Designing Confirmation Mechanisms and Error Recover Techniques in a Railway Information System for Spanish

R. SAN-SEGUNDO, J.M. MONTERO, J. FERREIROS, R. CÓRDOBA and J.M. PARDO
Grupo de Tecnología del Habla. Departamento de Ingeniería Electrónica. UPM
E.T.S.I. Telecomunicación. Ciudad Universitaria s/n,
28040 Madrid, Spain,
{lapiz|juancho|jfl|cordoba|pardo}@die.upm.es

Abstract

In this paper, we propose an approach for designing the confirmation strategies in a Railway Information system for Spanish, based on confidence measures obtained from recognition. We also present several error recover and user-modelling techniques incorporated in this system. In the field evaluation, it is shown that more than 60% of the confirmations were implicit ones. This kind of confirmations, in combination with fast error recover and user modelling techniques, makes the dialogue faster, obtaining a mean call duration of 204 seconds.

Introduction

For designing an automatic spoken dialogue system, confirmation mechanisms and error recover techniques constitute an important sub-problem. These mechanisms have a relevant influence over the general performance of the system. Good confirmation mechanisms can avoid asking the user some questions to validate the recognized words, making the dialogue faster. When the system does not confirm those words (because they have high recognition confidence), this assumption can fail. Because of this, fast error recover techniques are required. In this paper, we propose an approach for designing the confirmation strategies and implementing error recover and user-modelling techniques in a Railway Information system.

During last decade, the performance of spoken dialogue system for traveling information has improved substantially. One spoken dialogue project is the DARPA Communicator http://fofoca.mitre.org (Pellom, 2000; Rudnicky, 2000) that enables to access information about airline flights, hotels and rental cars. In Europe, one important project concerning railway information is ARISE (Lamel, 2000; Baggia, 2000).

1 Confirmation Mechanisms

For designing the Confirmation Mechanisms in Spoken Dialogue Systems, it is necessary to describe the confirmation strategies considered, to obtain confidence measures from the recognition module and to define a relationship between confidence values and confirmation strategies (Sturm, 1999).

1.1 Confirmation Strategies

Depending on the number of items to confirm:

- One item ("Have you said Madrid?").
- Several items ("Do you want to go from Madrid to Sevilla?").

Depending on the possibility to correct:

- *Explicit confirmation*: the system confirms one or several item values through a direct question. ("I understood you want to depart from Madrid. Is that correct?")
- *Implicit confirmation*: the system does not encourage the user to correct, it only reports about the recognition result. ("You leave from Madrid. Where are you arriving at?")
- *Semi-implicit confirmation*: it is similar to the implicit confirmation, but the user can correct ("You want to leave from Madrid. In case of error, say correct, otherwise, indicate your arrival city"). This confirmation allow error recover (Lavelle, 1999), but it is not very friendly for the user. In Section 3, we describe the CORRECT command that permits same functionality without increasing the prompt.
- *No confirmation*: the system does not provides feedback about the recognized value (e.g. in yes/no questions).

- *Item value rejection* and repeat question: when the confidence is low, the system does not present the value to the user and repeats the question ("Sorry, I can not understand. Where are you departing from?").

1.2 Confidence Measures in recognition

The recognition module used in our system is a large vocabulary telephone speech recognizer, that can recognize isolated words and simple expressions such as "On Monday", "Next week" or "In the morning". The recognizer is based on a hypothesis-verification approach. The best features for confidence annotation are concerned with the verification step and are based on (Macías, 2000):

- *First Candidate Score:* acoustic score of the best verification candidate.
- *Candidate Score Difference:* difference of acoustic scores between the 1st and 2nd verification candidates.
- *Candidate Score Mean and Variance:* average score and variance over the 10 best candidate words.
- *Score Ratio:* difference between the score of the phone sequence (hypothesis stage) and the score of the best candidate word (verification stage).

All features are divided by the number of frames. We have used a Multi-Layer Perceptron (MLP) to combine the features. On a database with 2,204 cases, 39.1% wrong recognized words are detected at 5% false rejection rate, reducing the minimum classification error from 15.8% (recognition error rate) to 14.0%.

1.3 Confirmation Mechanisms design

For designing the confirmation mechanism, it is necessary to plot the correct words and errors distributions as a function of the confidence value and to define different confidence thresholds (Figure 1).

We have defined 4 levels (3 thresholds):

1. *Very High confidence:* the number of correctly recognized words is much higher than the number of errors.
2. *High Confidence:* the number of correctly recognized words is higher than the number of errors.
3. *Low confidence:* Both distributions are similar. The system is not sure about the correctness of the recognized word.
4. *Very Low confidence:* In this case, there are much more errors than correctly recognized words, so we reject the recognized word and the system asks again.

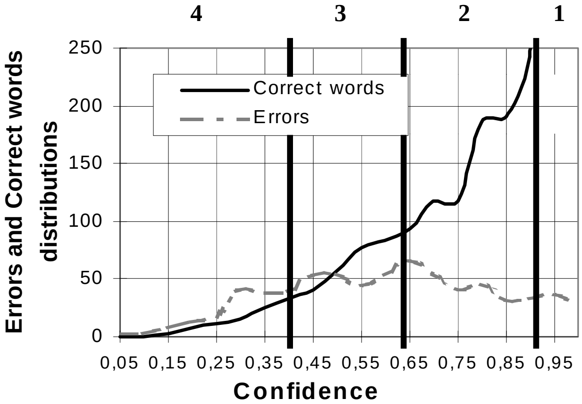

Figure 1.Errors and Correct words distributions vs. confidence.

For the departure and arrival cities step, we define CL(D) and CL(A) as departure/arrival city Confidence Levels (CL). Depending on the CL, we implement the following confirmation strategies:

- CL(D)=1 & CL(A)=1: *implicit confirmation of both items.* "You want to travel from Madrid to Sevilla. When do you want to leave?"
- CL(D)=2 or CL(A)=2: *explicit confirmation of both items.* "Do you want to travel from Madrid to Sevilla?"
- CL(D)=3 or CL(A)=3: *explicit confirmation of each item.*
- CL(D)=4 or CL(A)=4: *rejection of both items.*

Under hard conditions, the system asks the user to spell the city name (San-Segundo, 2000). For the spelled name recognizer, new confidence analysis is necessary and new thresholds are defined. These thresholds are used in the same way as described above.

1.4 Confirmation sentences design

a) *Implicit confirmation:* these sentences consist of two parts: the system presents the value of the item and then it asks the new item ("You want to travel from Madrid to Sevilla, When do you want to leave?"). This structure permits to make independent design for confirmation sentences and for item questions. When there is a semantic relationship between the item to confirm and next item, we could join both sentences in just one. For example:

S: "Which month do you want to leave?"
U: "July"
S: "Which *day of July* do you want to leave?"

Joining sentences produces a shorter dialogue.

b) *Semi-implicit confirmation:* the structure comprises three parts: the item presentation, the command for correction and the prompt for the next item. The sentence usually is very long and unfriendly.

c) *Explicit confirmation:* the structure consists of one sentence where the system explicitly asks whether the recognized item value is correct or not ("Do you want to leave on July the 18^{th}?"). This sentence can be divided in two: the first one presents the item value, and the second one asks about the correctness of the value ("You want to leave on July 18^{th}. Is that correct?"). In this case, as for implicit confirmations, independent analysis are possible.

In order to make the system friendly to the user, the confirmation sentence must be designed depending on the user answer. If he/she says "*this Monday*" to specify the departure date, it is better to use the sentence "Do you want to leave *this Monday*?" instead of "Do you want to leave on *February, 19th*?" for confirming it.

2 Error recover mechanisms

When an automatic system uses implicit confirmations, it is necessary to define some mechanisms to permit the user to recover from system errors:

START OVER: this command permits the user to start from scratch. Instead of resetting all the items, our system begins confirming groups of items explicitly (dialogue steps). When one group is not confirmed, the system starts from that point:

S: "The selected option is an Intercity train... "
U: "Start over"
S: "Let us start the call again. Do you want to go from Madrid to Barcelona?"
U: "Yes"
S: "Do you want to travel on July the 19^{th}?"
U: "No"
S: "Do you want to leave this week, next week or later?"

CORRECT: When the system makes a mistake and takes a wrong item value as right (in an implicit confirmation), the user can correct the system by saying this command at any point of the dialogue. In this case, the system asks for the last introduced item again. For example:

S: "On which month do you want to leave?"
U: "May, please"
S: "Which day of March?"
U: "Correct"
S: "Which month do you want to leave?"
U: "May"

3 User-Modelling

The user modelling technique applied is based on (Veldhuijzen van Zanten, 1999). We consider 4 user skill levels. Depending on current level, the prompt sentences are clearer (p.e they contain more information about how the user should answer, possible values, etc...) or the system provides more or less information per time unit. The levels we have considered are:

- **1st level.** The prompts explains how to interact with the system, the asked item, the possible accepted values and the way of specifying one of the values. (for the period of the day: "Please speak after the tone. Say the period of the day you want to travel in; in the morning, in the afternoon or in the evening.")
- **2nd level.** The prompts include the requested item, the accepted values for this item and the way to specify them. ("Say the period of the day you want to travel in; in the morning, in the afternoon or in the evening.")
- **3rd level.** Only the required item is included in the prompt. ("Say the period of the day you want to travel in.")
- **4th level.** The user knows everything and we can relax the question ("When do you want to leave?").

Current level depends on the initial state and the number of errors or positive confirmations along the interaction. In our case, the system starts at the 2^{nd} level (after providing a general explanation about how to interact with the system). When several errors (or positive confirmations) occur, it decreases (or increases) this level. The number of errors or positive confirmations that forces a change depends on current level. Thus, the system adapts dynamically the interactions to the user skill.

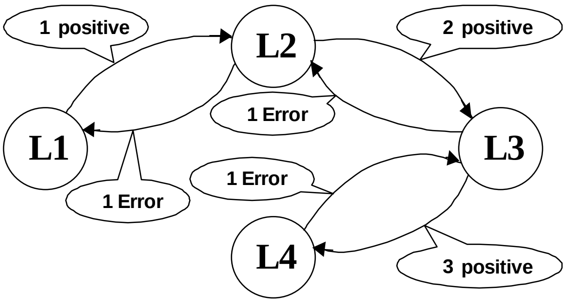

Figure 2. Diagram of transitions between levels in the user-modelling technique.

Example:
[The system is in the level 3]
S: "Say the period of the day you want to travel in."
U: "After lunch"
[The system recognizes "in the evening"]
S: "Have you said in the evening?"
U: "No"
[The system decreases the level from 3 to 2]
S: "Say the period of the day you want to travel in; in the morning, in the afternoon or in the evening."
U: "In the afternoon"

4 Field evaluation

In the field evaluation 30 users called the system and completed 4 scenarios (120 calls). The main measures are shown in table 1:

Table 1. Average global measures calculated by the system from users calls.

Measure	Value
Mean call duration (seconds)	204
% of implicit confirmations	61.3
% of explicit confirmations	38.7
START-OVER commands per call	0.08
CORRECT commands per call	0.43

Because of the high recognition rate and the confidence measures, the percentage of implicit confirmation (vs. explicit confirmations) has been relevant (61.3%). These confirmations have made the dialogue faster. In this analysis, we have not included the yes/not questions (they are not confirmed because the high accuracy) or the rejected answers. When an implicit confirmation is corrected by the user (less than 10% of the times), we have considered it as an explicit confirmation. Other important aspect is that the CORRECT command avoids the user starting from the beginning (the START OVER command is not frequently used).

5 Conclusions

In this paper, we have introduced an approach for designing confirmation mechanisms, based on recognition confidence measures and we propose several error recover and user-modelling techniques. This strategies have permitted to develop a fully automatic system for railway information with a good acceptability from the users (3.9 out of 5). The confirmation mechanism designed can be easily extended to a mixed-initiative systems with continuous speech recognition and understanding. In this case, confidence measures for semantic concepts are required (San-Segundo, 2001).

References

Baggia, P., Castagneri, G., Danieli, M., 2000. "Field trails of the italian ARISE train timetable system". Speech Communication. V31, No 4 pp 356-368.

Lamel, L., Rosset, S., Gauvain, J.L., Bennacef, S., Garnier-Rizet, H., Prouts, B., 2000. "The LIMSI ARISE system". Speech Communication. V31, No 4 pp 339-355.

Lavelle, A.C., Calmes, H., Pérennov, G., 1999."Confirmation strategies to improve correction rates in a telephonic inquiry dialogue system". EUROSPEECH, Budapest, Hungary.

Macías-Guarasa, J. Ferreiros, J., San-Segundo, R., Montero, JM., and Pardo, JM., 2000. "Acoustical and Lexical Based Confidence Measures for a Very Large Vocabulary Telephone Speech Hypothesis-Verification System". ICSLP. Beijing, China.

Pellom, B., Ward, W., Pradhan, S. 2000. "The CU Communicator: An Architecture for Dialogue Systems". ICSLP, Beijing, China.

Rudnicky, A., Bennet, C., Black, A., Chotomongcol, A., Lenzo, K., Oh, A., Singh, R. 2000. "Task and domain specific modelling in the Carnegie Mellon Communicator system". ICSLP, Beijing, China.

San-Segundo, R., Pellom, B., Hacioglu, K., Ward, W., and Pardo, JM., 2001. "Confidence measures for Spoken Dialogue Systems". ICASSP, Salt-Lake-City, Utah, USA.

San-Segundo, R., Colás, J., Ferreiros, J., Macías-Guarasa, J., Pardo, J.M., 2000. "Spanish Recognizer of continuously spelled names over the telephone". ICSLP, Beijing, China.

Sturm, J., den Os, E., and Boves, L., 1999 "Issues in Spoken Dialogue Systems: Experiences with the Dutch ARISE System". Proceedings of ESCA Workshop on Interactive Dialogue in MultiModal Systems. Kloter Irsee, Germany, 1-4.

Veldhuijzen van Zanten, G., 1999. "User modelling in adaptative dialogue management". EUROSPEECH, Budapest, Hungary.