The use of speech recognition technologies in the drafting of the verbatim reports of proceedings at the Italian Chamber of Deputies

Ms. Simonetta Zorzi, I, Head of the Office for summary reports of special Committees.
-0-0-0-0-

For more than five years now, the Chamber of Deputies has been acquiring an inclusive experience in the field of ASR (Automatic Speech Recognition) applied to report drafting techniques. Form many years the Chamber's Reports Department (former Shorthand Department), and in particular Ms Simonetta Zorzi, who is the official in charge of co-ordinating the project for the part falling within her competence, have been cooperating steadily with the IT Department within the framework of the project "Vocal Interaction/Assisted Transcription Systems", managed by the IT engineer Enrico Boccalaro.

The project is aimed at testing ASR technology, then applying it to an operational team of verbatim reporters. In this almost unexplored field, scientific and technological progress had been lagging behind for a long time. Then, in the last two years it has speeded up dramatically, in connection with the developments achieved in other areas of research, among which those related to mobile telephones. The first objective has thus been that of pinpointing and then testing Speech Recognition systems available on the market, in the first place those produced by Dragon (today Lernout & Hauspie), IBM and Philips. Due to many reasons, among which fast response and recognition accuracy, the IBM technology based on "discrete speech" was selected, and in 1998 the IT Department has substantially customized the system (user's interface and linguistic background), taking into account the suggestions that had emerged during the testing phase.

During the training courses organised later within some Departments of the Chamber of Deputies (among which the Reports Department), the system has shown satisfactory levels of performance from the point of view of accuracy and reliability, even though only dictionaries containing 20.000 words were available at the time. Nevertheless, we have realized that such an unnatural dictation system, which requires pauses between words and even the spelling of apostrophes, was not yet ready to be effectively implemented, taking also into account the huge size of parliamentary reporting, covering every year thousands of hours of sittings.

1999 marked the turning point, when it was possible to switch from the "discrete speech" mode to the "continuous speech" mode, by which it became possible to dictate a text in a more fluent way. This system, as opposed to the previous one, privileges a fluent and fast dictation, and penalises pauses and uncertainties.

Nevertheless, we soon realized that the products available on the market were not suitable for a daily use by a large number of parliamentary reporters. For this reason, we developed a project to implement the so-called LVCSRS (Large Vocabulary Continuous Speech Recognition System), thus overcoming some of the limits of the commercial products.

During the year 2000, the development of CameraVox was almost completed. This systems features some unique characteristics, like, first of all, a language model of continuous speech inspired by the parliamentary documentation going from the 10th to the 13th Parliament and including more than 350 millions words. Furthermore, a Dictionary of Continuous Parliamentary Speech of 80,000 words was added, together with a Vocal Proof Error Analyser, which consists of a text analyser working on a statistical background of morphological voice errors and on a deterministic base of Italian grammar rules. During the same year the system was introduced with all the stenographers working in the Standing Committees of the House.

In Italy, as in many other European countries, with the sole exception of Germany, it has become rather difficult to recruit staff with the required shorthand professional skills. Taking this into account, at the end of the 13th Parliament, a number of training courses on parliamentary reporting were started, meant to a group of about 30 documentalists (without specific skills), who were trained in the voice drafting of parliamentary records. The new method was introduced at the beginning of the new 14th Parliament.

The courses lasted four months. Soon after the CameraVox training period, these members of the staff were specifically trained in parliamentary reporting. Training consisted in exercising on parts of previous sittings and dictating the reports with the new system. Dictation has to be reasoned but also as faithful as possible with the speech pronounced by the Member, while taking into due account all the procedural aspects of the sitting.

Already at the beginning of the 14th Parliament, once training ended, this personnel was able to provide for the overall first integral version of the proceedings of both the Floor of the House and the Standing Committees.

As far as the Floor only is concerned, as of its first sitting, which took place on May 30, 2001, 1,000 work-shifts were performed in the first two months of work, totalling about 100 hours of sittings (in the past Parliament the average number of reported hours of sittings in both the Floor and the Committees amounted to 2,000 hours per year).

The daily monitoring of the quantitative and qualitative bulk of results reached so far, also taking into account the various difficulties emerged, allows us to express a positive judgement on the reliability of the large-scale use of the CameraVox system.

In fact, drafting times and accuracy of the texts produced met the standards envisaged. The same applies to the reporting of Committees, which features different

production criteria.