



Article Natural Language Understanding for Navigation of Service Robots in Low-Resource Domains and Languages: Scenarios in Spanish and Nahuatl

Amadeo Hernández ¹^(b), Rosa María Ortega-Mendoza ^{2,*}^(b), Esaú Villatoro-Tello ³, César Joel Camacho-Bello ¹^(b) and Obed Pérez-Cortés ²

- ¹ Laboratorio de Inteligencia Artificial, Universidad Politécnica de Tulancingo, Tulancingo 43629, Hidalgo, Mexico; amadeo.hernandez2115004@upt.edu.mx (A.H.); cesar.camacho@upt.edu.mx (C.J.C.-B.)
- ² Instituto de Ciencias Básicas e Ingeniería, Universidad Autónoma del Estado de Hidalgo,
 - Mineral de la Reforma 42184, Hidalgo, Mexico; obed_perez@uaeh.edu.mx Idiap Research Institute, 1920 Martigny, Switzerland
- * Correspondence: rosa.ortega@uaeh.edu.mx

3

Abstract: Human-robot interaction is becoming increasingly common to perform useful tasks in everyday life. From the human-machine communication perspective, achieving effective interaction in natural language is one challenge. To address it, natural language processing strategies have recently been used, commonly following a supervised machine learning framework. In this context, most approaches rely on the use of linguistic resources (e.g., taggers or embeddings), including training corpora. Unfortunately, such resources are scarce for some languages in specific domains, increasing the complexity of solution approaches. Motivated by these challenges, this paper explores deep learning methods for understanding natural language commands emitted to service robots that guide their movements in low-resource scenarios, defined by the use of Spanish and Nahuatl languages, for which linguistic resources are scarcely unavailable for this specific task. Particularly, we applied natural language understanding (NLU) techniques using deep neural networks and transformers-based models. As part of the research methodology, we introduced a labeled dataset of movement commands in the mentioned languages. The results show that models based on transformers work well to recognize commands (intent classification task) and their parameters (e.g., quantities and movement units) in Spanish, achieving a performance of 98.70% (accuracy) and 96.96% (F1) for the intent classification and slot-filling tasks, respectively). In Nahuatl, the best performance obtained was 93.5% (accuracy) and 88.57% (F1) in these tasks, respectively. In general, this study shows that robot movements can be guided in natural language through machine learning models using neural models and cross-lingual transfer strategies, even in low-resource scenarios.

Keywords: natural language understanding; intent classification; slot filling; deep learning models; service robots; low-resource domains; Nahuatl and Spanish utterances

MSC: 68T50

1. Introduction

Nowadays, the interaction between humans and robots in daily life has become a common activity, mainly to accomplish useful tasks for humans [1]. An important challenge is enabling robots to execute instructions conveyed in natural language (e.g., commands). This form of communication simplifies the process for non-expert users in robotics to instruct the machines on the task to be performed [2]. Recently, to face this challenge, solutions based on natural language processing (NLP) have been proposed [3,4]. In particular, approaches based on natural language understanding (NLU) techniques have been



Citation: Hernández, A.; Ortega-Mendoza, R.M.; Villatoro-Tello, E.; Camacho-Bello, C.J.; Pérez-Cortés, O. Natural Language Understanding for Navigation of Service Robots in Low-Resource Domains and Languages: Scenarios in Spanish and Nahuatl. *Mathematics* 2024, *12*, 1136. https://doi.org/10.3390/ math12081136

Academic Editors: Víctor Manuel Darriba Bilbao, Alexander Gelbukh and Alvaro Rodrigo

Received: 27 February 2024 Revised: 30 March 2024 Accepted: 6 April 2024 Published: 10 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). proposed and evaluated as promising alternatives [5,6]. Overall, these approaches enable the identification and extraction of parameters from commands provided to robots, such as units of measurement, names, dates, and locations.

Frequently, NLU approaches in this field follow a supervised machine learning perspective, causing dependence on labeled training datasets. Therefore, one of the main challenges is the availability of training data in different languages. In this context, humanrobot interaction faces complex challenges because many languages worldwide lack training corpora or linguistic resources for their processing (e.g., dictionaries and part-of-speech taggers). Furthermore, there are domains in which only limited resources are available [7,8]. Particularly, work with indigenous languages faces challenges such as morphological complexity and limited training data, as well as dialectal and orthographic variation [9]. To solve these problems, some efforts have emerged to develop methods or tools involving automatic translation and automatic speech recognition. For instance, for the machine translation task, diverse contributions were presented at the First Workshop on Natural Language Processing for Indigenous Languages of the Americas 2021 [10]. However, understanding commands for robots in these languages is a low-resource domain that has received little attention.

In this paper, we evaluated NLU approaches to map movement instructions emitted in natural language for service robots, specifically Spanish and Nahuatl, into an intermediate representation that can be easily translated to robot actions. The fundamental challenges in NLU involve two tasks: identifying the intended action (intent classification) and identifying the parameters associated with that action (slot filling) [11,12]. We explored these tasks in Spanish and Nahuatl languages by evaluating deep neural architectures such as biLSTM (Bidirectional Long Short-Term Memory) [13] and biGRU (Bidirectional Gated Recurrent Unit) [14]. Since this scenario has low annotated data, we also evaluated large pre-trained language models such as BETO and multilingual BERT [15]. Additionally, we studied the use of zero [16,17] and few-shot [18,19] learning techniques to face the scarcity of labeled data. As part of the methodology, we created a dataset containing labeled movements in Spanish and Nahuatl to address the movements of mobile robots.

It is worth mentioning that the integration of technology into indigenous communities is important for the development of a country. Designing applications specifically for Nahuatl speakers enhances technology accessibility within their communities, contributes to the preservation of the language, and reduces its marginalization. Particularly, Nahuatl is a prominent indigenous language in Mexico with approximately 1,651,958 speakers across at least 16 states [20]. To date, this domain has a low availability of linguistic resources for these languages. For example, in the case of Nahuatl, monolingual and parallel datasets are important to develop NLU-based methods; however, they are scarce [21].

The main contributions of this work are summarized as follows:

- Evaluation of different deep neural network models and pre-trained transformerbased language models for identifying commands and parameters of movements robot using low-resources scenarios from an NLU perspective: In particular, we explored instructions specified in Spanish and Nahuatl languages, where this task lacks annotated data. To our knowledge, using an indigenous language, especially Nahuatl, has not been evaluated in this domain. Therefore, we provide the first extensive study of the use of Nahuatl to guide service robots.
- A new labeled parallel corpus in Spanish and Nahuatl containing movement instructions for mobile robotics: While this dataset was created for intent classification and slot-filling tasks, it can be utilized to investigate various NLP tasks. This dataset will be publicly available for investigation purposes and to motivate further research on this topic.
- An analysis of few-shot learning strategies adapted for classifying Nahuatl instances in a low-resource cross-lingual scenario: To our knowledge, this strategy has not been explored for identifying commands for a mobile robot in Nahuatl.

• An analysis of the robustness of the models: We analyze the models' performance when errors are considered intentionally as a simulation of their potential application.

The rest of the document is organized as follows: Section 2 presents related work and theoretical background. Section 3 introduces the data collection used in this research. Section 4 exposes the implementation of the classification models. Section 5 presents the experiments and results. Section 6 analyzes the impact of pre-trained word vectors applied in the experiments and evaluates the robustness of the models to noise. Finally, conclusions and future work are presented in Section 7.

2. Related Works and Theoretical Background

2.1. Natural Language to Interact with Robots

The interaction between humans and robots has gained importance recently, marking a significant change in human–machine communication. Adopting natural language instead of low-level machine languages has been a considerable focus in this area [1]. Significant efforts have been made to translate natural language into commands for robots. For example, in [22], the authors used a semantic analyzer to address the problem of converting natural language commands in English into robot command language (RCL). For instance, the instruction "take a left, then the next right" is converted into the commands (do-sequentially <turn-left> <turn-right>). The authors utilized the SemEval 2014 dataset, which consists of 3409 English commands. The sequences with the highest probability were obtained using a Hidden Markov Model tagger, achieving an overall accuracy of 92.45%. In [1], an encoder–decoder RNN was used to translate natural language movement commands into a set of precise trajectories, which include commands and related parameters to the robot movements. The authors used a dataset consisting of 1600 unique entries in English, obtaining an average accuracy of 79.23% and 73.65% for single and multiple movement commands within a sentence, respectively.

Lately, deep neural networks have been evaluated in this field [23,24]. An example is the [25] study, where a seq2seq neural network with LSTM was implemented to translate Spanish commands into RCL, using a corpus resulting in 3785 sentences after expanding seven original sentences. Another significant development is the ability of robots to understand English navigation instructions as described in [26], where a dataset of 706 natural language navigation instructions was used to achieve 56% accuracy in simulation systems and 44.83% in field systems. Additionally, [27] demonstrated how a robotic planner can understand English scripts to move and manipulate objects in a continuous configuration space, achieving a 72% success rate on a test set of 657 statements.

2.2. Natural Language Understanding Approaches

Recently, NLU-based solutions have been used to address the challenge of comprehending human utterances in interactions with robots. The main goal is to understand the meaning and intention of instructions expressed in natural languages. NLU typically includes two tasks to provide a semantic analysis of user expressions: intent classification and slot filling.

2.2.1. Intent Classification and Slot-Filling Tasks

The intent classification task (hereafter denoted as IC) determines the intention or purpose behind a statement or query made by a user in natural language [28]. It can be modeled as a sentence classification task, where each utterance is labeled with an intent [29]. For example, in the phrase "the vehicle runs at 3 km/h", the intent is to run, which can be mapped to a specific command of the intermediate language to control a mobile robot (e.g., forward). Alternatively, in the slot filling task (hereafter referenced as SF), it is often focused on identifying relevant information to understand the context and intention of the user (i.e., arguments of the intents) [28]. Commonly, it uses IOB (inside–outside–beginning tagging) to identify and extract specific and structured information from a text [30] with

the aim of more deeply understanding the information contained in the text. For instance, in the following sentence, "the vehicle runs at 3 km/h", the slot is "3 km/h".

These tasks can be formally described as follows. Given an input utterance $x = (x_1, x_2, ..., x_T)$, where x_i are the tokens, and T is the total number of tokens in the utterance, the IC task is defined as the classification that assigns a correct intent label y^{intent} to the whole utterance x. The SF task is a token-level sequence labeling that assigns a corresponding slot label to each token x_i in the utterance, that is, $y^{slot} = y_1^{slot}, y_2^{slot}, \ldots, y_T^{slot}$ [11]. This task is commonly used to extract the parameters (e.g., values, quantities, and units) or specific information contained in the statements.

The two mentioned tasks have traditionally been tackled independently [31]. However, more recently, several authors have devised strategies for addressing both assignments concurrently [32,33] (hereafter denoted as the joint task). The aim is modeling and exploiting the dependencies between the tasks—intents and slot filling—to improve performance using a joint architecture. On the contrary, conducting each task in an independent manner involves using and training architectures separately [34].

Different research works have tackled tasks related to intent classification and parameter detection. For example, in [35], a probabilistic graphical model is presented for understanding natural language commands provided to autonomous systems that perform mobile navigation in semi-structured environments. The performance reported was 86% (accuracy) on an English corpus. On the other hand, [36] tackled the problem of giving instructions in English, focusing on commands that model actions and their arguments. For these tasks, the authors evaluated RNN and LSTM architectures. They also created a dataset using the RoboCup's generator (https://github.com/RoboCupAtHome/gpsr_command_generator, accessed on 5 September 2023). Similarly, the work presented in [37] conducted a comparative analysis of four natural language understanding models: Mbot, Rasa, LU4R, and ECG. The models were evaluated to comprehend commands for domestic service robots. This includes recognizing actions and any supplementary information within the commands. The results showed that Mbot and Rasa were models suitable for this task.

2.2.2. Common Evaluation Metrics

The IC task is typically approached as a supervised classification problem, where the common evaluation metrics are accuracy, precision, recall, and F1-score [38]. On the other hand, the SF task is commonly evaluated at an entity level using evaluation metrics from the Named Entity Recognition (NER) task [39] and following a multi-class classification setting. The IOB format represents the data labels as Begin, Inside, and Outside tags. In this regard, each slot type is considered as a classification label [38].

2.3. Deep Neural Networks and Large Pre-Trained Language Models

This section describes the theoretical background of models evaluated in this research. It is organized into two subsections: deep neural networks and transformer-based models.

2.3.1. Deep Neural Networks

The architectures of neural networks are widely used in natural language processing tasks and other applications related to sequence analysis. In general, deep neural models use vectors of numbers from data, commonly word embeddings [40]. We introduce word embeddings and some deep neural architectures in the following paragraphs.

Word Embeddings

These word vectors are numeric vectors representing words in a lower-dimensional space and allowing words with similar meanings to have similar representations [41]. Hence, they are commonly used to capture semantic and contextual information. Generally, these vectors can be learned from unlabeled data by applying techniques such as Word2vec or FastText. The Word2vec model uses two main architectures to train the word vectors:

Continuous Bag of Words (CBOW) and Skip-gram. The FastText model is an approach based on the Skipgram model to learn word representations by considering subword information [42]. Nowadays, it is a common practice to feed neural networks with word vector representations, using them as the first data-processing layer [43]. Furthermore, using pre-trained word embeddings has shown to be a powerful tool for different natural language processing tasks [44,45].

Bidirectional Long Short-Term Memory (biLSTM)

The bidirectional LSTM network is a type of LSTM that was developed to enhance the model's performance in data classification processes, specifically within sequence structures [46]. Bidirectional LSTMs make a backward and forward pass through the sequence before passing on to the next layer [47]. This network is commonly used in the Named Entity Recognition (NER) task [48].

Bidirectional Gated Recurrent Unit (biGRU)

A biGRU refers to a bidirectional variant of a GRU. It is often regarded as two RNNs stacked on top of each other [49]. Each recurrent unit in a Gated Recurrent Unit (GRU) network captures dependencies across different time scales. Like the LSTM unit, the GRU architecture has gating units that modulate the flow of information within the unit but without separate memory cells [50]. In NLP tasks, this architecture is important to understand the context of a word in a sentence.

Conditional Random Field (CRF)

The CRF-based methods have shown to be effective approaches for addressing sequence labeling tasks [51]. The CRF models focus on the sentence level rather than individual positions to find optimal sequences of labels. The inputs and outputs are directly connected, unlike LSTM and bidirectional LSTM networks, where recurrent memory cells are employed. It has been demonstrated that CRFs can achieve higher labeling accuracy overall [52].

2.3.2. Transformer-Based Pre-Trained Language Models

NLP advances recently incorporated transformer-based pre-trained language models, showing superior performance on various natural language processing tasks [53,54]. The transformer architecture is based on attention mechanisms [55], either the self-attention mechanism or Multi-Head Attention [56]. The successful performance achieved by this architecture in different NLP tasks has inspired the development of large pre-trained language models working under the "pre-training then fine-tuning" paradigm [15]. This implies that language models can be pre-trained on large unannotated corpora and fine-tuned for specific tasks on smaller supervised datasets. In this regard, Bert-based models have recently gained popularity.

Bert-Based Models

Recently, language-pre-trained transformer-based models have shown to be successful in several NLP tasks, for example, BERT [57], ROBERTA [58], etc. In particular, BERT, which stands for Bidirectional Encoder Representations from Transformers (BERT), is an architecture transformer-based neural network architecture. This pre-trained model was trained on large amounts of text. During its training, BERT learns to predict hidden words in a sentence due to its bidirectional contextualization. This pre-training technique allows BERT to capture complex semantic and syntactic relationships in language [57].

BERT can be adapted for tasks with a smaller and more specific dataset through a fine-tuning process [4,57]. In the following paragraphs, two pre-trained models applied in this research are briefly described:

• Multilingual BERT: Unlike BERT, which was trained on English data, Multilingual BERT (hereafter, denoted as mBERT) was trained on Wikipedia in 104 different lan-

guages [59]. Its ability to generalize in a cross-lingual manner [59] allows capturing shared linguistic knowledge across different languages [60].

 BETO: This is a pre-trained BERT-based language model for Spanish NLP tasks with a similar size to the BERT model and trained on a large Spanish corpus gathered from different sources [61]. BETO has shown its effectiveness in tasks in Spanish, such as text classification and named entity recognition, outperforming other pre-trained multilingual models [61,62].

Recently, several works have leveraged the properties of pre-trained models to study the interaction of robots and humans. In this context, Bucker et al. [63] developed a natural language interface for human–robot interaction, allowing users to adjust the trajectories of autonomous agents. To achieve this, they utilized BERT and CLIP (Contrastive Language-Image Pre-training) for fusing language data with trajectory information through multimodal transformers. The experiments revealed that this integration yields more intuitive interfaces than conventional approaches. It is important to mention that a few works address the use of languages other than English. For example, [64] developed a BERT-based model to tackle Turkish natural language understanding building a single model for both NLU tasks (intent and slot classification). However, little attention has been paid to processing utterances in indigenous languages such as Nahuatl.

Cross-Lingual Transfer Learning

The cross-lingual setting is a concept focused on performing detection tasks where there are few or non-existent training datasets in the target language [65]. It has been used to leverage data from higher-resource languages to enhance classification models' performance on low-resource languages. In recent years, large pre-trained transformerbased models have shown their effectiveness for tackling cross-lingual transfer learning to transfer knowledge from one language to another [66]. In this regard, two settings recently explored are zero-shot and few-shot cross-lingual learning.

In the zero-shot learning case, a pre-trained model is fine-tuned on task-specific supervised training data in the source language (e.g., English), and it is used for evaluating the task on another target-language test data [18,67,68], that is, not a single labeled instance exists for a target language [17]. Zero-shot cross-lingual has been used for different text classification tasks, such as news sentiment classification [69] and named entity disambiguation [70]. Recently, this ability has been exploited to tackle tasks involving low-resource languages [71].

On the other hand, cross-lingual few-shot learning aims to classify inputs based on only a limited number of samples (shot) [65]. Recently, few-shot crosslingual transfer has outperformed zero-shot learning on several text classification tasks [72]. Also, few-shot learning has shown skills to address the low-resource challenge in non-English languages [73]. The multilingual BERT model has shown cross-lingual capabilities, performing well as a zero-shot cross-lingual transfer model [74]. Some studies have shown sensitivity to the selection of few shots [18].

Recently, some methods have emerged for intent classification and slot-filling tasks on few-shot settings using benchmarks such as ATIS, CLINC, and SNIPS [75,76]. However, we did not find approaches under these scenarios that guide a robot in Nahuatl.

2.4. Evaluating the Robustness of Text Classification Systems with Noisy Data

Noise in text analysis has a significant impact on classifier performance. Hence, data cleaning is a crucial part of the data-processing cycle. However, unstructured text with noise is common in automatic environments, such as speech transcription or handwriting recognition [77]. In most NPL tasks, it is difficult to guarantee the absence of noise in the data. In this regard, the work in [77] simulated noise on benchmark datasets to study the effect of different types of noise on automatic text classification. On the other hand, the authors in [78] proposed methods to enhance the robustness of machine translation systems by emulating natural noise in clean data.

In particular, to enable a service robot to understand a motion command even when the commands contain grammatical and word recognition errors due to voice-to-text transcription by automatic speech recognition systems, it is necessary to train the classification model with a dataset that includes transcription errors [79]. In this regard, we evaluated the robustness of the models by introducing typographical errors into the dataset, simulating transcription errors due to Automatic Speech Recognition (ASR) systems [77].

3. Data Collections

We constructed a dataset comprising commands in Nahuatl and Spanish to conduct the experiments. This contains manually labeled phrases related to service robot navigation, for example, the command "carrito muévete lentamente por la calle" in Spanish and "tepostle xi nemi cayolic ipan octle" in Nahuatl, which correspond to the phrase "cart moves slowly down the street" in English. The dataset contains a total of 383 natural language navigation instructions for each language. The dataset includes the following fields: phrase, category, and entity tagging in IOB format (which is commonly used for labeling tokens tasks related to named entity recognition). Figures 1 and 2 show an example of one command in Spanish and Nahuatl in the datasets: "robot turns thirty-five degrees" and "tepoztle xi mollini ipan opochmaitl", respectively. In the figures, the entities and intentions are shown along with the IOB tagging. Particularly, in the IOB format, the symbols B, I, and O indicate the start of a named entity, a token that is inside a named entity, and a token that is not part of any named entity, respectively [80].

Phrase:	robot (<i>robot</i>	gira turns	treinta s <i>thirty-f</i> i	ive degrees	у 5)	cinco	grados
Entity tagging:	Ο	0	B-turn_	quantity	I-turn_quantity	I-turn_quantity	B -turn_unit
Category:	Girar	(Turr	ns)				

Figure 1. Example output for intent classification and slot filling of a natural language instruction in Spanish.

Phrase:	tepozt (<i>robot</i>	le xi turns	moll s <i>left</i>	isiuhca		
Entity tagging:	0	Ο	0	Ο	$B-angle_direction$	$B-velocity_quantity$
Category:	Girar	(Turi	ns)			

Figure 2. Example output for intent classification and slot filling of a natural language instruction in Nahuatl.

To build the dataset, we collected several commands in Spanish during mobile robotics competitions in university contests. These commands were translated manually into Nahuatl language by a Nahuatl speaker (this activity was performed by Amadeo Hernández, the first author of this paper, a native Nahuatl speaker from Hidalgo Huasteca in Mexico).

Figure 3 provides an overview of the labeling process. As observed, each instruction passes through two labeling processes: intents and IOB. The output of this process corresponds to the categorized expressions and the parameters labeled in IOB format.



Figure 3. Representation of the data-labeling process for intent classification and slot-filling tasks. Each instruction in the dataset is independently labeled for each task.

Some dataset statistics are given in Table 1. Since this is a parallel corpus, the Spanish and Nahuatl datasets have an equal number of instances. The dataset partitions were randomly generated, allocating 20% for testing, 15% for validation, and the remaining percentage for training.

Table 1. Some statistics of the generated datasets for Spanish and Nahuatl languages. The categories correspond to the intentions of the queries in the datasets.

Category							
Partition	Forward	Turn	Stop	Backward	Total		
Training	53	84	57	55	249		
Validation	16	20	11	10	57		
Test	17	32	15	13	77		
Total	86	136	83	78	383		

The entities or slots correspond to the quantities and units involved in the instructions of movements such as angles, coordinates, velocity, etc. Table 2 summarizes the entities in the proposed datasets.

Label	Category
angle	Turn angle
coordinate	Coordinates (x, y)
distance	Advance distance
turn	Turn in some direction
time	Move for a certain time
velocity	Move at a certain speed
0	Uncategorized entity

Table 2. Categories of entities used in the examples labeled in IOB format.

4. Model Training and Configuration

In this research, we applied NLU techniques to evaluate the understanding of navigation commands for service robots in low-resource domains. Specifically, we evaluated some NLU-based methods to identify and extract specific information, such as the intention behind the instruction (actions) and entity recognition (parameters necessary to act).

The models are focused on solving the intent classification and slot-filling tasks. We also tested a joint approach that combines components designed to perform both previously

mentioned tasks into a single network structure. The implementation of the models is described in the following subsections.

4.1. Models Setup for Intent Classification and Slot Filling

For the Intent Classification (IC) task, the SVM, biLSTM, biGRU methods, and the pre-trained models BETO [61] and mBERT [81] were tested. Word vectors were created to be used as inputs to the neural networks, including pre-trained word vectors from fastText [82] and Word2vec [83]. The softmax function was used in the final output layer to obtain the probability of the intent.

For the slot-filling (SF) task, several methods were tested, including biLSTM, biL-STM+CRF [84], biGRU, and the pre-trained models BETO and mBERT. The input to the networks consisted of word vectors created from the training dataset, and pre-trained word vectors from fastText and Word2vec were also applied. The output consists of one label for each word in IOB format [80]. The training of the networks is performed in the same way as in intent classification, using a subset of data from the Spanish and Nahuatl corpus. The models described were configured as follows:

- biLSTM and biGRU: 112 hidden layers, 64 units, dropout layer of 0.2, batch size of 64, and an Adam optimizer.
- BETO and mBERT: epochs 3, learning rate of 1×10^{-4} , and batch size of 16.

In the biLSTM network, pre-trained Word2vec (https://code.google.com/archive/p/ word2vec/, accessed on 18 October 2023) and fastText (https://fasttext.cc/, accessed on 9 October 2023) word vectors of 300 dimensions were applied, which were obtained using the Skip-gram model described in [42] with default parameters. The biLSTM network is augmented with the CRF layer to decode the best tag sequence for the input sentences [84].

4.2. A Joint Model for Intent Classification and Slot Filling

To consider the interaction between the two tasks, intent classification and slot filling, we implemented the following deep models:

- CNN-biLSTM. This joint model is based on a CNN-biLSTM architecture [31]. The model integrates a Convolutional Neural Network (CNN) configured with 240 filters, 64 hidden layers and dropout layer of 0.2 with a biLSTM to identify intents and extract the associated slots. The initial layer in the architecture of a model is a word embedding layer. This layer feeds to a CNN layer as a feature extractor. This CNN layer is succeeded by global max pooling and, subsequently, connected to a biLSTM module to capture contextual information of the sequences. The output from the LSTM feeds two branches. One branch generates target predictions by a TimeDistributed module, while the other is a fully connected layer focused on classification intentions.
- biLSTM+CRF. The first layer is the word representation through embeddings, and this layer is used as input to the biLSTM network. This network processes sequences in two directions, allowing it to capture both the preceding and following context of each word. The output of the biLSTM network is used as input to the CRF layer, which models sequential dependencies among entity label tags while considering the correlations between the tags [52].
- Joint model using BERT (joint-Bert). BERT is a language model based on the transformer architecture and has excelled in natural language processing (NLP) tasks. BERT takes the entire input into account, allowing it to understand queries better. The joint model simplifies intent classification (IC) and slot-filling (SF) tasks, as only one model needs to be trained and fine-tuned for both tasks [34]. The input to the model consists of sentences with annotations for intentions and slots, which are then converted into tokenized representations, and attention masks are generated. A pre-trained BERT model is used to fine-tune the model on the training data. The model's output sequences are used for intent classification and slot filling.
- Joint model using BETO (joint-BETO). This model follows the same architecture as the previous one; we only changed the pre-trained model by BETO.

5. Experiments and Results

In this section, we present a series of experiments to classify natural language expressions in Spanish and Nahuatl into commands with parameters that the service robot can execute to reach the movement specified by the user. To accomplish this objective, we address two tasks, IC and SF, performed independently (experiments 1 and 2) and using a joint architecture (experiment 3). Table A1 in Appendix A presents a summary of the experimental setup. It outlines the objective, language, task, and models evaluated for each experiment.

Figure 4 presents the general workflow for conducting experiments. During the training phase, we explored various architectures aimed at training a classifier. Specifically, we assessed deep neural networks (e.g., BiLSTM and BiGRU) and pre-trained BERT-based models through fine-tuning. In the prediction phase, the trained classifier assigns labels to new unlabeled utterances. While the diagram outlines the pipeline for both tasks, they are developed independently (with the exception of Experiment 3). Therefore, inputs (labeled utterances with intents or annotations in IOB format) and outputs (intents or IOB tags) depend on the task that is being evaluated.



Figure 4. Overview of the experimental framework. This is a general methodology that we followed to conduct the experiments for intent classification and slot-filling tasks.

5.1. Experimental Settings

5.1.1. Pre-Processing

For the experiments, the text was converted to lowercase. Characters that interfere with the sentence's meaning, such as stop words, special characters, spaces, and punctuation marks, were removed.

5.1.2. Baseline Methods

We considered a baseline method for each task: intent classification and slot filling. For the former, a traditional Bag of Words (BOW) was trained using an SVM as the classifier. It was configured with a linear kernel and a regularization parameter of two. On the other hand, a biLSTM model was used as a baseline method for the entity detection task. It was configured with 112 hidden states, 64 units, a batch size of 64, and the Adam optimizer.

5.1.3. Evaluation

We mainly reported accuracy and macro F1-score measures, which are widely used to assess the models' performance in the intent classification and slot-filling tasks.

5.2. Experiment 1: Intent Classification

This experiment is aimed at evaluating the intent classification task for detecting the intention of movement instructions for a service robot in Spanish and Nahuatl scenarios. To achieve it, we evaluated methods based on deep neural networks such as a biLSTM, a biGRU, and transformers (BETO and mBERT). Table 3 shows the performance obtained. The results obtained by the baseline approach are also shown for comparison purposes.

Table 3. Experimental results in intent classification. Several deep neuronal architectures were evaluated in Spanish and Nahuatl scenarios.

Method —	Spa	nish	Nah	uatl
	Accuracy	F1-Score	Accuracy	F1-Score
SVM (baseline)	0.9091	0.8912	0.8571	0.8338
biLSTM	0.9351	0.9158	0.9091	0.8933
biGRU	0.9481	0.9350	0.9214	0.9067
mBERT	0.9740	0.9610	0.8831	0.8597
BETO	0.9870	0.9724	-	-

The pre-trained model BETO obtained the best performance in Spanish, while the biGRU-based model performed better for Nahuatl. Compared to the baseline in Spanish, BETO surpassed it by 7.79% (accuracy), while in Nahuatl, the biGRU network achieved a absolute difference of 6.43% with the baseline. These results suggest that transformer-based methods have a wide understanding of the contexts in this task. In general, we observed that the methods based on transformers outperformed the baseline in classifying long utterances, compound sentences, or those containing compound words. For example, for the Nahuatl case, we observed that the baseline found difficulties in classifying the expression "quetzilnenemi", which means "to walk slowly". It should have been classified as "Forward"; however, it was categorized as "backward". This phrase is composed of "quetzi", which means slowly, and "nenemi", which means to walk.

It is important to point out that, to date, we have not found a transformer specifically trained for text classification in Nahuatl.

5.3. Experiment 2: Slot Filling

This experiment aims to evaluate deep neural architectures to detect action parameters in Spanish and Nahuatl commands (i.e., the slot-filling task). We evaluated the models biLSTM, biLSTM+CRF, biGRU, and the pre-trained transformers BETO and mBERT. The results are shown in Table 4. It is possible to note that the biGRU network achieves the best performance in Nahuatl scenarios, achieving values of F1-score of 89.92%.

Method —	Spa	nish	Nah	uatl
	Accuracy	F1-Score	Accuracy	F1-Score
biLSTM (baseline)	0.9598	0.8905	0.9602	0.8550
biLSTM+CRF	0.9350	0.9090	0.9221	0.8701
biGRU	0.9724	0.9185	0.9735	0.8992
mBERT	0.9872	0.9256	0.9518	0.8857
BETO	0.9910	0.9696	_	_

Table 4. Results of the slot-filling task. Different deep neural architectures were evaluated for theSpanish and Nahuatl datasets.

In general, the transformers-based models showed superior performance in Spanish, while the biGRU network obtained the best results for the Nahuatl language. We noted that in Spanish, the pre-trained model BETO achieved a performance of 96.96%, obtaining a difference of 4.4% and 7.91% over the mBERT model and the baseline, respectively. In contrast, for Nahuatl, the biGRU model outperformed the baseline by 4.42%. The results suggest that while using transformers is appropriate, important outcomes are achieved when they are pre-trained, specifically in the language under consideration. Additionally, the findings showed that recurrent neural architectures offer an adequate option for predicting parameters in movement instructions.

5.4. Experiment 3: Intent Classification and Slot Filling as a Joint Task

The purpose of this experiment is to evaluate the joint architecture that combines the components of intent classification and slot filling into a single network structure. The main idea is to try to improve the performance of each task by leveraging the information shared by these components. For this, we applied a combination of a CNN and a biLSTM (CNN+biLSTM), a biLSTM with a CRF layer (biLSTM+CRF), BETO, and mBERT. Table 5 shows the results.

Method –	Spa	nish	Nah	uatl
	IC (Accuracy)	SF (F1-Score)	IC (Accuracy)	SF (F1-Score)
CNN+biLSTM	0.9351	0.8337	0.8442	0.8153
biLSTM+CRF	0.9480	0.9185	0.9220	0.9022
joint mBERT	0.9649	0.9480	0.8772	0.7832
joint BETO	0.9825	0.9310	-	-

Table 5. Experimental results obtained when joint models were used to tackle both tasks (IC and SF).

Despite training the architecture using information from both tasks, the performance is reported separately. For the intent classification task in Spanish, the best performance was achieved by the BETO, with an accuracy of 98.25%. However, mBERT also showed a good performance of 96.49%, with a difference between both models of only 1.76%. This indicates that both models are feasible for application in the Spanish IC task. On the other hand, in Nahuatl scenarios, the biLSTM+CRF model achieved the highest performance, 92.20% and 90.22% for the IC and SF tasks, respectively. In general, we observed that, for this research, the joint models often do not have significant advantages over their individual counterparts.

6. Analysis

In this section, we deepen this study by analyzing the performance of the evaluated methods. Specifically, we first focused on exploring the impact of using different pretrained word vectors to feed the neural networks. Later, we evaluated the robustness of the methods when they were trained on noisy data, simulating real scenarios of speech recognition systems. The last rows of Table A1 in Appendix A present a summary of this section.

6.1. Comparison of Word Embedding Models for IC and SF

Recognizing the influence of embeddings on the performance of models, we investigated the utilization of different word embedding models in this section. Specifically, pre-trained embeddings from fastText (we used embeddings for Spanish from https://dl.fbaipublicfiles.com/fasttext/vectors-wiki/wiki.es.vec, accessed on 22 November 2023 and for Nahuatl https://dl.fbaipublicfiles.com/fasttext/vectors-wiki/wiki.na h.vec, accessed on 22 November 2023) and Word2vec (Embeddings for Spanish from http://vectors.nlpl.eu/repository/20/68.zip, accessed on 6 December 2023 and for Nahuatl https://sparknlp.org/2022/03/16/w2v_cc_300d_nah_3_0.html, accessed on 6 December 2023) were compared against embeddings generated for the problem at hand by using an embedding layer in the networks. The evaluation was conducted using the models based on biLSTM and biGRU neural networks (which are described in Section 4.1). Figure 5 shows the comparison results.



Figure 5. Comparison of the use of different word embeddings for intent classification and slot filling. The pre-trained embeddings were selected according to the language of the datasets. The embedding models were evaluated by feeding neural networks (biLSTM and biGRU). (**a**) Spanish Intent classification; (**b**) Nahuatl Intent classification; (**c**) Spanish Slot filling; (**d**) Nahuatl Slot filling.

In general, we observed that the neural models benefited from fastText embeddings in Spanish scenarios in both tasks. The pre-trained embeddings in this language covered more than 90% of the vocabulary. Conversely, for Nahuatl, models relying on an embedding layer outperformed those based on pre-trained word vectors. We attributed this performance difference to the limited coverage of pre-trained embeddings for the vocabulary of the dataset for this indigenous language, around 30% (coverage) for both Word2vec and fastText. Hence, the generated corpus-specific word embeddings obtained important improvements with respect to the other models. These findings suggest that pre-trained embeddings are effective for languages with ample coverage, but their applicability should be examined when dealing with indigenous languages, especially those with different variants within a family, such as Nahuatl.

6.2. Zero-Shot and Few-Shot Cross-Lingual for Nahuatl Scenarios

Nahuatl is most likely unseen by commonly used pre-trained models like mBERT. Hence, specific strategies are required to leverage the power of large language models to work effectively with this language and face the lack of annotated data in Nahuatl. Motivated by research demonstrating the cross-lingual transfer abilities of multilingual language models, we explored the use of mBERT to classify intents in Nahuatl in a cross-lingual setting. The main ideas in this experiment are as follows: (1) evaluating if the model can generalize well in a zero-shot cross-lingual setting, where no data from the target language are used in the fine-tuning process, and (2) investigating the performance of mBERT in cross-lingual few-shot fine-tuning. We adapted an experimental framework similar to previous research works [85,86]. That is, we defined Spanish and Nahuatl as source and target languages, respectively. For the zero-shot cross-lingual evaluation, only data from the Spanish corpus were used in the fine-tuning stage. Finally, we applied few-shot learning by adding the Nahuatl data (training partition) to the training set. These cross-lingual experiments were conducted using the pre-trained mBERT Model. The results are reported in Table 6.

Table 6. Results of the zero-shot and few-shot cross-lingual experiments for classifying intentionsin Nahuatl.

Settings	Accuracy	F1-Score
zero-shot	0.2597	0.2029
few-shot	0.9350	0.9221

The results show low performance in the cross-lingual zero-shot scenario. Since Nahuatl and Spanish are languages belonging to distinct families, this behavior is expected and consistent with several studies, indicating that the similarity of source-target languages is an important factor in obtaining the benefits of the transfer [17,59,87]. On the contrary, when Nahuatl instances were added in the training (few-shot scenarios), the results outperformed the counterpart (zero-shot) and even the performance of the models described in previous experiments (around 1%). It is evident that the introduction of Nahuatl instances in the training helps to generalize new patterns. In this case, we considered that cultural connections, some structural similarities between languages [66], the infiltration of Spanish into Modern Nahuatl, and the use of subword information by mBERT [88] are leveraged to transfer between models and to augment their generalization capabilities, which lead to better performance. In general, the results indicate that few-shot learning is an important strategy for dealing with this downstream task using the Nahuatl language.

Inspired by recent studies [65], we also analyzed the performance in k-shot crosslingual classification to explore the impact of varying sizes of training data. Therefore, *k* labeled samples of the target language per class were added to the training. We chose $k \in \{3, 5, 10\}$ to evaluate how the model generalizes to a new target language with limited labeled data per class. Since the performance may be influenced by the selection of *k* examples chosen for training, we present the average performance and standard deviation based on five distinct sample selections for each *k*. The idea is to evaluate if the training supporting other languages helps to reach better results in the target language.

Table 7 shows the results. We obtained insight into the size of the dataset of additional instances in the target language to enhance the classification process. The more instances in the target language added, the better the performance achieved. Hence, in the previous Table 6, the few-shot scenario performs best in this series of experiments utilizing all available instances belonging to the training partition from the Nahuatl dataset.

k	Accuracy	F1-Score
3	0.5350 ± 0.0372	0.4863 ± 0.0509
5	0.5532 ± 0.0270	0.5144 ± 0.0529
10	0.6961 ± 0.0491	0.6643 ± 0.0654

Table 7. Results of distinct sample selections in few-shot cross-lingual experiments for classifying intentions in Nahuatl.

6.3. Analysis of the Robustness against Noise

Considering the future automatic application of the proposed method, expressions could contain noise from automatic or manual transcription systems. Therefore, studying the effect of noisy datasets on the performance of the models used in experiments is important. In this experiment, we analyzed the robustness of the models against noise. For simulating transcription errors, we randomly introduced five common types of typographical errors at the character level in the datasets: insertion, deletion, substitution, duplication, and transposition of characters. For example, the word "lento" in Spanish ("slow" in English) could be altered according to the previous type of errors mentioned as "lentoo", "lente", "lentoo" or "lneto". Then, we applied a random error to *i*% of words in the sentence of each instance from the test dataset. We evaluated $i \in \{0, 10, 20, 30, 40, 50\}$. The evaluated models are those that achieved the best results in previous experiments for the Spanish and Nahuatl languages (Section 2.3).

Figure 6 reports the performance of the models in this experiment for the IC and SF tasks. In the former task, subtle differences in performance were observed in the presence of noise; however, as the number of errors increased, especially when the percentage was 50%, the advantage of BETO over the biGRU model became more noticeable. In the slot-filling task, the performance differences were more evident.



Figure 6. Results obtained when typographical errors were added to the Spanish dataset. The biGRU and mBERT models were evaluated in (**a**) intent classification and (**b**) slot-filling tasks.

Similar results were found in the Nahuatl scenarios in Figure 7. In general, the pretrained mBERT model exhibited remarkable performance stability in the presence of errors compared to the biGRU model. These results confirm the robustness of the large pre-trained models for tackling tasks on low-resource domains and their ability to handle situations with a higher degree of interference.



Figure 7. Results obtained when typographical errors were added to the Nahuatl dataset. The biGRU and mBERT models were evaluated in (**a**) intent classification and (**b**) slot-filling tasks.

7. Discussion

This research showed the feasibility of guiding robot movements using natural languages in scenarios with limited linguistic resources, such as those defined by commands expressed in Spanish and Nahuatl. The results revealed a superior performance of transformer-based models in this specific task compared to traditional deep neural networks such as GRU and BiLSTM. These models also demonstrated robustness against introduced errors, simulating their future automatic application.

Particularly, for scenarios in Nahuatl, the experiments showed that pre-trained embeddings are useful, but the varieties of Nahuatl make the task more challenging. In addition, we found that the absence of a large pre-trained language model poses a major obstacle. In response, this work proposes exploring alternative strategies, such as cross-lingual few-shot learning, where we observed that the results are mainly impacted by the selection of the instances added to the training dataset.

We hope that this work motivates the development of methods for integrating technology across indigenous languages. By prioritizing the development of such resources, we can help indigenous communities to benefit from technological advancements and preserve their languages. In this context, we have pointed out the importance of building pre-trained large language models in indigenous languages.

While the results were promising, our focus now shifts to assessing the ability of the models to generalize by the use of large datasets in domains defined by indigenous languages. In addition, it is recognized that the linguistic diversity within indigenous languages poses a significant challenge in generating automatic models, highlighting the need for further research in this area.

8. Conclusions

Our presented research evaluated the task of recognizing motion commands emitted in natural language in domains with limited linguistic resources. Specifically, we addressed the problem of identifying commands and their parameters for guiding the movements of service robots in the Spanish and Nahuatl languages. To achieve this goal, we implemented natural language understanding approaches by handling two typical tasks: intent classification and slot-filling detection (i.e., entity detection). As part of the methodology, a parallel dataset, which comprises motion instructions in Spanish and Nahuatl, was manually constructed, resulting in an additional contribution. Finally, we explored the capabilities of multilingual Bert for zero and few-shot learning for classifying movement utterances in Nahuatl.

To conduct the study, we implemented deep neural networks and Transformer-based models. Overall, we observed that directing the movements of a robot is feasible despite the limited availability of training data. The pre-trained language models, such as BETO and mBERT, stood out for their robustness in handling these tasks (intention classification

and slot filling) using the Spanish and Nahuatl languages. In addition, we noted that deep learning models fed by word embeddings performed better in Spanish. Particularly, for Nahuatl, it was important to generate personalized word vectors for the variant language at hand.

Regarding intent recognition in Nahuatl scenarios, we observed that cross-lingual few-shot learning strategies achieved good performance (using Spanish and Nahuatl as source and target languages, respectively), enhancing the process when the number of Nahuatl examples incorporated during the training increases, despite these languages being topologically different. This study aligns with a strategy to achieve access to technological advancements for indigenous communities. We hope that this study motivates further research into the field of automatic manipulation of service robots in indigenous languages and different low-resource languages.

For future work, we plan to augment the dataset with additional instances in Spanish and Nahuatl, aiming to enhance the generalization capability of the models. We are also interested in building other datasets using different indigenous languages, such as Otomi and Tepehua. This research will be directed toward classifying new instructions with parameters for service robots. Additionally, we plan to explore novel approaches to parameterize and optimize the evaluated models. Our future research directions will also involve assessing different multilingual transformer models (e.g., XLM-RoBERTa) to assess their suitability for this task. Finally, for cross-lingual transfer scenarios, we will study the selection of transfer languages to enhance the performance in Nahuatl scenarios related to the task.

Author Contributions: Conceptualization, R.M.O.-M., E.V.-T. and A.H.; methodology, R.M.O.-M. and E.V.-T.; software, A.H.; validation, A.H., E.V.-T., C.J.C.-B. and O.P.-C.; formal analysis, C.J.C.-B. and O.P.-C.; investigation, A.H. and R.M.O.-M.; resources, R.M.O.-M. and E.V.-T.; data curation, A.H.; writing—original draft preparation, A.H.; writing—review and editing, R.M.O.-M., C.J.C.-B., O.P.-C. and E.V.-T.; visualization, O.P.-C.; supervision, R.M.O.-M. and E.V.-T.; project administration, R.M.O.-M., E.V.-T. and A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by CONAHCYT under the scholarship CVU-307822.

Data Availability Statement: The data supporting the findings of this study are available upon request to the first author (A.H.).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Summary of Experiments

In this research, we conducted a series of experiments, which are presented in Sections 5 and 6. Table A1 provides a summary of the characteristics of each of them.

Table A1. Experimental setup for the intent classification (IC) and slot-filling (SF) tasks. The experiments were conducted using utterances expressed in Spanish (Sp) and Nahuatl (Na) languages.

Section	Purpose	Language	Task	Embeddings (emb.)	Evaluated Models
Section 5.2. Exp. 1	Evaluating the intent classifica- tion task for detecting the inten- tion of movement instructions	Sp and Na	IC	Generated, BERT-based	biLSTM, biGRU mBERT, BETO
Section 5.3. Exp. 2	Evaluating deep neural architec- tures to detect action parameters in commands	Sp and Na	SF	Generated, BERT-based	biLSTM, biLSTM+CRF biGRU mBERT, BETO

Section	Purpose	Language	Task	Embeddings (emb.)	Evaluated Models
Section 5.3. Exp. 3	Evaluating the joint architecture that combines the components of intent classification and slot fill- ing into a single network struc- ture	Sp and Na	IC and SF	Generated, BERT-based	CNN+biLSTM, biLSTM+CRF joint mBERT, joint BETO
Section 6.1 Use of different word embeddings models	Investigating the impact of the use of different pre-trained word embeddings models	Sp and Na	IC and SF	Pre-trained (fastText and Word2vec)	biLSTM, biGRU
Section 6.2 Zero-shot and few-shot cross-lingual	Exploring a cross-lingual setting to classify intents in Nahuatl	Na	IC	BERT-based	mBERT
Section 6.3 Robustness against noise	Exploring the robustness in noisy conditions that simulate real scenarios of speech recogni- tion systems	Sp and Na	IC and SF	Generated, BERT-based	biGRU mBERT, BETO

Table A1. Cont.

References

- Kahuttanaseth, W.; Dressler, A.; Netramai, C. Commanding mobile robot movement based on natural language processing with RNN encoderdecoder. In Proceedings of the 2018 5th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand, 17–18 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 161–166.
- Pramanick, P.; Barua, H.B.; Sarkar, C. DeComplex: Task planning from complex natural instructions by a collocating robot. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 6894–6901.
- Joseph, E. From Virtual to Real: A Framework for Verbal Interaction with Robots. In Proceedings of the Combined Workshop on Spatial Language Understanding (SpLU) and Grounded Communication for Robotics (RoboNLP), Minneapolis, MN, USA, 6 June 2019; pp. 18–28. [CrossRef]
- 4. Qiu, X.; Sun, T.; Xu, Y.; Shao, Y.; Dai, N.; Huang, X. Pre-trained models for natural language processing: A survey. *Sci. China Technol. Sci.* 2020, *63*, 1872–1897. [CrossRef]
- 5. Mensio, M.; Bastianelli, E.; Tiddi, I.; Rizzo, G. A Multi-layer LSTM-based Approach for Robot Command Interaction Modeling. *arXiv* **2018**, arXiv:1811.05242.
- Walker, N.; Peng, Y.T.; Cakmak, M. Neural Semantic Parsing with Anonymization for Command Understanding in General-Purpose Service Robots. In *RoboCup 2019: Robot World Cup XXIII*; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 337–350. [CrossRef]
- Neuraz, A.; Campillos-Llanos, L.; Burgun-Parenthoine, A.; Rosset, S. Natural language understanding for task oriented dialog in the biomedical domain in a low resources context. *arXiv* 2018, arXiv:1811.09417.
- 8. Magueresse, A.; Carles, V.; Heetderks, E. Low-resource Languages: A Review of Past Work and Future Challenges. *arXiv* 2020, arXiv:2006.07264.
- Littell, P.; Kazantseva, A.; Kuhn, R.; Pine, A.; Arppe, A.; Cox, C.; Junker, M.O. Indigenous language technologies in Canada: Assessment, challenges, and successes. In Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–26 August 2018; pp. 2620–2632.
- Mager, M.; Oncevay, A.; Rios, A.; Ruiz, I.V.M.; Palmer, A.; Neubig, G.; Kann, K. (Eds.) Findings of the AmericasNLP 2021 shared task on open machine translation for indigenous languages of the Americas. In Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas, Online, 11 June 2021.
- Louvan, S.; Magnini, B. Recent Neural Methods on Slot Filling and Intent Classification for Task-Oriented Dialogue Systems: A Survey. In Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain (Online), 8–13 December 2020; pp. 480–496. [CrossRef]
- 12. Mastel, P.M.; Namara, E.; Munezero, A.; Kagame, R.; Wang, Z.; Anzagira, A.; Gupta, A.; Ndibwile, J.D. Natural Language Understanding for African Languages. In Proceedings of the 4th Workshop on African Natural Language Processing, Kigali, Rwanda, 1 May 2023.
- 13. Huan, H.; Guo, Z.; Cai, T.; He, Z. A text classification method based on a convolutional and bidirectional long short-term memory model. *Connect. Sci.* 2022, 34, 2108–2124. [CrossRef]
- 14. Zhu, Q.; Jiang, X.; Ye, R. Sentiment Analysis of Review Text Based on BiGRU-Attention and Hybrid CNN. *IEEE Access* 2021, 9, 149077–149088. [CrossRef]
- 15. Wang, H.; Li, J.; Wu, H.; Hovy, E.; Sun, Y. Pre-Trained Language Models and Their Applications. *Engineering* **2023**, 25, 51–65. [CrossRef]

- 16. Ojha, A.K.; Malykh, V.; Karakanta, A.; Liu, C.H. Findings of the LoResMT 2020 Shared Task on Zero-Shot for Low-Resource languages. In Proceedings of the 3rd Workshop on Technologies for MT of Low Resource Languages, Suzhou, China, 4–7 December 2020; Karakanta, A., Ojha, A.K., Liu, C.H., Abbott, J., Ortega, J., Washington, J., Oco, N., Lakew, S.M., Pirinen, T.A., Malykh, V., et al., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2020; pp. 33–37.
- Lauscher, A.; Ravishankar, V.; Vulić, I.; Glavaš, G. From Zero to Hero: On the Limitations of Zero-Shot Language Transfer with Multilingual Transformers. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Online, 16–20 November 2020; Webber, B., Cohn, T., He, Y., Liu, Y., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2020; pp. 4483–4499. [CrossRef]
- 18. Zhao, M.; Zhu, Y.; Shareghi, E.; Vulić, I.; Reichart, R.; Korhonen, A.; Schütze, H. A Closer Look at Few-Shot Crosslingual Transfer: The Choice of Shots Matters. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Online, 1–6 August 2021; Zong, C., Xia, F., Li, W., Navigli, R., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2021; pp. 5751–5767. [CrossRef]
- Karimi Mahabadi, R.; Zettlemoyer, L.; Henderson, J.; Mathias, L.; Saeidi, M.; Stoyanov, V.; Yazdani, M. Prompt-free and Efficient Few-shot Learning with Language Models. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland, 22–27 May 2022; Muresan, S., Nakov, P., Villavicencio, A., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2022; pp. 3638–3652. [CrossRef]
- 20. INEGI. Population 3 years of age and over who speak the Nahuatl indigenous language by state, according to sex, 2010 and 2020 census years. In *Population and Housing Census*; INEGI: Aguascalientes, Mexico, 2020.
- 21. Gutierrez-Vasques, X.; Mijangos, V. Low-resource bilingual lexicon extraction using graph based word embeddings. *arXiv* 2017, arXiv:1710.02569.
- Thenmozhi, D.; Seshathiri, R.; Revanth, K.; Ruban, B. Robotic simulation using natural language commands. In Proceedings of the 2017 International Conference on Computer, Communication and Signal Processing (ICCCSP), Chennai, India, 10–11 January 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–4.
- Medina-Santiago, A.; Morales-Rosales, L.A.; Hernández-Gracidas, C.A.; Algredo-Badillo, I.; Pano-Azucena, A.D.; Orozco Torres, J.A. Reactive Obstacle–Avoidance Systems for Wheeled Mobile Robots Based on Artificial Intelligence. *Appl. Sci.* 2021, 11, 6468. [CrossRef]
- 24. Molina-Leal, A.; Gómez-Espinosa, A.; Escobedo Cabello, J.A.; Cuan-Urquizo, E.; Cruz-Ramírez, S.R. Trajectory Planning for a Mobile Robot in a Dynamic Environment Using an LSTM Neural Network. *Appl. Sci.* **2021**, *11*, 10689. [CrossRef]
- Bonilla, F.S.; Ugalde, F.R. Automatic translation of spanish natural language commands to control robot comands based on lstm neural network. In Proceedings of the 2019 Third IEEE International Conference on Robotic Computing (IRC), Naples, Italy, 25–27 February 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 125–131.
- Wang, C.; Yang, Q.; Guan, G.; Jiang, C.; Chen, Z. Implicit Action Inference for Natural Language Navigation Instruction Translation. In *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 1851–1859.
- Kuo, Y.L.; Katz, B.; Barbu, A. Deep compositional robotic planners that follow natural language commands. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 4906–4912.
- 28. Tur, G.; Hakkani-Tür, D.; Heck, L. What is left to be understood in ATIS? In Proceedings of the 2010 IEEE Spoken Language Technology Workshop, Berkeley, CA, USA, 12–15 December 2010; pp. 19–24. [CrossRef]
- Xu, W.; Haider, B.; Mansour, S. End-to-End Slot Alignment and Recognition for Cross-Lingual NLU. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Online, 16–20 November 2020; pp. 5052–5063. [CrossRef]
- Tjong Kim Sang, E.F.; De Meulder, F. Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition. In Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003, Edmonton, AB, Canada, 31 May–1 June 2003; pp. 142–147.
- 31. Weld, H.; Huang, X.; Long, S.; Poon, J.; Han, S.C. A Survey of Joint Intent Detection and Slot Filling Models in Natural Language Understanding. *ACM Comput. Surv.* 2022, *55*, 1–38. [CrossRef]
- Chen, Y.; Luo, Z. Pre-Trained Joint Model for Intent Classification and Slot Filling with Semantic Feature Fusion. Sensors 2023, 23, 2848. [CrossRef] [PubMed]
- 33. Ni, P.; Li, Y.; Li, G.; Chang, V. Natural language understanding approaches based on joint task of intent detection and slot filling for IoT voice interaction. *Neural Comput. Appl.* **2020**, *32*, 16149–16166. [CrossRef]
- 34. Chen, Q.; Zhuo, Z.; Wang, W. BERT for Joint Intent Classification and Slot Filling. arXiv 2019, arXiv:1902.10909.
- Tellex, S.; Kollar, T.; Dickerson, S.; Walter, M.; Banerjee, A.; Teller, S.; Roy, N. Understanding natural language commands for robotic navigation and mobile manipulation. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 7–11 August 2011; Volume 25, pp. 1507–1514.
- 36. Martins, P.H.; Custódio, L.; Ventura, R. A deep learning approach for understanding natural language commands for mobile service robots. *arXiv* **2018**, arXiv:1807.03053.

- Kramer, E.R.; Sáinz, A.O.; Mitrevski, A.; Plöger, P.G. Tell Your Robot What to Do: Evaluation of Natural Language Models for Robot Command Processing. In *Proceedings of the RoboCup 2019: Robot World Cup XXIII, Sydney, NSW, Australia, 8 July 2019;* Chalup, S.K., Niemüller, T., Suthakorn, J., Williams, M.A., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2019; Volume 11531, pp. 255–267. [CrossRef]
- Larson, S.; Leach, K. A Survey of Intent Classification and Slot-Filling Datasets for Task-Oriented Dialog. arXiv 2022, arXiv:2207.13211.
- Grishman, R.; Sundheim, B. Message Understanding Conference- 6: A Brief History. In Proceedings of the COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics, Copenhagen, Denmark, 5–9 August 1996.
- 40. Fathi, E.; Maleki Shoja, B. Chapter 9—Deep Neural Networks for Natural Language Processing. In *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications*; Gudivada, V.N., Rao, C., Eds.; Handbook of Statistics; Elsevier: Amsterdam, The Netherlands, 2018; Volume 38, pp. 229–316. [CrossRef]
- Mikolov, T.; Yih, W.T.; Zweig, G. Linguistic Regularities in Continuous Space Word Representations. In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, Georgia, 9–14 June 2013; pp. 746–751.
- Bojanowski, P.; Grave, E.; Joulin, A.; Mikolov, T. Enriching Word Vectors with Subword Information. *Trans. Assoc. Comput. Linguist.* 2017, 5, 135–146. [CrossRef]
- 43. Wadud, M.A.H.; Mridha, M.F.; Rahman, M.M. Word Embedding Methods for Word Representation in Deep Learning for Natural Language Processing. *Iraqi J. Sci.* 2022, *63*, 1349–1361. [CrossRef]
- Adoma, A.F.; Henry, N.M.; Chen, W. Comparative Analyses of Bert, Roberta, Distilbert, and Xlnet for Text-Based Emotion Recognition. In Proceedings of the 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 18–20 December 2020; pp. 117–121.
- 45. Ogunleye, B.; Dharmaraj, B. The Use of a Large Language Model for Cyberbullying Detection. *Analytics* **2023**, *2*, 694–707. [CrossRef]
- Aydoğan, M.; Karci, A. Improving the accuracy using pre-trained word embeddings on deep neural networks for Turkish text classification. *Phys. A Stat. Mech. Its Appl.* 2020, 541, 123288. [CrossRef]
- Plank, B.; Søgaard, A.; Goldberg, Y. Multilingual Part-of-Speech Tagging with Bidirectional Long Short-Term Memory Models and Auxiliary Loss. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Berlin, Germany, 7–12 August 2016; pp. 412–418. [CrossRef]
- 48. Yalova, K.; Yashyna, K.; Ivanochko, I. Neural approach for named entity recognition. CEUR Workshop Proc. 2021, 2853, 118–126.
- 49. Lopez, M.M.; Kalita, J. Deep Learning applied to NLP. arXiv 2017, arXiv:1703.03091.
- 50. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv* 2014, arXiv:1412.3555.
- 51. Wei, T.; Qi, J.; He, S.; Sun, S. Masked Conditional Random Fields for Sequence Labeling. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 6–11 June 2021; Toutanova, K., Rumshisky, A., Zettlemoyer, L., Hakkani-Tur, D., Beltagy, I., Bethard, S., Cotterell, R., Chakraborty, T., Zhou, Y., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2021; pp. 2024–2035. [CrossRef]
- 52. Huang, Z.; Xu, W.; Yu, K. Bidirectional LSTM-CRF Models for Sequence Tagging. *arXiv* 2015, arXiv:1508.01991.
- 53. Kalyan, K.S.; Rajasekharan, A.; Sangeetha, S. AMMUS: A Survey of Transformer-based Pretrained Models in Natural Language Processing. *arXiv* **2021**, arXiv:2108.05542.
- 54. Kotei, E.; Thirunavukarasu, R. A Systematic Review of Transformer-Based Pre-Trained Language Models through Self-Supervised Learning. *Information* **2023**, *14*, 187. [CrossRef]
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is All You Need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 4–9 December 2017; NIPS'17; pp. 6000–6010.
- 56. Islam, S.; Elmekki, H.; Elsebai, A.; Bentahar, J.; Drawel, N.; Rjoub, G.; Pedrycz, W. A comprehensive survey on applications of transformers for deep learning tasks. *Expert Syst. Appl.* **2024**, 241, 122666. [CrossRef]
- Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, MN, USA, 2–7 June 2019; pp. 4171–4186. [CrossRef]
- Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv 2019, arXiv:1907.11692.
- 59. Pires, T.; Schlinger, E.; Garrette, D. How Multilingual is Multilingual BERT? In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019; Korhonen, A., Traum, D., Màrquez, L., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2019; pp. 4996–5001. [CrossRef]
- Rajaee, S.; Pilehvar, M.T. An Isotropy Analysis in the Multilingual BERT Embedding Space. In Proceedings of the Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, 22–27 May 2022; Muresan, S., Nakov, P., Villavicencio, A., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2022; pp. 1309–1316. [CrossRef]

- 61. Cañete, J.; Chaperon, G.; Fuentes, R.; Ho, J.H.; Kang, H.; Pérez, J. Spanish Pre-Trained BERT Model and Evaluation Data. In Proceedings of the PML4DC at ICLR 2020, Addis Ababa, Ethiopia, 26–30 April 2020.
- Fernández-Martínez, F.; Luna-Jiménez, C.; Kleinlein, R.; Griol, D.; Callejas, Z.; Montero, J.M. Fine-Tuning BERT Models for Intent Recognition Using a Frequency Cut-Off Strategy for Domain-Specific Vocabulary Extension. *Appl. Sci.* 2022, 12, 1610. [CrossRef]
- Bucker, A.F.C.; Figueredo, L.F.C.; Haddadin, S.; Kapoor, A.; Ma, S.; Bonatti, R. Reshaping Robot Trajectories Using Natural Language Commands: A Study of Multi-Modal Data Alignment Using Transformers. In Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 23–27 October 2022; pp. 978–984.
- 64. Buyuk, O. Joint intent detection and slot filling for Turkish natural language understanding. *Turk. J. Electr. Eng. Comput. Sci.* **2023**, *31*, 844–859. [CrossRef]
- 65. Mozafari, M.; Farahbakhsh, R.; Crespi, N. Cross-Lingual Few-Shot Hate Speech and Offensive Language Detection Using Meta Learning. *IEEE Access* 2022, *10*, 14880–14896. [CrossRef]
- 66. Wang, Z.; Mayhew, S.; Roth, D. Cross-Lingual Ability of Multilingual BERT: An Empirical Study. In Proceedings of the 8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, 26–30 April 2020.
- Huang, K.H.; Ahmad, W.; Peng, N.; Chang, K.W. Improving Zero-Shot Cross-Lingual Transfer Learning via Robust Training. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Online, Punta Cana, Dominican Republic, 7–11 November 2021; Moens, M.F., Huang, X., Specia, L., Yih, S.W.T., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2021; pp. 1684–1697. [CrossRef]
- 68. Winata, G.; Wu, S.; Kulkarni, M.; Solorio, T.; Preotiuc-Pietro, D. Cross-lingual Few-Shot Learning on Unseen Languages. In Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Online, 20–23 November 2022; He, Y., Ji, H., Li, S., Liu, Y., Chang, C.H., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2022; pp. 777–791.
- Pelicon, A.; Pranjić, M.; Miljković, D.; Škrlj, B.; Pollak, S. Zero-Shot Learning for Cross-Lingual News Sentiment Classification. *Appl. Sci.* 2020, 10, 5993. [CrossRef]
- 70. Barrena, A.; Soroa, A.; Agirre, E. Towards zero-shot cross-lingual named entity disambiguation. *Expert Syst. Appl.* 2021, 184, 115542. [CrossRef]
- 71. Aepli, N.; Çöltekin, Ç.; Van Der Goot, R.; Jauhiainen, T.; Kazzaz, M.; Ljubešić, N.; North, K.; Plank, B.; Scherrer, Y.; Zampieri, M. Findings of the VarDial Evaluation Campaign 2023. In Proceedings of the Tenth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2023), Dubrovnik, Croatia, 5 May 2023; Scherrer, Y., Jauhiainen, T., Ljubešić, N., Nakov, P., Tiedemann, J., Zampieri, M., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2023; pp. 251–261. [CrossRef]
- Schumacher, E.; Mayfield, J.; Dredze, M. Cross-Lingual Transfer in Zero-Shot Cross-Language Entity Linking. In Proceedings of the Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, Online, 1–6 August 2021; Zong, C., Xia, F., Li, W., Navigli, R., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2021; pp. 583–595. [CrossRef]
- 73. Winata, G.I.; Madotto, A.; Lin, Z.; Liu, R.; Yosinski, J.; Fung, P. Language Models are Few-shot Multilingual Learners. In Proceedings of the 1st Workshop on Multilingual Representation Learning, Punta Cana, Dominican Republic, 11 November 2021; Ataman, D., Birch, A., Conneau, A., Firat, O., Ruder, S., Sahin, G.G., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2021; pp. 1–15. [CrossRef]
- Wu, S.; Dredze, M. Beto, Bentz, Becas: The Surprising Cross-Lingual Effectiveness of BERT. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; Inui, K., Jiang, J., Ng, V., Wan, X., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2019; pp. 833–844. [CrossRef]
- 75. Wu, T.W.; Su, R.; Juang, B. A Label-Aware BERT Attention Network for Zero-Shot Multi-Intent Detection in Spoken Language Understanding. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Online, Punta Cana, Dominican Republic, 7–11 November 2021; Moens, M.F., Huang, X., Specia, L., Yih, S.W.T., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2021; pp. 4884–4896. [CrossRef]
- Greco, A.; Roberto, A.; Saggese, A.; Vento, M. Efficient Transformers for on-robot Natural Language Understanding. In Proceedings of the 2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids), Ginowan, Japan, 28–30 November 2022; pp. 823–828. [CrossRef]
- Agarwal, S.; Godbole, S.; Punjani, D.; Roy, S. How Much Noise Is Too Much: A Study in Automatic Text Classification. In Proceedings of the Seventh IEEE International Conference on Data Mining (ICDM 2007), Omaha, NE, USA, 28–31 October 2007; pp. 3–12. CrossRef]
- Vaibhav, V.; Singh, S.; Stewart, C.; Neubig, G. Improving Robustness of Machine Translation with Synthetic Noise. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, MN, USA, 2–7 June 2019; pp. 1916–1920. [CrossRef]
- 79. Tada, Y.; Hagiwara, Y.; Tanaka, H.; Taniguchi, T. Robust understanding of robot-directed speech commands using sequence to sequence with noise injection. *Front. Robot. AI* **2020**, *6*, 144. [CrossRef] [PubMed]
- Ramshaw, L.; Marcus, M. Text Chunking using Transformation-Based Learning. In Proceedings of the Third Workshop on Very Large Corpora, Cambridge, MA, USA, 30 June 1995.
- 81. Sanh, V.; Debut, L.; Chaumond, J.; Wolf, T. DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. *arXiv* 2019, arXiv:1910.01108.

- 82. Grave, E.; Bojanowski, P.; Gupta, P.; Joulin, A.; Mikolov, T. Learning Word Vectors for 157 Languages. In Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan, 7–12 May 2018.
- 83. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* 2013, arXiv:1301.3781.
- Haihong, E.; Peiqing, N.; Zhongfu, C.; Meina, S. A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019; pp. 5467–5471. [CrossRef]
- Choi, H.; Kim, J.; Joe, S.; Min, S.; Gwon, Y. Analyzing Zero-shot Cross-lingual Transfer in Supervised NLP Tasks. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 9608–9613. [CrossRef]
- 86. Dürlich, L.; Reimann, S.; Finnveden, G.; Nivre, J.; Stymne, S. Cause and Effect in Governmental Reports: Two Data Sets for Causality Detection in Swedish. In Proceedings of the LREC 2022 workshop on Natural Language Processing for Political Sciences, Marseille, France, 20–25 June 2022; Afli, H., Alam, M., Bouamor, H., Casagran, C.B., Boland, C., Ghannay, S., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2022; pp. 46–55.
- Kim, H.; Komachi, M. Enhancing Few-shot Cross-lingual Transfer with Target Language Peculiar Examples. In Proceedings of the Findings of the Association for Computational Linguistics: ACL 2023, Toronto, ON, Canada, 9–14 July 2023; Rogers, A., Boyd-Graber, J., Okazaki, N., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2023; pp. 747–767. [CrossRef]
- 88. Castillo, N.D. Nahuatl: The influence of Spanish on the Language of the Aztecs. J. Am. Soc. Geolinguist. 2012, 38, 9–23.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.