

DIDACTIC EXPERIENCES IN TEACHING SPEECH RECOGNITION.

prof. Maria Luisa Corti Crippa. I

Italian State approved teacher. Involved in teaching the reporting techniques (language - stenotype - speech recognition)

The improvements happened in the last 4 years in the computer field, have brought speech recognition technology to a good level of speed and accuracy that encourage additional interest from new users.

Mr. Vesselin Tzecov at the Mannheim Congress in 1981 forecast speech recognition to become a reality at the end of the nineties. This forecast is timely became true. In 1995 at the Amsterdam Congress, Ing. Giorgio Lazzari of Trento (Italy) presented an advanced project of speech recognition software; this software was speaker independent for the Italian language. This software was developed and it is used for special application (medicine and registration of the landed properties).

This year in the programme of the Congress Session additional information on this technology will be spread.

I have had the great opportunity to follow this development in my quality of Managing director of a reporting company in Milano from 1981 to 1988. In that years, in conjunction with my partners, we approached several soft wares in order to find out the practical possibility to use this technology in addition to the traditional shorthand, stenotype and keyboard techniques.

Till 1998 this were only personal experiences, from that year on a rapid growth in the processors' speed as well as the possibility of larger RAM quantity, considerably improved the development of speech recognition software mainly IBM and Dragon systems.

At this point I personally, with other business partners, began using speech recognition and, later on, a first training course was held in Milano. I would like to point out at this stage that my main activity for more than 30 years was teaching shorthand - typewriting - accountancy and word processing. I got the Italian state approval for teaching in the high schools. Thanks to this experience I was able to figure training of speech recognition technology in order to get good results for a professional use in different fields of social and economic activities.

As a matter of fact we can mainly envisage the professional use of this technology by the following persons:

Secretaries (in a broad sense) who have to produce texts of good quality in short time

Executives who want to prepare drafts of their reports and speeches in a quick manner. An assistant can eventually carry out the final revision and formatting.

Reporters (be they court reporters or parliamentary or assemblies reporters) can prepare their reports from tapes or digital recordings. Practical experiences (mainly in the States and currently going on also in Italy) have shown that with increased training and with high efficient computers it is possible to have a real time reporting as well as close captioning of live events.

For all people involved in the afore mentioned activities, speech recognition can be considered as an integration to their

abilities, a good approach to this technology is highly essential in order to get practical results within a reasonable time-limit.

Based on my experience, I can soundly state that after a first approach of about 1 hour, the student is immediately reassured that he/she can obtain significant results, since to day speech recognition (we are using Dragon Naturally speaking professional) needs only 20 minutes for the first training. After this first training the student can easily dictate plain texts at a speed of 60/70 words per minute (about 150 syllables) with at least 94% of accuracy.

From this starting point the teacher can easily lead student toward professional results, which are mainly based on the traditional parameters of accuracy and rapidity. These goals are the same that steno and typewriting teachers have followed during their traditional teaching. For example: In steno (or stenotype) teaching, the writer must take care of avoiding deformation of signs or strokes, which can lead to misinterpretation, the same applies to speech recognition where you have to stress the good way of pronouncing words in order to get the right results.

In this learning we can easily reuse our experiences in controlling progressive results, investigating different domain of text by using the same training material that we have prepared for improving speed skills with steno and stenotype.

I found particularly important to have students dictating a text by listening to tapes with progressive speed.

With this method, in my training courses, with students ranging between 25 and 45 years of age, I was able to have them writing 100 words per minute (225 syllables) in about 90 hours of practice. The accuracy was in any case more than 95 % and the students practiced the ability to quickly correct the mistakes.

For professional reporting activity this speed and accuracy need to be further improved, and additional time and attention must be devoted since one must reach (in the Italian language) a speed of at least 130 words per minute (300 syllables) and for close captioning one must consider at least 150 words per minute (350 syllables). This very high training must be very well planned, but if we consider that at speech recognition is completely managed by computer technology, we can easily forecast integration with distance learning systems as well as with motivating exercises recorded from live broadcastings.

As a final word I can easily say that we, as teachers, have new opportunities arising from speech technology, which does not only means to produce good text in reduced time, but also the ability to command with voice any computer programme. Microsoft is planning to introduce speech-recognition as a standard feature (at least for the English language) in the new operating system (XPI), which will be soon released.

In short we must not skip this new opportunity, which is in line with our past, and brings us into the future.