

SYSTEM DEMONSTRATION GOALGETTER: GENERATION OF SPOKEN SOCCER REPORTS

Mariët Theune and Esther Klabbers
IPO, Center for Research on User-System Interaction,
Eindhoven University of Technology

Abstract - In this paper we describe a demonstration of the GoalGetter system, which generates spoken soccer reports (in Dutch) on the basis of tabular data. Two types of speech output are available. The demo runs via the web. It includes the possibility of 'creating your own match' and having GoalGetter generate a report on this match.

1. About the system

The GoalGetter system is a Data-to-Speech system which generates spoken soccer reports (in Dutch) on the basis of tabular data. The system takes as input data about a soccer match that are derived from a Teletext page.¹ The output of the system is a spoken, natural language report conveying the main events of the match described on the Teletext page.

GoalGetter was developed on the basis of D2S, a generic system for the creation of Data-to-Speech systems. The general architecture of D2S is represented in Figure 1. It consists of two modules, the *Language Generation Module* (LGM), and the *Speech Generation Module* (SGM). The LGM takes data as input and produces *enriched text*, i.e., text which is annotated with prosodic markers indicating pitch accents and intonational boundaries. This text is sent as input to the SGM, which turns it into a speech signal.

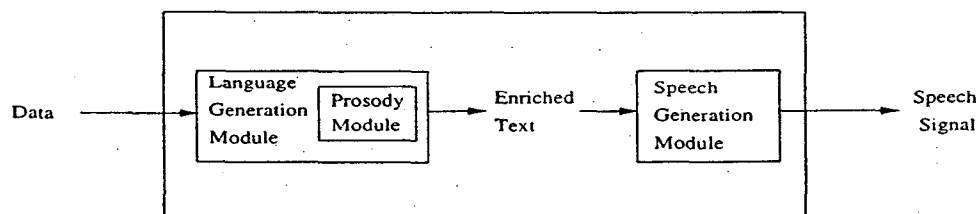


Figure 1: Global architecture of D2S

Language generation in the LGM is done using syntactic templates, which are syntactic tree structures containing slots for variable expressions. The selection and ordering of the templates and the filling of their slots depend mainly on conditions on the discourse context, which is represented in a Discourse Model. In order to achieve variation in the generated texts, the system generates a number of different expressions for each piece of information. If various expressions are equally suitable given the current context, one is chosen at random. For a detailed description of the syntactic template technique we refer to [Van Deemter & Odijk 97] and [Theune et al. 97].

In the Prosody module, accentuation and phrasing are determined using information about the syntactic structure of a sentence and about its context. Phrases expressing 'new' or 'contrastive' information are accented while those expressing 'given' information are deaccented. Speech generation in the SGM is done

1. Teletext is a system with which information is broadcast along with the television signal and decoded in the receiver. The information is distributed over various 'pages', most of which contain textual information, but some contain tables.

by means of either speech synthesis or an advanced form of phrase concatenation, using different prosodic versions of otherwise identical phrases. Which version is chosen depends on the prosodic markings provided by the LGM. For Dutch speech synthesis, we use the phonetics-to-speech system SPENGI (SPeech synthesis ENGIne) developed at IPO, which employs PSOLA-based diphones.

2. Outline of the demonstration

The demonstration of the GoalGetter system runs via the web and consists of three parts. First we present an example report that has been generated off-line by GoalGetter. This allows us to prepare an (annotated) English translation of the Dutch text in advance. (In the other parts of the demonstration, preparing a translation is not possible since reports will be generated on-line and it is impossible to predict the exact contents of the reports, due to the variation mentioned above. In those cases we provide a rough translation on the spot.) Second, we generate a report on-line, using an existing input table. Finally, we fill an input table with data selected by those attending the demonstration, and generate a report expressing these data. All modules depicted in Figure 1 are included in the demonstration. For each 'generation round', first a plain text version of the generated report is shown, and then the enriched text with the prosodic markers. Finally, the system's speech output is played.²

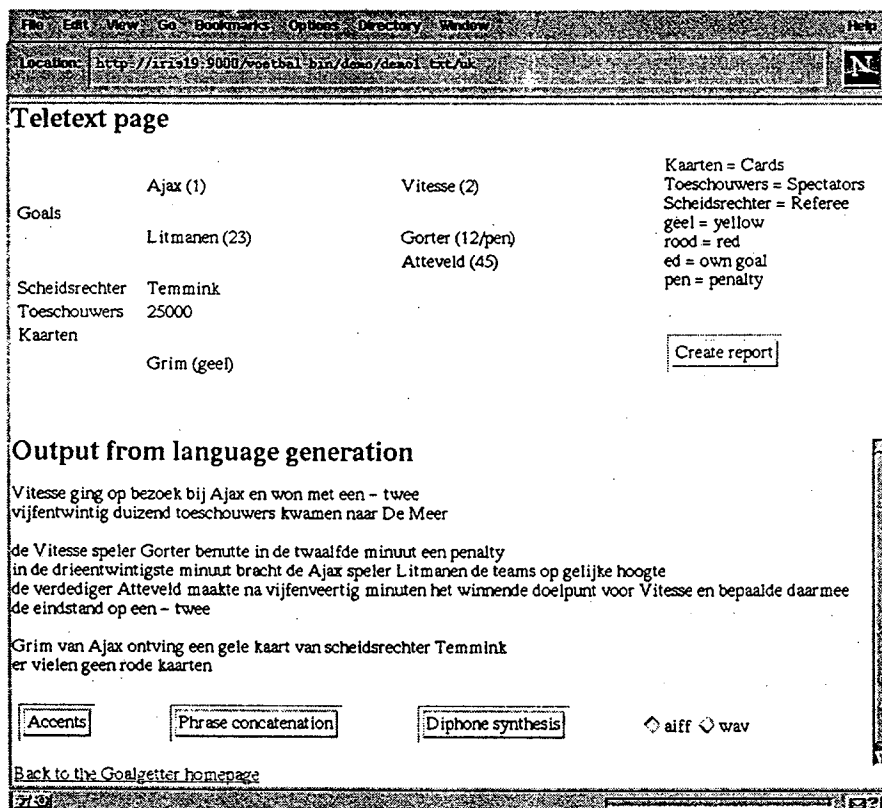


Figure 2: Screenshot showing Teletext page and generated report

Example report - In order to give a first impression of the GoalGetter system, we start the demonstration by

2. Due to the fact that a few phrases are missing from the phrase database, phrase concatenation is only available in the first part of the demonstration. The output from diphone synthesis is available in all parts.

showing an example 'Teletext page'³ and a report that was generated on the basis of this page in both plain and enriched text format. A written English translation will be provided. We play a spoken version of the report using both speech synthesis and phrase concatenation.

The screenshot in Figure 2 shows a Teletext page and the plain text version of a generated report.⁴ Since the screenshot was taken from the English language version of the web demonstration, to the right of the Teletext page we can see a translation table of the words and abbreviations occurring on the page. Below this table, there is a 'Create report' button, which can be used to generate a new report on the basis of the same Teletext page.

Below the report, there are three more buttons. Clicking the first replaces the plain text version of the report with the enriched text version, shown in Figure 3.⁵ The other two buttons can be clicked to play the audio files that were created when the report was generated, using phrase concatenation and diphone synthesis respectively. The sound files can be either in aiff or in wav format.

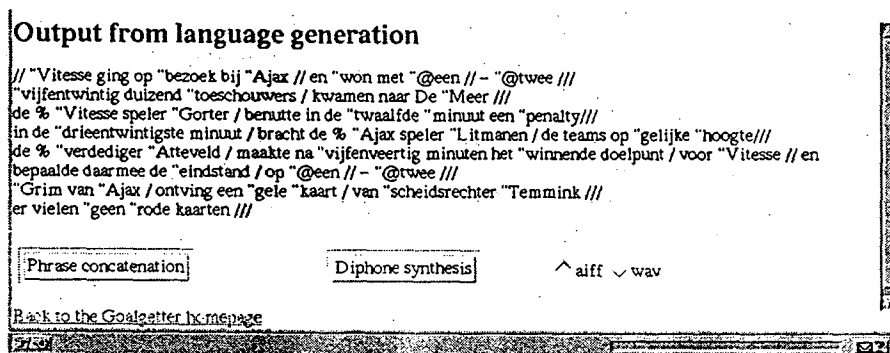


Figure 3: Partial screenshot showing generated report in enriched text format

Generate a report on-line - The next step is to generate a report on-line. We use the same Teletext page as before, which gives us the opportunity to point out the variation in the generated reports. Although the written report (including prosodic markers) is generated almost instantaneously, the fact that the sound files have to be created at the same time causes a small delay of about five seconds. When the report has been generated and one of the speech buttons is clicked, the corresponding sound file has already been created and only has to be read.

Creating a match - Finally, the web demonstration allows one to define the input for GoalGetter by filling an empty input table according to one's own preferences. We use this feature to 'create a match' using data selected by those present at the demonstration. On the basis of these data, a 'Teletext page' is created and reports can be generated in the manner explained above.

In Figure 4 we can see how an input table is filled. First, general information about the match is specified, such as the teams involved, the referee and the number of spectators. The teams and the referee are chosen from a predefined list, via a pull-down menu. After the general information has been determined, the main

3. In the web demonstration, we use simulated Teletext pages which are based on the original ones.
4. Translation: *Vitesse visited Ajax and gained a 1-2 victory. Twenty five thousand spectators came to De Meer. The Vitesse player Gorter kicked a penalty home in the 12th minute. In the 23rd minute, the Ajax player Litmanen leveled the score. After 45 minutes, the defender Atteveld scored the winning goal and thereby brought the final score to 1-2. Grim from Ajax received a yellow card from referee Temmink. There were no red cards*
5. In the enriched text, accents are indicated by double quotes, and intonation boundaries of different strengths by a number of slashes. The symbols @ and % indicate specific pronunciations for certain words.

events of the match can be specified in the 'Goals and cards' section. This is done by selecting a player from one of the teams, selecting the event this player was involved in (goal, own goal, penalty, yellow or red card), specifying the minute the event occurred in (for cards this is optional), and then clicking on the Add button to add the 'event' to the input table. This action can be repeated until an interesting match has been created. Then clicking the 'Create the Teletext page' button will produce a page similar to the one shown in Figure 2, and a report can be generated.

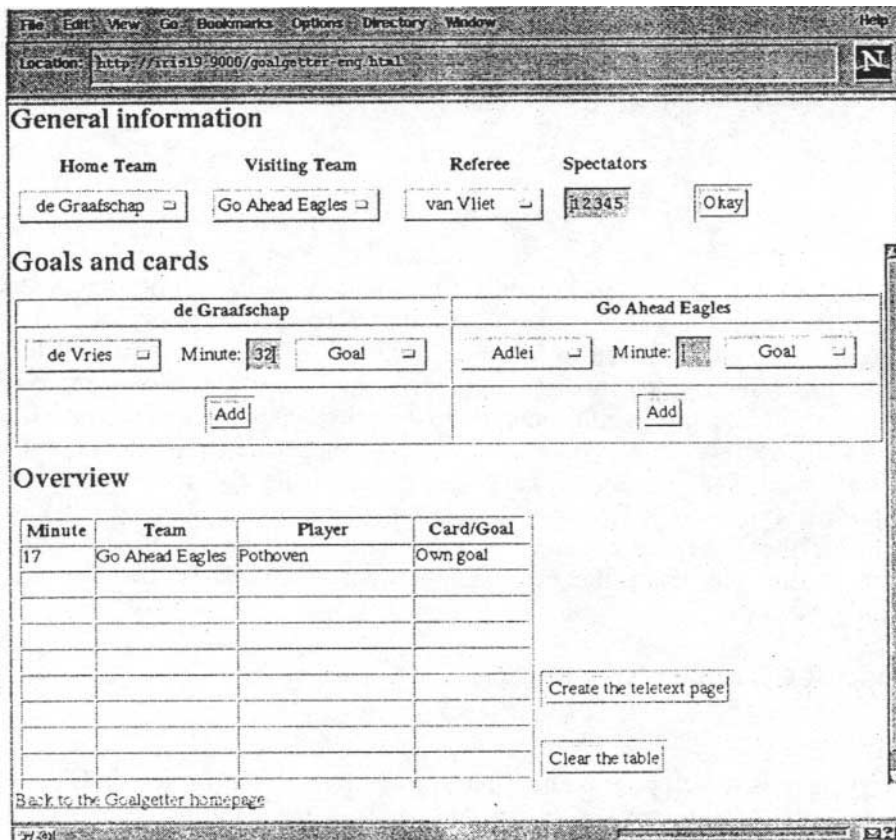


Figure 4: Screenshot showing the filling of an input table

3. Requirements

The GoalGetter demonstration can be run via the web. The web browser that is used should support frames and tables. (In the case of Netscape, this means that version 2.0 or higher is required.) In order to make the sound files audible, a sound card must be available. When running the demonstration on a PC, wav format should be chosen for the sound files, while for a Unix machine aiff-format is more suitable. The options of the web browser should be set so that audio files are played (e.g., using `sfplay`) instead of being saved to disk.

References

- [Van Deemter & Odijk 97] Van Deemter, K. and J. Odijk, 1997. Context Modelling and the Generation of Spoken Discourse. *Speech Communication* 21(1/2), 101-121.
- [Theune et al. 97] Theune, M., E. Klabbers, J. Odijk and J.R. de Pijper, 1997. From Data to Speech: A Generic Approach. Manuscript, URL: <http://www.tue.nl/ipo/people/theune/manuscript.ps.Z>.