EVIDENCE OF LATERALIZATION CUES IN GRAND AND UPRIGHT PIANO SOUNDS

Federico Fontana
Dipartimento di Scienze
Matematiche, Informatiche e Fisiche
Università di Udine, Italy
`federico.fontana@uniud.it`

Federico Avanzini
Dipartimento di Informatica
Università di Milano, Italy
`federico.avanzini@unimi.it`

Stefano Papetti
Institute for Computer Music
and Sound Technology
Zürcher Hochschule der Künste,
Switzerland
`stefano.papetti@zhdk.ch`

ABSTRACT

In a previous experiment we have measured the subjective perception of auditory lateralization in listeners who were exposed to binaural piano tone reproductions, under different conditions (normal and reversed-channel listening, manual or automatic tone production by a Disklavier, and disclosure or hiding of the same keys when they were autonomously moving during the automatic production of a tone.) This way, participants were engaged in a localization task under conditions also involving visual as well as proprioceptive (that is, relative to the position and muscular effort of their body parts) identification of the audio source with the moving key, even when the binaural feedback was reversed. Their answers, however, were clustered on a limited region of the keyboard when the channels were not reversed. The same region became especially narrow if the channels were reversed. In this paper we report about an acoustic analysis of the localization cues conducted on the stimuli that have been used in the aforementioned experiment. This new analysis employs a computational auditory model of sound localization cues in the horizontal plane. Results suggest that listeners used interaural level difference cues to localize the sound source, and that the contribution of visual and proprioceptive cues in the localization task was limited especially when the channels were reversed.

1. INTRODUCTION

The assumption that performing pianists can localize a note around the position of the key generating the corresponding tone looks convincing, to the point that the industry of digital pianos has sometimes relied on it [1, 2]. This assumption, however, has no explanation if only the sound coming from the vibrating strings that have been struck by the hammer is taken into account. This mechanical energy, in fact, is almost entirely transmitted from the strings to the soundboard which, in its turn, dissipates acoustic energy into the air through radiation patterns depending on the modal characteristics of the soundboard shape and material [3]. Different soundboard regions are associated to specific frequency components that can be radiated; the net result is that, once a key is pressed, the pianist is enveloped by a soundfield that does not propagate from any definite position in the keyboard.

There is, however, a possible explanation for the localization of a piano tone in the proximity of the corresponding key. This explanation does not consider the sound diffusion from the soundboard, but rather the initial acoustic transient produced by a keystroke before the respective strings start to vibrate. If this transient would be the result of sound onsets originating from the key position, furthermore having enough intensity and temporal extension to activate the auditory precedence effect [4], then sufficient conditions could hold for perceptually locking the entire sound around the same position.

A detailed study on the mechanical origin of piano transients [5] remains noncommittal about whether such transients can enable the auditory precedence effect. This study considers the first 20-25 ms of sound coming from the instrument *before* the hammer hits the string, containing in particular the "thump" originating when a key bottom bumps against the keybed at the end of its fly. This bump has the characteristic of a point-wise impact against a rigid part of the piano, so in principle it could enable the precedence effect. However, the mechanical wavefront radiating from the impact point propagates much faster along the keybed than does the companion airborne pressure wave originating from the same point. The outcome of this race between mechanical and pressure waves is uncertain, hence it is reasonable to assume that the acoustic energy radiating from the impact point arrives at the listener's ears as part of a beam of wave transients propagating from the whole keybed, thus being unable to bring precise cues of key localization to a pianist.

2. PREMISE TO THE ACOUSTIC ANALYSIS

Multi-channel measurements made in a silent room over the keyboard of a Seiler 1849 playing a note C4 suggest that the precedence effect cannot be enabled by that piano [6]. Those measurements in fact show too little relative delays among channel onsets for assuming the existence of significantly different intensity transients at the pianist's ears. From these measurements one should conclude that piano tones cannot be localized precisely, since they bring related spatial cues neither during the initial transient nor when the soundboard resonates during tone decay. In favor of this conclusion talks also a test, in which scrambled versions of multi-channel reproductions of piano tones were presented to pianists via a loudspeaker array [6]. In this test, only the stimuli obtained from random scrambling or from swapping half of the channels of the reproduced soundfield were judged as being of lower quality. Conversely, less disruptive channel reconfigurations led to tone reproductions which were reported to preserve the quality of both the instrument and the acoustic scene.

On the other hand pianists are exposed to a causal invariance occurring each time they play, as every tone inevitably follows from the action of their fingers at a specific keyboard position. Furthermore, pianists can see the (again invariant) position in space of the key(s) they are pressing. Is this proprioceptive and/or visual invariance alternative or supplementary to any auditory localization process? In a previous study [7] we have experimented on this hypothesis, achieving results that are summarized in Fig. 1 and Fig. 2. Both figures report for sound source positions estimated by pianists, while listening through headphones respectively to non-reversed and reversed reproductions of binaural recordings of five A notes—the source was a DC3 M4 grand Yamaha Disklavier, and the dummy listener was a KEMAR mannequin. In both cases such tones were passively listened to while in front of the same piano with its fall board closed (auditory localization), listened to while observing the Disklavier moving the corresponding key through its electro-mechanical actuators, or finally actively listened to while pressing the same key.

An ANOVA analysis among conditions suggested two facts, which can both be appreciated from visual inspection of Fig. 1 and Fig. 2.

1. When the audio channels were not reversed, the auditory system had no reason to get confused while processing possible localization cues existing in the notes: in this case, participants were progressively supported by visual and then somatosensory cues in localizing a tone toward the corresponding key.

2. On the contrary, when the audio channels were reversed the auditory localization process (if any) had to be contradictory. In this case participants to a good extent reversed the sound source location, by keeping the five note positions almost constant across the keyboard and furthermore with no apparent role of vision and the somatosensory system.

It must be noted that the experiment suffered from several limitations, and for this reason it must be considered a pilot study: headphones prevented listeners to reinforce the localization process through head movement [8]; besides this, suspicion existed about a possible learning effect affecting the participants who first attended the active playing task and later localized the tones through passive listening with and without visual feedback.

In spite of these limits, Fig. 1 and Fig. 2 suggest that some form of auditory localization may have taken place during the trials, especially when the sound channels were not reversed. Holding doubts about the existence of sufficient time delays between channels capable of eliciting the precedence effect, we decided to analyze intensity differences in the auditory stimuli used in the experiment, in an aim to uncover the existence of lateralization cues.

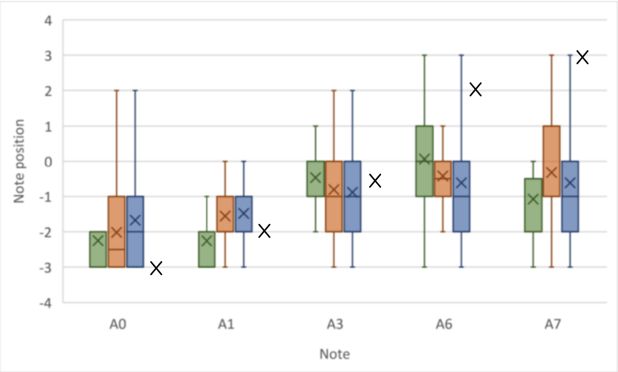

Figure 1. Key vs. perceived note position with non-reversed audio channels on a grand Disklavier. Right to left boxplots for each note localization: auditory (fall board closed); auditory-visual (fall board open, key self-moving); auditory-visual-proprioceptive (key pressed by pianist). Note positions are normalized in the range [-3,3] and marked with ‘X’: A0=-3; A1=-2; A3=-0.5; A6=2; A7=3.

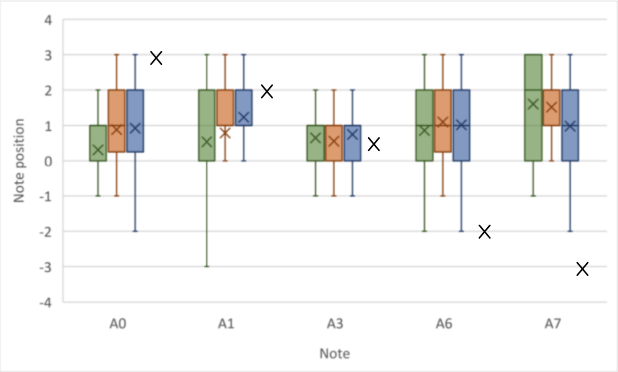

Figure 2. Key vs. perceived note position with reversed audio channels on a grand Disklavier. Right to left boxplots for each note localization: auditory (fall board closed); auditory-visual (fall board open, key self-moving); auditory-visual-proprioceptive (key pressed by pianist). Note positions are normalized in the range [-3,3] and marked with ‘X’: A0=3; A1=2; A3=0.5; A6=-2; A7=-3.

3. ANALYSIS OF PIANO TONES

The analysis has been conducted by borrowing tools from computational psychoacoustics. In particular, a decoder of temporally fluctuating interaural disparities [9] and a known binaural cross-correlation model [10] have been employed. These models process binaural sounds in sub-bands, and then yield a measure of lateralization as a function of time. Both are implemented, along with other models, in the Auditory Modeling Toolbox, a free software library distributed for the Octave and Matlab simulation environments [11].

Figure 3 shows Interaural Level Differences (ILD) in dB decoded by the interaural disparity model [9] during the first 0.5 s of each sound. Both pianos had the lid semi-open during the recording session. Sounds producing the left column consist of the stimuli A0, A1, A3, A6, A7 for the experiment reported in Sec. 2. Sounds producing the right column have been included for comparison, and were selected from binaural recordings of a Disklavier upright piano model DU1A. Each plot contains the ILD encoded in the gammatone sub-band available from the model, which was nearest to the fundamental frequency of the tone (thin black line, see also the box inside each plot), along with the ILD obtained by summing [1] the outputs coming out from all gammatone filters (bold gray line). The sums, hence, provide an estimate of the running ILD experienced broadband by the participants during the experiment.

4. DISCUSSION

In all notes occupying the left column, the sub-band ILD initially wanders and then stabilizes around the mid point. Stabilization, if any, occurs within the first tens of milliseconds, which are most responsible for the localization process. Each plot, however, shows a longer temporal window to give also account of subsequent changes in the localization cues brought by the sub-bands.

This behavior is caused by the changing pressure field radiated by the soundboard within that band. The corresponding plots in the right column exhibit slightly less ripple. Such reduced dynamics can be ascribed to the smaller soundboard of the upright piano, and consequent minor intensity of the sound radiated from it in the low frequency. Furthermore the grand piano was recorded in a room having higher reverberation than the upright piano room and, hence, more prominent reflections from the wall increasing the unpredictability of the ILD. Tests made on other sub-bands confirm the erratic behavior of the ILD also outside the fundamental frequency range of each tone.

Besides ripple, the sub-band ILD of the upright piano exhibits a generalized offset toward the left direction, which generally decreases with increasing octave and hence with the frequency. This offset might have been caused by the cornered position the upright piano was located inside the room which, in spite of its low reverberation, certainly gave rise to standing waves during the recording sessions.

The broadband ILD of the upright piano is definitely positioned to the left. In absence of anechoic measurements of the same piano, this ILD can be explained once again by its positioning inside the room. Besides this offset, the comparison between the two datasets adds reliability to the ILD values extracted by the measurements of the grand piano, confirming that the latter produces a similar meanwhile more dynamic soundfield. The greater dynamics is testified by an analysis of the tones using the binaural cross-correlation model, centered around a center frequency of interest ($f_c = 26$ in both plots) [10]. Fig. 4 and Fig. 5 in fact confirm the larger variability of the ITD in the grand piano for instance during the first 0.5 s on note A6.

The broadband ILD encoded in the grand piano sounds also points to the left direction, however with ripples that become generally more evident with increasing frequency. In note A3, such ripples almost immediately reverse the ILD. One main responsibility for the polarization of the ILD around the left part of the grand piano may be ascribed to the peculiar shape of its soundboard, which is asymmetric toward the same part: this asymmetry implies that more energy is radiated from the left, especially concerning the lower modes which can resonate only across larger regions of the soundboard surface. On the other hand we have no anechoic measurements of the grand piano either, hence it cannot be excluded that the offset toward the left of the ILD was the result of room effects, too.

4.1 Comparison with subjective data

A similarity between the objective data and perceived lateralization certainly exists on all notes, with the exception of tone A3 that according to the computed ILD should have been localized rightmost by the participants. In particular, visual and proprioceptive cues seem to have played a role in the localization of the tones, whose key positions fell into the region where the sound came from.

In other words, participants may have refined the localization using visual and proprioceptive information only when the tone approximately came from the same direction where the moving key was located, otherwise disregarding the visual and proprioceptive cues. In favor of this possibility plays also the fact that participants located all tones approximately on the same position when the audio channels were reversed. In fact, according to the computational analysis we made on the tones, the auditory localization of the channel-reversed sounds was substantially as accurate as the localization participants made when the channels were not reversed.

We conclude that the perceived direction of piano tones may be affected by the ventriloquism effect. This effect normally occurs in individuals who, while they are listening to speech in presence of a speaker, tend to identify the sound source with the speaker herself [13]. Furthermore, experiments have concluded that this effect is enabled within reasonably limited angles between the speaker and sound source position.

It is generally accepted that the “results [of multisensory experiments] generalize to real life only when they reflect automatic perceptual processes, and not response strategies adopted to satisfy the particular demands of laboratory tasks” [14]. In our experiment, it is reasonable to think that listeners integrated the positional information coming from different modalities only when the auditory channels were not reversed. As soon as the ecological validity of the listening scenario disappeared, then the visual and proprioceptive information lost their perceptual significance.

---

[1] This requires to linearize all outputs, record their sign, and then convert their sum back to dB.

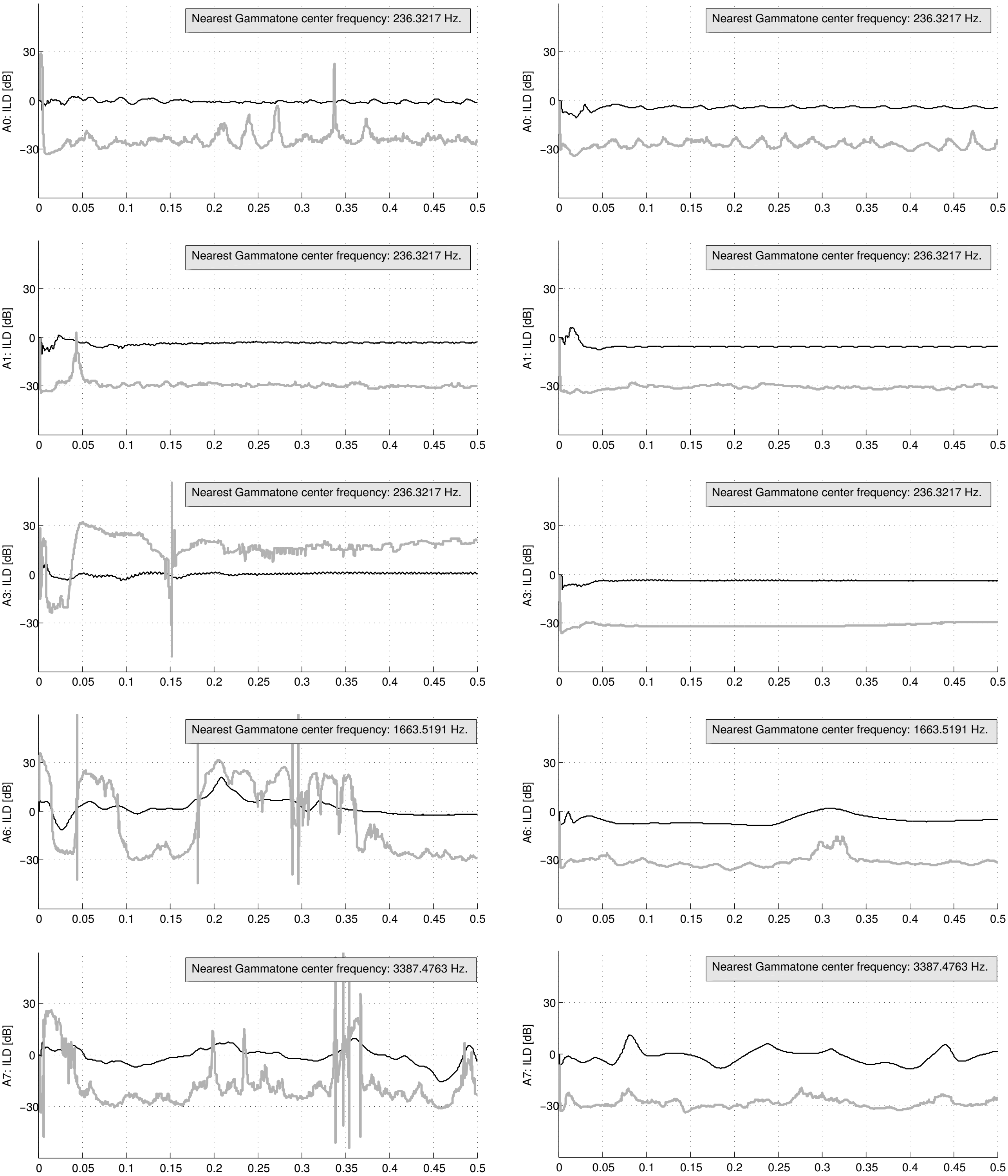

Figure 3. Left: Interaural Level Differences decoded from the stimuli leading to Fig. 1 and Fig. 2 [7], lid semi-open on a grand Disklavier. Right: Interaural Level Differences decoded from companion test stimuli [12], lid semi-open on an upright Disklavier. Each plot shows in thin black line the ILD decoded from the gammatone sub-band which is nearest to the tone fundamental frequency; in bold gray line the ILD summed across all gammatone sub-bands.

5. CONCLUSIONS

In this paper a computational psychoacoustic analysis has been made over data that had been previously used in a multisensory localization experiment, in which listeners were asked to localize piano tones also in presence of visual and somatosensory cues. The analysis suggests that listeners processed the multisensory information only when the different modalities could be coherently integrated: in this case the visual and proprioceptive cues contributed to refine the localization even if the sound source position did

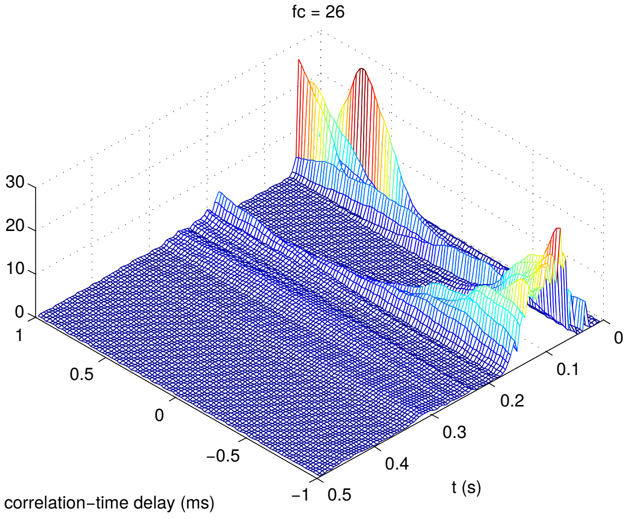

Figure 4. Binaural cross-correlation during the first 0.5 s of note A6 in the grand piano.

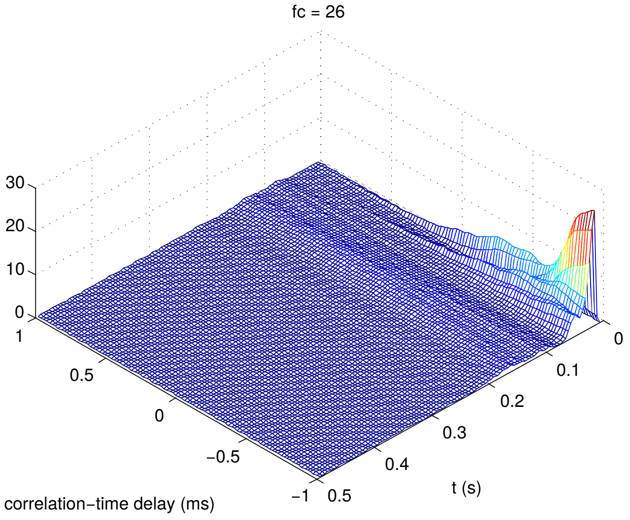

Figure 5. Binaural cross-correlation during the first 0.5 s of note A6 in the upright piano.

not match with that of the key. Conversely, the integration did not take place when the tone and key positions significantly differed.

This result resembles conclusions that have been drawn in several multisensory experiments involving the auditory channel, particularly the ventriloquist effect. At this point, the claim that the ventriloquist effect affects also piano performers should be validated settling pianists in more ecological playing conditions. For this reason we are planning to repeat the experiment, by avoiding headphone listening and binaural reproduction of recorded sounds. Also, it is true that this work certainly needs to be refined by conducting further investigations under an acoustically controlled environment.

6. REFERENCES

[1] K. Yamana, “Electronic musical instrument having a pan-pot function,” US Patent 4 577 540 A, Aug., 1986.

[2] R. D. Lawson, “Loudspeaker system for electronic piano,” US Patent 5 789 693, Aug., 1998.

[3] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*. New York: Springer-Verlag, 1991.

[4] A. D. Brown, G. C. Stecker, and D. J. Tollin, “The precedence effect in sound localization,” *J. of the Association for Research in Otolaryngology*, vol. 16, no. 1, pp. 1–28, 2015.

[5] A. Askenfelt, “Observations on the transient components of the piano tone,” *STL-QPSR*, vol. 34, no. 4, pp. 15–22, 1993, web available at http://www.speech.kth.se/qpsr.

[6] F. Fontana, Y. D. Pra, and A. Amendola, “Sensitivity to loudspeaker permutations during an eight-channel array reproduction of piano notes,” in *Proc. SMAC/SMC 2013*, Stockholm, Sweden, Jul. 30 - Aug. 3 2013.

[7] F. Fontana, D. Scappin, F. Avanzini, M. Bernardi, D. Bianco, and G. Klauer, “Auditory, visual and somatosensory localization of piano tones: A preliminary study,” in *Proc. 14th Sound and Music Computing Conference (SMC2017)*, Espoo, Finland, Jul.5–8 2017, pp. 254–260, available at http://smc2017.aalto.fi.

[8] P. T. Young, “The rôle of head movements in auditory localization,” *J. Experimental Psychology*, vol. 14, no. 2, pp. 95–124, 1931.

[9] M. Dietz, S. D. Ewert, V. Hohmann, and B. Kollmeier, “Coding of temporally fluctuating interaural timing disparities in a binaural processing model based on phase differences,” *Brain Research*, vol. 1220, pp. 234–245, 2008, active Listening. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0006899307022068

[10] W. Lindemann, “Extension of a binaural cross-correlation model by contralateral inhibition. i. simulation of lateralization for stationary signals,” *The Journal of the Acoustical Society of America*, vol. 80, no. 6, pp. 1608–1622, 1986.

[11] P. L. Søndergaard and P. Majdak, “The auditory modeling toolbox,” in *The technology of binaural listening*. Springer, 2013, pp. 33–56.

[12] F. Fontana, S. Papetti, H. Järveläinen, and F. Avanzini, “Detection of keyboard vibrations and effects on perceived piano quality,” *The Journal of the Acoustical Society of America*, vol. 142, no. 5, pp. 2953–2967, 2017.

[13] C. E. Jack and W. R. Thurlow, “Effects of degree of visual association and angle of displacement on the” ventriloquism” effect.” *Perceptual and motor skills*, vol. 37, no. 3, pp. 967–979, 1973.

[14] B. De Gelder and P. Bertelson, “Multisensory integration, perception and ecological validity,” *Trends in cognitive sciences*, vol. 7, no. 10, pp. 460–467, 2003.