A Speech-Act Oriented Approach for User-Interactive Editing and Regulation Processes Applied in Written and Spoken Technical Texts

Christina Alexandris

National University of Athens,
Athens, Greece
calexandris@gs.uoa.gr

Abstract. A speech-act oriented approach for Controlled Language specifications is presented for the implementation in a user-interactive HCI system for the editing process and for the regulation of written and, subsequently, spoken technical texts for Modern Greek. Sublanguage-specific and sublanguage independent parameters are used targeting to "Precision", "Directness" and "User-friendliness", based on the criteria of Moeller, 2005 for the success and efficiency of spoken Human-Computer Interaction, on the Utterance Level, the Functional Level and the Satisfaction Level.

Keywords: Controlled Language, Speech Act, Technical texts, Task-oriented dialog, prosodic modeling.

1 Introduction

The present study concerns a speech-act oriented approach for Controlled Language specifications in a user-interactive HCI system for the editing process and for the regulation of written and, subsequently, spoken technical texts. The language concerned is spoken and written Modern Greek, a language with a limited tradition in the production of technical texts and also a language into which a large number of technical texts are translated, the translation process often impacting the quality of the text produced. A basic problem that is addressed is the handling of ambiguity and difficulties in comprehensibility in instructive technical texts, especially when targets such as immediate and successful implementation, as well as security must be achieved.

The use of Controlled Languages has become a common approach for the efficient handling of the syntax and the lexicon of technical texts, with a tradition in languages such as English, German and French. In Controlled Languages, the identification of the functions of sentence types in the technical texts is data-driven and defined by the specific sublanguage of the technical text concerned, often in combination to the related speech act or speech acts expressed in the sentence.

J.A. Jacko (Ed.): Human-Computer Interaction, Part II, HCII 2009, LNCS 5611, pp. 645–653, 2009.

2 The Speech Act – Oriented Approach

The proposed strategy is based on an agent-oriented approach for the analysis of language [4] and targets to the success and efficiency of spoken Human-Computer Interaction, on the Utterance Level, the Functional Level and the Satisfaction Level, according to the criteria of Moeller, 2005.

The criteria on the Utterance Level (Question-Answer-Level) include informativeness, intelligibility and metacommunication handling [7]. The ease of use limits and/or functional limits, initiative and interaction control, processing speed/smoothness are classified as criteria concerning the Functional Level (System Capabilties) [7]. The criteria in respect to the Satisfaction Level include perceived task success, comparability of human partner and trustworthiness [7]. These principles and targets are summarized here as "Precision", "Directness" and "User-friendliness".

2.1 Speech Acts for Technical Texts and Task-Oriented Dialog

Here, the speech-act oriented approach for Controlled Language specifications is based on the (1) Task-oriented Dialog Speech Acts according to Heeman et al., 1998 and on the (2) Speech Acts for Technical Texts according to Lehrndorfer, 1996.

Specifically, the proposed speech-act oriented approach for Controlled Language specifications is based on the Speech Acts "Direction/Command", "Precaution/Safety" and "Description of Product/Device", defined as Speech Acts for Technical Texts [6].

For spoken technical texts constituting task-oriented dialogs, the Speech Acts for Technical Texts are combined with the Speech Acts for Task-oriented Dialogs [5], namely "Yes/No Question", "Request", "Check", "Confirm", "Inform", "Acknowledge" and "Filled-Pause".

2.2 Sublanguage-Specific and Sublanguage-Independent Parameters

The Speech Acts for Technical Texts define the type of action to be executed in respect to the sublanguage-specific and sublanguage-independent parameters of the Controlled Language. In a user-interactive system for editing technical texts with the use of a Controlled Language, the sublanguage-specific and sublanguage-independent parameters may be presented to the user in the form of questions and/or suggestions. In a Task-oriented Dialog system for spoken technical texts, the proposed sublanguage-specific and sublanguage-independent parameters can be used in the design phase of a Dialog System, in a user-interactive editing tool to be applied automatically in the implementation phase.

The scheme in which the General Guidelines can be described is characterized by the differentiation between rules that are (a) technical texts sublanguage-independent and (b) technical texts sublanguage-dependent parameters. An additional differentiation is the differentiation among General Guidelines related to (1) Syntax and (2) General Guidelines related to the Lexicon. In turn, the General Guidelines related to Syntax can be differentiated in (a) sublanguage-independent and (b) sublanguage-dependent rules. Similarly, the General Guidelines related to the Lexicon can be differentiated in (a) sublanguage-independent and (b) sublanguage-dependent rules.

For the language concerned, sublanguage-independent parameters concern sentence-length, order of information presented within a sentence, choice of verbs according to the speech-act indicated and the form of encoding information describing time, space (movement), quality and quantity. Sublanguage-specific parameters for Modern Greek involve the choice of processing or not processing "mixed categories" (for instance, participles), the choice of processing or not processing anaphora and the choice of verbs that are not related speech acts, but to the sublanguage-specific terminology.

3 The Sublanguage-Independent Parameters as a Controlled Language

3.1 Speech Acts for Technical Texts and Sentence Form

In respect to the sublanguage-independent parameters, sentence length is determined by the type of conjunction used for the breaking-down of the sentences into shorter units, to ensure clarity and intelligibility. The function of the conjunctions (for example, temporal, causative etc.) is linked to the Speech Act for Technical Texts related to each conjunction [6], a relation that is observed to be applicable in Modern Greek as well (Table 1). Specifically, it is observed that in Modern Greek the breaking-down process is determined by the conjunctions expressing the Speech Act "Direction/Command" [6]. In contrary, this process is seldom initiated by conjunctions linked to the Speech Act "Description of Product/Device" [6], since this process most often results to alterations in the information content of the original sentence.

Table 1. Relation of Sentence Content, Speech Act and "Breaking-Down" process (Modern Greek)

Sentence Content	Speech Act	Modification
Telicity	Direction/Command	breaking-down
Cause	Direction/Command	breaking-down
	Description of Product/Device	breaking-down
Condition	Direction/Command	breaking-down
	Description of Product/Device	no modification
Manner	Description of Product/Device	no modification
Contradiction	Description of Product/Device	no modification

Specifically, for the Speech Act Direction/Command, the breaking-down process is applied if the tasks to be performed are a serial sequence. For example, in the sentence "Deactivate the device and pull the plug", the sentence will be broken down into the following sentences: "Deactivate the device. Pull the plug". If the tasks are to be executed simultaneously, the sentence initiated by a temporal conjunction is preserved, as in the case of the sentence "While you press the green button, you can speak through the microphone" (translation from Modern Greek, with proximity to original syntactic structure). In the case of temporal overlapping of the tasks to be performed, the broken-down sentences are linked by the "and" conjunction, as in the

example “Hold pieces B-4 and B-4 under the surface and screw them there” (translation from Modern Greek, with proximity to original syntactic structure).

Additionally, the Speech Acts for Technical Texts also determine the order of the elements appearing in the sentences (Table 2). For example, in sentences expressing the Speech Act “Direction/Command” the first element is the verb, as in the case of the utterance “Choose the Microsoft-Word file you wish to open and press the button Open” (translation from Modern Greek, with proximity to original syntactic structure). In this case, the verb is positioned in the beginning of the sentence, regardless of whether the verb constitutes sublanguage-specific terminology or also expresses a speech act. Typical variations of the verbal element include verbs with an infinitive and verbs with a temporal expression. In Speech Acts expressing Precaution/Safety, the first element is the negation, as in the case of the utterance “Don’t push with force”. Typical variations of the negation include negations followed by (or following) a temporal expression and negations following attention markers such as “Attention” and “Caution”. A noun usually appears as the first element in sentences expressing the Speech Act “Description of Product/Device”, as in the example “The small circular plug is located at the back of the unit”.

Table 2. Relation of first element in sentence and Speech Act (Modern Greek)

First element	Variations:	Speech Act
Verb	Verbal expression Verb with infinitive Verb with temporal expression	Direction/Command
Negation	Negation with temporal expression Attention marker with negation	Precaution/Safety
Noun	-	Description of Product/Device

3.2 Speech Acts for Technical Texts and Lexicon

In respect to the lexicon, verbs that express a speech act in Technical Texts are defined as a specific sublanguage-independent group. In the domain of the Technical Texts, these verbs are identified as Directives, related to the Speech Acts “Command/Requirement”, “Prohibition”, “Permission”, “Possibility” and “Recommendation” [6]. The recommended respective verbs for Modern Greek ensuring a one-to-one mapping to the respective Speech Act and discouraging ambiguity are the following: Command/Requirement is mapped to the verbs “pr´epi-na” (“must”-infinitive) or “xri´azete” (“needed”), Prohibition, Permission and Recommendation are explicitly stated in the verbs “apagor´evete” (“prohibited”), “epitr´epete” (“allowed”) and “prot´inete” (“is recommended”) respectively, whereas Possibility is mapped to the verbs “bor´i” (“can”) or “´ine dinat´on” (“is possible”). In the previous examples, the transcriptions of the Greek expressions are purely phonological and are not based on morphological and orthographic criteria.

A similar preference strategy is applied to negations, where the particle “´mi” (“don’t”) as well as the negative expression with the temporal semantic content “po´te” (“never”) are usually used to negate full sentences and are, therefore,

preferred as opposed to the negative particle "΄dhen" ("not", "does not"), mainly used to negate verbs. As an additional sublanguage-independent parameter, the achievement of precision and clarity in information describing time, space (movement), quality and quantity is targeted with the use of adverbial modifiers [2], as in the expressions "precisely under", "exactly before", "completely flat" and "exactly two".

The proposed parameters are oriented toward the targets of the criteria of informativeness, intelligibility and metacommunication handling on the Utterance Level [7], and may be expressed here as "Precision", and "Directness".

4 The Speech Act – Oriented Approach in Spoken Technical Texts

4.1 Relation of Speech Acts and Steps in Task-Oriented Dialog Structure

In spoken Technical Texts constituting task-oriented dialogs, the content of the utterances produced by the Conversational Agent is related to the Speech Acts for Task-oriented Dialogs [5], and modeled according to the above-described specifications for the written Technical Texts (Table 3). The Speech Acts for Task-oriented Dialogs involve speech acts related to user-input recognition ("Acknowledge"), confirmation of user-input "Confirm", checking task completion/task success requested or activated by user ("Check"), providing user with necessary information or informing user about data requested by user, task success/failure or current status of process/system ("Inform") and handling of waiting time ("Filled Pause") [5]. The System may ask the user to provide specific input ("Request") and expect the user's response ("Respond") [5]. For reasons of efficiency, in many dialog systems, a considerable percentage of the questions asked by the System constitute "Yes/No Questions" ("Yes/No Question") requiring a "Yes" or a "No" as an answer from the user ("Yes/No Answer") [5].

In dialog systems for spoken technical texts, steps in the dialog structure may be related to more than one Speech Act. Specifically, steps in the dialog structure involving the recognition of the user's answer and/or keyword recognition in user-input may be related to the "Acknowledge", "Request" or "Y/N Question" Speech Acts, as in the respective examples of utterances produced by the System (or System's Conversational Agent", namely "You have chosen the "Abort" option" ("Acknowledge"), "Please enter the requested date. Please press "1" ("Request") [8] and "Do you wish to execute the program?" ("Y/N Question").

Problems in the processing of user-input and/or errors in the keyword recognition in user-input may be related to the both the Speech Acts "Check" and "Request", as in the example of the produced utterance "Input cannot be processed", "Your input cannot be processed. Please repeat" ("Check")/("Request").

Input provided by the user that does not constitute a "Yes/No Answer" or is not related to keyword recognition (Free input) can be followed by the Speech Acts "Check", "Inform" or "Request" as in the respective examples of utterances produced by the System "We assume that you have completed the process" ("Check") [8], "You still have 30 seconds to file your complaint" ("Inform") and "Please add any further information you consider important" ("Request") [8].

The Speech Acts "Confirm" and "Inform" may concern the closing of the dialog between System and User, as shown in the respective examples "Your entry has been successfully registered" ("Confirm") [8] and "Your entry has been registered as No IE-6780923478" ("Inform"). Waiting time for the processing of user-input or for the completion of a process is handled by appropriate messages produced by the System such as "Please wait for two seconds" [8], identified as a "Filled Pause" Speech Act.

Table 3. Relation of Step in Task-oriented Dialog Structure and Speech Act

Step in Dialog Structure	Example	Speech Act
Answer / Keyword Recognition	Do you wish to execute the program?	Y/N Question
Problems or errors in Answer / Keyword Recognition	Your input cannot be processed.	Inform
Free Input	You still have 30 seconds to file your complaint	
Close Dialog	Your entry has been registered as No IE-6780923478.	
Answer / Keyword Recognition	Please enter the requested date Please press "1" Please choose "Confirm", "Repeat" or "Abort"	Request
Problems or errors in Answer / Keyword Recognition	Your input cannot be processed: Please repeat	
Free Input	Please add any further information you consider important	
Problems or errors in Answer / Keyword Recognition	Input cannot be processed.	Check
Free Input	We assume that you have completed the process	
Close Dialog	Your entry has been successfully registered	Confirm
All steps in Dialog Structure	Please wait for two seconds	Filled Pause
Answer / Keyword Recognition	You have chosen the "Abort" option	Acknowledge

4.2 Prosodic Modeling and Speech Acts for Task-Oriented Dialog

Prosodic modeling of the utterances related to the Speech Acts for Task-oriented Dialogs is based on the use of prosodic emphasis on the sublanguage-specific elements constituting the most important information in the sentence's semantic content, as well as sublanguage-independent elements such as negations and elements expressing time, space (movement), quality and quantity [1].

Specifically, prosodic emphasis on the negations and elements expressing time, space (movement), quality and quantity is used for the achievement of Precision [1],

while prosodic emphasis on sublanguage-specific expressions and terminology is used for the achievement of comprehensibility resulting to Directness (Table 4).

Table 4. Relation of prosodic emphasis in Task-oriented Dialog Speech Acts and the purpose of utterance in spoken technical text (Modern Greek)

Parameter type	Elements receiving prosodic emphasis in Task-oriented Dialog Speech Acts	Purpose
Sublanguage-independent	spatial, temporal, quantitative expressions expressions related to manner and quality	Achievement of precision
Sublanguage-specific	Sublanguage-specific lexicon, expressions and terminology	Achievement of directness

For example, for the efficient handling of semantic content and/or for precision and directness in the interactions, the words “yes”, “no”, “packaging”, “execute”, “code”, (sublanguage-specific expressions), “two minutes”, “thirty seconds” (quantity - time), and “cannot” (negation) receive prosodic emphasis in the respective sentences: “SYSTEM: Please answer the following questions with a “yes” or a “no” Was there a problem with the packaging?”, “SYSTEM: “Do you wish to execute the program?” (Speech Act: Yes/No Question), “SYSTEM: What is the code of the container?” (Speech Act: Request), “SYSTEM: Wait for two minutes” (Speech Act: Filled Pause), “SYSTEM: “You still have 30 seconds to file your complaint” (Speech Act: Inform), “SYSTEM: Your input cannot be processed” (Speech Act: Inform/Check).

Here, we note that all translations from Modern Greek are rendered with proximity to original syntactic structure.

4.3 Prosodic Modeling and Non Task-Oriented Speech Acts

Utterances produced by the Conversational Agent that are not directly related to the Speech Acts for Task-oriented Dialogs [3], such as the Speech Acts “Thank” or “Apologize”, are not subjected to a sublanguage-independent strategy for prosodic modeling. In this case, prosodic modeling is sublanguage-specific, possibly empirical up to a certain extent, and also related to the style of communication chosen for the HCI system concerned.

The style of communication chosen is relate to the User-model of the Dialog System and targets to User-friendliness. For example, for the achievement of User-friendliness in the interactions [3], the words “sorry”, “correctly”, “thank” and “additional” receive prosodic emphasis in the respective sentences [1]: “SYSTEM: I’m sorry, I was not able to understand you correctly” (Speech Act “Apologize”, followed by an error message) [8] and “I thank you for the additional input” (“Thank”).

The speech-act oriented approach in the steps of the dialog structure for spoken technical texts are targeted to meet the requirements of “Precision”, “Directness” and “User-friendliness”, summarizing the criteria of informativeness, intelligibility and metacommunication handling on the Utterance Level (Question-Answer-Level) [7], the Functional Level (initiative and interaction control) [7] and the Satisfaction Level (perceived task success, comparability of human partner and trustworthiness) [7].

5 Applications

For the achievement of "Precision" and "Directness" in written technical texts, the above-presented sublanguage-independent parameters used as a Controlled Language may be applied automatically or in a user-interactive tool for editing technical texts. In the user-interactive tool, the proposed sublanguage-specific and sublanguage-independent parameters may be presented in the form of questions and/or suggestions to the user. Sublanguage-specific parameters allow the user to make decisions to meet the needs and requirements of the sublanguage concerned. These may occur in form of questions towards the user, or may be implemented automatically as a sublanguage-specific module interacting with the sublanguage-independent parameters.

For the achievement of "Precision", "Directness" and "User-friendliness" in spoken technical texts, the proposed sublanguage-specific and sublanguage-independent parameters can be used in design phase of the Dialog System. This can be achieved with or without a user-interactive editing tool, with the purpose of the application in the implementation phase.

6 Further Research

The efficiency of the sublanguage-independent parameters in the proposed speech act-oriented approach remains to be evaluated in a larger number of technical applications. The Speech Acts for Technical Texts according to Lehrndorfer, 1996, involving the German language and the Task-oriented Dialog Speech Acts according to Heeman et al., 1998, concerning the English language, the combination of which is proposed and described here, is observed to be compatible with the features of technical texts in Modern Greek. The compatibility of the above-described Speech Acts in respect to other languages or other language groups remains to be investigated.

Furthermore, the effect of the proposed sublanguage-independent parameters may also be tested in task-oriented but not technical applications, such as customer services and e-learning.

References

1. Alexandris, C.: Word Category and Prosodic Emphasis in Dialog Modules of Speech Technology Applications. In: Botinis, A. (ed.) Proceedings of the 2nd ISCA Workshop on Experimental Linguistics, ExLing 2008, Athens, Greece, August 2008, pp. 5–8 (2008)
2. Alexandris, C.: Show and Tell: Using Semantically Processable Prosodic Markers for Spatial Expressions in an HCI System for Consumer Complaints. In: Jacko, J.A. (ed.) Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environment, vol. 4552, pp. 13–22. Springer, New York (2007)
3. Alexandris, C., Fotinea, S.-E.: Discourse Particles: Indicators of Positive and Non-Positive Politeness in the Discourse Structure of Dialog Systems for Modern Greek. International Journal for Language Data Processing "Sprache & Datenverarbeitung" 2004(1-2), 19–29 (2004)
4. Hausser, R.: A Computational Model of Natural Language Communication, Interpretation, Inference and Production in Database Semantics. Springer, Berlin (2006)

5. Heeman, R., Byron, D., Allen, J.F.: Identifying Discourse Markers in Spoken Dialog. In: Proceedings of the AAAI Spring Symposium on Applying Machine Learning to Discourse Processing, Stanford (March 1998)
6. Lehrndorfer, A.: Kontrolliertes Deutsch: Linguistische und Sprachpsychologische Leitlinien für eine (maschniell) kontrollierte Sprache in der technischen Dokumentation. Narr, Tuebingen (1996)
7. Moeller, S.: Quality of Telephone-Based Spoken Dialogue Systems. Springer, New York (2005)
8. Nottas, M., Alexandris, C., Tsopanoglou, A., Bakamidis, S.: A Hybrid Approach to Dialog Input in the CitzenShield Dialog System for Consumer Complaints. In: Proceedings of HCI 2007, Beijing China (2007)