PROTECTING MINERS' HEARING WHILE FACILITATING COMMUNICATION

R. E. Bou Serhal, *J. Voix
École de technologie supérieure
1100 Rue Notre-Dame Ouest
Montréal, QC H3C 1K3
*(*Corresponding author: jeremie.voix@etsmtl.ca)*

T. Falk
Institut national de la recherche scientifique
800, de La Gauchetière Ouest
Portail Nord-Ouest, bureau 6900
Montréal, H5A 1K6

PROTECTING MINERS' HEARING WHILE FACILITATING COMMUNICATION

ABSTRACT

Many miners are exposed to dangerously high levels of noise on a daily basis. Over the past 15 years, a continually increasing number of occupational hearing loss has been reported from the mining community in the United States, of which, more than 95% is attributed to prolonged noise exposure. Although the noise exposure levels may differ between coal miners and metal and non-metal miners, in the absence of noise control at the source, the solution is the same: use of personal hearing protection devices (HPD). While protecting the miners' hearing it is also essential to no longer hinder their ability to communicate. With access to an advanced HPD that is customized to the miners' ears we are able to combine these two requirements. Using an intra-aural instantly custom molded HPD miners are protected from high levels of noise. The HPD is equipped with wireless capabilities, and contains both a speaker and an In-Ear Microphone (IEM). Therefore, the miners' speech may be captured from inside the ear and transmitted to the remote listener. This IEM signal is relatively noise-free since it is isolated from the background noise. The IEM speech signal, however, is "boomy" and is missing some high frequency content, making fricative consonants hard to understand. Nonetheless, the IEM speech signal is correlated with the natural speech signal and may be manipulated through statistical techniques to more closely resemble natural speech. By improving the intelligibility and quality of the IEM signal, numerous applications may be enabled. One use of the enhanced IEM signal will be for radio communication. Using wireless radio communication in a noisy mining environment is sometimes the only practical and affordable solution to allow communication between miners equipped with personal hearing protection devices. Traditionally, one of the weaknesses of such wireless radio communication lies in the lack of designating receivers: whether they are the intended receiver or not, all those carrying a radio receiver are subjected to the broadcasted signal. The current work will detail a new concept of a "radio-acoustical virtual environment" where the radio signal will only be received by miners within a given spatial range, such range depending on the user's vocal effort as well as the ambient and perceived background noise levels.

KEYWORDS

Hearing Protection Devices, Communication, Wireless, Radio

INTRODUCTION

Miners are among over 30 million workers in North America who are exposed to excessive levels of noise that put them at risk of losing their hearing (National Institute of Occupational Safety and Health, 1998). A study of metal and non metal miners across the united states reported over 95% of hearing loss reported to the Mine Safety and Health Administration (MSHA) was caused from prolonged noise exposure (Valoski, 1997). This is unfortunate as noise-induced hearing loss is a serious yet preventable health hazard. The Occupational Safety and Health Administration (OSHA) proposes the following three methods of protecting workers from noise exposure (Katz et al., 2009):

1. engineered reduction of the noise
2. limiting exposure time
3. use of personal hearing protection

The mining environment and current practices have made it difficult to prevent hazardous noise exposure to miners. Noise control i.e. the engineered reduction of noise is expensive and requires the attention of the higher management. New materials and enclosures have been developed to decrease the noise levels of some equipment. Even making sure that equipment is well maintained can aid in controlling the noise at the source (McBride, 2004). However, noise control can only go so far in limiting noise exposure as some exposure to loud noise, such as the impact from a drill bit, are inevitable. Limiting exposure time has also been unsuccessful. The allowable limit as set by the National Institute of Occupational Safety and Health (NIOSH) is 85 dBA for eight hours exposure (Berger, 2003). Yet studies have shown that 80% of U.S miners are exposed to a Time-Weighted Average (TWA) over 85 dBA, and of those, 25% are exposed to a TWA exceeding 90 dBA (McBride, 2004). The final solution is the use of personal Hearing Protection Devices (HPD). HPDs come in many different shapes and sizes and can be made from a variety of materials. The two main types of HPDs are intra-aural i.e. earplugs, and supra-aural i.e. earmuffs (Berger, 2003). There are a couple of points to consider when discussing HPDs: the comfort and the effectiveness of the personal HPD. Using HPDs that are comfortable to wear for an extended period of time is vital because an uncomfortable fit is more likely to drive the user to remove the HPD. It is also important to properly wear HPDs because an improper fit leads to misrepresented attenuation, causing the user to be unknowingly unprotected. Both of these issues may be resolved by a custom molded HPD that allows for a way to monitor the real attenuation inside the ear (Voix and Laville, 2009). The problems that arise with the use of HPDs in mining environments is twofold: the acoustical environment of mining and, as a consequence, difficulties in communication. Depending on the the Signal-to-Noise Ratio (SNR), HPD's can be detrimental to communication. Fernandes (2003) reports that in environments with +5 dB SNR and +10 dB SNR , wearing hearing protection decreases the intelligibility of speech. However, at -5 dB and -10 dB SNR, wearing hearing protection can increase speech intelligibility by up to 10%. Therefore, for environments where noise is intermittent, such as mining, wearing HPDs deteriorates communication and users are more likely to seek out forms to better communicate. Currently there are several different ways that are used to communicate in noise while using HPDs, one could:

1. Remove the HPD: get closer to a listener and adjust vocal effort to communicate. Removing an HPD to communicate is problematic as the effectiveness of HPDs is greatly reduced with non-continuous use (Berger, 2003). It also requires the miners to be in close proximity of one another to communicate.
2. Use passively filtered HPD: flat attenuation HPDs could be beneficial for speech communication as they do not attenuate high frequencies as much as other HPDs. However, in noise, these HPDs are not as effective as they usually do not provide sufficient attenuation. In quiet, they also decrease speech intelligibility, which would compel the wearer to remove the HPD for communication.
3. Use a hand-held radio device over HPDs: use of a walkie-talkie allows for distance communication with multiple people while remaining stationary. Using a hand-held radio overcomes the problem of proximity but still requires the removal of the HPD.
4. Use of a communication headset: usually an earmuff with a miniature loudspeaker and an external boom microphone. The voice picked up by the boom microphone is transmitted through either a wired or wireless network to a remote listener. Although these are the best current alternative, these headsets still present the following inconvenience: the external microphone will not only pick up the user's voice but it will as well pick up the background noise, which dramatically affects intelligibility.

Another issue associated with using any kind of radio transmitter, is that it does not distinguish a receiver and all communication is sent to everyone on the same radio channel. Therefore, the users' radio is often flooded with irrelevant conversation that could be annoying and somewhat loud and thus contributing to the noise dose. For underground miners, using radio communication in general is problematic. Electromagnetic wave propagation in underground mines is complex, rendering wireless communication a difficult task (Moutairou et al., 2009). Clearly there is a need for a device that provides good noise attenuation as well as good communication without compromising the performance of one for the other.

Proposed Approach

We propose a new concept called "Radio Acoustical Virtual Environment" (RAVE) in which miners can achieve intelligible communication without hindering their hearing protection. RAVE uses an advanced intra-aural instantly custom molded HPD, shown in Figure 1, equipped with an In-Ear Microphone (IEM), a miniature loudspeaker, a Digital Signal Processor (DSP), an Outer-Ear Microphone (OEM) and Wireless Radio (WR) capabilities. Such a device can capture a somewhat undisturbed speech signal from inside the ear (referred to as IEM speech). Because the signal captured originates from bone conducted vibrations, it lacks higher frequencies. Thus, the IEM signal must first be enhanced in its high frequency content. Once enhanced, the IEM signal is coded and sent to an appropriate radius of listeners based on the acoustical features of the produced speech and the level of background noise. Figure 2 demonstrates the anticipated performance of RAVE. If a miner is speaking at 70 dBA SPL in a quiet environment the radio signal will be transmitted to anyone within a 20 m radius. As the level of noise increases and the vocal effort of the speaker remains constant the transmitting distance will decrease. Therefore, in an extremely noisy environment the transmitting distance of the radio will only be 5 m to compensate for such phenomena as the Lombard effect.

This paper introduces the concept of RAVE and the methodology involved in realizing such a protocol. The next section discusses different techniques available for the enhancement of the IEM speech signal followed by the concept of vocal effort coding. Finally, limitations and promising avenues are discussed.

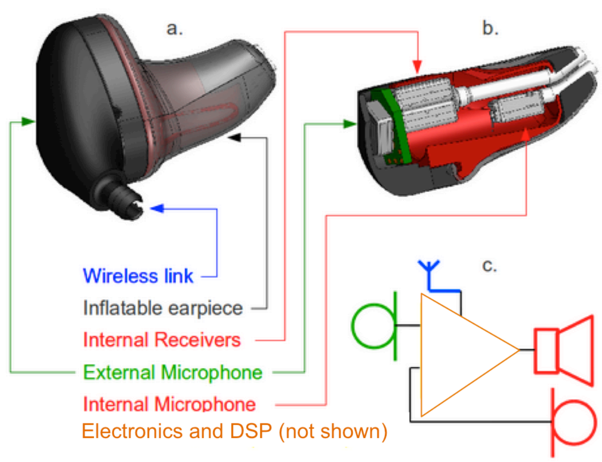

Figure 1 – Overview of digital custom earpiece (a), its electroacoustical components (b), and equivalent schematic (c).

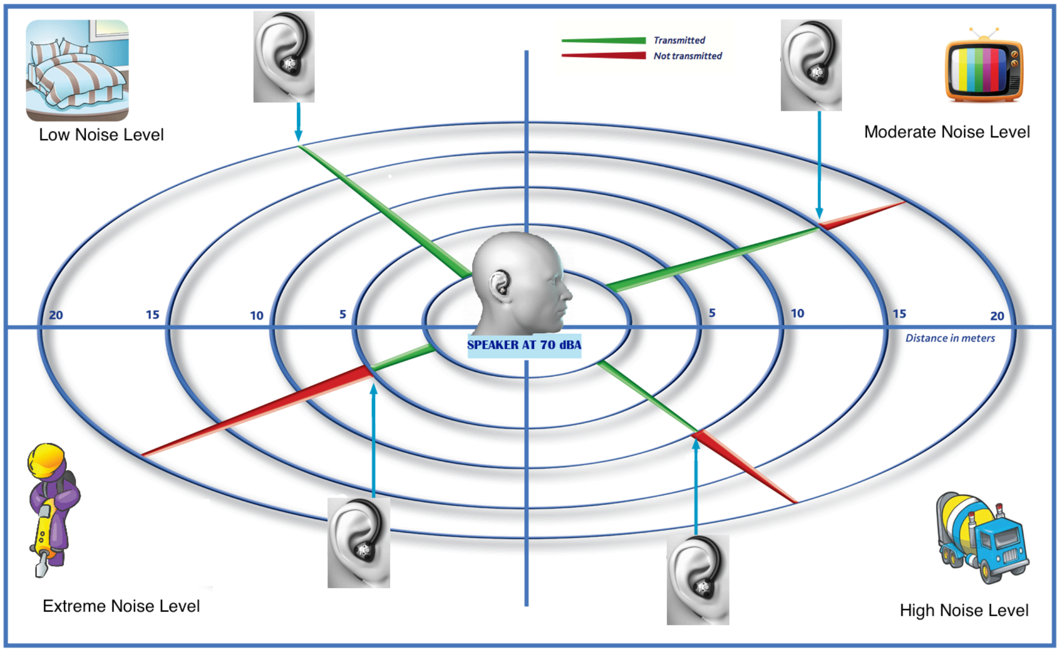

Figure 2 – Illustration of functionality of RAVE. The green and red lines represent the areas where the signal is transmitted and not transmitted, respectively.

ENHANCEMENT OF THE IEM SPEECH

When speech is captured conventionally (with a boom microphone), to be sent over a radio network in a noisy environment, it is disturbed and contains the noise picked up by the exposed microphone, even when using a directional microphone. On the other hand, capturing speech from inside the protected ear allows for the transmission of a less-disturbed speech signal that will not require extra de-noising, usually achieved by the electronics within the radio. When the ear canal is blocked by an in-ear device, there is a regeneration of the speech inside the ear canal and one experiences what is called the *occlusion effect* (Berger, 2003). The *occlusion effect* allows for the capturing of speech inside the ear, which is useful in noisy environments. Because of cranial bone conduction, this signal is "boomy", containing most of its energy in the lower frequencies while missing important high frequency content (Bernier and Voix, 2010). The difference between the frequency content of the IEM speech and the OEM speech (referred to as REF) of the utterance /u/, for a male speaker, is demonstrated in Figure 3. From Figure 3, it we notice that above 1.8 kHz the IEM signal is missing important high frequency content. As a consequence of the IEM signal's limited bandwidth, fricative consonants such as /s/ and /f/, and nasals such as /n/ and /m/ are unintelligible. The IEM signal is thus perceived as having lower quality and intelligibility than "free air speech", or speech that is recorded near the mouth. To solve this, the IEM signal could be expanded using Bandwidth Extension (BWE) of the speech signal. Many different BWE techniques exist, and the proper choice depends on the desired results and available resources. BWE can range from spectral estimation and expansion through excitation signal extension, to Vector Quantization (VQ) and codebook mapping. Iser et al. (2008) give a good review of the basics of such techniques (Iser et al., 2008). In the past, the need for BWE arose because of the limited bandwidth of the telephone network. The narrow bandwidth of a telephone is about 3.5 kHz leaving some significant parts of human speech unrepresented. In this context, wideband signals refer to signals that can represent the entire vocal range while narrowband signals can only represent a limited part of the vocal range. With access to an IEM and an OEM, BWE can be used for our purposes by treating the IEM signal as the narrowband signal and the free-air speech captured by the OEM as the wideband signal. All available techniques for BWE are listed in Figure 4. It is important to assess the resources available to choose a practical and efficient technique with good performance. Some things to consider are the computational complexity and cost of the algorithm, power consumption and whether the algorithm will be speaker dependent or speaker independent. *Excitation signal extension* and *spectral envelope expansion* could be used for speaker independent BWE. Quality may be increased with speaker dependent techniques using spectral envelope expansion at a cost of some practicality. When speaker dependent algorithms are used the miner must train the algorithm. Although speaker dependent algorithms may lead to better quality reconstructed speech, they are less robust when compared to speaker independent algorithms. Small variations in speech that may be caused by a common cold may lead to undesirable results. This could be palliated by making the algorithm re-trainable. However, this is impractical and may lead miners to abandoning the use of the device. It is thus important to evaluate such adverse effects and assure that the BWE algorithm used is practical, efficient, and reliable.

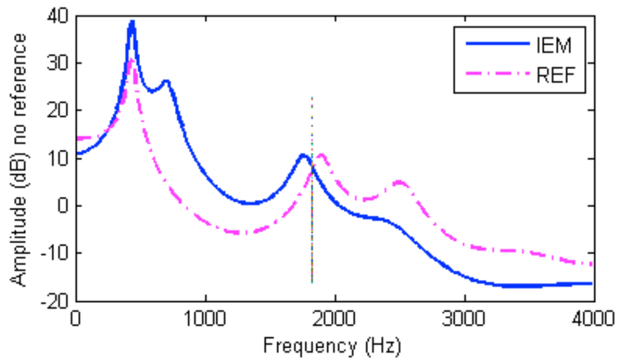

Figure 3 – IEM vs. REF spectral envelopes of the utterance /u/ from the word 'canoe', showing the increased low frequency content and the missing high frequency content.

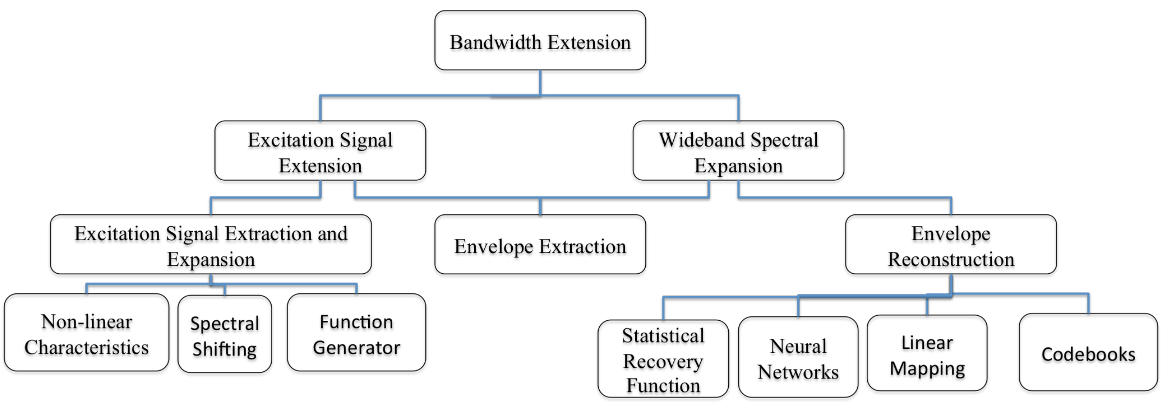

Figure 4 – Classification of different bandwidth extension techniques applicable to in-ear microphone signal pickup inside miners' ears

VOCAL EFFORT CODING

In this section we discuss the various vocal modes and their relationship with physical distance between a speaking miner and a listening miner. Naturally, human beings adjust their vocal effort to compensate for changes in their environment. One can whisper a confidential message, call out for a meeting or shout out for help. It is important to distinguish "vocal effort" from "vocal level". The latter suggests a change in Sound-Pressure Level (SPL) while vocal effort involves a lot more than just changes in SPL (Traunmüller and Eriksson, 2000). Zhang et al. classified 5 speech modes: (1) whispered, (2) soft, (3) neutral, (4) loud, and (5) shouted (Zhang and Hansen, 2007). Each of these speech modes is characterized by its deviations from the neutral speaking condition. Many studies have been done to characterize each speech mode as to enhance speaker recognition systems and other applications. In particular, whispered and shouted speech require the most dramatic change in excitation (Zhang and

Hansen, 2007) and have thus received a lot of attention. Our interest lies mostly with the shouted speech mode and the changes in acoustical features that occur. As documented by many, as the vocal effort increases so does the fundamental frequency, *F0*. Another widely accepted change in the formants is the increase of the first formant, *F1* (Liénard and Di Benedetto, 1999; Elliot, 2000; Garnier et al., 2008). Liénard et al. (1999), however, also claim an increase in the second formant, *F2*, for females but this has not yet been widely accepted. Shouted sentences have increased initial *F0* slope but a decreased final *F0* (Fux et al., 2011) and a decreased spectral slope (Zhang and Hansen, 2007). They are longer in duration which is caused by longer word duration, but have a decreased silence duration (Zhang and Hansen, 2007). Typically, shouted speech is detected based on *F0*, *F1* and the spectral tilt (Nanjo, 2009). A summary of these changes can be seen in Table 1.

Traunmüller et al. describe vocal effort as *"the quantity that ordinary speakers vary when they adapt their speech to the demands of an increased or decreased communication distance"* (Traunmüller and Eriksson, 2000). As distance increases so does the vocal effort. In fact, Brungart et al. report that as distance doubles the intensity increases by 8 dB, while Liénard et al. report that *F0* increases at 3.5 Hz/dB (Fux et al., 2011; Liénard and Di Benedetto, 1999). Distance, however, is not the only time we adjust our vocal effort. When our ability to hear our own voice changes, as a result of background noise for example, our vocal effort changes (Junqua, 1993). This is known as the *Lombard* effect. Although Lombard speech may share some characteristics with shouted speech, it is unique and cannot be treated the same way as shouted speech. Speakers vary their vocal effort based on the spectrotemporal properties of the background noise. In fact, significant differences of adjustments in the presence of white noise and babble noise have been reported (Traunmüller and Eriksson, 2000). A summary of the acoustical changes cause by Lombard speech as found by Junqua (1993) is shown in Table 1.

Bringing together the acoustical changes caused by vocal effort and those caused by the Lombard effect, will bring about a relationship between vocal effort while wearing HPD in noise and intended communication distance. Scheduled tests on a group of normal-hearing human subjects will be the starting point in the data collection involved in achieving the aforementioned relationship. Once adequate data collection is reached, we envision producing a relationship as portrayed in Table 2. The green blocks represent distances where speech is intelligible for the given vocal effort and residual background noise level under the HPD. The yellow blocks represent areas of reduced intelligibility or areas where intelligibility is achieved only with reinforcement from facial cues or gestures. Red blocks represent areas where speech is unintelligible. Note, the numbers in this table are strictly for illustrative purposes and do not yet come from research data. Once this table is compiled, the vocal effort of the speaker may be coded and sent to an appropriate radius of intended listeners through an *ad-hoc* radio system such as cognitive radios (Li et al., 2011).

Table 1 – Summary of acoustical differences between shouted speech and Lombard when compared to neutral speech .

Acoustical Feature	**Shouted Speech**	**Lombard Speech**
F0	Increased frequency	Increased frequency (more dominant in males)
F1	Increased frequency	Increased frequency (more dominant in females)
F2	Increased frequency (females only)	Increased frequency (females only)
Sentence Duration	Increased duration	Increased duration
SPL	Increased level	Slightly increased level

Table 2 – Illustrative table of relationship between vocal effort and communication distance in the presence of background noise while wearing HPD.

		Residual Background Noise (dBA SPL)				
		<60	60-70	70-80	80-90	>90
Vocal effort of speaker	Whispered	2 m	unintelligible	unintelligible	unintelligible	unintelligible
	Soft	4 m	1 m	reduced intelligibility	reduced intelligibility	unintelligible
	Neutral	15 m	8 m	1 m	reduced intelligibility	unintelligible
	Loud	20 m	10 m	1 m	reduced intelligibility	unintelligible
	Shouted	40 m	20 m	10 m	5 m	unintelligible

DISCUSSION

The "Radio-Acoustical Virtual Environment" discussed will allow miners to communicate without the need to remove their HPDs and without having to move closer to their listener. Undisturbed speech from inside the ear canal will be captured and transmitted over wireless radio to the remote listener. The transmitted signal will only be received by listening miners within a given spatial range, this range depending on the speaking miner's vocal effort and background noise level. This solves most of the issues that are currently faced by miners trying to communicate and protect their hearing, however, a few problems persist and require further discussion. Access to the auditory platform shown in Figure 1, can open up the door to a more adaptive hearing protection device.

Previously, we mentioned that wearing HPDs in quiet environments decreases intelligibility and with the current design of RAVE this problem persists. In this case, we could take advantage of the OEM and the DSP by utilizing them to monitor the environmental SPL (Mazur and Voix, 2012). If the level is safe, the internal speaker could be used to reproduce what is picked up by the OEM and bypass the HPD. If the OEM registers that the levels are unsafe then no bypass occurs and the HPD functions as previously discussed. It would be useful to have a way to manually enable the bypass of the HPD and allow the signal picked up by the OEM to pass through for communication between those wearing the HPD described and those that are not. Another issue to consider when the environment is quiet is the annoyance caused by the *occlusion effect*. In noise, we depend on the *occlusion effect* for communication, which is not problematic because the high levels of noise counteract the predominance of the *occlusion effect*. However, when trying to communicate in quiet, even when the HPD is bypassed, one's own speech is predominantly what is heard which makes it annoying for the speaker. To solve this, an active *occlusion effect* reduction system can be implemented (Bernier and Voix, 2012). The last foreseen difficulty is the use of a wireless radio for distant communication in underground mining. Research in this area is growing and many new protocols are developing. Advancements in this area (Ndoh and Delisle, 2005; Srinivasan, Ndoh, and Kaluri, 2005; Moutairou et al., 2009) could be further investigated to be implemented with our radio system, to offer the most efficient radio system available.

CONCLUSIONS

RAVE already addresses many of the issues that are faced by miners communicating in noise and is thus a better alternative to what is presently available. Good hearing protection is currently achieved at the cost of decreased communication while good communication is achieved at the cost of jeopardizing good hearing protection. Providing miners with satisfactory hearing protection and communication is still difficult and requires the compromise of one or the other. In this paper, we propose a new distance sensitive protocol that provides intelligible speech to miners wearing hearing protection. Using changes in acoustical features of speech the vocal effort will be coded and the speech signal will be sent in a way that mimics a natural acoustical environment. Providing miners with such a device will enhance their work experience and potentially promote the use of HPDs in noisy environments.

ACKNOWLEDGMENTS

The authors would like to thank Sonomax Technologies Inc. and its Sonomax-ÉTS Industrial Research Chair in In-Ear Technologies (CRITIAS) for their financial support. The authors also acknowledge the financial support from l'Équipe de recherche en sécurité du travail et en analyse des risques industriels (ÉREST) as well as from MITACS- Accelerate research internship program.

REFERENCES

Berger, E.H. 2003. *The Noise Manual*. AIHA.

Bernier, Antoine, and Jérémie Voix. 2010. "Signal Characterization of Occluded Ear Versus Free-air Voice Pickup on Human Subjects." *Canadian Acoustics* Vol. 38 (Num. 3): pp. 78–79.

Bernier, Antoine, and Jérémie Voix. 2012. "Une Nouvelle Protection Auditive Adaptée Aux Musiciens." *Fameq* 27 (1): 39–41.

Elliot, Jennifer. 2000. "Comparing the Acoustic Properties of Normal and Shouted Speech: A Study in Forensic Phonetics." In *8th Aus. Int. Conf. Speech Sci. & Tech*, 154–159.

Fernandes, J. C. 2003. "Effects of Hearing Protector Devices on Speech Intelligibility." *Applied Acoustics* 64: 581–590.

Fux, Thibaut, Gang Feng, and Véronique Zimpfer. 2011. "Talker-to-listener Distance Effects on the Variations of the Intensity and the Fundamental Frequency of Speech." *Cognition*: 4964–4967.

Garnier, Maëva, Lucie Bailly, Marion Dohen, Pauline Welby, and Hélène Lœvenbruck. "An Acoustic and Articulatory Study of Lombard Speech : Global Effects on the Utterance": 2–5.

Garnier, Maëva, Joe Wolfe, Nathalie Henrich, and John Smith. 2008. "Interrelationship Between Vocal Effort and Vocal Tract Acoustics : a Pilot Study" 2: 3–6.

Iser, Bernd, Gerhard Schmidt, and Wolfgang Minker. *Bandwidth Extension of Speech Signals*. ISBN 978-0-387-68898-5. doi:10.1007/978-0-387-68899-2.

Junqua, J C. 1993. "The Lombard Reflex and Its Role on Human Listeners and Automatic Speech Recognizers." *The Journal of the Acoustical Society of America* 93 (1) (January): 510–24.

Katz, Jack, Larry Medwetsky, Robert Bukard, and Linda Hood. 2009. *Handbook of Clinical Audiology*. 6th ed. Lippincott Williams & Wilkins.

Li, Jun, Yifeng Zhou, Louise Lamont, and Francois Gagnon. 2011. "A Novel Routing Algorithm in Cognitive Radio Ad Hoc Networks." In *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, 1–5.

Liénard, J S, and M G Di Benedetto. 1999. "Effect of Vocal Effort on Spectral Properties of Vowels." *The Journal of the Acoustical Society of America* 106 (1) (July): 411–22.

Mazur, Kuba, and Jérémie Voix. 2012. "Development of an Individual Dosimetric Hearing Protection Device." In *INTER-NOISE*.

McBride, David I. 2004. “Noise-induced Hearing Loss and Hearing Conservation in Mining.” *Occupational Medicine (Oxford, England)* 54 (5) (August): 290–6. doi:10.1093/occmed/kqh075.

Moutairou, Manani, Student Member, Gilles Y Delisle, and Michel Misson. 2009. “Coverage Efficiency of Narrow-Band Wave Propagation in Mining Environments” 51 (2): 391–400.

Nanjo, Hiroaki. 2009. “Acoustic-based Security System : Towards Robust Understanding of Emergency Shout” 1: 725–728. doi:10.1109/IAS.2009.121.

National Institute of Occupational Safety and Health. 1998. “Occupational Noise Exposure”. U.S Department of Health and Human Services.

Ndoh, M., and G.Y. Delisle. 2005. “Propagation Characteristics for Modern Wireless System Networks in Underground Mine Galleries.” *First International Workshop on Wireless Communicaion in Underground and Confined Area*.

Srinivasan, K., M. Ndoh, and K. Kaluri. 2005. “Advanced Wireless Networks for Underground Mine Communications.” *First International Workshop on Wireless Communicaion in Underground and Confined Area*.

Traunmüller, H, and A Eriksson. 2000. “Acoustic Effects of Variation in Vocal Effort by Men, Women, and Children.” *The Journal of the Acoustical Society of America* 107 (6) (June): 3438–51.

Valoski, Michael P. 1997. “REPORTED NOISE-INDUCED HEARING LOSS AMONG MINERS.” *Applied Occupational and Environmental Hygiene* 12 (12): 1055–1058.

Voix, Jérémie, and Frédéric Laville. 2009. “The Objective Measurement of Individual Earplug Field Performance.” *The Journal of the Acoustical Society of America* 125 (6) (June): 3722–32. doi:10.1121/1.3125769.

Zhang, Chi, and John H L Hansen. 2007. “Analysis and Classification of Speech Mode : Whispered Through Shouted.” In *Proceedings of the Interspeech*, 2289– 2292.