



JEFF ALLEN WRITES:

Speech technology after the hoax

The speech technology industry has been growing exponentially over the past few years as speech-enabled applications expand into new areas of influence. With the development of new technologies, it is now possible to treat data content entry through the use of voice-driven processes rather than relying on key-entry or other traditional methods. In other words, it is now possible to capture and process speech in ways that were never possible before.

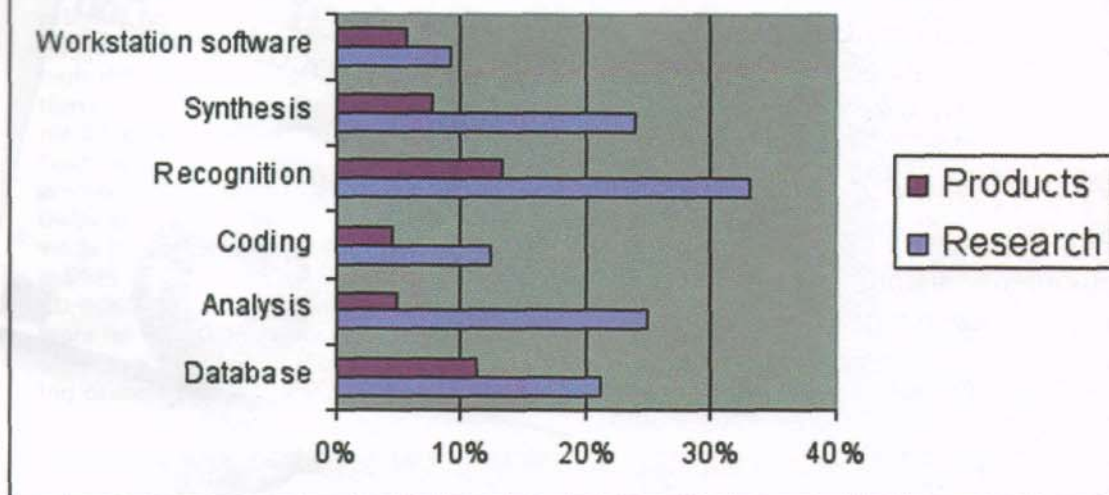
The following list includes just some of the types of systems and applications that now utilise speech technologies: airline information; bill payment system; call center; call processing; chat and e-mail; customer service; desktop financial applications; e-mail address retrieval; fax readers; Internet translation service; Internet voice portal server; logistics productivity; medical record software; medical speech dictation software; messaging systems; mobile organisers; multi-lingual training authoring system; music identification service; PC games; personal finance package; post office sorting system; self-service travel shopping; speech-to-speech translation; telecom assistant; time management scheduling for service merchants; virtual personal assistant software; voice dialing; voice verification; voice-activated auto attendant; voice-activated voice mail; voice-activated wearable computers; voice-controlled cockpit; voice-controlled e-mail; voice-controlled robotic service; voice-dialing directory; voice-driven Internet browser; voice-enabled corporate

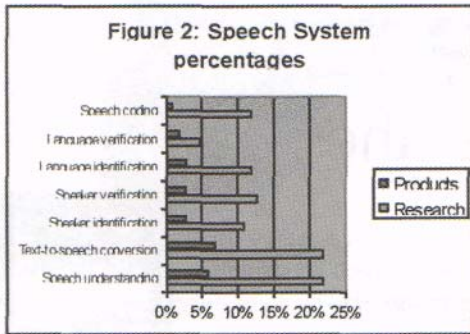
routing and information; voice-interactive navigation system (source: Speech Recognition Update newsletter, monthly issues: January 1999 - March 2000).

The types of systems and applications mentioned above reflect a large number of varied fields and domains that include: airlines and aerospace; audio products; automotive; broad-based information and commerce; construction; customer interaction; direction finding; educational software; employee self-service; finance and banking; foreign exchange quotes; health care and medical; law and judicial services; logistics; military and police; music; mutual fund trading; pharmaceutical; postal service; power utilities; property firms; railway information; restaurant information; sales/distribution; spoken language proficiency assessment; stock exchange; surveyor; telecom information; telecommunications; travel; u-access discount phone shopping; warehouse (stock) picking. It is obvious that speech processing has an ongoing significant impact on business, educational, and personal needs. In 1999 and 2000, a number of e-mail-based surveys covering many areas of language processing were conducted within the European Commission funded LE4-8335 (LRSPP) project. These survey results were collected and shared publicly in reports and presentations up through mid-2000. A portion of those results concerning the data and resource



Figure 1: Speech Processing percentages



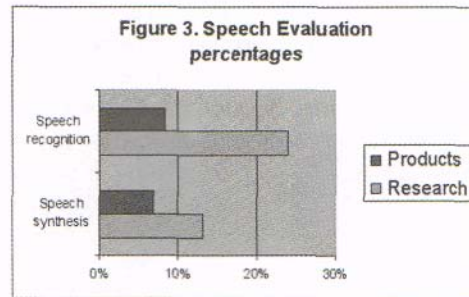


language needs were provided in my column article in IJLD issue 5. In the present column article, I would like to focus on the speech processing needs.

The survey methodology and results have already been fully described and presented elsewhere in full detail (Allen, 2000; Allen and Choukri, 2000). Of the nearly 1000 potential respondents who were contacted directly, a response rate of 25.3% was obtained. The information below reflects the responses of 250 respondents of which approximately 1/3 of respondents are involved in speech technologies.

One section of the survey aimed at gathering information about the type of work being conducted in the Speech domain, including basic research as well as commercial product development. From results that are given in Figure 1, we see that up to 1/3 of all surveyed respondents are actively involved in speech recognition; the other areas of speech R&D (speech synthesis; speech databases; speech analysis; speech coding; speech workstation software) having slightly lower percentages.

The survey also focused on determining current and near-future needs for data to develop and train the following types of systems: speech understanding; text-to-speech conversion; speaker identification; speaker verification; language identification; language verification; speech coding (Figure 2). Again, these percentages given are based upon the total number of respondents of the general survey.



In general, a greater number of users in the speech field are working on or with speech understanding (including speech recognition and speech dictation) and text-to-speech conversion systems. Most of the other types of speech systems each represent 10-15% of respondents. The assessment of speech recognition and speech synthesis are also areas that were surveyed. The results obtained (Figure 3) show that a number as high as 25% of all respondents seek speech-type data and resources for the evaluation of speech recognition systems for research purposes whereas there is a lower percentage for research-based speech synthesis. Commercial speech product evaluation is lower.