Voice recognition products—an occupational risk for users with ULDs?

N. R. Williams

Background Voice recognition systems (VRS) allow speech to be converted both directly into text—which appears on the screen of a computer—and to direct equipment to perform specific functions. Suggested applications are many and varied, including increasing efficiency in the reporting of radiographs, allowing directed surgery and enabling individuals with upper limb disorders (ULDs) who cannot use other input devices, such as keyboards and mice, to carry out word processing and other activities.

Aim This paper describes four cases of vocal dysfunction related to the use of such software, which have been identified from the database of the Voice and Speech Laboratory of the Massachusetts Eye and Ear Infirmary (MEEI).

Method The database was searched using key words ‘voice recognition’ and four cases were identified from a total of 4800.

Results In all cases, the VRS was supplied to assist individuals with ULDs who could not use conventional input devices. Case reports illustrate time of onset and symptoms experienced.

Conclusion The cases illustrate the need for risk assessment and consideration of the ergonomic aspects of voice use prior to such adaptations being used, particularly in those who already experience work-related ULDs.

Key words Software; speech; upper limb disorder; vocal dysfunction; voice recognition.

Setting All four cases were patients referred to the Voice and Speech Laboratory of the Massachusetts Eye and Ear Infirmary (MEEI), Boston, MA, for evaluation. This tertiary centre receives referrals from physicians across Boston and the wider Massachusetts area. Patients undergo an initial assessment consisting of history, clinical examination and investigations such as acoustic testing, electroglossotography and aerodynamic testing. Cases are reviewed in a multidisciplinary case conference and a treatment plan developed. Post-treatment assessment may be undertaken at the referral centre closer to the patient’s home or at MEEI. From the database of >4800 clinical patients referred to the centre, the cases were identified by use of the search words ‘voice recognition’ and their analysis represents a retrospective review.

Case report 1
A 27-year-old female computer programmer started using interrupted voice recognition software (VRS) following diagnosis of carpal tunnel syndrome. She reported a sudden onset of symptoms in September 1997 associated with VRS use 8 h/day. Symptoms included sore throat after 1–2 h use and after work when speaking with friends, a gravel like quality of voice and straining to speak on several occasions per day. The symptoms had never occurred previously, but once developed were reported to be worse at the end of the working day. She neither smoked nor drank any alcohol, but had regular exposure to an air-conditioned working environment. Her voice use at home consisted of normal conver-
sation, talking on the phone and singing. At work, she used the VRS for 8 h/day, 5 days/week and also spoke in normal conversation and infrequently as a lecturer. Socially, she had normal conversation, sometimes extended conversations and would infrequently talk over noise (e.g. in a noisy bar). The symptoms had no impact on her socially or on her ability to do her job.

Endoscopic evaluation revealed hyperfunction of the pharyngeal muscle, with intermittent antero-posterior (AP) glottal tension and hyperfunction of the intrinsic musculature during normal respiration.

She was treated with voice therapy, which raised the intensity of her voice from 63 to 68 dB, altered its average frequency from 169 to 206 kHz, decreased jitter and shimmer and improved subglottic pressures. The result of the voice therapy was to produce a voice which sounded normal to the listener. At re-evaluation at the laboratory, the clinical opinion was that she now had a normal larynx and voice.

**Case report 2**

A 31-year-old female disabled writer described sudden onset of a tired voice, hoarseness and difficulty talking over people, which coincided with the start of use of an (unspecified) voice-activated computer system. The patient had had a single episode of a voice problem previously, which had fully resolved. Symptoms at presentation included a hoarse quality to the voice, vocal fatigue, frequent throat clearing and dry throat, trouble speaking out loud and throat irritation. There was no daily or seasonal variation. Past medical history revealed tonsillectomy and adenoidectomy, asthma, multiple chemical sensitivities and repetitive stress injury. She was known to be allergic to wheat, tomatoes and dairy products. There was no history of current alcohol consumption or smoking, but she worked in an air-conditioned environment. Her voice use at home consisted of normal conversation, singing and use of a speaker-phone. At work, she would talk in normal and extended conversation, talked over noise and lectured, in addition to using the VRS. The symptoms did not affect her socially, but she found difficulty using the computer because of lack of recognition due to her voice problems. The patient was treated with voice therapy, but the database did not describe any results or re-evaluation.

There was no daily or seasonal variation in symptoms. She suffered no allergies and had no alcohol or smoking history; however, she was regularly exposed to air conditioning. She used her voice at home for normal speech and talking on the phone, infrequently having extended conversations, talking over noise and singing. The vocal demands at work consisted of 9 h/day, 5 days/week, using an unspecified speech recognition system, normal conversation, talking in some extended conversations on the phone and infrequently talking over noise. The symptoms produced no impact on her socially, but she felt she could work longer hours if her voice were stronger.

Investigations revealed oedema of both vocal cords, severe AP glottal compression and moderate to severe AP supraglottal compression with anterior translocation of the pharynx when speaking. The symptoms were consistent with bilateral vocal fold oedema and vocal hyperfunction. She was treated with short-term voice therapy, focusing on improvement of co-ordination of respiration and phonation, improving voice placement and improving respiratory support for speaking. Formal re-evaluation was not performed at the laboratory, but clinically she improved with treatment.

**Case report 3**

A 25-year-old female computer programmer presented with gradual-onset vocal fatigue and increased throat clearing. It was observed at assessment that she spoke too long on the single breath, with a resulting gravel-like voice. There were no previous voice problems, but she did have a history of repetitive strain injury and migraines.

**Case report 4**

A 39-year-old female assistant director–researcher/writer using continuous speech VRS presented with a 6 week history of vocal fatigue, odynophobia, soreness and tightness in her throat and tongue. She also reported pain on palpation of the thyrohyoid space. There was no past history of previous problems, but she suffered from hypothyroidism, Ménière’s disease and repetitive strain injury. The latter had initiated her use of the VRS—for 7 h/day, 5 days/week. There was no seasonal or daily variation of her symptoms and no history of alcohol consumption or smoking. She did not work in an air-conditioned environment.

At home, she used her voice in normal and infrequently extended conversation, talking over noise and on the telephone. At work, she usually had normal conversation and infrequently extended conversation, talking over noise and talking on the phone. The impact of her symptoms socially caused her to avoid talking to people and at work she found it difficult to use her computer system, again because of problems with recognition.

On investigation, endoscopy revealed bilateral supraglottal oedema, gastro-oesophageal reflux and vocal hyperfunction. She was treated with individual voice therapy, focusing on decreasing harsh glottal attacks, improving co-ordination of respiration and phonation and achieving loudness better for respiratory support and a comprehensive vocal hygiene programme to reinforce...
reflex precautions. Again, she was not re-evaluated at the laboratory, but had a clinical improvement at discharge.

Discussion
The outline of information collected for each patient illustrates important points in the documentation of a history with regard to patients with voice dysfunction. It is essential that social and hobby use of voice is recorded, that symptoms of gastro-oesophageal reflux, alcohol consumption and smoking history are sought and that impact of symptoms is documented. As with other occupational conditions, voice disorders may also show a pattern of worsening at the end of the working day and towards the end of a working week and are relieved by weekends and holidays.

The four case reports describe individuals who had been supplied with VRS as an aid following the development of a work-related ULD or repetitive strain injury (‘RSI’). No other cases of VRS-related voice disorder were found on the database of >4800 patients. There was a preponderance of reporting sudden onset soon after the equipment was first used and a variety of symptoms, with a hoarse voice and sore throat being the most common at presentation. Subjective symptoms were confirmed by objective findings on nasal endoscopy and testing, as described above.

The technology and applications of VRS are rapidly developing. Early systems only worked with interrupted speech, but the development of more advanced software has meant that systems can now be used with more natural, fluent, continuous speech. Unfortunately, the type of VRS was only recorded in two of the four cases—one occurred with the interrupted and one with the fluent type product.

A review of the medical literature (Medline 1966–2001) revealed very few papers concerning the medical and occupational health aspects of VRS use or reports of ill-health. However, there were many papers describing the potential and early use of VRS in a variety of clinical settings, such as in reporting of radiographs in radiology departments [1,2], in ophthalmic surgery [3], in pathology reporting [4] and in clinical outpatient departments [5]. In all of these instances, the reason for moving to VRS was efficiency gains, as most people speak faster than they can type. One author, Gale [2], sounded a note of caution that apparent cost savings were offset by lost productivity by the radiologist because of the time needed to learn to use the system effectively. However, as the systems develop, the time spent teaching the system to recognize the voice and commands is likely to reduce and VRS will be more commonly seen in workplaces, especially in health care, as a routine input tool rather than as at present, where in most workplaces it is used primarily to assist individuals with ULDs in computer work.

The literature review identified two papers which reported on potential health problems related to VRS use. Kambeeyanda et al. [6] reported potential problems with voice and VRS use in 1997. They carried out a preliminary study using data gathered from survey dissemination and clinical studies and sought experience from users on the Internet. They concluded that individuals with ULDs were most susceptible to voice problems with VRS and postulated that the mechanism may be that when they used discrete (or interrupted) speech recognition systems, they had a tendency to maintain constant pitch, volume and inflection, thus keeping the vocal musculature tense and in a fixed position. Maintaining this position for an extended length of time leads to fatigue and, eventually, injury. They recommended that VRS users were informed about abnormal speech patterns, vocal hygiene, warming up and cooling down in a similar way to prior to exercising other muscles in the body and used a variety of input devices, not just VRS. Their paper reported that the commonest symptoms were hoarseness and sore throats, the same symptoms as the four cases reported here.

The second paper, by Haxer et al. [7], describes the evaluation and treatment of a single case of vocal dysfunction following VRS use and reports that the authors’ opinion is that the use of VRS makes high vocal demands.

The four cases from the MEEI database all occurred in individuals with ULDs, who had been supplied with the VRS to aid their disability. At the time of the study, VRS was largely supplied to this group of workers and its use on grounds of efficiency was only emerging. Should we therefore be surprised that reported cases of VRS seem to be occurring in individuals with established ULDs? Probably not, given that this is currently the main target group of consumers actually using the product. But alternative explanations could also be that individuals with one ULD or overuse syndrome have a ‘susceptibility’ to overuse syndromes generally, or they may have different behaviours relating to thresholds to present at clinics requesting treatment. It would be useful to be able to calculate prevalence rates for the condition in VRS users; however, there are no data available on number of users, number of VRS systems sold, or on the numbers of users per system.

At a practical level, there are some important points for practising occupational physicians, as follows.

1. Occupational voice disorders can occur through overuse in relation to VRS.
2. Users need to be given information and trained not only in the setting up and use of equipment, but also in warming up their voices, not shouting, changing pitch and flexion of voice, maintaining adequate hydration and avoiding irritation of the airways, e.g.
through tobacco smoke. These recommendations have been commonplace for a number of years for professional voice users such as opera singers and actors. Users also need to be aware of the need to take adequate breaks and to report symptoms such as sore throats and hoarseness.

3. VRS may be advocated as a major step forward as a solution for individuals with disabilities [8], but developing a voice disorder causes difficulties for the system in recognizing an individual’s voice and in operating effectively. Inability to use keyboard, mouse and VRS can render a user unable to do their work at all and so risk-assessment and monitoring following initiation of use would be sensible precautions.

Conclusion

The author believes this is the earliest reported case series of voice problems arising in users of VRS. Occupational health advisors need to consider vocal repetitive tasks in their ergonomic assessments, particularly where VRS is provided in the presence of a pre-existing ULD. It is also recommended that training, early recognition, reporting and monitoring systems are also put into place when VRS is in use.

Acknowledgements

The author would like to thank Drs R. Hillman and G. Bunting of MEEI and Harvard Medical School for their assistance. The author travelled to Boston on a Golden Jubilee Travelling Fellowship awarded by the Society of Occupational Medicine in May 2001.

References