criteria. Also, undetermined states can be used as an element of re-checking the requirements, in order to remove any excesses that existed during their initial formulation. For instance, if an application strictly demands a multilingual framework, then the selected natural language configuration may fail to satisfy this specific criterion. However, a granular requirements analysis often reveals that absolute multilingualism constitutes an over-specification for the target operational context. By recalibrating this system's parameter from multilingual to monolingual, the operational constraints are optimized, thereby enabling the formal acceptance of the criterion within the evaluation algorithm.

4.2.2. Lexicon development method

As previously noted, the language model developed by the authors allows for additional grammatical vectors, such as tenses, persons, and inflections. Consequently, the lexicon encodes additional morphosyntactic features that will be utilized by the parser to derive critical semantic information. This information structure is highly compact compared to LLMs but is mathematically vital for SLMs and CNLs. Building upon previous work by Giachos et al. [21], where an algorithm for learning a new word was established, the lexicon is developed from scratch. Structurally, the lexicon is implemented as a relational database comprising interconnected relational tables that the parser can access sequentially by searching for the semantic set of each word, ultimately resolving the semantics of the entire incoming sentence, including the temporal and spatial parameters. Figure 4 shows the structural architectural diagram of the method that structures and places the parts of a word in the tables of the lexicon database.

To motivate this architectural choice, it must be highlighted that in the Greek language, the verb's inflection (ending)

Figure 2
Algorithm for checking the adequacy of a selected natural language in terms of nationality

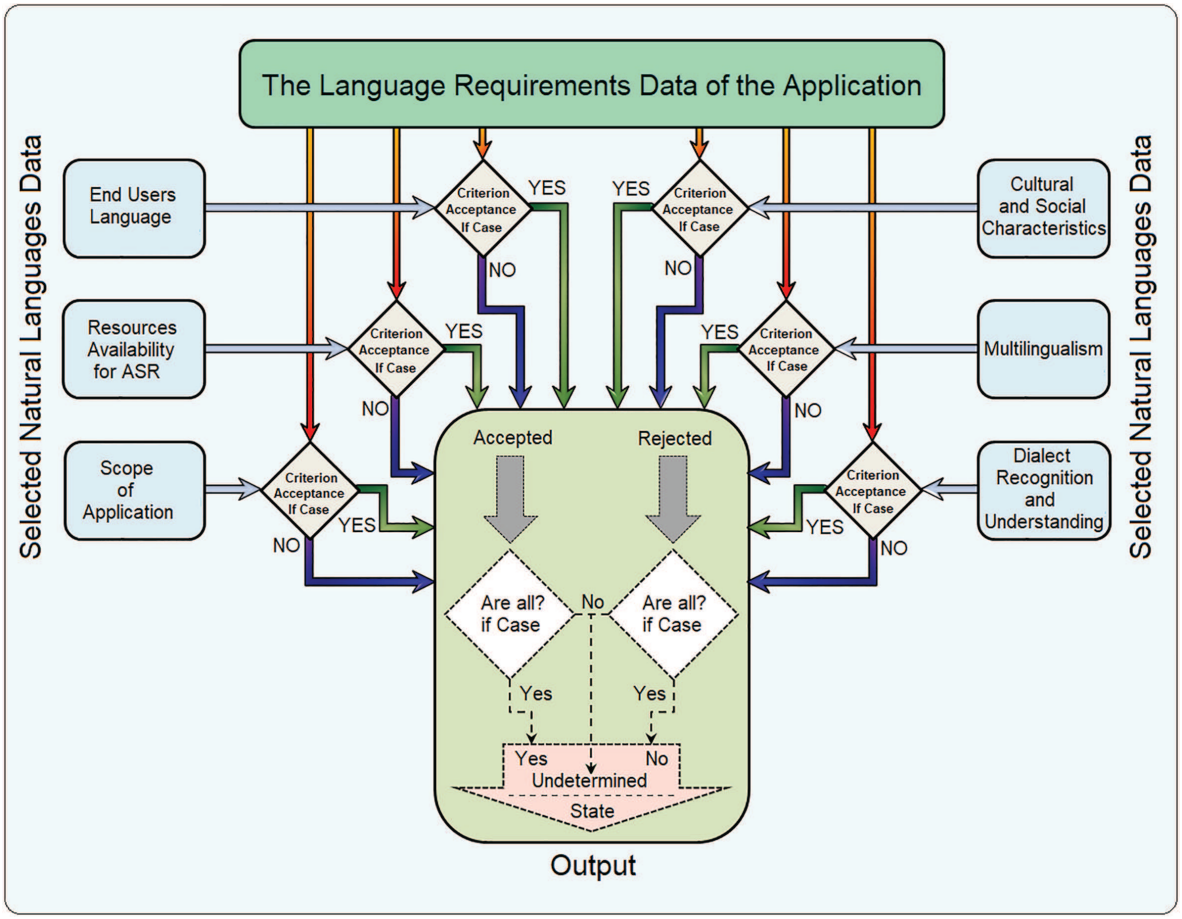


Figure 3
Pseudocode representation of the algorithm for checking the adequacy of a selected natural language

```
ALGORITHM LanguageRequirementsValidator

// Inputs: The language requirements data of the application
           for the selected natural language
VARIABLES:
    endUsersLangMatch, asrResourcesMatch, appScopeMatch,
    culturalCharMatch, multiLingMatch, dialectRecognMatch: BOOLEAN
    acceptedFlags, rejectedFlags: ARRAY OF BOOLEAN
    allAccepted, allRejected: BOOLEAN
    finalOutput: STRING

BEGIN
    // Initialize empty arrays for decision flags
    acceptedFlags ← []
    rejectedFlags ← []

    // 1. Evaluate Selected Natural Languages Data (Left side of the flowchart)
    // Criterion: End Users Language
    IF checkCriteria(EndUsersLanguage) == "YES" THEN
        APPEND TRUE TO acceptedFlags
    ELSE
        APPEND FALSE TO rejectedFlags
    ENDIF

    // Criterion: Resources Availability for ASR
    IF checkCriteria(ResourcesAvailabilityForASR) == "YES" THEN
        APPEND TRUE TO acceptedFlags
    ELSE
        APPEND FALSE TO rejectedFlags
    ENDIF

    // Criterion: Scope of Application
    IF checkCriteria(ScopeOfApplication) == "YES" THEN
        APPEND TRUE TO acceptedFlags
    ELSE
        APPEND FALSE TO rejectedFlags
    ENDIF

    // 2. Evaluate Selected Natural Languages Data (Right side of the flowchart)
    // Criterion: Cultural and Social Characteristics
    IF checkCriteria(CulturalAndSocialCharacteristics) == "YES" THEN
        APPEND TRUE TO acceptedFlags
    ELSE
        APPEND FALSE TO rejectedFlags
    ENDIF

    // Criterion: Multilingualism
    IF checkCriteria(Multilingualism) == "YES" THEN
        APPEND TRUE TO acceptedFlags
    ELSE
        APPEND FALSE TO rejectedFlags
    ENDIF

    // Criterion: Dialect Recognition and Understanding
    IF checkCriteria(DialectRecognitionAndUnderstanding) == "YES" THEN
        APPEND TRUE TO acceptedFlags
    ELSE
        APPEND FALSE TO rejectedFlags
    ENDIF

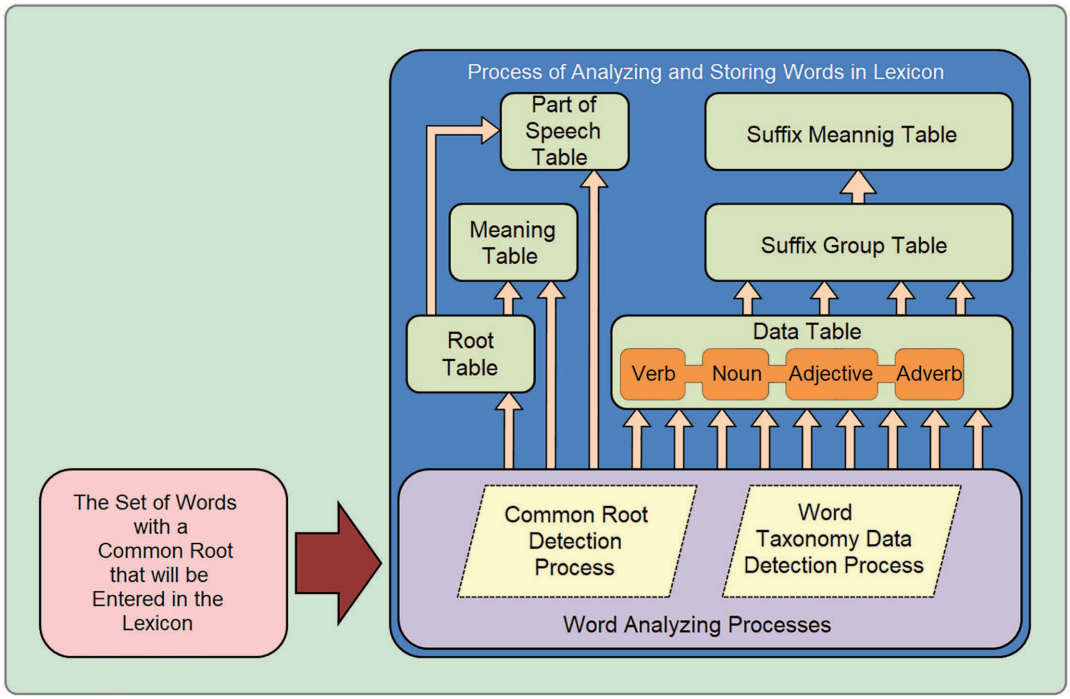
    // 3. Final Evaluation and Output Generation (Central Output Block)
    allAccepted ← checkAllTrue(acceptedFlags)
    allRejected ← checkAllTrue(rejectedFlags)

    IF allAccepted == TRUE THEN
        finalOutput ← "YES (Accepted)"
    ELSE IF allRejected == TRUE THEN
        finalOutput ← "NO (Rejected)"
    ELSE
        finalOutput ← "Undetermined State"
    ENDIF

    OUTPUT finalOutput

END
```

Figure 4
Method diagram of analyzing & storing words in lexicon



explicitly denotes the person, just as the adjective's inflection denotes gender. This introduces a higher degree of morphological complexity and rule-dependency compared to English, where verbs and adjectives maintain relatively invariant forms. Furthermore, in Greek, subject pronouns can be omitted (a pro-drop language phenomenon), where the subject is syntactically implied through verb inflections.

The input to the algorithm is a set of words sharing a common root and core semantic meaning across various parts of speech (e.g., a verb conjugated in three persons in singular and plural, its derivative adverbs, imperative forms, or an adjective in three genders yielding adverbs and nouns). Therefore, the initial clustering of word sets required for communication takes precedence. This is defined as an "initial placement" because the system's architecture natively supports dynamic lexicon expansion [21]. The algorithm initiates execution as soon as a set of words reaches its input stream.

The data processing pipeline is executed across six sequentially interconnected tables as follows:

1) "root_table" is the first stage. The common root (the invariant part of the words describing the core meaning) is identified. Crucially, to optimize parser execution, the roots are indexed and ranked in descending order based on character length (from most letters to fewest) rather than alphabetically. This structural prioritization prevents false-positive substring matches during parsing.
2) "meaning_table" is the second table. In the corresponding position, the semantic interpretation of the word is entered, acting as the direct interface linking the lexical token to the meaning of the root, which the NLU unit will analyze later.
3) "part of speech table" is the table that declares the default grammatical category of the root.
4) "data_table" is the fourth table that mathematically encodes the semantic and syntactic properties of the root; this table stores a strict 9-digit vector (D1-D2-D3- D4-D5-D6-D7-D8-D9). The structural encoding of this vector is defined as follows:

   a. D1 . . . D4 (Syntactic Multi-functional Indicators): These digits indicate whether the root can generate a verb (D1), noun (D2), adjective (D3), or adverb (D4). A value of 0 denotes structural inability, while 1 denotes ability. Specifically, for the verb position (D1), a multi-tiered mapping is applied: 1 = transitive verb requiring one object, 2 = ditransitive verb requiring two objects, 3 = intransitive movement/shift verb between two spatial nodes, and 4 = transitive movement/shift verb requiring two objects.
   b. D5 . . . D8 (Suffix Group Pointers): These digits map the root to its corresponding morphological suffix paradigm group (indexed in the "suffix_group_table") for each of the four parts of speech, respectively.
   c. D9 (Spatial Semantic Flag): Indicates spatial properties, where 0 represents no spatial context, while 1 explicitly flags a geographic or spatial position description.
   For example, a vector value of 3101xxxx0 formally denotes a root that can function as a movement verb, requires an object, can transition into a noun, cannot form an adjective, can form an adverb, is linked to specific suffix groups (xxxx), and contains no spatial references.

5) "suffix group_table" is the next table that stores the groups of morphological endings clustered by structural and declension similarity, accommodating the complex inflections of the Greek language. It should be noted that in the Greek language, as in most languages, words do not have the same suffixes (endings) for the same part of speech or for the number. Thus, the suffixes are grouped in this based on their similarity.
6) "suffix group_table" is the last table. In each position of this table, specific codes are noted that indicate what the suffix

of its corresponding position represents. This table is needful for NLU, aiming to create the semantic substance of the sentence/command. In Table 1, the codes that characterize the grammatical status of the corresponding suffixes (endings) are presented.

Table 1
Codes corresponding to the grammatical status indicated by the suffix

Suffix meaning code	Part of speech, grammatical status
1PS	Verb, 1st person singular present
2PS	Verb, 2nd person singular present
3PS	Verb, 3rd person singular present
1PP	Verb, 1st person plural present
2PP	Verb, 2nd person plural present
3PP	Verb, 3rd person plural present
IMPP	Verb, imperative present
IMPA	Verb, imperative aorist (indefinite)
ADVT	Adverb, manner (shows a way)
ADVP	Adverb, indicates place
ADJM	Adjective, gender masculine
ADJF	Adjective, gender feminine
ADJF	Adjective, gender neuter
NM	Noun, gender masculine
NF	Noun, gender feminine
NN	Noun, gender neuter

Outside the lexicon but part of the grammar are the prepositions of the Greek language, the articles, the particles, the pronouns, the numbers, and the numerals.

5. Experimental Results

This section demonstrates the practical application, algorithmic validation, and operational performance of the proposed methodology within the context of the authors' iHRI architecture. To evaluate the execution efficiency of both the language adequacy checking and lexicon development processes, a targeted Greek linguistic corpus—comprising representative action verbs and spatial terms—is utilized as a functional test case. The following subsections detail the dataset structure and system configuration, present the execution results across the relational database entities, and provide a comprehensive methodological evaluation of the framework's structural adequacy and computational efficiency.

5.1. Language's nationality experimental setup

According to the method of subsection 4.2.1 and taking into account the application for which the lexicon is intended, the requirement for each criterion must be defined. Correspondingly, from the selected language, which here is Greek, the offer for each criterion must be given. For each criterion, two data points are derived. The first (required data) represents the linguistic requirements necessary for the application under the given criterion. The second (provided data) reflects the selected language, particularly its integration or offering within the context of that criterion. Based on the comparison of the two data (comparison result), the criterion is met or not, as shown below:

1) End users' language
   Required data: The language is addressed to a robotic application or is in the development stage through a research process in the Greek University community. The language of the end users is Greek.
   Provided data: The language is Greek.
   Comparison result: Acceptable.
2) Availability of resources for voice recognition (ASR)
   Required data: Sufficient resources for the Greek language chosen.
   Provided data: The ASR that was integrated supports the Greek language.
   Comparison result: Acceptable.
3) Target market or scope
   Required data: It's going for two research prototypes, which are "Vehicle for Operational Uses" and "Virtual Assistants for people with brain damage"; in both cases, the national language of the users is sufficient since these are research prototypes. Moreover, a patient with brain damage, who is examined in the second case, does not require being highly educated and multilingual.
   Provided data: The national language, which is Greek.
   Comparison result: Acceptable.
4) Cultural and social characteristics
   Required data: As mentioned in the previous criterion, the requirements in this scope are sufficient for the nationality of the language alone.
   Provided data: The national language, which is Greek.
   Comparison result: Acceptable.
5) Multilingualism
   Required data: Similar to the above two criteria, you do not require multilingualism due to the scope of application.
   Provided data: The national language, which is Greek.
   Comparison result: Acceptable.
6) Dialect recognition and understanding
   Required data: Although in the Greek territory there are several language idioms for many places, it is not a problem because the iHRI used allows learning new words, in the event that the user mentions a word in the local dialect. Also, in the current research phase, there is greater standardization in the handling of the system, so the national language in its basic form is sufficient.
   Provided data: The national language, which is Greek.

All criteria based on the method have been accepted, so the language that has been selected is characterized as sufficient and suitable for the prescribed use.

5.2. Dataset storage and reproduction

According to the method of subsection 4.2.2, the Greek words "προχωρώ" and "στρίβω," which are verbs by default, and the word "αριστερός," which is an adjective by default, will be examined and stored. In English, they are interpreted as "I move forward" and "I turn," respectively, for the first two verbs, and "left" (masculine gender) for the adjective. Then, in Tables 2–4, some versions of the words that may arise and are desired to be included in the lexicon are presented. At the same time, their writing in Latin characters, the part of speech, their grammatical status, and their interpretation in the English language are given.

Table 2
Set of words based on the "*προχωρώ*" verb that are going to be formed and stored in the lexicon

Word in Greek	Word in Latin characters	Part of speech by default	Grammatical status	Interpretation in English
προχωρώ	prochoro	Verb	1st person singular present	I move forward
προχωράς	prochoras	Verb	2nd person singular present	You move forward
προχωρεί	prochori	Verb	3rd person singular present	He/She/It moves forward
προχωράμε	prochorame	Verb	1st person plural present	We move forward
προχωράτε	prochorate	Verb	2nd person plural present	You move forward
προχωρούν	prochoroun	Verb	3rd person plural present	They move forward
προχώρα	prochora	Verb	Imperative present	Move forward now
προχώρησε	prochorise	Verb	Imperative indefinite	Moved forward
προχωρώντας	prochorontas	Adverb	Shows a way	Moving forward

Table 3
Set of words based on the "*στρίβω*" verb that are going to be formed and stored in the lexicon

Word in Greek	Word in Latin characters	Part of speech by default	Grammatical status	Interpretation in English
στρίβω	strivo	Verb	1st person singular present	I turn
στρίβεις	strivis	Verb	2nd person singular present	You turn
στρίβει	strivi	Verb	3rd person singular present	Ye/She/It turns
στρίβουμε	strivoume	Verb	1st person plural present	We turn
στρίβετε	strivete	Verb	2nd person plural present	You turn
στρίβουν	strivoun	Verb	3rd person plural present	They turn
στρίβε	strive	Verb	Imperative present	Turn now (continuous)
στρίψε	stripse	Verb	Imperative indefinite	Turn (simple imperative)
στρίβοντας	strivontas	Adverb	Shows a way	Turning

Table 4
Set of words based on the "*αριστερός*" adjective that are going to be formed and stored in the lexicon

Word in Greek	Word in Latin characters	Part of speech by default	Grammatical status	Interpretation in English
αριστερός	aristeros	Adjective	Gender masculine	The left one (masculine)
αριστερή	aristeri	Adjective	Gender feminine	The left one (feminine)
αριστερό	aristero	Adjective	Gender neuter	The left one (neuter)
αριστερά	aristera	Adverb	Indicates place	Left

If the particle "*να*" ("to" in English) is inserted before the first six verbs (Table 2 and Table 3) that are in the definite declension, and then the verbs become in the subjunctive. Correspondingly, if the particle "*θα*" ("will" in English) is inserted, then they become in the future tense. The example could include the past tense or even the plural in the imperatives, but it would start to become very complex. In any case, the set of words is an initial convention that, as already mentioned, we accept simplicity for the sake of CNL and always within the limits of the application. For the applications that the authors mention, the above is sufficient, without however excluding further expansion. The following is a description of the algorithm's implementation:

The set of words enters the initial process of finding a common root. The unit recognizes as a root the "*προχωρ-*" (prochor-) or "*στρί-*" (stri-) or "*αριστερ-*" (arister-), respectively, which it registers in the root table named "root_table." It should be noted again that the ranking in the table is not based on alphabetical order but on the root with the most letters to the one with the fewest.

When the three words above go through this specific process of subsection 4.2.2, the following results in Table 5 will emerge. It should be noted that this table does not show the contents of the lexicon's database tables, but rather the organization of words into a common root, suffixes, and general root meanings. It is considered necessary to have a discussion about the segmentation of words before presenting the data entered in the database tables, in order to make the process easier to understand. Tables 5 and 6 help in this direction. The data of the database tables are shown

Table 5
Results for three sets of words in lexicon format

	Suffixes for any part of speech				Interpretation in
Root	Verb	Noun	Adjective	Adverb	English
proxor-	o, as, i, ame, ate, oun, a, ise	–	–	ontas	Move forward
stri-v-ps-	o, is, i, oume, ete, oun, e, e	–	–	ontas	Turn
arister-	–	–	os, i, o	a	Left

Table 6
Results for three set of words with added adjective suffixes for the verbs, in lexicon format

	Suffixes for any part of speech				Interpretation
Root	Verb	Noun	Adjective	Adverb	in English
proxor-i-	o, as, i, ame, ate, oun, a, ise	–	menos, meni, meno	ontas	Move forward
stri-v-ps-	o, is, i, oume, ete, oun, e, e	–	menos, meni, meno	ontas	Turn
arister-	–	–	os, i, o	a	Left

in another table below. For the understanding of all readers, they are given in Latin characters and translated into English.

In the second word of Table 5, it is observed that there are some letters ("ps" corresponds to the Greek letter "ψ") that are surrounded by dashes. The parser ignores these when checking to locate the root of the word it is examining. That is, from the content "stri-v-ps-", the parser receives only "stri." However, when the word is analyzed, the word will contain one of the two or neither between the root and the suffix—a letter that is not included in the root nor in the suffix, for example, "stri-v-..." or "stri-ps-..." or "stri...". Note again that the suffixes in the groups are followed by the codes in the "suffix meaning_table" table that indicates the grammatical status of each one separately. To the verbs presented above, the adjectives that arise from them could be added, which are "proxorimenos," "proxorimeni," "proxorimeno" (is moved forward; in the three genders) and "strimenos," "strimeni," "strimeno" (is turned; in the three genders). Here we observe that the endings are the same in both sets, but only in the root "stri-" is the "i" present, while in "prochor-", it is not. In this case, the two roots need a different group of endings. To avoid this, we can put "i-" next to the root "prochor-" (prochor-i-), and the group of endings becomes "menos," "meni," "meno" for both roots. The result is what follows in Table 6 and changes only in terms of one root and the endings and not the meaning.

The lexicon's tables, after analysis and entry, are shown in Table 7.

Figure 5 shows a schematic representation of the contents of the dictionary tables. It also shows how the roots are connected to the meaning and to their default part of speech. For the second root of the table, further connections are shown with the remaining tables, which ultimately compose the entire set of words of this root that have been entered.

Next, reversing the process, we observe the composition of all words from one position of the "root_table." Choosing the 2nd position, it is returned:

Root = prochor(i)-
Meaning = move forward
part of speech_table = verb
data_table (D1- D2- D3- D4- D5- D6- D7- D8- D9) = 301120450 which mean:

1) D1 = 3: A verb can be derived that indicates a shift or movement between two points;
2) D2 = 0: No noun can be derived;
3) D3 = 1: An adjective can be derived;
4) D4 = 1: An adverb can be derived;
5) D5 = 2: The endings of the verbs are in position 2 of the "suffix group_table";

Table 7
The complete six-table relational schema of the lexicon, populated with the comprehensive data from the three word sets

root_table	meaning_ table	part of speech_ table	data_table	suffix group_table	suffix meaning_table
[arister-, proxor-i-, stri-v-ps-]	[left, move forward, turn]	[adjective, verb, verb]	[001100360, 301120450, 101110450]	[{}, {o, is, i, oume, ete, oun, e, e}, {o, as, i, ame, ate, oun, a, ise}, {os, i, o}, {menos, meni, meno}, {ontas}, {a}]	[{}, {1PS, 2PS, 3PS, 1PP, 2PP, 3PP, IMPP, IMPA}, {1PS, 2PS, 3PS, 1PP, 2PP, 3PP, IMPP, IMPA}, {ADJM, ADJF, ADJN}, {ADJM, ADJF, ADJN}, {ADVT}, {ADVL}]

Figure 5
Lexicon's tables connection

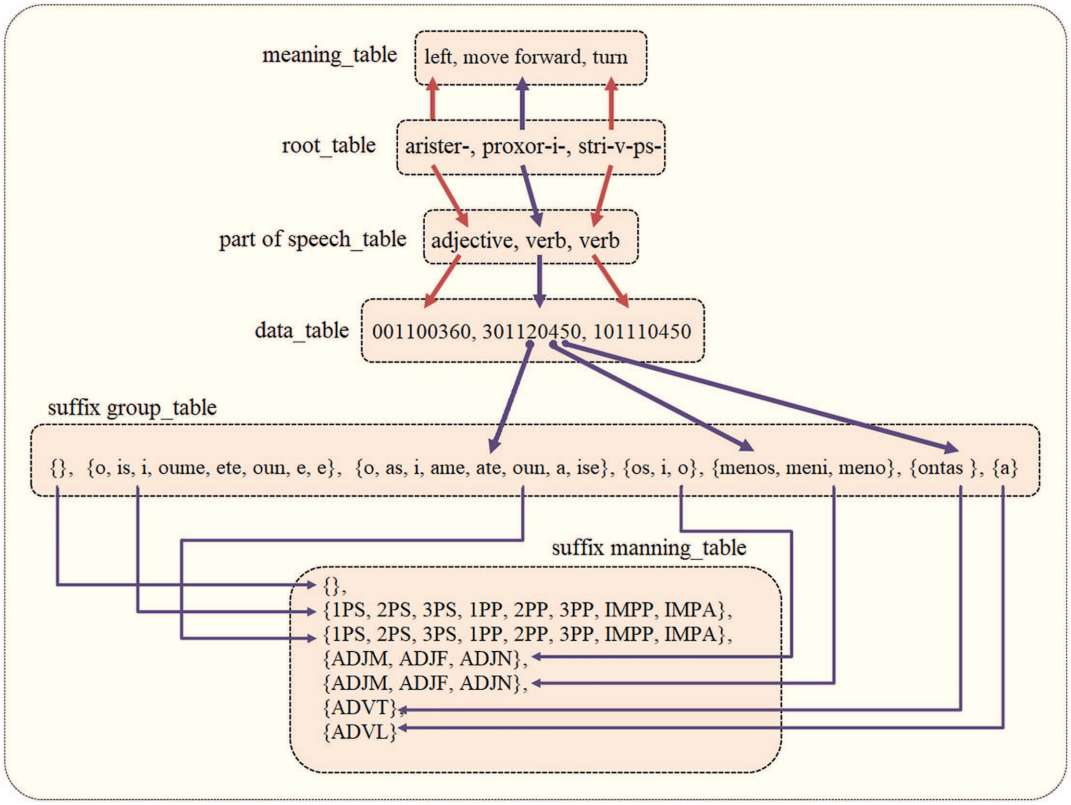


Table 8
All words reproduced from the 2nd entry position of the lexicon

Word	Part of speech	Suffix meaning code	Grammatical status	Interpretation in English
prochor(i)o	Verb	1PS	1st person singular present	I move forward
prochor(i)as	Verb	2PS	2nd person singular present	You move forward
prochor(i)i	Verb	3PS	3rd person singular present	He/She/It moves forward
prochor(i)ame	Verb	1PP	1st person plural present	We move forward
prochor(i)ate	Verb	2PP	2nd person plural present	You move forward
prochor(i)oun	Verb	3PP	3rd person plural present	They move forward
prochor(i)a	Verb	IMPP	Imperative present	Move forward now
prochor(i)ise	Verb	IMPA	Imperative indefinite	Moved forward
prochor(i)ontas	Adverb	ADVT	Shows a way	Moving forward
prochor(i)menos	Adjective	ADJM	Gender masculine	He is moved
prochor(i)meni	Adjective	ADJF	Gender feminine	She is moved
prochor(i)meno	Adjective	ADJF	Gender neuter	It is moved

6) D6 = 0: There are no endings for nouns;
7) D7 = 4: The endings of the adjectives are in position 4 of the "suffix group_table";
8) D8 = 5: The ending of the adverb is in position 5 of the "suffix group_table";
9) D9 = 0: The root does not describe place or space.

According to the above, Table 8 emerges.

It is very clear that there is a significant gain in space in the lexicon and several words can be supported, while maintaining several elements of the grammar of the natural language. In the words reproduced in parentheses, the "i" has been inserted because the method cannot know where it is inserted and where it is not. The correctness of the word arises during the analysis of an incoming word; that is, if the word "prochorame" arrives, then the word analyzer will find the root "prochor" and will determine that the ending "ame" is correct for the root. This is the word! On the contrary, in the word "prochorimeni," it will find the root "prochor"; it will determine that the ending "meni" is correct, but it will have an extra "i." Because the root has the notation "i-", it will fit it in this position, and the word will be correct!

5.3. Methodological validation and performance evaluation

The validity and accuracy of the proposed SLM framework were evaluated through an internal functional verification process, as detailed in subsections 5.1 and 5.2. The primary indirect evaluation of this methodology relied on the reverse reproduction of lexical tokens; the system successfully reconstructed all grammatical variants—including persons, inflections, and tenses—with zero error, using exclusively the encoded roots stored in the database. This flawless reconstruction confirms the structural reliability of the lexicon. Furthermore, the parser's capacity to process elliptical speech by dynamically retrieving missing context from the conversation history, rather than rejecting the commands, demonstrates that the model ensures the semantic precision necessary for robotic grounding and navigation. Since the current work focuses primarily on introducing the model's architectural infrastructure, extensive experimental validation with quantitative data in real-world robotic environments constitutes the immediate next step of this research.

In this context, empirical evaluations—such as latency measurements or metrics demonstrating improvements in ambiguity resolution and fault tolerance in safety-critical settings—pertain directly to the runtime execution of the overarching iHRI system itself. Consequently, such operational evaluations fall outside the scope of the present study, which is dedicated strictly to the isolated linguistic and structural methodology of the dictionary database.

6. Discussion and Conclusions

A designed system that will have communication capabilities needs to be supported by a language model. Natural language models can be small SLMs or large LLMs. LLMs are constantly developing with impressive results, even at the level of private dialogue between a human and a machine. Nevertheless, SLMs exist, and there are many applications that need them.

Through this work, both the characteristics that an SLM must have and some relevant criteria for the selection and control of the lexicon that it uses are highlighted. The selection of a lexicon regarding the basic language and the vocabulary cannot be a random choice. It is always selected with criteria, even if this seems to be done unintentionally.

A method that compares the system requirements with the capabilities of a selected lexicon (adequacy checking method) is given in this work, and it can classify it into three levels of suitability.

Also, a method for developing a lexicon for an SLM in the natural Greek language, which is the main subject of the work, is given. This method excels in that the lexicon is not simple standardized words in specific grammatical expressions. The lexicon that is developed comprises six tables of a simple database that are appropriately connected and produce, with simple procedures, much more information than a simple standard lexicon. The language model remains small, but the system uses more grammatical expressions and becomes friendlier. The fact that the system demonstrated a zero-error rate during the reverse reproduction test shows that we can successfully handle "heavy" grammar and inflections within a very light database structure. In practice, this means we can deploy reliable language processing directly on the robot's local hardware. As a result, the system stays fully capable of understanding conversational gaps and guiding the robot safely, even when operating in completely isolated environments without internet or cloud access.

As directions for future research, both the improvement of the lexicon and the further development of its selection and adequacy checking method are given:

1) The adequacy checking method is recommended to be studied through machine learning models. The result of the method could be a suitability percentage instead of three simple categories. In addition, the method can propose this criterion: if it is improved, the suitability percentage will also be improved.
2) In the lexicon development method, it is recommended to study the corresponding process for other natural languages. In addition, it is recommended to make the segmentation of words and endings more systematic, which will lead to even greater data compression.

In conclusion, this paper presents a method for designing and developing a lexicon for an SLM of the natural Greek language for robotic applications. This lexicon, in combination with the other elements contained in the language unit of an intelligent human–robot interaction that the authors have developed, provides additional linguistic capabilities. In addition, another method is given for checking the suitability of a lexicon selected for use in a robotic application. This method is based on the language selection criteria identified in the various indicative studies examined. The methods show good functionality, and the authors are led to the conclusion that there is scope for further research and improvement, as well as their application to other natural languages, especially those of rich declensional morphology.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Ioannis Giachos: Conceptualization, Methodology, Software, Investigation, Resources, Writing – original draft, Visualization. **Evangelos Papakitsos:** Conceptualization, Methodology, Investigation, Resources, Writing – original draft. **Nikoletta Samaridi:** Investigation, Resources, Data curation. **Michail Papoutsidakis:** Validation, Formal analysis. **Nikolaos Laskaris:** Writing – review & editing, Supervision.

References

[1] Marge, M., Espy-Wilson, C., Ward, N. G., Alwan, A., Artzi, Y., Bansal, M., . . . , & Yu, Z. (2022). Spoken language interaction with robots: Recommendations for future research. *Computer Speech & Language*, *71*, 101255. https://doi.org/10.1016/j.csl.2021.101255

[2] Reimann, M. M., Kunneman, F. A., Oertel, C., & Hindriks, K. V. (2024). A survey on dialogue management in human-robot interaction. *ACM Transactions on Human-Robot Interaction*, *13*(2), 22. https://doi.org/10.1145/3648605

[3] Rojo-Echeburúa, A. (2024). *Top 15 small language models for 2025*. Datacamp. https://www.datacamp.com/blog/top-small-language-models?dc_referrer=https%3A%2F%2Fwww.google.com%2F

[4] Giachos, I., Batzaki, E., Papakitsos, E. C., Kaminaris, S., & Laskaris, N. (2023). A natural language generation algorithm for Greek by using hole semantics and a systemic grammatical formalism. *Journal of Computer Science Research, 5*(4), 27–37. https://doi.org/10.30564/jcsr.v5i4.6067

[5] Pierce, M. (2023). Language and meaning in the age of modernism. C.K. Ogden and his contemporaries. *Language & History*, *66*(3), 292–294. https://doi.org/10.1080/17597536.2022.2162330

[6] Veizaga, A., Alferez, M., Torre, D., Sabetzadeh, M., & Briand, L. (2021). On systematically building a controlled natural language for functional requirements. *Empirical Software Engineering*, *26*(4), 79. https://doi.org/10.1007/s10664-021-09956-6

[7] Carvalho, A. (2022). REX: General-purpose CNL with code generation support. *Applied Sciences*, *12*(15), 7700. https://doi.org/10.3390/app12157700

[8] Listenmaa, I., Hanafiah, M., Cheong, R., & Källberg, A. (2021). Towards CNL-based verbalization of computational contracts. In *Proceedings of the Seventh International Workshop on Controlled Natural Language*.

[9] Wang, A. W., & Friginal, E. (2025). From tradition to innovation? A diachronic corpus analysis of simplified technical English in aviation. *Journal of English and Applied Linguistics*, *4*(2), 1. https://doi.org/10.59588/2961-3094.1168

[10] Steurs, F., & Leiden, I. N. T. (2021). Language is business! On the impact of linguistic applications in our society. *Canadian Journal of Netherlandic Studies Revue canadienne d'études néerlandaises*, *29*(30), 1–16.

[11] Cheng, L., Ben, P., & Qiao, Y. (2022). Research on automatic error correction method in English writing based on deep neural network. *Computational Intelligence and Neuroscience*, *2022*(1), 2709255. https://doi.org/10.1155/2022/2709255

[12] Caruso, S., Dodaro, C., Scudo, F. L., Maratea, M., & Reale, K. (2025). On the translation of ASP rules to (controlled) natural language sentences. In *Proceedings of the Workshops and Doctoral Consortium of the 41st International Conference on Logic Programming*.

[13] Tran, D., Yan, F., Yihun, Y., Tan, J., & He, H. (2021). A framework of controlled robot language for reliable human-robot collaboration. In *Social Robotics: 13th International Conference*, 339–349. https://doi.org/10.1007/978-3-030-90525-5_29

[14] Lengeling, J. C. (2021). Facilitating the application of controlled natural language (CNL) to standardize communication in logistics and supply chain management. In *Proceedings of the Seventh International Workshop on Controlled Natural Language*.

[15] Hegde, H., Vendetti, J., Goutte-Gattat, D., Caufield, J. H., Graybeal, J. B., Harris, N. L., . . . , & Mungall, C. J. (2025). A change language for ontologies and knowledge graphs. *Database*, baae133. https://doi.org/10.1093/database/baae133 *2025*

[16] Karim, H., Gupta, D., & Sitharaman, S. (2025). Securing LLM workloads with NIST AI RMF in the internet of robotic things. *IEEE Access*, *13*, 69631–69649. https://doi.org/10.1109/ACCESS.2025.3561235

[17] Jonek, M., Niermann, D., Petzoldt, C., Manns, M., & Freitag, M. (2023). Automated analysis of assembly processes in human-robot collaboration: Research approaches and challenges. *Procedia CIRP*, *120*, 1203–1208. https://doi.org/10.1016/j.procir.2023.09.149

[18] Law, W. T., Fan, K. W., Li, K. S., & Mo, T. (2023). Applied design and methodology of delivery robots based on human-robot interaction in smart cities. *EAI Endorsed Transactions on Smart Cities*, *7*(2), 1.

[19] Veeranna, T., Kumar, K. P., Aruna, M., Verma, A., & Kurian, T. (2024). Enhancing human-robot interaction through natural language processing and speech recognition techniques. In *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences*, 1–6. https://doi.org/10.1109/IC3TES62412.2024.10877529

[20] Araujo, H., Mousavi, M. R., & Varshosaz, M. (2023). Testing, validation, and verification of robotic and autonomous systems: A systematic review. *ACM Transactions on Software Engineering and Methodology*, *32*(2), 51. https://doi.org/10.1145/3542945

[21] Giachos, I., Papakitsos, E. C., Antonopoulos, I., & Laskaris, N. (2023). Systemic and hole semantics in human-machine language interfaces. In *2023 17th International Conference on Engineering of Modern Electric Systems*, 1–4. https://doi.org/10.1109/EMES58375.2023.10171635

[22] Ammar, S., & Bhiri, M. T. (2021). A formal approach combining event-B and PDDL for planning problems. In *Proceedings of the 16th International Conference on Software Technologies*, 261–268. https://doi.org/10.5220/0010577102610268

[23] Brizzio, M., Cordy, M., Papadakis, M., Sánchez, C., Aguirre, N., & Degiovanni, R. (2023). Automated repair of unrealisable LTL specifications guided by model counting. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 1499–1507. https://doi.org/10.1145/3583131.3590454

[24] Wertheim, O., & Brafman, R. I. (2025). Model-based AI planning and execution platforms for robotics. *AI Magazine*, *46*(3), e70034. https://doi.org/10.1002/aaai.70034

[25] Godwin-Jones, R. (2023). Presence and agency in real and virtual spaces: The promise of extended reality for language learning. *Language Learning & Technology*, *27*(3), 6–26. https://doi.org/10.64152/10125/73529

[26] Yemmene, P., & Besacier, L. (2019). Motivations, challenges, and perspectives for the development of an automatic speech recognition system for the under-resourced Ngiemboon language. In *Proceedings of the First International Workshop on NLP Solutions for Under Resourced Languages*, 59–67.

[27] Huang, G., & Moore, R. K. (2023). Using social robots for language learning: Are we there yet? *Journal of China Computer-Assisted Language Learning*, *3*(1), 208–230. https://doi.org/10.1515/jccall-2023-0013

[28] Besacier, L., Barnard, E., Karpov, A., & Schultz, T. (2014). Automatic speech recognition for under-resourced languages: A survey. *Speech Communication*, *56*, 85–100. https://doi.org/10.1016/j.specom.2013.07.008

[29] Toshniwal, S., Sainath, T. N., Weiss, R. J., Li, B., Moreno, P., Weinstein, E., & Rao, K. (2018). Multilingual speech recognition with a single end-to-end model. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, 4904–4908. https://doi.org/10.1109/ICASSP.2018.8461972

[30] Meelen, M., O'Neill, A., & Coto-Solano, R. (2024). End-to-end speech recognition for endangered languages of Nepal. In *Proceedings of the Seventh Workshop on the Use of Computational Methods in the Study of Endangered Languages*, 83–93. https://doi.org/10.18653/v1/2024.computel-1.12

[31] Mavridis, N. (2015). A review of verbal and non-verbal human–robot interactive communication. *Robotics and Autonomous Systems*, *63*, 22–35. https://doi.org/10.1016/j.robot.2014.09.031

[32] Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954

[33] Giachos, I., Batzaki, E., Papakitsos, E. C., Papoutsidakis, M., & Laskaris, N. (2024). Developing a natural language understanding system for dealing with the sequencing problem in simulating brain damage. *WSEAS Transactions on Biology and Biomedicine*, *21*, 138–147. https://doi.org/10.37394/23208.2024.21.14

[34] Andrist, S., Mutlu, B., & Tapus, A. (2015). Look like me: Matching robot personality via gaze to increase motivation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3603–3612. https://doi.org/10.1145/2702123.2702592

[35] Cirasa, C., Sapienza, A., Cantucci, F., Conti, D., & Falcone, R. (2026). Trust and accent: How speaker accent influences interaction with humanoid robots. *Applied Sciences*, *16*(9), 4342. https://doi.org/10.3390/app16094342

[36] Bruno, B., Amirova, A., Sandygulova, A., Lugrin, B., & Johal, W. (2023). Culture in social robots for education. In B. J. Dunstan, V. J. T. K. Koh, D. T. Tillman, & S. A. Brown (Eds.), *Cultural robotics: Social robots and their emergent cultural ecologies* (pp. 127–145). Springer International Publishing. https://doi.org/10.1007/978-3-031-28138-9_9

[37] Adel, H., Vu, N. T., Kraus, F., Schlippe, T., Li, H., & Schultz, T. (2013). Recurrent neural network language modeling for code switching conversational speech. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 8411–8415. https://doi.org/10.1109/ICASSP.2013.6639306

[38] Du, C., Li, H., Lu, Y., Wang, L., & Qian, Y. (2021). Data augmentation for end-to-end code-switching speech recognition. In *2021 IEEE Spoken Language Technology Workshop*, 194–200. https://doi.org/10.1109/SLT48900.2021.9383620

[39] Agro, M. T., Kulkarni, A., Kadaoui, K., Talat, Z., & Aldarmaki, H. (2026). Code-switching in end-to-end automatic speech recognition: A systematic literature review. In *Proceedings of the Fifteenth Language Resources and Evaluation Conference*, 9790–9812.

[40] Errattahi, R., El Hannani, A., & Ouahmane, H. (2018). Automatic speech recognition errors detection and correction: A review. *Procedia Computer Science*, *128*, 32–37. https://doi.org/10.1016/j.procs.2018.03.005

[41] Hirayama, N., Yoshino, K., Itoyama, K., Mori, S., & Okuno, H. G. (2015). Automatic speech recognition for mixed dialect utterances by mixing dialect language models. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *23*(2), 373–382. https://doi.org/10.1109/TASLP.2014.2387414

[42] Li, Q., Mai, Q., Wang, M., & Ma, M. (2024). Chinese dialect speech recognition: A comprehensive survey. *Artificial Intelligence Review*, *57*(2), 25. https://doi.org/10.1007/s10462-023-10668-0

[43] Torrent, T. T., Matos, E. E. D. S., Belcavello, F., Viridiano, M., Gamonal, M. A., Costa, A. D. D., & Marim, M. C. (2022). Representing context in FrameNet: A multidimensional, multimodal approach. *Frontiers in Psychology*, *13*, 838441. https://doi.org/10.3389/fpsyg.2022.838441

[44] Lehmann, J., Gattogi, P., Bhandiwad, D., Ferré, S., & Vahdati, S. (2023). Language models as controlled natural language semantic parsers for knowledge graph question answering. In *ECAI 2023: 26th European Conference on Artificial Intelligence*, *372*, 1348–1356. https://doi.org/10.3233/FAIA230411

[45] Wang, Y., Fodor, P., & Kifer, M. (2023). Knowledge authoring for rules and actions. *Theory and Practice of Logic Programming*, *23*(4), 797–811. https://doi.org/10.1017/S1471068423000169

How to Cite: Giachos, I., Samaridi, N., Papakitsos, E., Papoutsidakis, M., & Laskaris, N. (2026). Designing a Small Language Model Lexicon in Natural Language for Robotic Applications. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA62027336