URGENCY IN SPEECH WARNINGS

J. Edworthy, E. Hellier,
University of Plymouth, United Kingdom

Introduction It is well established that primary acoustic parameters such as pitch, speed, level, spectral content and other variables affect the perceived urgency of nonverbal auditory warnings (Momtahan, 1990; Edworthy, Loxley & Dennis, 1991; Hellier, Edworthy & Dennis, 1993; Haas & Casali, 1995; Haas & Edworthy, 1996). Findings that increases in, for example, pitch, speed and level increase the perceived urgency of nonverbal warnings are replicable and well substantiated. These findings can be implemented into the design of warnings systems. For example, prioritized warnings systems whereby more urgent-sounding warnings signal more urgent or immediate situations and hazards are ergonomically possible. Furthermore, with the development of intelligent computer monitoring, decision-making, and neural networks in particular, intelligent warning systems where warnings can be appropriately matched to the current urgency of a situation.
Today's availability of computing power has also made it increasingly possible to incorporate speech warnings systems rather than nospeech warnings systems into complex, noisy environments. While speech warnings systems may have some advantage in that they can be less ambiguous than nonspeech warnings systems, they are beset with other problems. For example, speech systems may be less intelligible than nonspeech systems in the same noise background. People may become over-reliant on speech systems because speech is directly understandable, but the system may still subject to false alarms. In some situations, messages make take longer to deliver than is appropriate when situations require immediate action, whereas a nonspeech system may be able to cut through more decisively and immediately. However, they do have advantages and the concept of perceived urgency in speech warnings has received research attention. Some of this research is reviewed in this paper, together with some practical observations as to how research findings might be integrated into speech warning system design.

Comparison of speech and nonspeech warnings Typically, an old and out-of-date nonspeech auditory warning system is replaced by a new, more technologically advanced, speech warnings system. Results of empirical comparisons of such systems will therefore invariably suggest that speech warnings systems outperform nonspeech systems. However, in terms of perceived urgency there appears to be little difference in the relative urgencies of speech and nonspeech warnings overall. Nonspeech warnings can be made different from one another through variation across acoustic design parameters such as temporal pattern, pitch contour, number of pulses and so on. Speech messages can have different content. Within a particular warning sound or message, urgency can be altered by altering pitch, speed and level (as well as other, secondary variables). In order to preserve intelligibility and believability, speech warnings cannot be as radically altered individually as can nonspeech warnings. However, if speech and nonspeech warnings are compared in an ergonomically rational manner, they are about as urgent as one another (Edworthy, Walters, Hellier & Weedon, 2000a). The main point is that it appears that the perceived urgency of speech messages can be manipulated in a similar way to nonspeech messages. This is important, because urgency mapping appears to be a performance advantage for both (Edworthy, Hellier, Walters, Weedon & Adams, 2000b).

Acoustic influences in perceived urgency in speech messages It is well known that emotion in the voice is conveyed through changes in the acoustic parameters of speech (Scherer, 1986), therefore we would expect urgency (which we can think of as a corollary of fear in the voice) to be conveyed through changes in the ways words and sentences are spoken. A number of studies show that the listener reliably interprets cues from the speaker, and that these cues are underpinned by acoustic variables. For example, studies by Barzegar & Wogalter (1998; 2000) show that if a speaker is asked to say a word, or sentence, in either an urgent or non-urgent way then this is picked up by the listener and their intended carefulness ratings follow (participants rate themselves as having higher intended carefulness when words and sentences are spoken in an urgent, rather than a nonurgent, way). Hollander & Wogalter (2000) applied some of the parameters known to affect perceived urgency (after Edworthy et al, 1991) to spoken warnings and found similar effects. One interesting finding was that the word 'Deadly' was relatively unaltered by acoustic manipulation, a finding which has been replicated (Hellier et al, 2002) and hints at the complex interaction between the semantics of the word and its acoustic presentation (see below).

Hellier et al (2002) replicated many of the earlier findings, and also carried out a detailed acoustic analysis of the frequency, intensity and speed of the utterances produced by the speakers. This analysis showed that 'urgent' and 'nonurgent' utterances produced by a male and female speaker differed considerably on all three variables, and that the magnitude of the differences in the acoustic variables mirrored the magnitude of the differences in perceived urgency ratings. We also showed that the perceived urgency of synthetic utterances could be reliably predicted and manipulated through applying systematic acoustic changes to the utterances heard, in a way that had been achieved earlier for nonspeech warnings (Hellier et al, 1993). Furthermore, as with nonspeech auditory warnings, detailed experimentation with a small set of warning words (Hellier, Weedon, Edworthy & Walters, 2000) shows that Steven's Power Law (Stevens, 1957) can provide a good fit for the relationship between changes in acoustic parameter and changes in urgency rating for some of the most fundamental parameters such as pitch and speed. Again, the word 'Deadly' showed itself to be less influenced by acoustic manipulation than other words.

Significant in these psychophysical scaling studies of speech warning words is that the Stevens' Power Law exponents derived tend to be higher than those derived for nonspeech warnings. Thus it takes only a small change in acoustic variable to produce a large effect on the perceived urgency of a word, whereas a larger change is necessary for nonspeech warnings. However, this does not necessarily mean that speech warnings are able to convey a larger range of urgencies than nonspeech warnings (Edworthy et al, 2000a suggests that they do not) as the potential range for those acoustic variables is much smaller for speech than for nonspeech stimuli. Large changes tend to produce utterances, which are unintelligible, or unbelievable, or unnatural-sounding. Thus although the effects are large, the range within which the designer can sensibly operate is quite small.

Semantic effects Although nonspeech warnings can potentially have semantic meaning, this is much more obvious with speech warnings. We can expect therefore for the meaning of the message to have some influence on the perceived urgency of speech. Research on the interaction between acoustics and semantics has tended to focus on a small set of warning signal words (words such as 'Deadly', 'Danger', 'Warning', 'Attention', 'Note' and so on) because it is already well established that in the visual domain these words are rated reliably and consistently different from one another. Words such as 'Deadly' and 'Danger' produce higher ratings of urgency or arousal strength than words such as 'Caution' and 'Note' when those words are presented in the same visual format (e.g. Wogalter & Silver, 1995). Studies also show that the ratings arrived at are fairly resistant to both the scaling method used and the

possibility of anchoring and bias effects (Hellier, Wright, Edworthy & Newstead, 2000a). Studies of speech warnings have shown similar rankings for these words when presented auditorily. For example, Edworthy et al (in press) showed that approximately the same ordering of words is found when the words are presented in a similar acoustic fashion (when words are presented through a synthetic voice, where the acoustic variables are controlled as far as is possible). Edworthy et al's study also demonstrate the presence of phonetic effects in warning signal words – for example the word 'Lethal' is rated more highly when presented visually than when presented auditorily. To some extent natural speakers can compensate for this lack of emphasis by speaking the word louder, but the word particularly suffers when presented synthetically.

Particularly interesting and relevant in the design of speech warnings systems is the interaction between semantics and acoustics. A replicable finding is that the word 'Deadly', which has unambiguous meaning, is more resistant to manipulation through acoustics than other words such as 'No'; the latter is a more plastic word than the former and so its urgency can be changed more dramatically through acoustic manipulation. Hellier, Weedon, Edworthy & Walters (2000b) showed that the Stevens' Power Law exponents relating urgency to acoustic changes were different for different words, suggesting again that semantics and acoustics interact. There is scope for much further research on the interaction between semantics meaning, acoustic presentation and phonetic structure in speech warning messages. Of interest also are longer speech messages (explored by Barzegar & Wogalter, 2000), as warning signal words are likely to be only one component of more complex speech warning messages.

Males versus female voices In the design of speech warnings systems, there is much practical interest in the selection of the sex of the speaker used. There is little evidence to suggest that one should be chosen over another for other than acoustic reasons. There appear to be minimal intelligibility differences between male and female voices when appropriately compared (for example when the voices are equally loud), the selection of the voice should depend on the noise spectrum within with the voice is to fit. For example, a noise spectrum with lots of high frequency noise would be better suited to a male voice, and a spectrum with a lot of low frequency noise would be better suited to a female voice.

Differences between male and female speakers have been observed in perceived urgency judgements, but these have been shown to be almost completely underpinned by acoustic variables. For example, a number of the studies (Barzegar & Wogalter, 1998, 200; Hellier et al, 2002; Edworthy et al, in press) found that female voices overall produced higher urgency ratings than male voices. In addition, the female voice typically produces a higher range of urgencies between 'urgent' and 'nonurgent' utterances. This effect seems largely to be a result of the higher pitch, and the greater pitch range, of the female voice. Males voices can compensate to some extent by speaking more loudly (some of the experimental evidence on this topic suggests such a conclusion), suggesting again that urgency judgements are underpinned by acoustic variables. As the female voice appears to cover a greater range of urgencies, this would suggest that, if there are no particular difficulties with the noise spectrum, that a female voice might be preferable to that of a male. This is particularly relevant as many of the transformations applied to speech in order to fit it to a spectrum and to make it intelligible (such as filtering, peak clipping etc) will tend to remove emotional nuance from both real and synthetic voices.

The findings for urgency also suggest that judgements exist along a continuum, mapping the acoustic continua that underpin them. This in turn suggests that for the acoustic components of speaker sex recognition that judgements might also take place along a continuum (Mullenix, Johnson, TopcuDurgun, & Farnsworth, 1995).

Summary and design guidelines Perceived urgency in speech warnings is predictable and manipulable in much the same way that it is for nonspeech warnings. Typically, only small changes in acoustic values produce large changes in perceived urgency, although the range over which the acoustic parameters can vary is small in comparison to nonspeech warnings. The urgency of speech message can be enhanced (or reduced) by altering its pitch, speed, and level. For some warning signal words, Stevens' Power Law exponents exist showing the precise relationship between acoustic changes and perceived urgency changes. The semantics of speech warnings also has an effect, which remains even when speech is presented in an acoustically controlled manner. There is a set of warning signal words which has been explored at length and for which there are known semantic effects. For some words, the way in which they are presented acoustically has a large influence on their perceived urgency while for other less plastic words acoustic effects are small. The differences between male and female voices, at least in terms of perceived urgency, seem to be completely underpinned by acoustic differences. Perceived urgency ratings of words spoken by females tend to produce a larger range of responses than those spoken by males, suggesting that there may be some advantage to using female voices in warnings systems, particularly when the voice is synthesized. However, the overlap between noise spectra and speech spectra should be the primary concern in deciding whether to use a male or a female voice in such systems.

Keywords: auditory warnings; speech communication; perceived urgency

References

Barzegar, R. S. & Wogalter, M.S. (2000). Intended carefulness ratings for voiced warning statements. Proceedings of the IEA 2000/HFES 2000 Congress, 3, 686-9. Human Factors and Ergonomics Society: Santa Monica, CA.

Barzegar, R.S. & Wogalter, M.S. (1998). Effects of auditorily-presented warning signal words on intended carefulness. In M.A. Hanson (Ed.), Contemporary Ergonomics 1998 (pp. 311-315). London: Taylor & Francis

Edworthy, J, Hellier, E J, Walters. K, Clift-Matthews, W & Crowther, M (in press) Acoustic, semantic and phonetic influences in spoken warning signal words. To appear in Applied Cognitive Psychology

Edworthy, J**,** Hellier, E., Walters, K, & Weedon, B. (2000a) Comparing Speech and Non Speech Warnings. Proceedings of the IEA 2000/HFES 2000 Congress, 3, 746-9. Human Factors and Ergonomics Society: Santa Monica, CA.

Edworthy, J, Hellier, E J, Walters, K, Weedon, B, & Adams, A S (2000b) The relationship between task performance, reaction time, and perceived urgency in nonverbal auditory warnings. Proceedings of the IEA 2000/HFES 2000 Congress, 3, 674-677. San Diego, California: Human Factors and Ergonomics Society.

Edworthy, J., Loxley, S. & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. Human Factors, 33(2), 205-231.

Haas, E. C. & Edworthy, J. (1996). Designing urgency into auditory warnings using pitch, speed and loudness. Computing and Control Engineering Journal, 7(4), 193-198.

Haas, E.C. & Casali, J.G. (1995). Perceived urgency of and response to multi-tone and frequency-modulated warning signals in broadband noise. Ergonomics, 38(11), 2313-2326.

Hellier, E J, Edworthy, J, Weedon, B, Walters, K & Adams, A S (2002) The perceived urgency of speech warnings: Semantics versus acoustics. Human Factors, 44(1), 1-17.

Hellier, E. J., Wright, D. B., Edworthy, J. & Newstead, S. (2000). On the stability of the arousal strength of warning signal words. Applied Cognitive Psychology, 14, 577-592

Hellier, E J, Weedon, B, Edworthy, J & Walters, K (2000) Using psychophysics to design speech warnings. Proceedings of the IEA 2000/HFES 2000 Congress, 3, 698-701. San Diego, California: Human Factors and Ergonomics Society.

Hellier, E.J., Edworthy, J. & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. Human Factors, 35,(4), 693-706.

Hollander, T.D. & Wogalter, M.S. (2000). Connoted hazard of voiced warning signal words: an examination of auditory components. Proceedings of the IEA 2000/HFES 2000 Congress, 3, 702-5. Human Factors and Ergonomics Society: Santa Monica, CA.

Momtahan, K L (1990) 'Mapping of psychoacoustic parameters to the perceived urgency of auditory signals'. Unpublished master's thesis, Carleton University, Ottawa, Ontario, Canada.

Mullenix, J W, Johnson, K A, TopcuDurgun, M, & Farnsworth, L M (1995) The perceptual representation of voice gender. Journal of the Acoustical Society of America, 98(6), 3080-95.

Scherer, K.R. (1986). Vocal affect expression: A review and a model for future research. Psychological Bulletin, 99(2), 143-165

Stevens, S S (1957) On the psychophysical law. Psychological Review, 64, 153-81.

Weedon, B., Hellier, E., Edworthy, J**.** & Walters, K. (2000) Perceived Urgency in Speech Warnings. Proceedings of the IEA 2000/HFES 2000 Congress, 3, 690-3. Human Factors and Ergonomics Society: Santa Monica, CA.

Wogalter, M.S. & Silver, N.C. (1995). Warning signal words: connoted strength and understandability by children, elders and non-native English speakers. Ergonomics, 34, 79-89.