A Sound Foundation Through Early Amplification

Proceedings of the 7th International Conference 2016

The evolution of wireless systems in pediatric settings

Evert Dijkstra, M.S.
Hans E. Mülder, Ph.D.

Abstract

Performance and usability of wireless systems have changed tremendously over the past years and improved the outcomes of children with hearing loss. By changing the technology from analog frequency-modulated (FM) to digitally adaptive systems on 2.4 GHz, frequency channel management was eliminated; thus, making the systems much easier to use. Simultaneously, context-dependent signal processing has proven to be very effective in increasing performance and ease of use of these state-of-the-art systems. This paper will describe the most important wireless and signal processing technologies that are currently employed in pediatric settings.

Introduction

Approximately 50 years ago, the first wireless systems for children with hearing impairment emerged (Victorian Collections, 2017). Despite their bulkiness and lack of user-friendliness, these systems demonstrated great speech recognition score improvements in adverse listening environments such as classrooms (Ross, 2003).

In these listening situations, the presence of reverberation and background noise can make it difficult to understand an individual speaker. Moreover, as sound travels away from the source it reduces in intensity, while the background noise remains relatively constant. Hence, the signal-to-noise ratio (SNR) decreases (Nabelek, 1972). The combination of reverberation, background noise, and distance from the speaker results in poor listening conditions, even for normal hearing individuals, but the impact is even greater on children with hearing loss (Nabalek & Pickett, 1974).

Indeed, the basic working principles of early wireless systems were similar to current state of the art equipment. The idea was to pick up the desired speech signal as close as possible to the source and then transmit this good quality signal wirelessly to the child. The wireless system "short-circuits" the acoustical path. The result is that, in the classroom, the pupil perceives the teacher's voice as being in very close proximity, with a high SNR, which is crucial for speech understanding.

As esteemed audiologist Mark Ross pointed out in various papers, wireless systems were and remain the "most significant educational tool developed for hearing impaired children ever developed since the advent of group and personal amplification devices " (Ross, 1992; Ross 1995). Despite rapid technological progress in digital hearing instrument technology, wireless systems remain the single most effective way to increase SNR, which in turn is the most important factor for speech understanding. The latter is also of utmost importance for children with hearing loss who are in the process of developing speech and language.

Although the basic principles have not changed with respect to the systems that were used approximately 50 years ago, modern wireless and digital signal processing technology has improved esthetics (which helps to improve acceptance), performance, and ease-of-use. Currently, modern wireless systems allow people with hearing loss to understand speech significantly better in noisy situations than their normal hearing peers (Thibodeau, 2014).

Arguably, one of the bigger changes in wireless technology was the move to digital transmission on the 2.4 GHz band. This paved the way for higher performance and (much) easier-to-use systems. In the next section, we will describe the basics of this digital transmission technology. In parallel, digital signal processing of wireless technologies contributed to significantly improved speech recognition performance by children. The main technologies in this field, collectively called "context dependent signal processing", will be summarized in the second section of this paper.

Digital adaptive wireless technology on 2.4 GHz

First of all, 2.4 GHZ is a band in the radio spectrum where certain rules apply for digital transmission that need to be fulfilled by any system in this band. These rules, however, do not guarantee optimal audiological performance of a wireless system. Here, a system (Roger) is described that is optimized for audiological performance, beyond what is standard for 2.4 GHz. In other words, not all systems on 2.4GHz operate in the same way or deliver comparable performance.

As illustrated in Figure 1, audio signals are digitized and packaged in very short (160 µs) digital bursts of codes (packets) and broadcasted several times, each at different channels between 2.4000 and 2.4835 GHz. Frequency hopping between channels, in combination with repeated broadcasting, avoids interference issues. End-to-end (including the delay in a common digital hearing aid) audio delay is well below 25 milliseconds, and these systems are tap-proof.

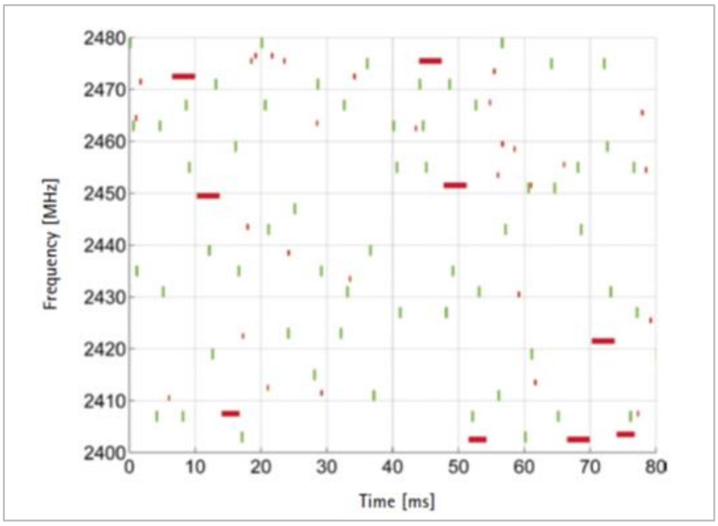

Figure 1. By hoping frequencies and repeated broadcast of audio packet, mutual interference can be minimized.

Sophisticated systems employ adaptive frequency hopping, which means only free channels are used. Two different and complimentary mechanisms are used to blacklist frequencies for transmission. First, the wireless microphone "sniffs" the presence of Wi-Fi or other interferers and changes the hopping sequence in such a way that the occupied channels are avoided. Second, the receivers regularly exchange the reception quality per frequency channel with the transmitting wireless microphones. When the reception quality of any of these channels is insufficient, the transmitter adapts its sequence accordingly. Finally, if an audio packet is not received correctly, intelligent packet loss concealment (PLC) algorithms on the receiver side 'fill in the blanks' to ensure sound quality and listening comfort. Simulation calculations with one system in stacked blocks of 8 x 8 x 2 meters (a worst-case scenario) revealed that no self-interference occurs. In other words, there is no limit to the number of systems that can be used simultaneously in one building, which was never possible with traditional FM systems. Indeed, painstaking frequency channel planning exercises have become obsolete as the system automatically chooses the optimal sequence(s). Note also that the sound quality of a digitally adaptive system remains the same (and independent with distance) until too many transmission errors occur (out-of-range). This is not true for a traditional FM system where the signal quality gradually decreases with distance until it is totally suppressed. Digitally adaptive systems can only be conceived in the 2.4 GHz band or higher, implying another important benefit over traditional FM systems. The wavelength at 2.4 GHz is an order of magnitude smaller than what is needed for the traditional FM systems, which allows smaller antennae that can be fully integrated in the transmitter housing no longer requiring an external microphone cable.

Digitally adaptive systems broadcast audio packets at different channels within the 2.4 GHz band (between 2.4000 and 2.4385 GHz), which means that different carrier wavelengths are selected. A receiver can receive such packets directly via line of sight, but also through a different path when the electromagnetic radio waves are reflected from the walls, the floor, or the ceiling. The different lengths of the different signal paths can enhance or reduce the electromagnetic field's strength at the position of the receiving antenna, depending on the phase and amplitude differences of the different waves. This interference behavior is dependent on the wavelength.

This so called "multi-path fading" (see Figure 2) can be mitigated by (1) transmission of the same signal several times on different frequencies, and/or (2) transmission of the same signal at different times, and/or (3) using two spatially separated radios. The first and second techniques, called *frequency diversity* and *time diversity* respectively, can be advantageously applied in miniaturized receivers, whereas the third technique, called *spatial diversity*, can only be used in larger systems such as sound-field loudspeaker units, which are also wireless receivers.

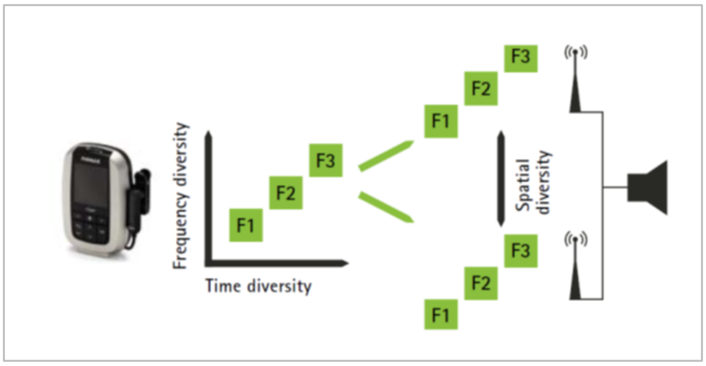

Figure 2. Mitigation of interference by time, frequency and space diversity.

The multi-path fading phenomenon can, if the systems are conceived properly, also be an advantage range-wise. The maximum communication distance is generally larger for bigger indoor rooms (e.g., auditorium, gym) than open green fields.

An interesting multi-fading mitigation protocol is depicted in Figures 3a and 3b. Figure 3a shows how a time-slot of 4 milliseconds is subdivided in 10 slots of 400 microseconds. In each of these small timeslots, we can now transmit small radio packets (in this particular case 320 bits in 160 microseconds) representing either audio or control data. In Figure 3b, different possible allocations of the slots are depicted. As an example, slot "0" serves as a beacon and transmits the pseudo random frequency hopping code to all devices in the network, whereas slot "1" collects all the link quality statistics of all devices. The slots "2" through "9" could be attributed to audio signals (first line shows a mono audio signal that is repeated three times, whereas the second line shows a stereo signal where the left and right signals are each repeated thrice. Finally, the last line shows four different audio signals that are each repeated twice).

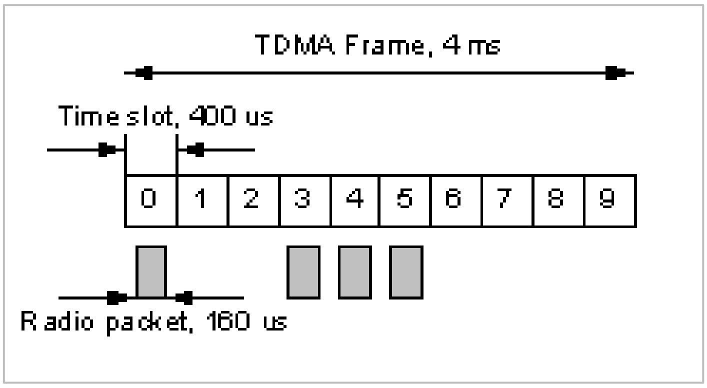

Figure 3a. Each of the small timeslots can transmit small radio packets representing either audio or control data.

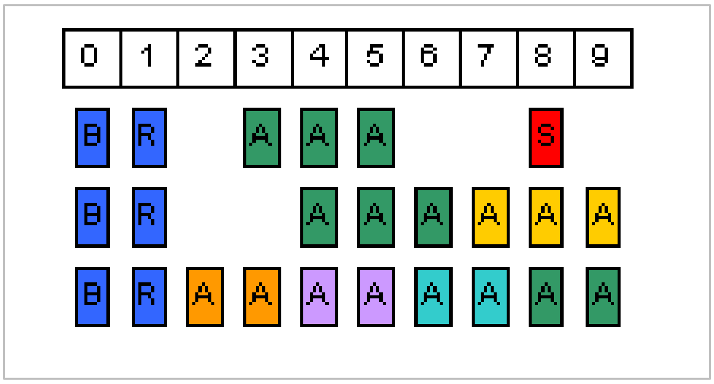

Figure 3b. Different possible allocations of the slots are possible.

Assuming that audio packets are repeated thrice, then this repetition takes place on three different frequencies and at three different moments in time (Figure 4), illustrating concretely what is meant by *frequency diversity* and *time diversity*. Note that the receiver(s) stop processing the repetitions once they receive an error-free packet, thus saving significantly on power consumption.

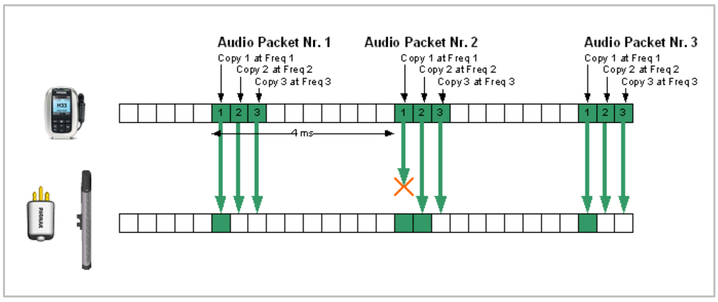

Figure 4. Each audio packet is transmitted three times at a different frequency and a different time, thus implementing frequency and time diversity.

In comparison, Bluetooth® technology requires all packets to be acknowledged by the receiver. If acknowledgement of a packet's reception does not arrive at the Bluetooth transmitter, the packet is broadcast again. This means that Bluetooth receivers are quasi-continuously transmitting back to the transmitter, which significantly increases power consumption at the Bluetooth receiver.

In the Bluetooth headset protocol, the audio delay is still acceptable (10 to 15 milliseconds), but the audio bandwidth is often limited (up to 3 to 4 kHz), unless one uses the 'wideband speech' feature of the hands-free profile version1.6, which can go up to 7 kHz. In the Bluetooth audio streaming protocol (A2DP), the bandwidth increases to 20 kHz, but the audio delay of approximately 100 ms prevents it from being suitable for live face-to-face communication. Only with special Bluetooth chips on both ends can this delay be reduced, to approximately 40 milliseconds.

All standardized Bluetooth profiles suffer from the fact that they can stream to only one receiver, this makes it unsuitable for educational purposes. In contrast, the educational system described in Figures 3 and 4 allows connection to an unlimited number of receivers (including soundfield). In addition, several transmitters (teacher microphone(s), pass-around microphone(s), and multi-media transmitters) can all be connected in the same network and transmit to the receiver(s). So instead of a network topology of 1 to 1, a topology of N to N has become possible, which reflects real-life situations much better where multiple talkers (e.g., teachers, children) are heard by multiple listeners.

Context dependent signal processing

Although digital communication systems with (adaptive) frequency hopping offer great benefit in terms of speech clarity, flexibility, and ease-of-use, speech recognition is not intrinsically better when compared to traditional FM systems (Mülder, 2011). To optimize speech intelligibility, sophisticated, context-dependent, audio signal processing schemes are required in the transmitter(s), the receiver(s), and possibly the hearing instruments (including cochlear implants).

The basic idea of a wireless system is to counteract the strong decrease in amplitude of speech level with increasing distance from the speaker (Boothroyd, 2003). Indeed, as the speech energy decreases as a function of distance, the SNR decreases to negative values at the listener's end. This can easily happen, even for distances as small as 2 meters (6 feet).

The *SNR advantage* refers to the benefit in SNR due to the use of a wireless system as compared to the situation without the wireless system. The value is derived by the SNR value obtained using the wireless signal transmission minus the SNR value that would be obtained without the wireless system (Platz, 2003).

On the other hand, the FM or "Roger" advantage (in the context of digital transmission, can be described as a "level" advantage) measures the relative level at a certain frequency of both signals when both the wireless signal and the hearing instrument microphone are active *at the same time*.

To avoid a varying level advantage as a function of external variables (e.g., exact microphone placement, head-movements), built-in compression in the transmitter is necessary. Various types and knee-points exist and might also vary as a function of the context (e.g. table microphone use, distant speech at low noise levels), but in most of the cases, the knee-points will be around 72 to 78 SPL providing a stable "level" advantage. Attack and release times need to be set appropriately by the designer (i.e., manufacturer) for optimal listening comfort and speech clarity.

In a non-adaptive system (a traditional FM system or a simple digital system), the compressed transmitter signal is mixed with the hearing instrument microphone. Typically, the mixing is done in such a way that a 10 dB level advantage for the wireless system is obtained under the following conditions: speaker's voice is picked up at 80 dB SPL by a lapel microphone that is 30 cm from the speaker's mouth and the listener is 2 meters away from the speaker (American Academy of Audiology, 2011). As the mixing of both signals takes place at the input of the hearing instrument or cochlear implant, both signals will be processed simultaneously and in the same way. The result of the whole chain (Figure 5) is that we find an improved SNR at the listener's ear and the obtained SNR is less than that at the speaker's position.

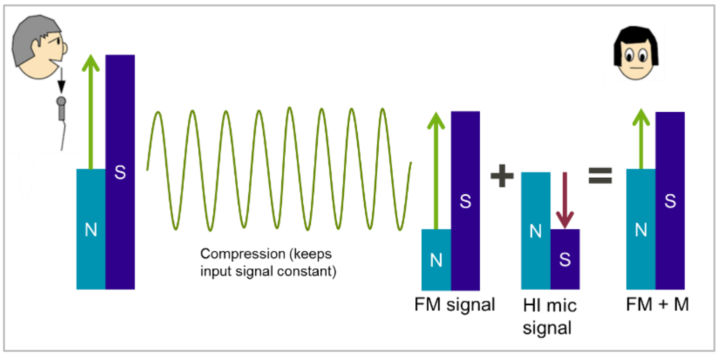

Figure 5. In a non-adaptive system, the SNR at the listener has improved but is less than at the speaker.

In an adaptive system, the gain of the wireless system is incrementally increased as a function of the environmental noise. As depicted in Figure 6, the result of this adaptation is that the SNR at the speaker's position will at all times be the same as at the listener's ear. It is as if this very good SNR is copied at the speaker and pasted at the listener, which is ideal.

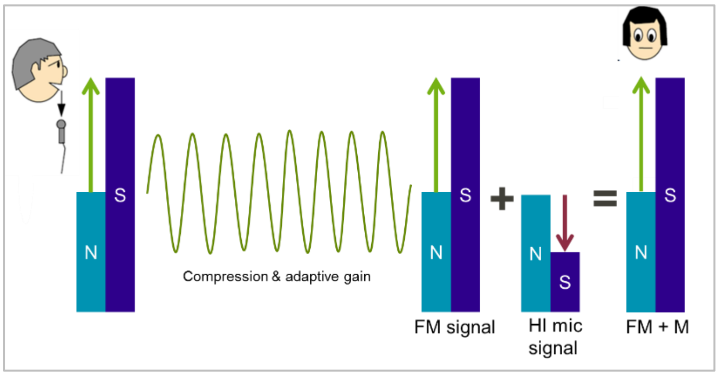

Figure 6. The SNR at the listener and the speaker are equal (copy-paste of SNR).

In many frequent-use cases, it is possible to detect reliably when someone is talking into the microphone(s). In that case, the gain of the wireless system is reduced, giving back the functionality of the hearing instrument alone.

The winning strategy to obtain a maximum SNR and, therefore, a best possible speech recognition score could be summarized as follows: (1) bring the transmitter as close as possible to the source and maximize the SNR by using high performance beam-formers in the transmitters, (2)adaptively mix the wireless microphone signal with the ear-level microphone of the hearing instrument by increasing the gain of the receiver in higher ambient noise levels (previously mentioned copy-paste of SNR), and (3) reduce the gain when no voice is present.

Applying this strategy has been instrumental in obtaining excellent speech recognition scores (Figure 7), as well as acceptance and appreciation by users. Note that the basic strategy was already implemented with "dynamic FM" technology (Thibodeau, 2010; Wolfe et al., 2009) and that further significant improvements could be achieved with Roger. The latter was due to further improved beam-formers (point 1 of the strategy) and more precise calculations for the adaptive mixing (point 2 of the strategy).

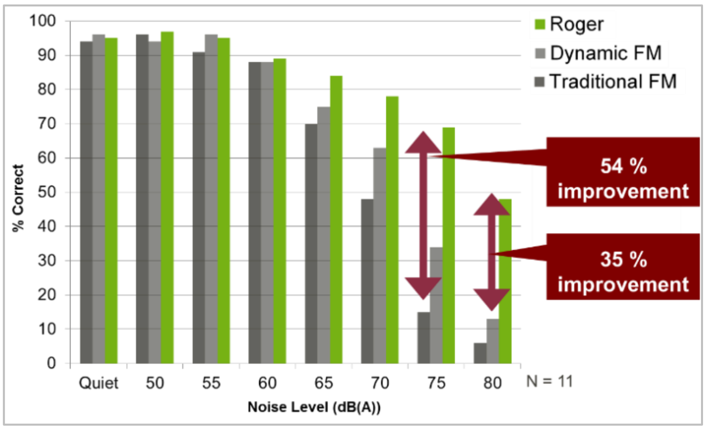

Figure 7. The adaptive signal processing is instrumental in obtaining high speech recognition scores.

Lapel-style microphones are successfully used by teachers in many classrooms around the world to ensure students with hearing loss or other hearing difficulties have clear access to the teacher's voice throughout the school day. However, teaching styles are becoming increasingly more dynamic and interactive with estimates of up to 34% of the school day involving peer or group discussion activities (Rich & Gigandet, 2016). Figure 8 shows the breakdown of teaching style and classroom activities from an internal Phonak study conducted in numerous schools across multiple countries (Feilner, 2016). Although dynamic and participatory styles of learning are becoming standard, these settings create acoustical challenges or even barriers to children with hearing loss. For them, it is sometimes impossible to hear and understand effectively.

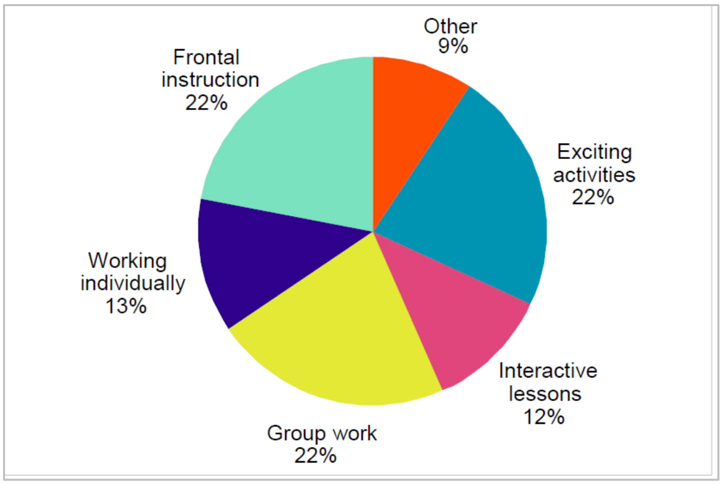

Figure 8: The breakdown of teaching style and classroom activities in multiple schools and multiple countries

To provide auditory benefit in non-frontal education environments, a new variant of the context-dependent signal processing principle has been developed. It is now possible to automatically activate Small Group mode when a teacher microphone is placed flat on a table or floor. This mode uses an array of three orthogonally placed omnidirectional microphones to create a beam in the direction of the primary talker (with the best SNR) and suppressing noise coming from other directions.

Small Group mode is designed so that the microphone can be placed in the center of 2 to 5 group members during group learning or listening activities. Specific signal characteristics, such as SNR and energy level, are analyzed and used to localize speech information and to identify the talker's direction. This allows the device to follow the conversation automatically by always emphasizing the active talker.

Often in group discussions, conversations can quickly move from one talker to the next. It is also not rare to observe children (or adults) interrupting each other. Given this, it is highly important that signal information is not lost during such challenging situations. The Small Group mode must be designed in such a way that it adapts smoothly during transitions between talkers and provides a pleasant sound quality without interruptions, even in the most challenging situations where people are talking at the same time and interrupting each other. Jones (2016) demonstrated that the Small Group mode improves understanding of multiple near-field talkers in noise by 20% compared to a transmitter with an omnidirectional pickup placed at the same spot. The improvement increases to over 30% compared to hearing instruments alone.

For interactive lessons, pass-around microphones have been proven to increase participation by children with hearing loss because they can better understand the topics brought forward by their peers. The use of these pass-around microphones is even more effective if used in combination with a sound-field system. In such cases, the sound-field system which provides acoustic benefit to all children (with or without hearing loss) also provides direct acoustical feedback to the children and teachers using the pass-around microphones, thus ensuring that the equipment is used properly.

Pass-around microphones also benefit from smart, context-dependent, signal processing. As an example, depending on whether the microphone is standing on the table or hand-held, one can change the microphone sensitivity, gain model, and beam forming characteristics. Also, if the microphone lies flat on a table it can be muted automatically.

In a digital adaptive wireless system on 2.4 GHz, teachers' and students' microphones can be, together with receivers and sound-field speakers, connected in a single multi-talker network (Wolfe, 2013). Although the number of voices that can be transmitted simultaneously to children with hearing loss could theoretically be high, it should nevertheless be limited in order to avoid a cacophony for people with hearing loss. For practical purposes, it appears that a simultaneous transmission of maximum two, out of N, voices is optimal.

Although the context-dependent signal processing can be improved further, it is interesting to note that one variant has been researched and rarely shown to be outperformed by manual control use (Mülder, 2015).

Conclusion

Wireless systems in educational settings have evolved tremendously over the past few years. First of all, the advent of miniaturized 2.4 GHz digital frequency hopping systems eliminated the tediousfrequency channel planning that was required with traditional FM technologies. Second, the development of context- dependent signal processing schemes allowed improved performance of these systems to unprecedented levels.

References

American Academy of Audiology (2011). Remote Microphone Hearing Assistance Technologies for Children and Youth from Birth to 21 Years. In *Clinical Practice Guidelines.*

Boothroyd, A. (2003). Room Acoustics and Speech Reception: A Model and Some Implications, *Access proceedings,* 21.

Feilner, M., (2016). New automatic operation of hearing instruments: Performance in school situations, *5th Phonak European Pediatric Conference, Berlin.*

G. De Ceulaer et al, (2015). Speech understanding in noise with the Roger Pen, Naida CI Q70 processor, and integrated Roger 17 receiver in a multi-talker Network, *European Archives of Oto-Rhino-Laryngology,* ISSN 0937-4477.

Jones, C. (2016). Legit Innovations for Teen Tech Users, 7*th International Phonak Pediatric Audiology Conference, Atlanta.*

Mülder, H. (2015). Comparison of Roger Pen automatic mode with manual mode, *Phonak Field Study News.*

Mülder, H. (2011). Digital or Dynamic FM? Which gives better speech understanding in noise, *Phonak Field Study News.*

Nabelek, A. (1972). Room acoustics and speech perception. In *Handbook of Clinical Audiology,* 4th edn, 624–636.

Nabelek, A., & Pickett J. (1974). Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners, *Journal Speech, Language and Hearing Research* 17(4):724–739.

Platz, R. (2003). SNR Advantage, FM Advantage and FM Fitting, Access Proceedings, Ch. 14.

Rich, S., & Gigandet, X. (2016). Roger Touchscreen Mic Small Group mode: Changing the dynamics of group activities in the classroom, *Phonak Insight*

Ross, M. (2003). FM Systems: A Little History and Some Personal Reflections, *keynote address Access Conference.*

Ross, M. (1992). Room Acoustics and Speech Perception. In, FM Auditory Training Systems: Characteristics, Selection and Use, Ross, M. (Ed).

Thibodeau, L. (2010). Benefits of Adaptive FM Systems on Speech Recognition in Noise for Listeners Who Use Hearing Aids, *The American Journal of Audiology,* 19(1):36-45.

Thibodeau, L. (2014). Comparison of speech recognition with adaptive digital and FM wireless technology by listeners who use hearing aids, *American Journal of Audiology*, 23:201-210.

Victorian Collections. (2017). FM Phonic Ear. Retrieved 12 July, 2017, from https://victoriancollections.net.au/items/56664ac52162f101d0bf144c

Wolfe, J., Schafer, E., Heldner, B., Mülder, H., Ward, E., & Vincent, B. (2009). Evaluation of Speech Recognition in Noise with Cochlear Implants and Dynamic FM, *Journal of American Academy Audiology,* 20:409–421.

Wolfe, J. (2013). Evaluation of Speech Recognition with Personal FM and Classroom Audio Distribution Systems, *Journal of Educational Audiology,* 19.

Acknowledgements:
Our greatest acknowledgement goes to all staff members of Phonak Communications in Murten (Switzerland) and Halmstad (Sweden) for their full dedication to wireless technology for people with hearing loss.

Authors

Evert Dijkstra, M.S.
Managing Director Phonak Communications
Phonak Communications
Evert.Dijkstra@Phonak.com
Murten, Switzerland

Hans E. Mülder, Ph.D.
Head of "Severe to Profound" category
Phonak Communications
Evert.Dijkstra@Phonak.com
Murten, Switzerland

Editors

Anne Marie Tharpe, Ph.D.
Chair, Phonak Research Advisory Board
Professor and Chair
Department of Hearing & Speech Sciences
Vanderbilt University School of Medicine
Nashville, Tennessee, United States

Marlene Bagatto, Ph.D.
Research Associate and Adjunct Research Professor
National Centre for Audiology
Western University
London, Ontario, Canada