Tutorial

Guidelines for Selecting Microphones for Human Voice Production Research

Jan G. Švec
Palacký University Olomouc, Olomouc, the Czech Republic

Svante Granqvist
Royal Institute of Technology, Stockholm, Sweden

Purpose: This tutorial addresses fundamental characteristics of microphones (frequency response, frequency range, dynamic range, and directionality), which are important for accurate measurements of voice and speech.
Method: Technical and voice literature was reviewed and analyzed. The following recommendations on desirable microphone characteristics were formulated: The frequency response of microphones should be flat (i.e., variation of less than 2 dB) within the frequency range between the lowest expected fundamental frequency of voice and the highest spectral component of interest. The equivalent noise level of the microphones is recommended to be at least 15 dB lower than the sound level of the softest phonations. The upper limit of the dynamic range of the microphone should be above the sound level of the loudest phonations. Directional microphones should be placed at the distance that corresponds to their maximally flat frequency response, to avoid the proximity effect; otherwise, they will be unsuitable for spectral and level measurements. Numerical values for these recommendations were derived for the microphone distances of 30 cm and 5 cm.
Conclusions: The recommendations, while preliminary and in need of further numerical justification, should provide the basis for better accuracy and repeatability of studies on voice and speech production in the future.

Key Words: voice, measurement, microphones, requirements

When creating a voice and speech laboratory and preparing measurements, one faces the question "Which microphone should be used?" Although microphones present basic means for registration of voice signals, there has not been enough information published on which microphones are or are not suitable for voice measurements. While there have been attempts to provide recommendations for the choice of microphones (Baken & Orlikoff, 2000; Bless et al., 1992; Laver, Hiller, & Beck, 1992; Spielman, Starr, Popolo, & Hunter, 2007; Švec, Šrámková, & Granqvist, 2009; Titze, 1995), so far there has been an insufficient explanation of the principles on which the recommendations should be based. This lack of information has led to a situation in which studies have been published with improperly chosen microphones, and the reported results have contained inherent errors.

In voice and speech research, the purpose of the microphone is to convert the sound pressure signal to an electric signal with the same characteristics (see Figure 1). However, most microphones are not developed for this purpose but rather for recording of music, performances, public address systems, broadcasting, and so on (AKG Acoustics, 2003; Howard & Murphy, 2007). Consequently, many of the microphones are not suited for accurate measurements of voice and speech. Voice measurements are often done by investigators with a nontechnical background who may not study the specialized technical literature on microphones and their characteristics. Also, the supportive technical personnel in clinics and institutes often do not have sufficient expertise in acoustics of voice and speech and therefore can hardly provide expert support in these issues.

The purpose of this tutorial is therefore to provide guidelines for selecting a microphone that is suitable for measurement of voice and speech. Specifications provided by the manufacturer are reviewed and put into relation with the characteristics of voice and speech. Finally, recommendations are formulated that can be used for selecting the proper microphone for voice and speech research.

To ensure accurate recording of voice and speech, we will consider three fundamental characteristics of sound: (a) *fundamental frequency* (F_0); (b) *timbre*—that is, the *sound spectrum*; and (c) the *pressure amplitude* measured via

FIGURE 1. Schematic illustration of a conversion of a sound pressure (voice) signal into an electric signal by means of a microphone. Microphone 1 keeps the waveform characteristics the same, while Microphone 2 alters the waveform. For voice production measurements, alteration of the waveform is undesirable. An omnidirectional measurement microphone and a cardioid performance microphone with a proximity effect, both placed at a 5-cm distance, were used to gather the data for this example.

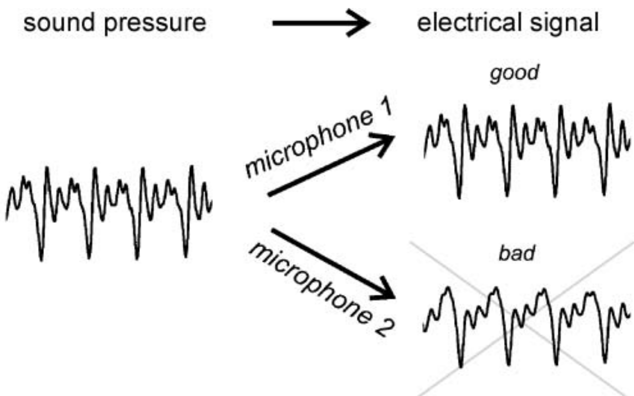
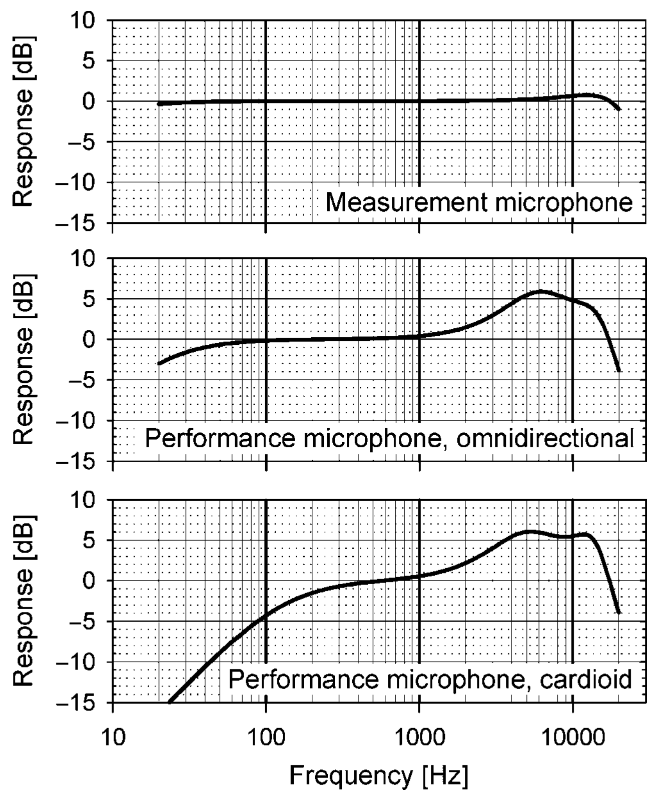


FIGURE 2. Frequency responses for three typical microphones: an omnidirectional measurement microphone, an omnidirectional studio microphone, and a cardioid microphone for stage use (here the response is valid for the far field). Note the broad presence peak in the 3–10-kHz region for the two latter microphones.



the *sound pressure level* (SPL). These three characteristics should ideally not be affected in the captured sound. While the F_0 of voice is normally well preserved in the captured sound, the accuracy of the captured timbre and SPL depends on the frequency response and the dynamic range of the microphone. An additional factor is the noise in the room, which can also influence the accuracy of the measurement. All these factors will be considered in the sections below.

Frequency Response and Range

Microphone transducers have an inherent frequency response curve mirroring their design and purpose. The ideal frequency response for measurements is flat (i.e., having the same sensitivity for sounds with different frequencies), because this results in a signal that has the same spectrum (timbre) as the original sound. No real microphone has a perfectly flat response, however. For example, it is common for studio/stage microphones to have a "presence peak" in the 3–10-kHz region, which amplifies the high frequencies of the sound more than its low frequencies and can make the sound perceptually more appealing (see Figure 2). For accurate measurement of SPL and timbre of voice, such coloration of the spectrum should be avoided. High-quality microphones intended for measurement purposes therefore do not show this presence peak.

The frequency range of the microphone response (i.e., the range at which the frequency response is sufficiently flat) should be broad enough for capturing the complete spectrum of the voice/speech sound, from the lowest to the highest frequency components of interest. In normal voice, the lowest frequency component is determined by the voice's F_0, which can go down to about 50 Hz in males (e.g., Leino, Laukkanen, Ilomäki, & Mäki, 2008; Sulter, Schutte, & Miller, 1995). In hoarse or creaky voices, however, there could be subharmonic or chaotic components of the sound, the frequencies of which could be well below 50 Hz (e.g., Herzel, 1996; Neubauer, Edgerton, & Herzel, 2004; Švec, Schutte, & Miller, 1996; Tokuda, Horáček, Švec, & Herzel, 2007). The microphone should therefore be ideally capable of capturing the lowest frequencies possible. In measurement microphones, a bottom limit of 10 Hz is often used. In these cases, however, care needs to be taken regarding the room acoustics and ambient noise because the captured signals can be polluted by low-frequency and infrasonic noise from the room and require special treatment.

In speech, the highest frequency components are produced in consonants such as /s/, in which the spectral maximum is centered around 7000–8000 Hz (Fant, 1959) and some of the sound components reach frequencies of 10000 Hz and higher. On vowels, the investigations have often been limited up to the frequencies of 5000 Hz, around which there is a spectral minimum (Dang & Honda, 1997; Ternström, 2008). Current research has shown, however, that there are also frequency components higher than 5000 Hz produced in vowels, which may be important for the perceived voice quality (Ternström, 2008). The ideal top limit for a microphone frequency response is thus considered to be around the highest frequencies perceivable by the human ear, that is, 16000–20000 Hz.

To specify the requirements for "flatness" of the frequency response, it is useful to take into account tolerances provided by national and international standards. The U.S. standard ANSI S1.15-1997/Part 1 (American National Standards Institute [ANSI], 2006), which is comparable to the international standard IEC 61094-1 (International Electrotechnical Commission [IEC], 1992), specifies two classes of accuracy (LS1 and LS2) for microphones designed specially for measurement purposes, that is, laboratory standard microphones (LS type). Here the frequency response is required to be flat within 2 dB for the frequency range of 10–8000 Hz (LS1-type microphones) or 10–20000 Hz (LS2-type microphones). The 2-dB tolerance is the maximum difference between the greatest and smallest sensitivity level within the required interval.

Considering the descriptions above, it can be recommended that the low-frequency limit of the microphone intended for voice measurement should be lower than the lowest produced frequency of voice, with 50 Hz being the maximum. The upper frequency limit of the microphone should be above the highest spectral frequency of interest, with 8000 Hz being the minimum. Between the low-frequency and upper frequency limits, the frequency response of the microphone should be flat, with the tolerance of 2 dB.

Some applications, such as inverse filtering of voice, require not only a flat frequency response of a microphone but also a flat phase response. This is important to preserve the exact shape of the waveform. The phase response is usually not provided by the manufacturer in the microphone data sheet. Flat phase response is usually guaranteed when the frequency response is flat, but the lower limit for phase response flatness occurs at a frequency that is about 10 times higher than the low-frequency limit of the microphone (Brüel & Kjaer, 1996, Figures 2.18 and 2.19). Therefore, the low-frequency limit of the microphone may be required to be a decade below the voice F_0 ($F_0/10$). Microphones of LS type (both LS1 and LS2, according to ANSI standard) have the low-frequency limit of 10 Hz. These allow inverse filtering of the captured voice signals with fundamental frequencies down to at least 100 Hz. In cases of inverse filtering where calibrated airflow values are also of interest, such microphones are still insufficient and are usually replaced by an airflow mask and sensors with a low-frequency limit down to DC levels, that is, 0 Hz (Rothenberg, 1973).

Directionality

Directionality is another important characteristic of microphones. A microphone that is omnidirectional has the same sensitivity regardless of the direction to the sound source. Directional microphones, on the other hand, respond differently to the sound coming from different directions. The directionality of a microphone is typically shown by its directivity pattern, or polar plot (see Figure 3).

Probably the most common directivity pattern for directional microphones is the cardioid (Figure 3b), even though other patterns do exist. The cardioid microphone picks up well the signal coming from the front but suppresses the signal coming from other directions. The graph shows that the cardioid microphone has the lowest sensitivity for sounds originating from the 180° direction—that is, from the back. This has been found useful for suppressing the ambient noise and reverberation sound in the room, as illustrated in Figure 4.

While the suppression of the ambient noise is advantageous for measurement of the voice signal, directional microphones have some serious frequency response issues that need to be considered before using them. The problem is that the directional pattern is achieved by making the microphone sensitive to the pressure gradient, which in turn is proportional to the air particle velocity, rather than to the sound pressure only (Eargle, 2001; Merhaut, 1980). Directional microphones therefore do not necessarily record the actual sound pressure. A consequence of this is the so-called *proximity effect,* which is covered in the next section. Microphones for precision

FIGURE 3. Examples of polar plots of three types of microphones: (a) omnidirectional, (b) cardioid, and (c) hypercardioid. The hypercardioid directionality can vary between designs, in particular with respect to the level of the lobe directed toward 180°.

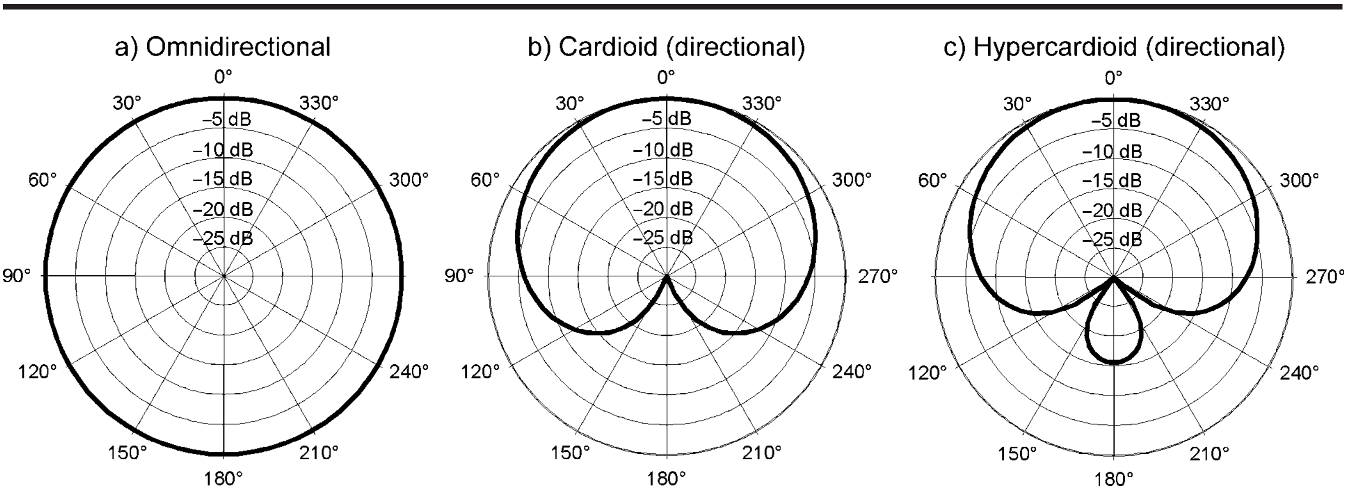
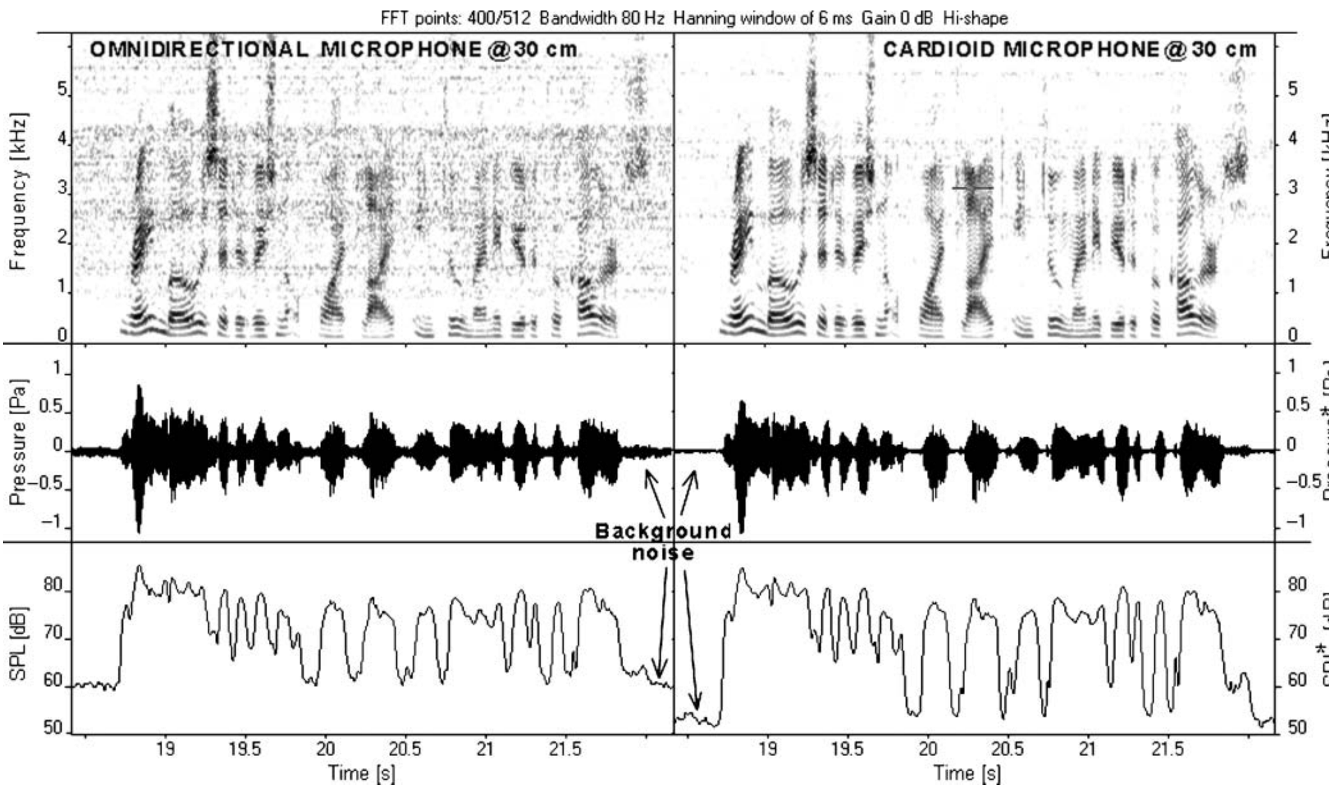


FIGURE 4. Speech signal recorded at 30 cm simultaneously with an omnidirectional and directional (cardioid) microphone in the presence of an ambient noise. The cardioid microphone lowers the effect of the ambient noise, as demonstrated in its cleaner spectrogram and lower SPLs during the speech pauses. (*In a strict sense, the signal from the cardiod microphone is not proportional to the sound pressure actually present in the room but rather approximates the pressure as it would have occurred in a less noisy environment.)



measurement of sound pressure, such as those incorporated in sound level meters, are omnidirectional.

Omnidirectionality of a microphone is ensured only up to a certain frequency limit above which even "omnidirectional" microphones become directional. This frequency limit depends on the size of the tip of the microphone (for a 1-in. tip, the limit is around 5 kHz; for smaller sizes, the frequency is higher), and it is a result of sound diffraction and interference. When the size of the tip of the microphone becomes comparable to the wavelength of the sound, the microphone becomes less sensitive to sounds coming from the side and back of the microphone. The frequency response of the microphone at high frequencies then varies with the angle of incidence. Measurements show that for a 0.5-in. membrane microphone, the level at 10 kHz and 70° drops by approximately 3 dB, and at 20 kHz by approximately 7 dB compared to frontal incidence due to this effect (Brüel & Kjaer, 1996). For a microphone with twice the membrane diameter (1 in.), the same effect can be observed at half the frequencies (5 and 10 kHz; Brüel & Kjaer, 1996). In terms of frequency response and directionality, a smaller tip is therefore preferable, but usually a compromise is needed because a very small tip mostly results in a high noise level of the microphone.

This implies that for accurate measurements of high-frequency components, the microphone should be oriented so that a maximally flat response is achieved. The optimal direction varies between microphones, however, so the manufacturer's data sheet needs to be consulted in order to find this angle for each microphone model. Generally, there are two basic microphone orientations (Brüel & Kjaer, 1984, 1996): (a) toward the mouth of the speaker, which is used with the microphones optimized for sound coming from one direction (so-called "free-field" or "frontal incidence field" microphones), and (b) at the angle of approximately 70°, which is applicable for microphones optimized mainly for measurements of noise coming from all directions (so called "diffuse field" or "random incidence field" microphones). Such diffuse-field microphones can be found in sound level meters fulfilling the ANSI S1.4-1983 standard (ANSI, 1985). The sound level meters fulfilling the international standard IEC 61672-1 (IEC, 2002) may have both a random incidence and free-field response. For voice measurements, microphones optimized for the free-field should be preferred—these are directed toward the sound source, that is, the mouth. Some high-quality microphones come with replaceable grids to achieve flat response for either frontal or diffuse incidence; in these cases, the grid for frontal incidence should be used.

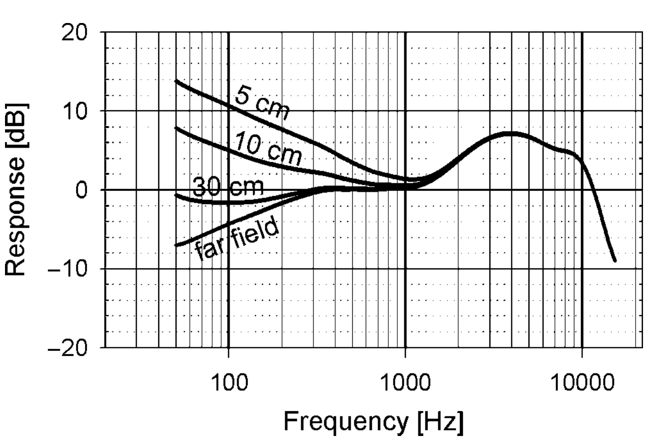
Proximity Effect

The frequency response of an omnidirectional microphone does not depend on the mouth-to-microphone distance. Directional microphones, on the other hand suffer from the *proximity effect*. This effect boosts the lower frequencies when such a microphone is positioned close to the mouth (see Figure 5).

The proximity effect is not simple to compensate for, since the level of the boost depends on the distance between the mouth and the microphone. However, some microphones have built-in compensation for the proximity effect, so that their frequency response becomes approximately flat at a particular distance. If a different distance is used, the sound waveform will become altered (recall Figure 1), and there will be a frequency response error, as illustrated in Figure 5. When closer than the reference distance, the microphone will boost low frequencies, whereas at distances farther than the reference, the low frequencies will be suppressed. The low-frequency boost or suppression in a response of a cardioid microphone is up to 6 dB per halving or doubling the distance, respectively (Merhaut, 1980).

Unfortunately, many manufacturers do not specify the reference distance for directional microphones (Šrámková, 2008; Švec et al., 2009) or they measure headset microphones at a distance far away from the source. This means that even if the microphone data sheet indicates a flat frequency response, the actual response in the recordings may have a strong boost of the lower frequencies. While this effect has been explored by singers for changing their voice timbre during performance, the variable frequency response is not desirable for accurate measurements of the voice signal. It is therefore important to be aware of the variability of the frequency response in directional microphones and to use only the distance for the measurement at which the frequency response is flat. If the distance is not known and the microphone is directional, the microphone should not be used for SPL and spectral measurements of voice and speech.

FIGURE 5. Proximity effect of a typical cardioid microphone. The low-frequency components of the registered spectrum are boosted when the mouth-to-microphone distance diminishes. This particular microphone has the flattest response at the distance of 30 cm.



Dynamic Range

Microphones have a usable dynamic range that is limited by their internal noise at low levels and distortion (clipping) at high levels (Brüel & Kjaer, 1996; DPA Microphones, 2007; see Figure 6). The noise level is typically shown in the microphone specifications as an equivalent A-weighted sound level and is given in dBA units (calibrated for the standard reference value of 20 μPa, corresponding to 0 dB). The spectrum of the noise is approximately white for high frequencies. At low frequencies, there is usually some increase of the noise. This increase is, however, of little importance because its level is usually lower than that of the room noise, which typically exhibits a similar behavior.

To ensure accurate measurement of voice and speech, the internal noise level of the microphone (as well as the room noise level) should be considerably lower than the level of the softest phonations expected. For clean recordings, the microphone noise level is recommended to be minimally 15 dB below the voice levels (see Figure 6). This corresponds approximately to a noise signal with amplitudes at least ⅕ lower than the amplitudes of the voice signal (see Figure 7). The same condition holds also for the background noise levels. The requirement of a signal-to-noise ratio of at least 15 dB has been adopted as a standard in classroom acoustics (ANSI, 2002) and is used here for simplicity. For some applications, however, even higher signal-to-noise ratios (30 dB or more) may be required—for instance, in perturbation measurements, noise levels less than 30 dB below the voice levels were found to have significant influence on the measured jitter and shimmer values of voice (Deliyski, Shaw, Evans, & Vesselinov, 2006; Perry, Ingrisano, Palmer, & McDonald, 2000).

The upper limit of the dynamic range of the microphone is specified as the SPL that results in total harmonic distortion (THD) of 3% (Brüel & Kjaer, 1996). An even stricter requirement of 1% THD is specified by the ANSI S1.15-1997 standard (ANSI, 2006). For voice and speech recordings, the level of this upper microphone limit should be equal to or higher than the levels of the loudest phonations (see Figure 6).

FIGURE 6. Dynamic range of a microphone and minimum requirements based on the dynamic range of voice. THD = total harmonic distortion.

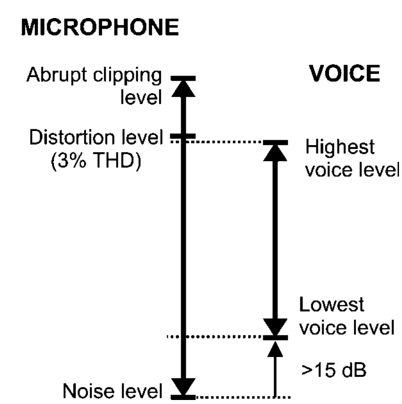


FIGURE 7. The influence of noise on the pressure waveform: (a) a clean periodic signal with no noise—that is, with infinite signal-to-noise ratio (SNR); (b–e) the same waveform under the conditions when pink noise is added such that the noise level is (b) 20 dB, (c) 15 dB, (d) 10 dB, or (e) 5 dB weaker than the signal; and (f) when the noise has the same level as the signal. Notice the increased perturbation of the waveform when the noise level is increased. For accurate measurement of voice, the signal is recommended to be at least 15 dB stronger than the internal noise of the microphone and the background noise (cases a–c).

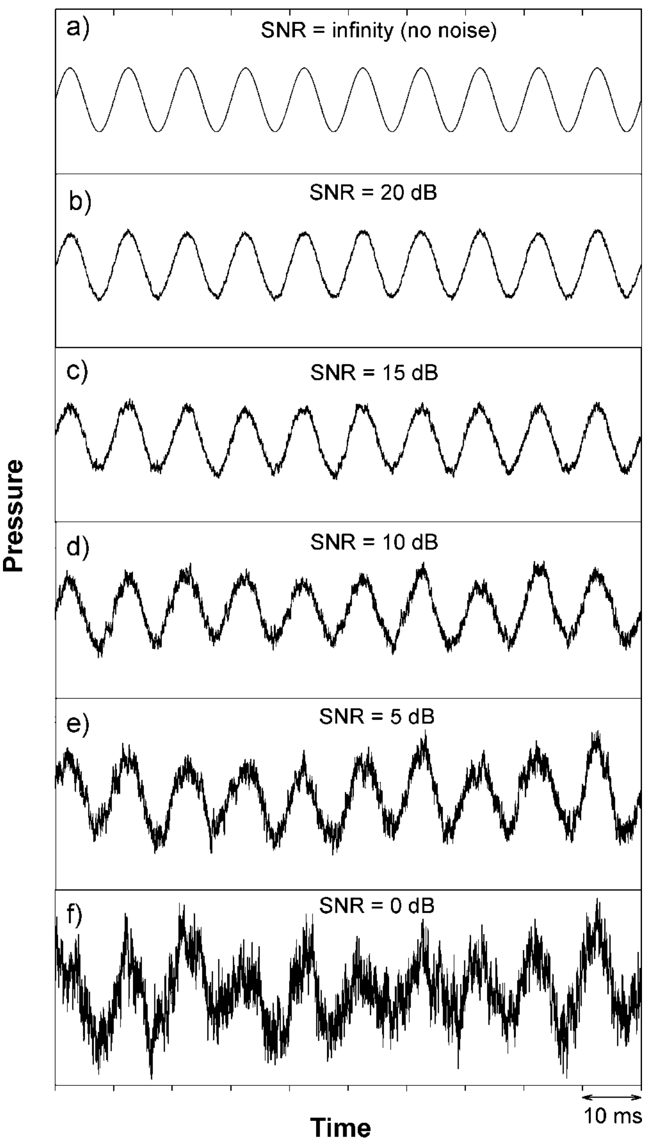
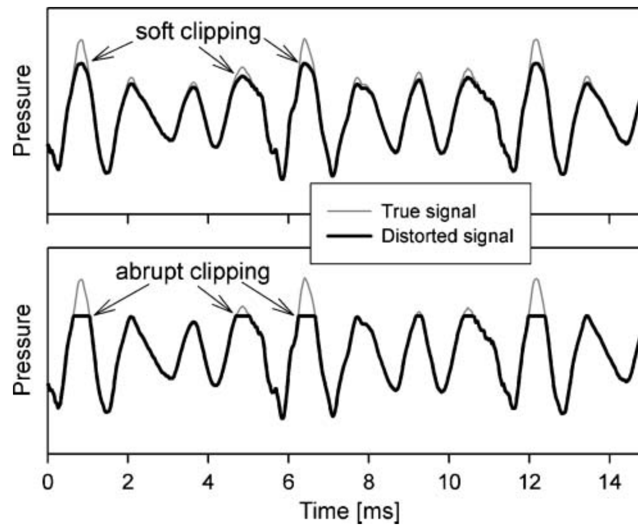


FIGURE 8. Microphone distortion: soft and abrupt clipping. The upper panel illustrates a waveform with soft clipping distortion as it may look when a microphone is overloaded. The abrupt clipping (lower panel) can occur in a microphone amplifier and is easier to detect visually. In this illustration, only the positive peaks are affected, but examples where both or only the negative peaks are affected are common.



When exceeded, distortion levels increase, and the peaks of the captured signal become progressively clipped. Two types of clipping—*soft* and *abrupt*—could be distinguished. Soft clipping occurs when high acoustic pressures deform the microphone membrane beyond linear construction limits. It can be somewhat difficult to notice because the peaks are smoothly lowered (see Figure 8, top) and the distortion may not be audible in the sound. Abrupt clipping, on the other hand, originates mostly in electronic amplifiers and is clearly recognizable via flattened peaks of the captured waveform (see Figure 8, bottom). Such a distortion is usually audible in the sound. Typically, it occurs at higher levels than soft clipping (DPA Microphones, 2007).

Dynamic Range Versus Microphone Distance

To specify numerical values for the low and high limits of the dynamic range of the microphone, it is important to consider the distance at which the microphone is going to be used. With increasing distance from the mouth, the voice signal gets weaker (i.e., the voice level decreases). Here, we will consider two microphone positions: (a) at a 30-cm distance in front of the mouth, which is the position recommended by the Union of European Phoniatricians for stand-mounted microphones (Schutte & Seidner, 1983); and (b) at a 5-cm distance to the side of the mouth, which is the "close-miking" position often used with head-mounted microphones. According to the "distance law," voice levels decrease by 6 dB when the mouth-to-microphone distance is doubled. At the distance of 5 cm, the levels are approximately 15 dB higher than at 30 cm from the mouth. (A caution is needed here, however, because placing the microphone in the proximity of the head may lead to artifacts and interference effects. The distance of 5 cm is considered here because of its frequent use in practice, but it deserves more investigation and verification.)

What should be the maximum noise level of the microphone for these measurement distances? The softest phonation levels have been reported around 40 dBA at a 30-cm distance (Heylen, Wuyts, Mertens, De Bodt, & Van de Heyning, 2002; Hunter, Švec, & Titze, 2006; Leino et al., 2008; Ma et al., 2007; Schneider & Bigenzahn, 2003; Sulter et al., 1995), which corresponds to approximately 55 dBA at a 5-cm distance from the mouth. Dedicated measurements indicate that in extreme cases the softest sustained phonations can even approach levels of 30 dBA at 30 cm, or 45 dBA at 5 cm (Šrámková, 2010). Taking into account the 15-dB signal-to-noise ratio rule mentioned before, we can formulate the requirement for the internal equivalent noise level of the microphone (and for the room noise level) to be below 15 dBA or 30 dBA when intended to be used at a 30- or 5-cm distance, respectively. Information on the actual equivalent internal noise level of the microphone can be found in the specifications provided by the microphone manufacturer.

The requirements for the microphone noise levels are valid also for the ambient noise levels. Also, here the signal-to-noise ratio of 15 dB requires the noise level to be below 15 dBA when the microphone is placed at the 30-cm distance and below 30 dBA when placed at 5 cm. The ambient noise levels can be measured with a sound level meter; in regular office rooms, these levels are often more than 40 dBA, which makes evaluations of the softest phonations there problematic. Sometimes, the conditions may be improved by careful high-pass filtering of the signals, but this method is out of the scope of this article.

The upper dynamic limit of the microphone can be derived from the highest levels reported for voice. In extreme cases, such as in shouting or loud operatic singing, these levels may reach values between 120 and 130 dB at the 30-cm distance (Angerstein & Neuschaefer-Rube, 1998; Leino et al., 2008; Šrámková, 2010; Sulter et al., 1995), which corresponds to the levels of approximately 135–145(!) dB at the distance of 5 cm. Many commercial microphones cannot record such high acoustic pressures faithfully and will severely distort these loud sounds. Distortion alters voice spectrum and also causes underestimation of high sound levels. Since some type of distortion (i.e., soft clipping) can be difficult to notice in the recorded waveform, it is important to study the manufacturer's specifications for the maximum acceptable level of the microphone to avoid these problems.

Transducer Type

Microphone transducers convert acoustic pressures into an electric signal. The most common transducer types are electret, condenser, and dynamic (AKG Acoustics, 2003; Howard & Murphy, 2007). The properties of the dynamic transducers are considerably different from the electret and condenser ones. In a dynamic transducer, the electric signals are produced through electromagnetic induction. The deflections of a microphone membrane due to varying sound pressure are transferred into movements of an induction coil or a ribbon in a permanent magnetic field, inducing electric signals (AKG Acoustics, 2003; Baken & Orlikoff, 2000; Howard & Murphy, 2007; Merhaut, 1980). The dynamic transducers have the advantage of not requiring a power supply, but their frequency response is, on average, considerably worse than that of the electret and condenser transducers.

The electret and condenser microphones typically have a flatter frequency response. The electret and condenser transducers both require power, either from a battery or by phantom power from the microphone amplifier. In these microphones, the microphone membrane acts as a capacitor plate; when deflected due to acoustic pressures, the voltage of the capacitor changes accordingly, causing an electric signal (AKG Acoustics, 2003; Baken & Orlikoff, 2000; Howard & Murphy, 2007; Merhaut, 1980). The difference between the electret and condenser transducer types is in the way the capacitor is polarized: The condenser microphone achieves the polarization entirely through an externally applied voltage, while the electret microphone uses a permanently polarized ferroelectric material (Elko & Harney, 2009; Sessler & West, 1962). These technicalities are, however, of little concern for the user. Practical consequences are that the electret microphones are less expensive than the condenser ones. It has been assumed that the performance of the electret microphone may deteriorate over time, but current measurements have shown only very small changes of their sensitivity after more than 10 years of use when handled carefully (Yasuno & Miura, 2006). Apart from the disadvantage of power requirement, the electret and condenser microphones outperform dynamic microphones in almost every aspect, for measurement purposes. The ANSI S1.15-1997 standard (ANSI, 2006) specifies laboratory standard microphones to be of condenser type.

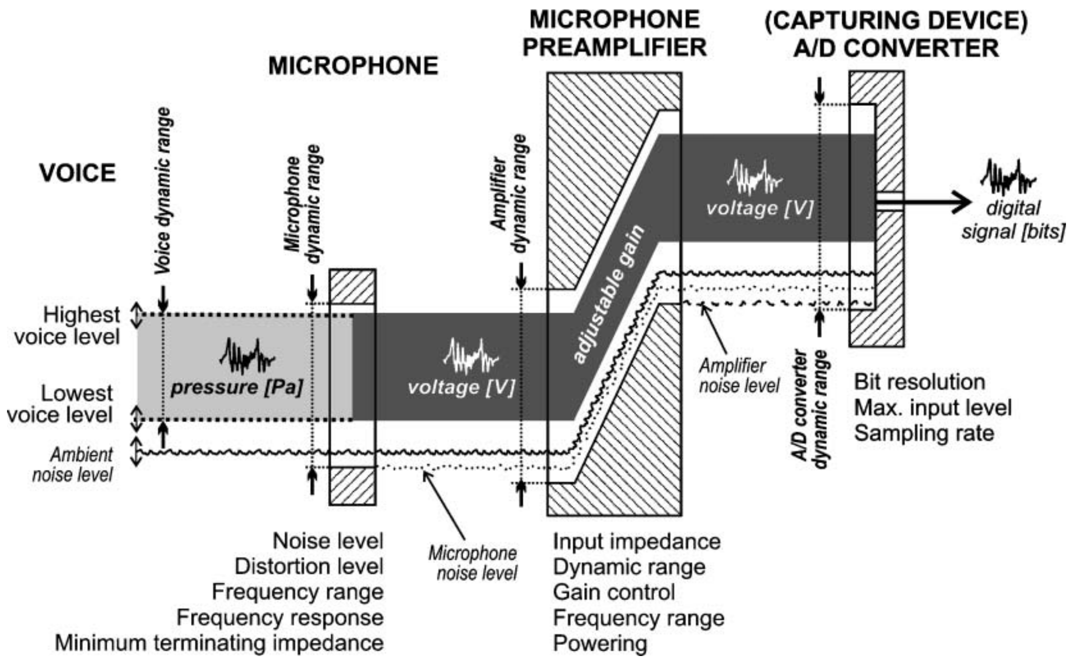
Microphone Preamplifier

To capture the sound, a microphone needs to be connected to a microphone preamplifier and to a capturing device (see Figure 9). The function of the preamplifier is to adjust the relatively weak voltage levels of the microphone signal (around 1 mV) to the standard line levels (around 1 V), which allow the sound to be captured by a standard sound-recording device (such as a digital recorder or computer). There are many microphone preamplifiers on the market, and their properties vary considerably. It is beyond the scope of this article to deal with the electronics connected to the preamplifiers. We will, however, briefly mention their most important parameters—that is, the input impedance, dynamic range, gain, frequency range, and powering.

Each microphone requires a minimum load impedance (i.e., *input impedance*) of the preamplifier. This is normally listed in the microphone specifications as the "minimum terminating impedance" of the microphone. If the impedance of the preamplifier is too low, the sensitivity of the microphone may be lowered, and more serious, the low-frequency response of the microphone may become impaired. For instance, miniature, head-mounted electret microphones often require a minimum terminating impedance of around 5 kΩ, while the input impedance of many microphone preamplifiers is only around 2 kΩ, thus not allowing optimal transmission of the microphone signals. The microphone's and preamplifier's data sheets should therefore be consulted to verify that the input impedance is at least as high as the minimum terminating impedance required by the microphone.

The dynamic range of the preamplifier should correspond to the dynamic range of the microphone. When the dynamic range of the preamplifier is smaller, it will either add noise to the microphone signal or introduce clipping at high levels. The dynamic range of the preamplifier should therefore also be at least 15 dB greater than the maximum dynamic range of voice, to avoid pollution by noise and clipping, as in the

FIGURE 9. The path of the sound signal through the microphone and the preamplifier to the digital capturing device. The parameters that should be well selected in voice measurement are listed below each of the equipment parts. The goal is to keep the F_0, spectrum, and level of the voice signal unperturbed as well as to keep the noise level well below the signal level. A/D = analog-to-digital.



case of the microphone (recall Figures 7 and 8). For a microphone having a dynamic range of 115 dB (i.e., 15–130 dB), the preamplifier is expected to operate over at least the same dynamic range. Many preamplifiers do not satisfy this condition, making the measurements over the total dynamic range of voice a nontrivial task. Some authors have bypassed this problem by registering voice simultaneously via two microphones placed at different distances (e.g., 5 cm and 30 cm) and using the closer one for measuring the soft phonations and the distant one for the loud phonations (Pabon, 2007; Šrámková, 2010). For the sake of brevity, such special solutions are not taken up in detail here.

Normally, preamplifiers have a gain knob to adjust the level of the signal. The gain allows a user to adjust the signal levels so that they match the levels of the capturing device. The adjustment should be done so that the loudest sounds do not overload the preamplifier (causing signal clipping) but are close to the upper limit of the preamplifier's dynamic range (Ternström & Granqvist, 2010). Some preamplifiers do not provide high enough attenuation to avoid clipping at very high sound levels, even though the microphone can handle the signal. In that case, such a preamplifier is not suitable to be used with the microphone. The basic rule is that the gain of the preamplifier should allow the signal level to be adjusted so that the maximum voice levels are slightly below the maximum recordable levels of that device.

As with microphones, the frequency range of the preamplifier should be large enough to capture the whole spectrum of voice from the lowest frequencies to the highest spectral components of interest, and the frequency response should be flat over this range. Some preamplifiers include frequency equalizers or bass/treble knobs allowing modification of the sound spectrum (Howard & Murphy, 2007). For measurement purposes, such modifications should be avoided.

As mentioned in the previous section, condenser and electret microphones need to be powered. Most frequently, the power is delivered through the microphone cable from the so-called "phantom power supply" unit, which provides a standardized voltage of 48 V. Many preamplifiers contain such a phantom power unit. Some measurement microphones require a specialized power supply that is not compatible with 48-V phantom power. Other microphones, especially the least expensive ones, use a battery for power. Battery operation can be problematic, since the health of the battery often affects the microphone sensitivity. Supplying power from the preamplifier is more reliable. The choice between standardized 48-V phantom power and other specialized solutions is largely determined by the choice of microphone. One advantage with the specialized solution is that the interfacing is taken care of by the same manufacturer, and thus compatibility can be guaranteed for power, impedance as well as levels. Purchasing a preamplifier together with a microphone made by the same manufacturer guarantees their compatibility and is therefore usually a safe and good solution.

Digital Capturing Device for Microphone Signals

For capturing the voltage signal coming from the microphone preamplifier with a digital recording device, an analog-to-digital (A/D) converter needs to be used (see Figure 9).

A/D converters are embedded in digital sound-capturing devices, including sound cards in computers (Ternström & Granqvist, 2010). Here, we will mention three important parameters of the A/D converter that are most closely related to the microphone characteristics: bit resolution, maximum input level, and sampling rate.

The digital dynamic range of the A/D converter is determined by its bit resolution (Watkinson, 1998). To capture the complete dynamic range of voice and ensure the minimum of 15-dB signal-to-noise ratio, the dynamic range of about 115 dB is needed (corresponding to the range of 15–130 dB at 30 cm determined earlier for microphones). According to the information provided in Table 1, this indicates that a universal A/D converter for voice signals should have a resolution of 20 bits or more. Many consumer-grade audio devices (including computer sound cards) offer resolution of only 16 bits, corresponding to 96-dB maximum theoretical range (Ternström & Granqvist, 2010). This means that such devices cannot record the extreme ranges of human voices without adjusting the gain. There have been attempts to reduce these demands by, for example, splitting a single microphone signal into two channels with different gains, or using an automatic attenuator capable of switching the gain at specific levels (e.g., the Kay Pentax Voice Range Profile Program Model 4326). The details of such special solutions go beyond the scope of this article. The good news is that recent high-quality professional sound-recording devices offer the resolution of 24 bits, which corresponds to a theoretical dynamic range of 141 dB. Even though the actual dynamic range of the electronics rarely exceeds 120 dB, it is sufficient to cover the whole dynamic range of voice up to its very extremes without gain adjustment.

A/D converters have a maximum input level above which the signal gets clipped. In professional audio equipment, the maximum input level is usually around 7 V (European Broadcasting Union, 1979). In consumer-grade audio equipment, this value is usually smaller—around 1 V. As mentioned in the previous section, the maximum voice levels coming out of the preamplifier should be adjusted so that they are close to but not exceeding the maximum input level of the A/D converter. For this purpose, the recording equipment usually offers an indicator (such as a red color) that warns about clipping. When recording, clipping should always be avoided.

The sampling rate (also called the sampling frequency, or F_S) should be at least twice as high as the frequency of the highest spectral component of interest to digitally capture the whole spectrum of voice, according to Shannon's (1949) sampling theorem. The highest frequency captured by an A/D converter is called the Nyquist frequency (F_N) and is equal to half the sampling rate ($F_N = F_S/2$; Allaby & Allaby, 1999). This means that if the voice range up to 8 kHz is to be captured, a sampling rate of at least 16 kHz is required. For capturing voice signal components up to 16–20 kHz (upper limit of the human hearing range), sampling rates of at least 32–40 kHz are needed. Standard audio recorders usually operate at sampling rates of 44.1 kHz or 48 kHz, which allow for capturing sounds over the whole hearing range of humans. Using these standard sampling rates is often preferable, since nonstandard rates may introduce sampling rate conversion artifacts. In special cases, such as in highly accurate frequency perturbation measurements, even higher sampling rates may be advantageous (Titze, 1995; Titze, Horii, & Scherer, 1987).

When sending the signal from the microphone preamplifier to the audio-recording device, the socket designated as "line in" is used. Some recording devices offer a "mic in" socket, which is intended to be used as an input for a microphone signal without a preamplifier. This indicates that the device has a microphone preamplifier integrated in the equipment. In the case of using the mic in socket, the embedded preamplifier should be checked for parameters as described in the previous section.

TABLE 1. Bit resolution and corresponding theoretical dynamic range of an A/D converter.

Bit resolution	Theoretical dynamic range
8 bits	45 dB
12 bits	69 dB
16 bits	93 dB
18 bits	105 dB
20 bits	117 dB
24 bits	141 dB

Note. Based on Watkinson (1998, p. 246, triangular dither, Formula 8.2).

Recommendations

Based on the information provided above, we can formulate the following recommendations for microphones:

1. The noise level (i.e., the low dynamic limit) of the microphone should be at least 15 dB below the softest produced voice level. The same criterion should be considered for the ambient noise level.
2. The upper dynamic limit of the microphone (i.e., the 3% THD level) should be at least as high as the loudest produced voice level.
3. The low-frequency limit of the microphone should be lower than the lowest produced frequency of the voice.
4. The upper frequency limit of the microphone should be higher than the highest spectral frequency of interest.
5. The frequency response of the microphone between the low and upper frequency limit should be flat.
6. Directional microphones should be used for SPL and spectral measurements only at the distance at which the frequency response is flat, to avoid the proximity effect. That distance should be found in the microphone specifications. If the distance is not known, the microphone is not considered suitable for the SPL and spectral measurements of voice and speech.

These recommendations ensure that the F_0, spectrum, and SPL of the voice are not considerably affected in the captured sound. Table 2 summarizes the microphone recommendations and provides the corresponding numerical values for the two microphone distances (30 cm and 5 cm) described earlier (for different distances, the values should be interpolated from the data). The values are considered for two voice ranges: (a) the extreme range—that is, the overall range of voice up to its reported extreme limits; and (b) the limited range of voice, which encompasses the voice and speech in the common situations but not for extreme phonations or for frequencies above 8 kHz.

There may be specialized measurements for which these numerical requirements may still not be enough. However, in some cases (such as for measurements only at comfortable voice levels), these requirements may be too strict.

Discussion

Based on our general recommendations, it can be said that different phonation tasks pose different demands on microphones. For simple measurements of comfortable phonations, the requirements are different than for advanced measurements of very loud or very soft phonations. Furthermore, head microphones positioned close to the mouth (distance around 5 cm) are expected to have different specifications than microphones mounted on a stand at a distance of 30 cm (see Table 2). The requirements also imply that measurements of F0 perturbations have different demands (i.e., signal-to-noise ratio of 30 dB or more; Deliyski et al., 2006; Perry et al., 2000) than measurements of the voice SPL or voice spectrum (i.e., flat frequency response of the microphone).

While the best selection of microphone strongly depends on the purpose of the recording, it is often desirable to have a microphone that works well for several purposes. Microphones that cover the entire frequency and dynamic range of the voice are considerably more expensive than microphones that cover a limited part of this range. While inexpensive microphones can work well for a given purpose, there is always the risk that the microphone will be used outside its capabilities. In this perspective, microphones of high quality typically turn out to be cost-effective in spite of their high price tag.

An Internet search on characteristics of microphones available on the market in 2008 revealed that many commercial microphones did not fulfill the recommendations offered in Table 2 (Šrámková, 2010; Švec et al., 2009). In some head-mounted microphones, the upper dynamic limit was around 130 dB, which is too low to capture the loudest voice at the distance of approximately 5 cm. The frequency response of many microphones was not sufficiently flat and exhibited a "presence peak" (i.e., level gain of up to 7 dB at frequencies around 3–10 kHz). In directional microphones, the reference distance for the flattest response was often not provided. This indicates that the task of selecting a microphone should not be taken lightly.

Laboratory measurement microphones of the LS type (according to ANSI standards) guarantee that, except for noise level, all the recommendations in Table 2 are met. The LS1-type microphones guarantee the parameters for the limited range, whereas the LS2-type microphones satisfy the recommendations also for the extreme range. The maximum noise levels of the microphones are not prescribed by the ANSI standard and should be checked in the specification sheet of the microphone; many of the microphones of LS type fulfill the noise recommendations from Table 2.

How do our recommendations compare with the recommendations of other authors? For voice perturbation measurements, Titze (1995, p. 28) offered the following recommendations on microphones:

> For type 1 signals for which a perturbation measure of the order of 0.1% is to be extracted to 10% accuracy, the following recommendations are made:
>
> a. A professional-grade condenser microphone (omnidirectional or cardioid) with a minimum sensitivity of –60 dB should be used (Titze & Winholtz, 1993).
>
> b. For steady vowel utterances, the mouth-to-microphone distance can be held constant and less than 10 cm (preferably 3–4 cm) in order to avoid an artificial wow and to

TABLE 2. Preliminary recommendations for microphones intended for voice and speech measurements.

Microphone specifications	Recommendation	Extreme range 30 cm	Extreme range 5 cm[a]	Limited range 30 cm	Limited range 5 cm[a]
Noise level	At least 15 dB below softest voice level[b]	≤15 dBA	≤30 dBA	≤30 dBA	≤45 dBA
Maximum level (3% total harmonic distortion)	Above maximum voice level	≥130 dB	≥145 dB	≥120 dB	≥135 dB
f_L: lower frequency limit (–2 dB)	Below the lowest frequency of voice[c]	≤10 Hz	≤10 Hz	≤50 Hz	≤50 Hz
f_U: upper frequency limit (–2 dB)	Above the highest spectral frequency of interest	>16 kHz	>16 kHz	>8 kHz	>8 kHz
Flatness between f_L and f_U (except a gain above 5 kHz)	Flat	≤2 dB	≤2 dB	≤2 dB	≤2 dB
Maximum gain between 5 kHz and f_U	No gain[d]	≤1 dB	≤3 dB	≤3 dB	≤5 dB

[a]The suitability of the mouth-to-microphone distance of 5 cm for accurate measurements of voice needs verification due to uncertainties of sound radiation in the proximity of the head.
[b]For perturbation measures, the 15-dB signal-to-noise ratio (SNR) may not be sufficient, and 30 dB SNR should instead be considered (Deliyski et al., 2006; Perry et al., 2000).
[c]Inverse filtering may require this limit to be about a decade (1/10) below the F_0 of voice.
[d]For the head-mounted microphones mounted on the side of the mouth, it might be advantageous to have a presence peak, because of the loss of high frequencies at the side of the head (Cabrera et al., 2002; Dunn & Farnsworth, 1939; Marshall & Meyer, 1985). There are, however, no standards on how large this peak should be, and therefore our recommendation is to use a microphone at 30 cm in front of the mouth for measurements in which the high-frequency content above 5 kHz is critical.

maintain a high signal-to-noise ratio; a miniature head-mounted microphone is recommended (Winholtz & Titze, [1997]).

These recommendations do not conflict with the recommendations provided in this article. It should be noted, however, that perturbation measurements do not pose as high demands on the spectral properties of the sound (hence the proximity effect of the cardioid microphone is not of much concern here) or on the dynamic range of the microphone, since the voice is usually produced at comfortable levels. A high signal-to-noise ratio is, however, very important for perturbation measures because the noise signal can cause a considerable contamination of the voice signal (recall Figure 7).

The very small microphone distance of 3–4 cm recommended by Titze (1995) for perturbation measurements has a beneficial effect of increasing the signal level with respect to the noise level (of the microphone as well as the room), thus improving the signal-to-noise ratio. At distances smaller than 20 cm, however, the microphone is usually positioned to the side of the mouth, which may cause problems for accurate spectral and SPL measurements of voice and speech because voice has been found to radiate less high-frequency energy (above 1 kHz) to the side compared with the front of the mouth (Cabrera, Davis, Barnes, Jacobs, & Bell, 2002; Dunn & Farnsworth, 1939; Marshall & Meyer, 1985). Also, small changes of microphone position at such very close distance may cause nonnegligible changes in the measured SPL. And in the case of loud voice, this close position requires that the microphone be capable of recording very high sound levels (potentially even up to 147 dB) without adding distortion to the signal. The placement of the microphone for voice recordings is therefore an issue that deserves more investigation. In the interest of coherence of the topic, we plan to address this problem in a separate article.

Schutte and Seidner (1983) recommended the room noise levels to be lower than 40 dBA when the microphone is put on a stand and placed at 30 cm distance from the lips. When the microphone is head-mounted at 10 cm distance from the lips, Dejonckere et al. (2001, p. 78) specified that "the recordings should be made ideally in a sound-treated room, but a quiet room with ambient noise < 50 dB is acceptable." In light of our analysis, these recommendations would be acceptable only for voice produced at comfortable or high levels, exceeding 55 dBA SPL at 30 cm, or 65 dB SPL at 10 cm. For soft voices approaching phonation threshold levels of 30 dBA at 30 cm, much lower noise levels are required according to our recommendations—that is, 15, 25, or 30 dBA when the microphone is placed at 30, 10, or 5 cm, respectively.

A cardioid microphone has been considered by some authors to be the best choice when voice is measured in clinics (Baken & Orlikoff, 2000). Cardioid microphones typically suppress the noise level in the room by 5 dB (recall Figure 4), which can be helpful in busy clinics, especially when perturbation measurements are of interest. On the other hand, the proximity effect of the microphone may distort the spectral measurements, such as the soft phonation index parameter in the commonly used Multi-Dimensional Voice Program analysis software (Kay Elemetrics, 1999; see Deliyski, 1993; Muñoz, Mendoza, Fresneda, Carballo, & López, 2003; Roussel & Lobdell, 2006), the H1/H2 parameter (Björkner, 2008; Salomao & Sundberg, 2008), or the alpha ratio (Frokjaer-Jensen & Prytz, 1976; Ilomäki, Laukkanen, Leppänen, & Vilkman, 2008; Laukkanen, Ilomäki, Leppänen, & Vilkman, 2008; Master, De Biase, Chiari, & Laukkanen, 2008; Waaramaa, Laukkanen, Alku, & Vayrynen, 2008) and cause inaccuracy even in calibrated measurements of voice SPL. According to our recommendations, the use of cardioid and other directional microphones is problematic when they are not positioned at the distance for which the microphone frequency response is flat.

Before using a microphone for measurement purposes, it is therefore important to study its specifications. These specifications are normally supplied in the documentation accompanying the microphone and can usually also be found on the manufacturer's website. In microphones fulfilling the LS standard, the specifications should be accurate. In microphones that are not classified (i.e., not fulfilling the LS1 or LS2 standard), the accuracy and tolerances of the general microphone specifications provided by the manufacturer in the data sheet or on the web may, however, be questionable—there has not been enough information to answer this question. Some manufacturers, however, do supply individually measured accurate specifications for their microphones. Such microphones can be considered reliable for measurement purposes. For unclassified microphones without individually measured characteristics, it may be worth having them tested and measured by a specialist. Certainly, when the characteristics are not known, the microphone should not be considered suitable for voice measurement.

Conclusion

Despite the fact that voice and speech measurements are carried out routinely for clinical and research purposes, the subject of microphone selection has not received sufficient attention in the voice and speech literature. In this article, we have attempted to lay down some fundamental principles to guide the selection of microphones. While these recommendations can be considered preliminary, they provide a basis for improvement of accuracy of the measurements. It is our hope that improved knowledge about microphones and their characteristics will allow researchers to make more accurate measurements of voice and speech in the future.

Acknowledgments

The study was supported by the Grant Agency of the Czech Republic, Project GACR 101/08/1155, and by the Wenner-Gren Foundation in Sweden. The research was supported also by the COST Action 2103 Advanced Voice Function Assessment. The authors would like to thank Skype for the connection. Both authors made an equal contribution to the article.

References

AKG Acoustics. (2003). *The ABC's of AKG: Microphone basics & fundamentals of usage*. Nashville, TN: Author.

Allaby, A., & Allaby, M. (1999). Nyquist frequency. In *A dictionary of earth sciences*. Retrieved from www.encyclopedia.com/doc/1O13-Nyquistfrequency.html.

American National Standards Institute. (1985). *Specification for sound level meters* (ANSI S1.4-1983). Melville, NY: Acoustical Society of America.
American National Standards Institute. (2002). *Acoustical performance criteria, design requirements, and guidelines for schools* (ANSI S12.60-2002). Melville, NY: Acoustical Society of America.
American National Standards Institute. (2006). *Measurement microphones. Part 1: Specifications for laboratory standard microphones* (ANSI S1.15-1997/Part 1 [R2006]). Melville, NY: Acoustical Society of America.
Angerstein, W., & Neuschaefer-Rube, C. (1998). Sound pressure level examinations of the calling and speaking voice in healthy persons and in patients with hyperfunctional dysphonia. *Logopedics Phoniatrics Vocology, 23,* 21–26.
Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego, CA: Singular.
Björkner, E. (2008). Musical theater and opera singing—why so different? A study of subglottal pressure, voice source, and formant frequency characteristics. *Journal of Voice, 22,* 533–540.
Bless, D. M., Baken, R. J., Hacki, T., Fritzell, B., Laver, J., Schutte, H. K., . . . Hurme, P. (1992). International Association of Logopedics and Phoniatrics (IALP) Voice Committee discussion of assessment topics. *Journal of Voice, 6,* 194–210.
Brüel & Kjaer. (1984). *Measuring sound*. Naerum, Denmark: Author.
Brüel & Kjaer. (1996). *Microphone handbook*. Naerum, Denmark: Author.
Cabrera, D., Davis, P., Barnes, J., Jacobs, M., & Bell, D. (2002). Recording the operatic voice for acoustic analysis. *Acoustics Australia, 30,* 103–108.
Dang, J., & Honda, K. (1997). Acoustic characteristics of the piriform fossa in models and humans. *The Journal of the Acoustical Society of America, 101,* 456–465.
Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., . . . Woisard, V. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *European Archives of Oto-Rhino-Laryngology, 258,* 77–82.
Deliyski, D. D. (1993). Acoustic model and evaluation of pathological voice production. *Proceedings from EUROSPEECH '93: Third European Conference on Speech Communication and Technology* (pp. 1969–1972). Retrieved from www.isca-speech.org/archive/eurospeech_1993/e93_1969.html.
Deliyski, D. D., Shaw, H. S., Evans, M. K., & Vesselinov, R. (2006). Regression tree approach to studying factors influencing acoustic voice analysis. *Folia Phoniatrica et Logopaedica, 58,* 274–288.
DPA Microphones. (2007). *Calculating the dynamic range of a microphone*. Retrieved from www.dpamicrophones.com.
Dunn, H. K., & Farnsworth, D. W. (1939). Exploration of pressure field around the human head during speech. *The Journal of the Acoustical Society of America, 10,* 184–199.
Eargle, J. (2001). *The microphone book*. Boston, MA: Focal Press.
Elko, G. W., & Harney, K. P. (2009). A history of consumer microphones: The electret condenser microphone meets micro-electro-mechanical-systems. *Acoustics Today, 5,* 4–13.
European Broadcasting Union. (1979). *The E.B.U. Standard peak-programme meter for the control of international transmission* (Tech. 3205-E, 2nd ed.). Brussels, Belgium: Author.
Fant, G. (1959). Acoustic analysis and synthesis of speech with applications to Swedish. *Ericsson Technics, 1,* 1–106.
Frokjaer-Jensen, B., & Prytz, S. (1976). Registration of voice quality. *Brüel & Kjaer Technical Review, 3,* 3–17. Retrieved from www.bksv.com/doc/technicalreview1976-3.pdf.
Herzel, H. (1996). Possible mechanisms of vocal instabilities. In P. J. Davis & N. H. Fletcher (Eds.), *Vocal fold physiology: Controlling complexity and chaos* (pp. 63–75). San Diego, CA: Singular.
Heylen, L., Wuyts, F. L., Mertens, F., De Bodt, M., & Van de Heyning, P. H. (2002). Normative voice range profiles of male and female professional voice users. *Journal of Voice, 16,* 1–7.
Howard, D. M., & Murphy, D. (2007). *Voice science, acoustics, and recording*. San Diego, CA: Plural.
Hunter, E. J., Švec, J. G., & Titze, I. R. (2006). Comparison of the produced and perceived voice range profiles in untrained and trained classical singers. *Journal of Voice, 20,* 513–526.
Ilomäki, I., Laukkanen, A. M., Leppänen, K., & Vilkman, E. (2008). Effects of voice training and voice hygiene education on acoustic and perceptual speech parameters and self-reported vocal well-being in female teachers. *Logopedics Phoniatrics Vocology, 33,* 83–92.
International Electrotechnical Commission. (1992). *International standard: Measurement microphones—Part 1: Specifications for laboratory standard microphones* (IEC 61094-1:1992). Geneva, Switzerland: Author.
International Electrotechnical Commission. (2002). *Electroacoustics: sound level meters—Part 1: Specification* (IEC 61672-1). Geneva, Switzerland: Author.
Kay Elemetrics. (1999). Multi-Dimensional Voice Program Model 5105 (Version 2.0) [Computer software]. Lincoln Park, NJ: Author.
Laukkanen, A. M., Ilomäki, I., Leppänen, K., & Vilkman, E. (2008). Acoustic measures and self-reports of vocal fatigue by female teachers. *Journal of Voice, 22,* 283–289.
Laver, J., Hiller, S., & Beck, J. M. (1992). Acoustic waveform perturbations and voice disorders. *Journal of Voice, 6,* 115–126.
Leino, T., Laukkanen, A. M., Ilomäki, I., & Mäki, E. (2008). Assessment of vocal capacity of Finnish university students. *Folia Phoniatrica et Logopaedica, 60,* 199–209.
Ma, E., Robertson, J., Radford, C., Vagne, S., El-Halabi, R., & Yiu, E. (2007). Reliability of speaking and maximum voice range measures in screening for dysphonia. *Journal of Voice, 21,* 397–406.
Marshall, A. H., & Meyer, J. (1985). The directivity and auditory impressions of singers. *Acustica, 58,* 130–140.
Master, S., De Biase, N., Chiari, B. M., & Laukkanen, A. M. (2008). Acoustic and perceptual analyses of Brazilian male actors' and nonactors' voices: Long-term average spectrum and the "actor's formant." *Journal of Voice, 22,* 146–154.
Merhaut, J. (1980). *Theory of electroacoustics*. New York, NY: McGraw-Hill.
Muñoz, J., Mendoza, E., Fresneda, M. D., Carballo, G., & López, P. (2003). Acoustic and perceptual indicators of normal and pathological voice. *Folia Phoniatrica et Logopaedica, 55,* 102–114.
Neubauer, J., Edgerton, M., & Herzel, H. (2004). Nonlinear phenomena in contemporary vocal music. *Journal of Voice, 18,* 1–12.
Pabon, P. (2007). Voice Profiler (Version 4.0) [Computer software]. Utrecht, the Netherlands: Author.
Perry, C. K., Ingrisano, D. R., Palmer, M. A., & McDonald, E. J. (2000). Effects of environmental noise on computer-derived voice estimates from female speakers. *Journal of Voice, 14,* 146–153.
Rothenberg, M. (1973). A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *The Journal of the Acoustical Society of America, 53,* 1632–1645.

Roussel, N. C., & Lobdell, M. (2006). The clinical utility of the soft phonation index. *Clinical Linguistics & Phonetics, 20,* 181–186.

Salomao, G. L., & Sundberg, J. (2008). Relation between perceived voice register and flow glottogram parameters in males. *The Journal of the Acoustical Society of America, 124,* 546–551.

Schneider, B., & Bigenzahn, W. (2003). Influence of glottal closure configuration on vocal efficacy in young normal-speaking women. *Journal of Voice, 17,* 468–480.

Schutte, H. K., & Seidner, W. (1983). Recommendation by the Union of European Phoniatricians (UEP): Standardizing voice area measurement/phonetography. *Folia Phoniatrica et Logopaedica, 35,* 286–288.

Sessler, G. M., & West, J. E. (1962). Self-biased condenser microphone with high capacitance. *The Journal of the Acoustical Society of America, 34,* 1787–1788.

Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE, 37,* 10–21.

Spielman, J., Starr, A. C., Popolo, P. S., & Hunter, E. J. (2007). *Recommendations for the creation of a voice acoustics laboratory* [NCVS Online Technical Memo No. 7, Version 1.4]. Retrieved from www.ncvs.org/e-learning/technical.html.

Šrámková, H. (2008). *Technické požadavky pro akustickou registraci hlasu a řeči* [Technical requirements on acoustic registration of voice and speech] (Unpublished bachelor's thesis). Palacký University Olomouc, Olomouc, Czech Republic.

Šrámková, H. (2010). *Vyšetření dynamického rozsahu akustického tlaku u lidského hlasu* [Examination of dynamic ranges of sound pressures in human voice] (Unpublished master's thesis). Palacký University Olomouc, Olomouc, Czech Republic.

Sulter, A. M., Schutte, H. K., & Miller, D. G. (1995). Differences in phonetogram features between male and female subjects with and without vocal training. *Journal of Voice, 9,* 363–377.

Švec, J. G., Schutte, H. K., & Miller, D. G. (1996). A subharmonic vibratory pattern in normal vocal folds. *Journal of Speech and Hearing Research, 39,* 135–143.

Švec, J. G., Šrámková, H., & Granqvist, S. (2009). Basic requirements on microphones for voice recordings. In C. Manfredi (Ed.), *Models and analysis of vocal emissions for biomedical applications: 6th International Workshop* (pp. 157–160). Firenze, Italy: Firenze University Press.

Ternström, S. (2008). High fidelity on the line: Please say "aah." *Echoes Newsletter of the Acoustical Society of America, 18,* 4–5.

Ternström, S., & Granqvist, S. (2010). Personal computers in the voice laboratory: Part two—audio devices. *Logopedics Phoniatrics Vocology, 35,* 98–102.

Titze, I. R. (1995). *Workshop on acoustic voice analysis. Summary statement*. Salt Lake City, UT: National Center for Voice and Speech. Retrieved from www.ncvs.org/museum-archive/sumstat.pdf.

Titze, I. R., Horii, Y., & Scherer, R. C. (1987). Some technical considerations in voice perturbation measurements. *Journal of Speech and Hearing Research, 30,* 252–260.

Titze, I. R., & Winholtz, W. S. (1993). Effect of microphone type and placement on voice perturbation measurements. *Journal of Speech and Hearing Research, 36,* 1177–1190.

Tokuda, I. T., Horáček, J., Švec, J. G., & Herzel, H. (2007). Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments. *The Journal of the Acoustical Society of America, 122,* 519–531.

Waaramaa, T., Laukkanen, A. M., Alku, P., & Vayrynen, E. (2008). Monopitched expression of emotions in different vowels. *Folia Phoniatrica et Logopaedica, 60,* 249–255.

Watkinson, J. (1998). *The art of sound reproduction*. Woburn, MA: Focal Press.

Winholtz, W. S., & Titze, I. R. (1997). Miniature head-mounted microphone for voice perturbation analysis. *Journal of Speech, Language, and Hearing Research, 40,* 894–899.

Yasuno, Y., & Miura, K. (2006). Sensitivity change with long-term preservation and practical use of electret condenser microphone. *Acoustical Science and Technology, 27,* 302–304.

Received September 25, 2009
Revision received February 8, 2010
Accepted June 3, 2010
DOI: 10.1044/1058-0360(2010/09-0091)

Contact author: Jan G. Švec, Palacký University Olomouc, Faculty of Science, Department of Experimental Physics, Laboratory of Biophysics, 17. listopadu 12, 771 46 Olomouc, Czech Republic. E-mail: svecjang@gmail.com or jan.svec@upol.cz.