Processing speech acts: spoken communication for aircraft maintenance

Savvas Chatzipanayiotidis, Christina Alexandris
National and Kapodistrian University of Athens, Greece

https://doi.org/10.36505/ExLing-2023/14/0007/000601

Abstract

The identification of speech act types and speech act content constitutes the basis for the construction and implementation of a flexible template-based, slot-filling framework for spoken technical texts in the domain of aircraft maintenance for its direct deployment or its subsequent integration in neural network approaches and/or as training data. Keywords extracted from the content of each spoken utterance (620 expressions, approximately 2800 spoken utterances) are integrated to the "generic-intent" slot of the template-based, slot-filling framework, which is, in turn, linked to the respective Speech Act. The mandatory recognition of two keywords in complimentary relation (x,y) within the framework is observed to produce correct results for Greek spoken Technical texts.

Keywords: speech acts, spoken technical texts, template-based slot-filling framework, commands, aircraft maintenance, NLU

Introduction

Spoken technical texts pose multiple challenges for their processing and possible translation. In the case of commercial and military aircraft maintenance, correctness, clarity, and precision are of critical importance for securing safety, precision, efficiency, and direct deployment as well as their understandability for non-native speaker and/or translation in different languages. These features and targets are directly linked to the development of Controlled Languages in the aircraft industry, a strict, domain-specific framework considered a traditional strategy for the processing of spoken task-specific utterances and technical texts (Wojcik and Holmback. 1996, Lehrndorfer, 1996). Specifically, typical practices for processing spoken input in both monolingual and multilingual applications involve template-based, slot-filling strategies (Jurafsky and Martin, 2022), from which recent approaches with the use of neural networks are developed (Okur et al., 2023). Here, we focus on the data obtained from Greek spoken technical texts for aircraft maintenance and technical support. The data is presented corresponds to the analysis stage for subsequent processing and integration in Human-Computer Interaction applications involving speech.

The identification of speech act types and speech act content constitutes the basis for the construction and implementation of a flexible template-based, slot-filling framework for multilingual spoken technical texts in the domain of aircraft maintenance for its direct deployment or its subsequent integration in neural network approaches and/or as training data. The presented slot-filling framework, targeted to process English, German and Greek, is designed to process spoken input, taking into account the varying degrees of quality of Speech Recognition (ASR) in many languages as well as the factor of the noisy environment in the domain of aircraft maintenance.

Keyword and speech act recognition approach

Speech act recognition (1), in combination with the recognition of the appropriate word groups contributes both to the speed and to the correctness of processing spoken input. In addition, the present approach also focuses on the minimum possible size of the sets in keyword recognition (2), taking into account the varying degrees of quality of ASR systems. The features (1) and (2) constitute basic elements of the proposed template-based, slot-filling framework which is based on keyword recognition with keywords in the form of tuples (x,y).

In particular, the present approach is based on the mandatory recognition of two keywords in complimentary relation (x,y). According to the evaluation performed on the data presented below, this relation is observed to ensure the correct identification of speech act type as well as the correct identification of content type in the spoken utterances. Furthermore, the recognition of two keywords in complimentary relation accounts for cases in which speech act type and content type coincide or overlap, a possibility in some types of commands and utterances concerning safety regulations and alerts. The keywords extracted from the content of each spoken utterance are integrated to the "generic-intent" slot of the template-based, slot-filling framework. The information content of the "generic-intent" slot is linked to the respective Speech Act.

Data and information management

The keywords integrated into the template-based, slot-filling framework are related to 620 expressions from approximately 2800 spoken utterances extracted from data was collected both from and available datasets of transcribed Greek spoken technical texts and from transcribed spoken dialogs for aircraft maintenance and technical support (Greek). The 620 expressions are mapped into 21 distinct tasks and related Speech Acts. The Speech Acts are divided into three (3) general categories, namely "Command-Question", "Information-Answer" and "Emergency". The "Command-Question" category contains the speech acts – listed in order of frequency - "Command", "Wh-Question", "Y/N-Question", "Requirement-Command" and "Request", as

demonstrated in the example in Figure 1. The "generic-intent" slot of the template-based, slot-filling framework contains information types (classes and subclasses) corresponding to the intention of the Speaker and the respective Illocutionary Speech Act. These are the following: (flights: (flight schedule), (flight tracking)), (maintenance-task: (aircraft inspection), (aircraft relocation), (fueling), (cleaning), (sampling), (testing), (repair), (replacement), (material inspection)), (resources: (aircraft availability), (aircraft location), (personnel for task), (spare part)), (safety measures: (safety measures-request)) (manual-references: (instructions)) and (maintenance-report: (report-request)).

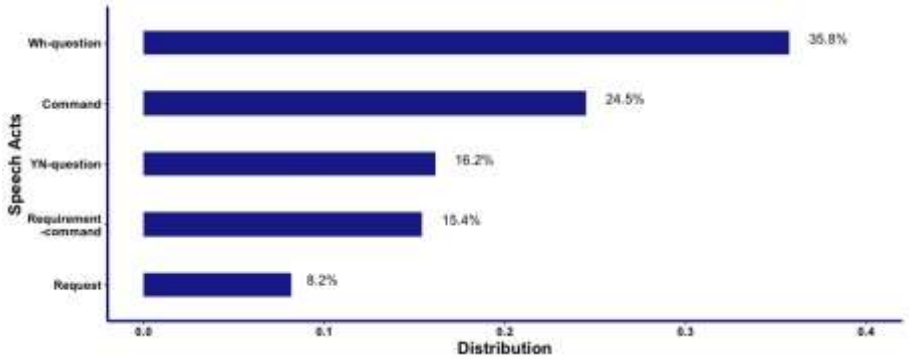

Figure 1. Speech Act types (Command-Question) category: order of frequency.

The "Information-Answer" category contains the speech acts "Positive Report", and "Negative Report" in order of frequency. In this case, the "generic-intent" slot contains only the (maintenance-report: (report-request)) information type. In the "Emergency" category there is only one Speech Act "Declaration", and the "generic-intent" slot contains only the (emergency-declaration) information type. The spoken utterances of this special category are clustered together with the utterances of "Information-Answer" category because they can be generalized as of "Information" type. However, the utterances of "Emergency" type can be encountered in any possible interaction and, hence, interrupt the ongoing dialogue flow and provoke an immediate action. Figure 2 illustrates the distribution of the Speech Acts from these two categories in order of frequency. Most of the respective keywords constituting the actual information type belong to the following general groups: "Wh-Questions", "Quantitative Information", "Other Questions" and "Written Documentation". As stated above, the correct identification of utterance content is based on the mandatory recognition of two keywords in complimentary relation (x,y), corresponding to information type and speech act category. For example, the utterance "Do you know the coordinates of the aircraft?" corresponds to the information in the slots {flights: flight tracking (aircraft)} linked to the speech act "Y/N-Question" - and expected Y/N-Answer". In the spoken Greek utterance, the two keywords in complimentary

relation (x,y) are the verb "(you)know" ("Γνωρίζεις") in sentence-initial "Y/N-Question" position and the term "coordinates" ("συντεταγμένες").

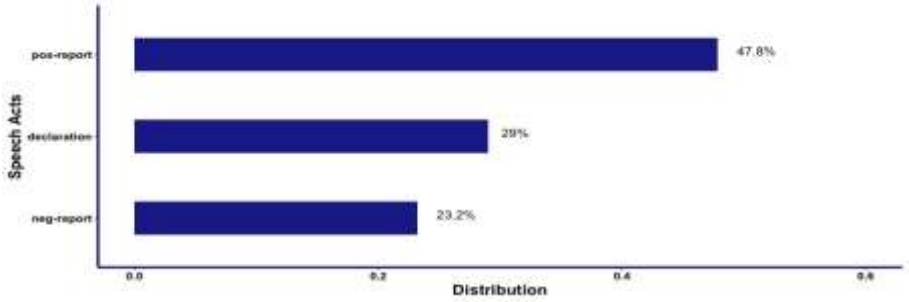

Figure 2. Speech Act types from the "Information-Answer" & "Emergency" categories with order of frequency.

Table 1. Speech acts with keywords in complimentary relation.

Utterance	Intent	Generic Intent	Specific Speech Act	Generic Speech Act Info
Γνωρίζεις τις συντεταγμένες του αεροσκάφους;	Flight tracking	flights	yn-question	YN-ANSWER

Conclusions and further research

The mandatory recognition of two keywords in complimentary relation (x,y) observed to produce correct results for Greek spoken Technical texts is intended to be compared to similar data, if available, in English and German. Data comparison for utterance content and speech act type in technically-related interactions in aircraft maintenance is targeted to be included in further research that may provide insights to possible upgrades in the proposed approach.

References

Jurafsky, D., Martin, J.H. 2022 (3rd ed.). Speech and Language Processing, an Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition.

Lehrndorfer, A. 1996. Kontrolliertes Deutsch: Linguistische und Sprachpsychologische Leitlinien für eine (maschinell) kontrollierte Sprache in der technischen Dokumentation. Tübingen, Germany: Gunter Narr Verlag.

Okur E., Kumar S.H., Sahay, S., et al. 2023. Natural Language Interactions in Autonomous Vehicles: Intent Detection and Slot Filling from Passenger Utterances.

Wojcik R.H., Holmback, H. 1996. Getting a controlled language off the ground at Boeing". In Proceedings of CLAW– 1996, 22–31. Leuven, Belgium.