Psychophysical comparison of the auditory and tactile perception: a survey

Sebastian Merchel1 · M. Ercan Altinsoy1

Received: 15 January 2020 / Accepted: 16 June 2020 / Published online: 21 July 2020
© The Author(s) 2020

Abstract
In this paper, the psychophysical abilities and limitations of the auditory and vibrotactile modality will be discussed. A direct comparison reveals similarities and differences. The knowledge of these is the basis for the design of perceptually optimized auditory-tactile human–machine interfaces or multimodal music applications. Literature data and own results for psychophysical characteristics are summarized. An overview of the absolute perception thresholds of both modalities is given. The main factors which influence these thresholds are discussed: age, energy integration, masking and adaptation. Subsequently, the differential sensitivity (discrimination of intensity, frequency, temporal aspects and location) for suprathreshold signals is compared.

Keywords Auditory-tactile perception · Absolute thresholds · Differential sensitivity

1 Introduction

The perception of vibrations at the skin and sound are often coupled in real life, e.g., while playing an instrument or listening to music with low frequency content. In these cases, the physical stimuli which excite both modalities are usually highly correlated. If new multimodal systems are designed, sound and vibrations can be influenced separately. Just think of the auditory and vibrotactile feedback of a button on a touch screen, or vibrotactile feedback of electronic music instruments, or bimodal devices for guidance of blind persons. For example, the authors developed and optimized systems for multimodal reproduction of music [64,65,67]. To this end, a vibration actuator was coupled to a surface in contact with the listener, e.g., an electrodynamic shaker mounted in a backpack, integrated in clothing or attached below a seat or floor. Audio reproduction was implemented with conventional loudspeakers or headphones. To generate appropriate music-related vibrations from the audio signal various signal processing approaches were compared. It was found that it is beneficial to consider the perceptual capabilities and limitations of both modalities in this design process. Therefore, knowledge of the fundamental characteristics of the auditory and vibrotactile sensory modalities was necessary. Many similarities can be found regarding psycho-physical characteristics, although the anatomy and physiology of both modalities are quite different. A good overview of the basic structure and functionality of the human hearing organ as well as the histology and physiology of the mechanoreceptive system including the neural processing in the somatosensory and auditory areas of the brain can be found in [59,86] and will not be described here.

The current survey aims to compare the sense of hearing and touch using data from psychophysical experiments. Special attention is given to the perception of vibrations in the frequency range where sound and vibration perception overlap: between a few Hertz and several hundred Hertz. The authors hope that this overview helps to design good auditory-tactile feedback that matches perceptually. This paper is based on the dissertation of the first author [63]. Reproduction is kindly permitted by the Shaker Verlag, Germany.

The perception of sound has been studied for several decades. The basic physical attributes of sound (e.g., intensity, frequency or location of a sound source) have been correlated to perceptual attributes like loudness, pitch or distance. Different effects like adaptation to loud signals
or masking characterize the auditory system. In contrast to our hearing, vibrations can be perceived at different parts of the body. Most vibrotactile studies focus on vibrations transmitted via hand and finger. However, the principal mechanoreceptors in the skin are similar at different body sites. In the overlapping frequency range of auditory and vibrotactile perception, vibrations are likely to stimulate mainly the Meissner and Pacinian mechanoreceptors which can be found all over the body [86], however, with varying populations and surrounding tissue mechanics. Nevertheless, data from different body sites is used for a general comparison.

A common measurement unit for sound is the sound pressure level, $L_{\text{SPL}}$. It is defined as the logarithmic ratio of the effective value of the sound pressure $p$ and has a reference value $p_0 = 20 \mu \text{Pa}$:

$$L_{\text{SPL}} = 20 \log \frac{p}{p_0} \text{ dB.}$$

A similar unit for measuring vibrations is the acceleration level $L_{\text{acc}}$. It is defined as the logarithmic ratio of the acceleration $a$ and has a reference value $a_0 = 1 \mu \text{m/s}^2$:

$$L_{\text{acc}} = 20 \log \frac{a}{a_0} \text{ dB.}$$

In contrast to sound pressure level, 0 dB acceleration level is not related to the perception threshold. Therefore, sensation level (the level above threshold) will be used to compare the auditory and vibrotactile modality directly. Please note that within this paper the term ‘vibrotactile’ will be sometimes abbreviated as ‘tactile’. However, the article will not discuss other types of tactile sensations (e.g., temperature).

### 2 Absolute sensitivity

A fundamental characteristic of a sensory modality is the absolute perception threshold. Minimum and maximum perceivable levels for auditory and tactile perception will be discussed in this section. Basic effects like energy integration, masking and adaptation will be compared.

#### 2.1 Sensation area

**Auditory** Sound can be heard between approximately 20 Hz and 20 kHz. Below 20 Hz the tonal sensation ceases, and below 10 Hz single cycles of the sound can be perceived [71]. The upper frequency limit depends strongly on the age of the subject. Figure 1 shows that the hearing is most sensitive to sound pressure between approximately 300 Hz and 7000 Hz. It becomes less sensitive for decreasing and increasing frequency. In addition, the figure shows estimates for the pain threshold and the annoyance threshold after Winckel [112]. The curves of equal subjective intensity (equal loudness contours) are plotted according to ISO 226:2003 [52]. They follow the threshold curve to some degree. It can be seen that they get closer toward lower frequencies. The auditory dynamic range is thus frequency dependent from 50 dB to more than 100 dB.

The hair cells in the cochlea can be regarded as the most sensitive mechanoreceptors of the human body. The minimum perceivable sound pressure causes only $10^{-10}$ m displacement in the inner ear, which corresponds roughly to the diameter of a hydrogen atom [86].

**Tactile** In comparison the vibrotactile sense is rather limited. Only frequencies up to approximately 1 kHz can be perceived via the mechanoreceptive system. Similar to the ear, the vibration sensitivity of the skin depends on frequency. Figure 2 shows the frequency dependent perception threshold on the thenar eminence adapted from Verrillo et al. [103]. It can be seen that the glabrous skin becomes more sensitive to the acceleration of its surface with decreasing frequency. Similar results were reported for various regions of the body [45]. It was found that the sensitivity depends on the distribution and density of the mechanoreceptors, with lower thresholds for areas with higher receptor density [56]. Hairy skin is approximately 10–20 dB less sensitive depending on frequency [101].

The curves of equal subjective intensity follow the threshold to some degree. Again a frequency dependence can be seen, with smaller dynamic ranges for frequencies above approximately 300 Hz. At frequencies below 200 Hz, vibrations more than 40–55 dB above threshold become very...
unpleasant or painful [70]. The dynamic range can thus be quantified between approximately 40–50 dB.

Similar curves of equal vibration intensity have been measured by the authors for seat vibrations using two different methods: magnitude estimation and intensity matching. Interestingly, the slight frequency dependence of the dynamic range could not be confirmed [66].

The growth of perceived intensity above threshold is another very important aspect when comparing the auditory and vibrotactile modality. Compared to audition, the increase in perceived magnitude is steeper with increasing level in the vibrotactile domain, particularly at low sensation levels. For a detailed discussion of this relevant topic, the reader is referred to [63] where a new perceptually motivated measurement was proposed to represent human vibration intensity perception: the perceived vibration magnitude $M$ in $\text{vip}$, comparable to auditory loudness $N$ in $\text{sone}$.

### 2.2 Age and gender

**Auditory** The threshold of hearing rises naturally with increasing age. This effect is referred to as presbyacusis and involves primarily frequencies above 3000 Hz. Figure 3 presents data that depicts the progression of hearing loss with age [89]. The data is averaged over men and woman, however, it has been shown that presbyacusis starts more gradual in women but grows faster once started [8]. In addition, noise-induced hearing loss (sociacusis) is a common phenomenon today.

**Tactile** Similar to hearing, age has a considerable influence on vibrotactile thresholds. The sensitivity for high frequencies decreases progressively with age [91,102]. Figure 3 illustrates the shift of the vibrotactile detection threshold for four age groups [98]. At higher frequencies, where the Pacinian system is predominant, a strong loss of sensitivity can be observed with increasing age. No effect was found for low frequencies.

In general, no gender differences were found for vibrotactile thresholds between men and women [61,99]. Only Gescheider reported that women are slightly more sensitive to high-frequency vibrations at the thenar eminence a few days before menstruation [38].

### 2.3 Energy integration

An other important characteristic of the auditory and vibrotactile modality, which has an influence on the threshold, is the ability to integrate energy. This is often discussed using the relationship between the duration and the threshold (or intensity) of a stimulus.

**Auditory** The auditory threshold of detection decreases with increasing duration up to a stimulus length of approximately 1 s. This holds true for various types of stimuli over a broad frequency range [27]. Figure 4 shows data from Plomp and Bouman [83] and Florentine [24] for a stimulus frequency of 250 Hz. The curves follow the prediction made by the theory of temporal summation, which was formulated by Zwislocki in 1960 [113].

**Tactile** Temporal energy integration can also be found in the vibrotactile domain, but only in the Pacinian system [32,33]. No temporal summation was found for low frequencies, e.g.
Fig. 4 Auditory and vibrotactile threshold shift as a function of burst duration after [24, 83, 95]. Data are plotted in dB re threshold of detection for the longest stimulus of each curve. In all cases, the stimuli frequency was 250 Hz. The vibrotactile stimuli were applied to the skin of the hand using different contactor sizes at 25 Hz [36]. Data after Verrillo [95] are plotted for comparison in Fig. 4. Stimuli with a frequency of 250 Hz were delivered to the glabrous skin of the palm using a large contactor (2.9 cm²). He measured a 3 dB reduction of threshold per doubling of duration up to a stimulus length of 300 ms, indicating a complete integration of energy. Similar curves were found at 100 Hz and 500 Hz, frequencies at which mainly the Pacinian corpuscles are responsive to vibration. The same trend was found in suprathreshold experiments [3]. Other experiments by the author with seat vibrations at 40 Hz, 80 Hz, 160 Hz and 320 Hz confirmed the above conclusions but are not plotted here for clarity [68]. The data agrees well with the curves found in the auditory domain in spite of fundamentally different biomechanical conditions of the tactile sense compared to hearing. It remains open if this suggests similar perceptual mechanisms or if it can be explained otherwise, e.g., by surrounding tissue mechanics.

Additional curves for smaller contactor sizes (0.05 cm² and 0.02 cm²) can be seen in Fig. 4 [95]. As the size of the stimulated area is reduced, the dependence of duration upon the threshold is accordingly reduced. Using smaller contact areas, more and more non-Pacinian receptors will be stimulated [86]. Consequently, the amount of temporal summation declines.

In addition, absolute vibrotactile sensitivity at higher frequencies depends strongly on the size of the stimulated area. It has been shown that for frequencies between 80 and 320 Hz (Pacinian channel) the threshold decreases with 3 dB per doubling of contact area at the thenar eminance of the hand [94, 96]. Similar results have been reported for the hairy skin at the forearm [97]. No effects were found for lower frequencies [36].

Until now, only a single stimuli has been examined. However, in everyday life, two or more simultaneous stimuli are not unusual. If subjects are asked to judge the combined intensity of two tones, the result is proportional to the overall energy if the frequencies lie within a critical band in audition. However, if frequency components outside the critical bandwidth are added, the perceived intensity grows much stronger and the sensation magnitudes of the individual components can be summed [20]. Interestingly, similar effects have been found in the vibrotactile domain. Evidence for energy integration within the Pacinian channel has been discussed above and addition of sensation magnitudes between mechano-receptive channels has been reported [60, 104]. It was therefore suggested that the Pacinian channel is analogous to a critical band in the auditory system [57].

2.4 Masking

If multiple stimuli are heard or felt in close temporal proximity, they might interfere. One such effect is the suppression of one stimulus by another, which is called masking.

Auditory Early experiments used two sinusoids as masker and test signal to investigate masking ([109] as cited by [73]). However, when both signals were close together in frequency, beats occurred and complicated the results. To avoid this problem, later studies used narrow band noise as masker. The shifted threshold for detecting a test tone at various frequencies in the presence of a masker with fixed center frequency and amplitude was determined. This masked threshold is sometimes called masked audiogram or masking pattern. It is strongly correlated with the excitation pattern the masker generates on the basilar membrane [10]. An exemplary masking pattern is shown in Fig. 5 with data from [13]. For the plotted curve, a 90 Hz wide band of masking noise is centered at 410 Hz with 40 dB SPL. A narrow masking region can be seen. However, for higher sensation levels, which are not plotted here, the masking pattern spreads especially towards the high-frequency side.

In general, auditory masking patterns are dependent on masker frequency, duration and level. They show steep slopes towards lower frequencies and less steep slopes towards higher frequencies on a logarithmic frequency axis. However, towards low sensation levels or low frequencies, masking patterns are getting more and more symmetrical [13, 93], as illustrated in Fig. 5. Interestingly, low frequency maskers (e.g., at 150 Hz) seem to have their maximum effect slightly shifted towards higher frequencies [93] and their masking pattern broadens significantly [14, 15].

In the above studies, masker and test signal have been presented to the same ear or both ears diotically. However, even
Fig. 5 Auditory and vibrotactile masked thresholds relative to unmasked condition as a function of frequency. The vibration masker were narrow band noises centered at 31.5 Hz, 63 Hz, 160 Hz and 275 Hz with fixed level approximately 25 dB above threshold. Data from Stamm, Altinsoy and Merchel [90] are plotted for whole-body vibrations (25 Hz noise bandwidth) and from Gescheider et al. [39] for vibrations at the thenar eminence (100 Hz noise bandwidth). For comparison, an auditory masking pattern is plotted for a 90Hz wide band of masking noise, centered at 410 Hz with 40 dB SPL [13]. Test stimuli were simultaneously presented sinusoids in all conditions for dichotic conditions masking was found [16,17]. Therefore, central processing must be involved in the masking process, since the masker is presented to one ear and the test signal to the other.

Even if the masker and the test signal are presented one after the other, masking effects have been reported. This is referred to as post-masking (forward masking) if the test signal comes slightly behind the masker, or pre-masking (backward masking) if the test signal precedes the masker as is illustrated in Fig. 6 using data from Elliott [16]. A 50 ms long white noise masker at 90 dB SPL was used to mask a 7 ms long test tone at 500 Hz. It can be seen that post-masking is active up to approximately 100 ms. Other studies reported slightly longer post-masking intervals, e.g. Jesteadt et al. [53] used tones from 125 to 4000 Hz and reported that more post-masking occurred at very low frequencies than at high frequencies. Pre-masking is believed to be much weaker. Some studies even showed, that pre-masking diminishes or almost disappears if subjects are highly trained [80].

Tactile Similar to audition, the detectability of a vibration might be reduced by another one. Again, this effect depends on frequency, intensity and timing of both stimuli. As in audition, masking increases as a function of increasing masker intensity and decreasing frequency separation. However, there is good evidence that the different mechano-receptive channels do not mask each other [39,57].

Fig. 6 Auditory and vibrotactile pre- and post-masking as a function of the gap between signal and masker. Data from Gescheider et al. [35] is plotted using a 250 Hz vibration masker at the thenar eminence with 20 dB sensation level. The test signal was also a 250 Hz vibration. For comparison, auditory data from Elliott [16] is plotted using a white noise masker at 90 dB SPL. The test signal was a tone at 500 Hz. Additionally, pre-masking is plotted after Oxenham and Moore [80] using a noise masker and a 6 kHz tone.

Vibrotactile masking patterns from Stamm et al. [90] and Gescheider et al. [39] are plotted in Fig. 5. Narrow band masking noise was simultaneously presented with sinusoidal test stimuli. Strong masking towards higher frequencies can be seen, which might be due to masking within the Pacinian channel. For decreasing frequencies lower than the masker, the threshold of the Pacinian channel might exceed the threshold of another tactile channel, e.g. RA1, which takes over and gradually reduces the masking effect [35]. In this sense, the overlapping vibrotactile channels could be regarded similar to overlapping auditory bands, however, with only few fixed filters. This would explain the strong asymmetry of vibrotactile masking patterns plotted here.

Thresholds might be elevated, even if two vibrations stimulate the body at different locations [37,42]. This is referred to as ‘lateral masking’ or ‘suppression’ and can be compared to dichotic masking discussed above. In both modalities neuronal and central processes seem to be involved in masking. However, the underlying mechanisms are not yet completely understood.

Similar to audition, masking is strongest for simultaneous stimulus presentation and decreases with increasing interval between test signal and masker [42,58]. This is illustrated in Fig. 6. Vibrotactile masking at the thenar eminence is plotted with data from Gescheider et al. [35] for a sinusoidal masker and test signal at 250 Hz. He found that the rate of decay of post-masking appears to be approximately the same than pre-masking, independent of masker type (sinusoidal or noise) and stimulated mechano-receptor. Compared to audi-
tion, temporal masking seems to be much more extended for vibrations at the skin. In addition, for hearing there is a stronger asymmetry towards post-masting.

If more than one stimuli is presented, also other changes in sensation have been reported. E.g., a stimuli can cause a subsequent one to appear more intense, with increasing intensity for decreasing time interval in-between the both. This is called enhancement and has been reported for short tone bursts in audition [114] and vibrotactile perception [104].

2.5 Adaptation and fatigue

In the previous section, masking, the ability of an intense stimulus to obscure a second weaker test stimulus, was described. In this section, the ability of a temporally extended stimulus will be discussed to gradually desensitize a sensory channel. This might result in the decline of apparent magnitude of a stimulus during presentation. Even some time after the stimulus has stopped, it might be harder to detect a test signal.

Auditory In audition it is often distinguished between adaptation and fatigue. Auditory adaptation refers to the decline in sensitivity within the first minutes of stimulus presentation [73]. However, this effect seems to be restricted to low sensation levels or high frequencies [50,107]. Auditory fatigue is often understood as the shift in threshold after excessive exposure to a fatiguing stimulus. This temporary threshold shift (TTS) is well known from rock music [12] and will be summarized in the following. The TTS generally increases with increasing intensity and duration of the fatiguing stimulus. Similar to masking, larger TTS have been found with decreasing frequency separation. Interestingly, fatigue effects are less marked at low frequencies, possibly due to the middle ear reflex [73]. After cessation of the fatiguing stimulus, hearing recovers from the TTS approximately proportional to the logarithm of the recovery time, if the TTS is not too large (e.g., < 40 dB) and exposure time is not too long (e.g., < 1 days) [69]. Such an exemplary TTS curve is plotted in Fig. 7 for 25 min of stimulation at 4kHz, a frequency where auditory fatigue is most effective.

Tactile Similar to audition, the absolute perception threshold for vibration increases and recovers over time due to prolonged stimulation. In vibrotactile literature, this effect is sometimes referred to as fatigue and sometimes as adaptation. The TTS increases again with increasing intensity and duration of stimulation. For intense stimulation over a longer period, recovery time can last up to several minutes. Compared to audition, generally much lower sensation levels are required for the effect to appear and much steeper slopes have been reported [7,29,40,108].

Fig. 7 Auditory and vibrotactile temporary threshold shifts during and after exposure to long-lasting stimulation. Data from Hahn [46,47] is plotted for vibratory stimulation of the Pacinian channel with different intensities and durations. For comparison, an exemplary temporary threshold shift for the auditory system is plotted after Miller [69].

Two exemplary TTS curves are plotted in Fig. 7 using data from Hahn [46,47]. The upper curve was measured using a large contact area on the fingerpad vibrating with 60Hz. Only 34 dB sensation level were necessary to reach 17 dB TTS after 25 minutes of exposure. However, the TTS recovered much faster compared to audition. The lower curve was measured using a small contact area on the fingerpad vibrating at 200 Hz at only 14 dB sensation level. Again steep rising and falling slopes can be seen. Like for masking, it is widely believed, that adaptation can not occur between different vibrotactile channels [47,51].

3 Differential sensitivity

Beside the absolute sensitivity, the smallest detectable changes of a stimulus are useful for a psychophysical comparison between the auditory and the tactile modality. Therefore, difference limen for intensity, frequency, duration and location will be discussed in the following.

3.1 Intensity discrimination

Auditory In Fig. 8 auditory just noticeable differences in level (JNDLs) are plotted against frequency after Florentine et al. [23] and Jesteadt et al. [54]. For high sensation levels (70dB and 80dB above threshold) the auditory system is very sensitive to intensity changes, with a differential threshold of only 0.5 dB to 1 dB. This holds true over a broad frequency range. However, for low sensation levels (30dB
Auditory and tactile just noticeable differences in level as a function of stimulus frequency. Sensation level is relatively more important at high frequencies than at low ones, where the curves tend to converge. Unfortunately, no data is known for frequencies below 250 Hz.

If a single frequency is selected, the difference limen can be replotted as a function of sensation level. Figure 9 shows different tactile studies, which measured difference limen as a function of level. Only Gescheider [41] reported a significant decrement of JNDL with increasing sensation level. Other studies found no effect. However, a smaller dynamic range [9] and much lower vibration frequencies [26,79] were tested. It is therefore difficult to compare the results.

### 3.2 Frequency discrimination

**Auditory** One of the fundamental characteristics of the auditory system is the ability to discriminate between different frequencies. Just noticeable differences in frequency (JNDFs) smaller than 1 Hz can be perceived at low frequencies. Figure 10 summarizes data from various laboratories [18,72,74,84,111]. It can be seen that the plotted auditory JNDFs are for tones with a minimum length of 200 ms. For shorter tones the JNDFs increase rapidly [19]. Again, an influence between 0.4 and 2.3 dB can be found in the literature [26,85,92], e.g., it has been shown that JNDFs for seat vibrations can be as small as 0.5 dB [62]. Similar studies [1,26,79] found slightly higher values and are summarized in Fig. 8. None reported a dependence of JNDL on stimulus frequency. Interestingly, the study with the lowest levels measured the highest JNDFs (Bellmann [1]). The study with the strongest vibrations revealed the lowest difference thresholds (Matsumoto et al. [62]). This suggests a similar dependence of JNDL on sensation level as in audition. However, few data exists to test this hypothesis. Figure 9 shows different tactile studies, which measured difference limen as a function of level. Only Gescheider [41] reported a significant decrement of JNDL with increasing sensation level. Other studies found no effect. However, a smaller dynamic range [9] and much lower vibration frequencies [26,79] were tested. It is therefore difficult to compare the results.
of sensation level on the difference limen was found, with higher level resulting in smaller JNDFs [111].

**Tactile** The tactile ability to discriminate between vibration frequencies is quite limited if compared to the auditory system. However, only few data exists for tactile JNDF. This might be due to the difficulty of eliminating concomitant cues, like intensity differences, during experiments. Studies with stimulation at the hand, buttocks and forearm are plotted for comparison in Fig. 10.

Goff [43] investigated sinusoidal stimulation at the fingertip. Five frequencies (25 Hz, 50 Hz, 100 Hz, 150 Hz and 200 Hz) were selected, and their magnitudes were adjusted to equal intensities (approximately 20 dB above the threshold). He found that the JNDF ranged from 8 to over 100 Hz, increasing with increasing reference frequency.

Rothenberg et al. [85] experimented with sinusoidal stimuli at the volar forearm. Frequencies between 25 and 250 Hz were evaluated. Their amplitudes were normalized to achieve a uniform subjective magnitude (approximately 14 dB above threshold). The results revealed difference limen ranging from 4 to over 75 Hz.

The ability to detect changes in seat vibration frequency was measured by Merchel [63] using frequencies between 20 and 90 Hz. Stimulus amplitudes were normalized to equally perceived intensity approximately 20 dB above threshold. The measured JNDF again increased with increasing frequency, from approximately 7–66 Hz.

### 3.3 Temporal discrimination

A further interesting aspect of both modalities is the ability to make temporal discriminations. Different stimuli and approaches have been used for investigations in the auditory domain, e.g. recognition of amplitude modulation [28] or identification of an increase in duration [10]. However, there are not many studies in the tactile domain. A lucid evaluation of temporal resolution is provided by the minimum detectible separation between two successive stimuli. This is referred to as gap detection threshold and will be used exemplarily for comparison in the following.

**Auditory** Numerous studies have investigated gap detection thresholds using different stimuli [21,22,25,44,49,75–77, 87]. The minimum auditory temporal resolution was found for clicks and broad noise. It is in the order of 2–3 ms. Exemplary data from Gescheider [31] and Plomp [82] is plotted in Fig. 11. It can be seen that also the gap detection threshold depends on sensation level and increases significantly towards lower intensities.

This is also true for sinusoidal excitation. Data from Moore et al. [78] is plotted in Fig. 12. At levels which are adequately audible, sinusoidal gap detection thresholds are roughly constant, but increase rapidly for levels close to the perception threshold. Minimum gap thresholds at about 17 ms for the 100 Hz stimulus and 6–9 ms for frequencies from 200 to 2000 Hz have been found. Slightly lower gap detection thresholds have been reported in other studies, e.g. 5 ms for 400 Hz by Shailer and Moore [88], which might be explained by different experimental procedures. No influence was found for embedding burst duration or temporal position of the gap [25,49].

**Tactile** Figures 11 and 12 compare tactile gap detection thresholds for noise, clicks and sinusoids delivered to the hand from different publications [11,31,34]. The minimum detectible gap between two tactile stimuli was found to be
both modalities. Such thresholds were obtained for noise and clicks with sensation levels of about 35 dB and for sinusoids with sensation levels of about 20 dB. For lower intensities, the minimum detectible gap increases similar to our hearing. Different to the auditory modality, vibratory gaps seem to be harder to detect between noise bursts than between sinusoidal bursts. Gap detection thresholds for noise and clicks were found to be 3–10 times higher for tactile perception than for hearing. In contrast, sinusoidal gap thresholds seem to be comparable between the modalities at low sensation levels. The reason for this behavior is not yet understood.

With increasing age, the ability to detect temporal gaps in vibration reduces marginally [11]. Similarly, a slight increase in auditory gap detection threshold with age has been reported [49,77]. However, it can be assumed that aging does not lead to severe reduction of temporal resolution in both modalities.

### 3.4 Location discrimination

Localization in the auditory and tactile domain is quite distinct. In audition, only two input signals from the ears are available to estimate the position of an auditory event somewhere in space. In contrast, mechanoreceptors are spatially distributed all over the body and tactile events are mostly perceived in proximity of the body.

**Auditory** The localization ability of the auditory system can be partially described using the minimum angle at which two sources can be separated. This minimum audible angle (MAA) depends on the character of the sound and the position of the source relative to the listener. For impulsive sounds in front of the listener, MAAs about 1° of the source relative to the listener. For impulsive sounds, MAAs were found to be 3–10 times higher for tactile perception than for hearing. In contrast, sinusoidal gap thresholds seem to be comparable between the modalities at low sensation levels. The reason for this behavior is not yet understood.

With increasing age, the ability to detect temporal gaps in vibration reduces marginally [11]. Similarly, a slight increase in auditory gap detection threshold with age has been reported [49,77]. However, it can be assumed that aging does not lead to severe reduction of temporal resolution in both modalities.

**Tactile** The spatial sensitivity of the tactile sensory system can be measured e.g. by a two point discrimination tasks, where two spatially separated stimuli are presented either simultaneously or shortly one after another. The subjects have to decide, whether they felt one or two contact points. Tactile spatial acuity varies significantly across the body surface. It was found that thresholds vary between about 1–2 mm and 45 mm depending on location on the skin [110]. Regions with high receptor density, e.g. the fingers, have low spatial discrimination thresholds, whereas areas with low receptor density, e.g. the back, show low spatial acuity. Interestingly, the perception of tactile distance seems to depend not only on the spatial separation, but also on the timing between two vibrotactile stimuli [4]. It has also been argued that the absolute vibrotactile localization ability depends on the position of stimulus sites relative to body landmarks, like the joints of the wrist or the elbow. E.g., the ability to localize vibrotactile stimuli on a linear array of tactors on the forearm is significantly better near the wrist and elbow, when compared to the localization for sites far from such natural anchor points [6]. Similar evidence for anatomically defined anchor points that provide localization referents was found on the abdomen [5].

If two spatially separated areas are stimulated, effects similar to auditory phantom source localization or precedence have been found [106], suggesting similar neural mechanisms for both modalities.

Some researchers even tried to reproduce the localization ability of our hearing system. If two spatially separated microphones are used to drive two vibration actuators mounted to the forearms [105] or fingertips [30], subjects can accurately localize sound sources after some training. Interestingly, many subjects reported that tactile sensations were projected out into space’ to match the position of the corresponding sound source. This decoupling of receptor location and perceptual event is known from vision and audition [92].

### 4 Summary

In this paper, basic psychophysical abilities and limitations of the auditory and vibrotactile modality are discussed in a comparative manner. The validity of such comparisons could be questioned because of different methodologies used in the reviewed papers. Different researchers pursued different questions at different times with different test participants (number, gender, age, …) and different equipment. However, general trends in the data can often be identified. If available, data from several studies are plotted on top of each other to check consistency. Sometimes not all available data are presented for reasons of clarity. Being aware of the variations between the compared studies, the authors believe that this comparison provides the background for the auditory-tactile design, e.g., of perceptually optimized human–machine interfaces or multimodal music applications. This sections summarizes the main similarities and differences between both modalities and discusses useful applications scenarios.

Both modalities show frequency dependent perception thresholds, but with different characteristics. When designing auditory-tactile feedback with the goal of equal intensity in both modalities, this disparity can be compensated by careful frequency equalization using the differences between the threshold curves. Compared to the sense of hearing, vibrotactile perception is restricted to low frequencies. At 20 Hz the usable amplitude range of both modalities is similar. How-
ever, with increasing frequency the auditory dynamic range increases rapidly, while the vibrotactile dynamic range seems to remain constant up to approximately 200 Hz. Compared to audition, the increase in perceived magnitude is steeper with increasing level in the vibrotactile domain, particularly at low sensation levels. If the target of a multimodal design is to match the perceived intensity of a stimuli in both modalities, e.g., for auditory-tactile button feedback of a touch screen, the dynamic range of one domain should be adapted, e.g., using a compressor for vibration processing.

Both modalities show severe impairment of sensitivity with increasing age. This effect has a similar tendency: it is stronger towards the upper frequency limit of each modality. However, around 250 Hz the age-induced threshold shift seems to be stronger for the sense of touch than for hearing. This is especially crucial in the context of auditory-tactile feedback design, since the vibrotactile dynamic range is considerably smaller than the auditory dynamic range. A vibrotactile threshold shift of 20 dB at 200 Hz almost halves the available amplitude range. In other words: vibrations which are strong for younger subjects, might not be perceived at all by the elderly. Again, dynamic compression in the tactile domain helps the designer to reduce this effect with the drawback of a decreased dynamic range. Because less impairment was reported in the vibrotactile domain below 40 Hz, it might be worth to consider this frequency range for a feedback design which is less dependent on age.

The auditory system is able to integrate energy over time for stimuli durations up to approximately 1 s. A similar temporal effect can be found in the vibrotactile system for sufficiently high frequencies and relatively large stimulation areas. In addition energy integration over space has been observed. From this it follows that the size of a vibrating contact area, e.g., the size of a vibrating smart watch, must be taken into account by the designer if the perceived intensities are to be matched in both modalities.

Both modalities show the ability of one stimulus to mask (or enhance) another. In comparison, in the vibrotactile modality broader masking patterns are excited around the masker frequency with strong masking towards higher frequencies. Also in time domain, the vibrotactile threshold is raised over a longer period around the duration of a masker. Strong masking in the vibrotactile modality suggests that, e.g., when designing a system for auditory-tactile music reproduction, it might suffice to reproduce the fundamental of a complex sound in the vibratory domain without changing the overall percept.

Temporary threshold shifts due to prolonged stimulation occur in both modalities. In audition high levels or long exposure times are necessary. In the vibrotactile domain, even small sensation levels result in a temporary threshold shifts, which, however, grows and recovers fast. This effect might be relevant for the designer in practical applications if strong background vibrations are present, e.g., at the steering wheel when driving a car.

The just noticeable differences in level for sound and vibration seem to be remarkably similar at low frequencies. However, the difference limen of tactile frequency discrimination are much higher compared to audition. This is very important in the context of multimodal design, since frequency information is one of the fundamental components of audio signals, resulting in pitch perception. This perceptual feature is only available to a very limited extent in the tactile domain.

Gap detection thresholds for sinusoidal stimuli are comparable in the tactile and the auditory system. However, this seems not to be the case for noises and clicks. The influence of the sensation level on auditory and tactile temporal resolving power is remarkably similar. Additionally, the gap detection thresholds are in the millisecond range, indicating good temporal resolution for both modalities. Sound and vibrations are therefore equally suitable for reproducing temporal information via a user interface. However, depending on the application, the different temporal acuity with different reproduction intensities must be taken into account.

It is difficult to compare the localization ability in both modalities. Auditory events can be perceived everywhere around the listener, however, resolution is quite limited. The spatial resolution of somatosensation is generally more detailed, but tactile events are restricted to the proximity of the skin. However, it has been demonstrated that the projection of tactile events towards a sound source is possible. Sensory substitution systems for the hearing impaired use the good location discrimination of the tactile system to encode information, e.g., the frequency of a sound, in order to overcome shortcomings in tactile frequency perception.

This article focused on the independent absolute and differential sensitivities of both modalities. It is important to note, however, that many multimodal illusions exist that exploit features of our audio and tactile perceptual abilities, e.g., the auditory–tactile loudness illusion [63]. A future article will explore these crossmodal interactions further.

Acknowledgements Open Access funding provided by Projekt DEAL.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence.
References

1. Bellmann MA (2002) Perception of whole-body vibrations: from basic experiments to effects of seat and steering-wheel vibrations on the passenger’s comfort inside vehicles. Ph.D. thesis, Carl von Ossietzky –University Oldenburg
2. Blauert J (1997) Spatial hearing—the psychophysics of human sound localization. MIT Press, Cambridge
3. Bochereau S, Terekhov A, Hayward V (2014) Amplitude and duration interdependence in the perceived intensity of complex tactile signals. In: International conference on human haptic sensing and touch enabled computer applications, Springer, Berlin
4. Cholewiak RW (1999) The perception of tactile distance: influences of body site, space, and time. Perception 28(7):851–875
5. Cholewiak RW, Brill JC, Schwab A (2004) Vibrotactile localization on the abdomen: effects of place and space. Percept Psychophys 66(6):970–987
6. Cholewiak RW, Collins AA (2003) Vibrotactile localization on the arm: Effects of place, space, and age. Percept Psychophys 65(7):1058–1077
7. Cohen L, Lindley S (1938) Studies in vibratory sensibility. Am J Psychol 51:44–63
8. Corso J (1959) Age and sex differences in pure-tone thresholds. J Acoust Soc Am 31(4):498–507
9. Craig JC (1972) Difference threshold for intensity of tactile stimuli. Percept. Psychophys. 11(2):150–152
10. Dooley GJ, Moore BCI (1988) Duration discrimination of steady and gliding tones: a new method for estimating sensitivity to rate of change. J Acoust Soc Am 84(4):1332–1337
11. van Doren CL, Gescheider GA, Verrillo RT (1990) Vibrotactile temporal gap detection as a function of age. J Acoust Soc Am 87(5):2201–2206
12. Drake-Lee AB (1992) Beyond music: auditory temporary threshold shift in rock musicians after a heavy metal concert. J R Soc Med 85(10):617–619
13. Egan JP, Hake HW (1950) On the masking pattern of a simple auditory stimulus. J Acoust Soc Am 22(5):622–630
14. Ehmer R (1959) Masking patterns of tones. J Acoust Soc Am 31(8):1115–1120
15. Ehmer RH (1959) Masking by tones vs noise bands. J Acoust Soc Am 31(9):1253–1256
16. Elliott L (1962) Backward and forward masking of probe tones of different frequencies. J Acoust Soc Am 34(8):1116–1117
17. Elliott L (1962) Backward masking: monotonic and dichotic conditions. J Acoust Soc Am 97(5):38–44
18. Fastl H (1978) Frequency discrimination for pulsed versus modulated tones. J Acoust Soc Am 63(1):275–277
19. Fastl H, Hesse A (1984) Frequency discrimination for pure tones at short durations. Acustica 56(1):41–47
20. Fastl H, Zwicker E (2007) Psychoacoustics: facts and models, 3rd edn. Springer, Berlin
21. Fitzgibbon PS, Wightman FL (1982) Gap detection in normal and hearing-impaired listeners. J Acoust Soc Am 72(3):761–765
22. Florentine M, Buus S, Geng W (1999) Psychometric functions for gap detection in a yes-no procedure. J Acoust Soc Am 106(6):3512–3520
23. Florentine M, Buus S, Mason C (1987) Level discrimination as a function of level for tones from 0.25 to 16 kHz. J Acoust Soc Am 81(5):1528–1541
24. Florentine M, Fastl H, Buus S (1988) Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking. J Acoust Soc Am 84(1988):195–203
25. Forrest TG, Green DM (1987) Detection of partially filled gaps in noise and the temporal modulation transfer function. J Acoust Soc Am 82(6):1933–1943
26. Forta NG (2009) Vibration intensity difference thresholds. Ph.D. thesis, University of Southampton
27. Garner W (1947) The effect of frequency and spectrum on temporal integration of energy in the ear. J Acoust Soc Am 19(5):808–815
28. Gelfand SA (1998) Hearing—an introduction to psychological and physiological acoustics, 3rd edn. Marcel Dekker, New York
29. Gescheider G, Wright J (1969) Effects of vibrotactile adaptation on the perception of stimuli of varied intensity. J Exp Psychol 81(3):449–453
30. Gescheider GA (1965) Cutaneous sound localization. J Exp Psychol 70:617–625
31. Gescheider GA (1967) Auditory and cutaneous temporal resolution of successive brief stimuli. J Exp Psychol 75(4):570–572
32. Gescheider GA (1976) Evidence in support of the duplex theory of mechanoreception. Sensor Process 1(1):68–76
33. Gescheider GA, Berryhill ME, Verrillo RT, Bolanowski SJ (1999) Vibrotactile temporal summation: probability summation or neural integration? Somatosens Motor Res 16(3):229–242
34. Gescheider GA, Bolanowski SJ, Chatterton SK (2003) Temporal gap detection in tactile channels. Somatosens Motor Res 20(3–4):239–247
35. Gescheider GA, Bolanowski SJ, Verrillo RT (1989) Vibrotactile masking: effects of stimulus onset asynchrony and stimulus frequency. J Acoust Soc Am 85(5):2059–2064
36. Gescheider GA, Joelson JM (1983) Vibrotactile temporal summation for threshold and suprathreshold levels of stimulation. Percept Psychophys 33(2):156–162
37. Gescheider GA, Verrillo RT (1982) Contralateral enhancement and suppression of vibrotactile sensation. Percept Psychophys 32(1):69–74
38. Gescheider GA, Verrillo RT (1984) Effects of the menstrual cycle on vibrotactile sensitivity. J Acoust Soc Am 36(6):586–592
39. Gescheider GA, Verrillo RT, van Doren CL (1982) Prediction of vibrotactile masking functions. J Acoust Soc Am 72(5):1421–1426
40. Gescheider GA, Wright JH (1968) Effects of sensory adaptation on the form of the psychophysical magnitude function for cutaneous vibration. J Exp Psychol 77(2):308–13
41. Gescheider GA, Zwislocki JJ, Rasmussen A (1996) Effects of stimulus duration on the amplitude difference limen for vibration. J Acoust Soc Am 100(4.1):2312–2319
42. Gilson R (1969) Vibrotactile masking: some spatial and temporal aspects. Percept Psychophys 5:176–180
43. Goff GD (1967) Differential discrimination of frequency of cutaneous mechanical vibration. J Exp Psychol 74(2.1):294–299
44. Green DM, Forrest TG (1989) Temporal gaps in noise and sinusoids. J Acoust Soc Am 86(3):961–970
45. Griffin M (1990) Handbook of human vibrations. Elsevier Academic Press, London
46. Hahn JF (1966) Vibrotactile adaptation and recovery measured by two methods. J Exp Psychol 71(5):655–658
47. Hahn JF (1968) Low-frequency vibrotactile adaptation. J Exp Psychol 78(4):655–659
48. Harris J (1963) Loudness discrimination. J Speech Hear Disorders Monograph Supplement Number II
99. Verrillo RT (1979) Comparison of vibrotactile threshold and suprathreshold responses in men and women. Percept Psychophys 26(1):20–24
100. Verrillo RT (1980) Age related changes in the sensitivity to vibration. J Gerontol 35(2):185–193
101. Verrillo RT (2009) Bioresponse to vibration. In: Havelock D, Kuwano S, Vorländer M (eds.) Handbook of signal processing in acoustics, pp 1185–1244
102. Verrillo RT, Bolanowski SJ, Gescheider GA (2002) Effect of aging on the subjective magnitude of vibration. Somatosens Motor Res 19(3):238–244
103. Verrillo RT, Fraioli AJ, Smith RL (1969) Sensation magnitude of vibrotactile stimuli. Percept Psychophys 6(6A):366–372
104. Verrillo RT, Gescheider GA (1975) Enhancement and summation in the perception of two successive vibrotactile stimuli. Percept Psychophys 18(2):128–136
105. von Békésy G (1957) Sensations on the skin similar to directional hearing, beats, and harmonics of the ear. J Acoust Soc Am 84(1):830–841
106. von Békésy G (1958) Funneling in the nervous system and its role in loudness and sensation intensity on the skin. J Acoust Soc Am 30(5):399–412
107. Ward WD (1966) Temporary threshold shift in males and females. J Acoust Soc Am 40(2):478–485
108. Wedell CH, Cummings SB (1938) Fatigue of the vibratory sense. J Exp Psychol 22(5):429–438
109. Wegel R, Lane C (1924) The auditory masking of one sound by another and its probable relation to the dynamics of the inner ear. Phys Rev 23:266–285
110. Weinstein S (1968) Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In: Kenshalo DR (ed) The skin senses. Charles C Thomas, Illinois, pp 195–218
111. Wier C, Jesteadt W, Green D (1977) Frequency discrimination as a function of frequency and sensation level. J Acoust Soc Am 61(1):178–184
112. Winckel F (1969) Nachrichtentechnik unter kybernetischen Aspekten. In: Handbuch für HF- und E-Techniker Bd. 8. Berlin, Germany
113. Zwislocki JJ (1960) Theory of temporal auditory summation. J Acoust Soc Am 3(8):1046–1060
114. Zwislocki JJ, Sokolich WG (1974) On loudness enhancement of a tone burst by a preceding tone burst. Percept Psychophys 16(1):87–90

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.