AUDITORY, VISUAL AND SOMATOSENSORY LOCALIZATION OF PIANO TONES: A PRELIMINARY STUDY

Federico Fontana, Debora Scappin
Università di Udine
Dept. of Math., Comp. Sci. & Phys.
federico.fontana@uniud.it
scappin.debora@spes.uniud.it

Federico Avanzini
Università di Padova
Dept. of Information Engin.
avanzini@dei.unipd.it

Mattia Bernardi, Devid Bianco, Giorgio Klauer
Conservatorio "Cesare Pollini"
mattiabernardi9@gmail.com
biancodevid@gmail.com
giorgio.klauer@gmail.com

ABSTRACT

The paper presents an experiment in which subjects had to localize headphone-reproduced binaural piano tones while in front of a grand Disklavier instrument. Three experimental conditions were designed: when the fallboard was closed the localization was auditory only; when the fallboard was open the localization was auditory and visual, since the Disklavier's actuated key could be seen moving down while the corresponding note was produced; when the listener actively played the note the localization was auditory, visual and somatosensory. In all conditions the tones were reproduced using binaural recordings previously acquired on the same instrument. Such tones were presented either transparently or by reversing the channels. Thirteen subjects participated in the experiment. Results suggest that if auditory localization associates the tone to the corresponding key, then the visual and somatosensory feedback refine the localization. Conversely, if auditory localization is confused then the visual and somatosensory channels cannot improve it. Further experimentation is needed to explain these results in relation with i) possible activation of the auditory precedence effect at least for some notes, and ii) potential locking of the sound source position that visual and/or somatosensory cues might cause when subjects observe a key moving down, or depress it.

1. INTRODUCTION

The industry of digital pianos has traditionally relied on the assumption that pianists are able to localize a tone while playing their instrument. The tone is often assumed to arrive at the listening point with an angle spanning the median plane and the key position of the corresponding note. This assumption has led to the design of several panning technologies: besides their use for improving a stereophonic loudspeaker reproduction of the recorded soundfield from an electronic instrument such as a digital piano [1], these technologies were felt especially necessary to render the same soundfield through headphones, or when the sound results from a synthesis model instead of a sample-based design [2].

This assumption finds no straightforward explanation in measurements of real piano soundfields. Contrarily to what one would initially expect, piano tones do not show radiation patterns that can be linked to the note position. These patterns, rather, depend on the modal characteristics of the soundboard, which absorbs almost all the energy that resonates within the strings and then radiates into the air according to its specific patterns of propagation [3]. Different soundboard regions radiate corresponding frequency components; the net result is that, once a tone is played, the pianist is enveloped by a soundfield having no relationship with the position in the keyboard of the corresponding note. This soundfield has been reproduced also in grand digital pianos mounting more expensive sound diffusion technologies [4].

Soundboard radiation prevails during the steady state, i.e., several tens of milliseconds after the hammer has excited the strings, which in their turn freely oscillate while gradually transferring their energy to the soundboard through the bridge. Conversely, before the hammer releases the strings the piano is in transient state. This transient does not admit simple description, as it incorporates several mechanisms starting exactly when a key is pressed. Most such mechanisms produce audible onsets until the soundboard is loaded with sufficient energy and the steady state is reached. Their audibility before the soundboard takes the lead hence defines a transient soundfield, that can be characterized only if the instrument components responsible for these onsets are clearly localized.

Characterizing a transient soundfield is key for understanding whether a piano emits localized onsets, capable of activating the auditory precedence effect [5]. It is known that this effect locks the perceived localization of a sound source to the angular position from where sufficient energy has arrived to the ears during the first milliseconds of a sound event. If this were true for piano tones, then the subsequent steady-state transformation of the acoustic field in a set of frequency-dependent sound sources would have no chance to unlock the perceived angular position of a piano tone. The question, then, becomes to understand if the onsets forming the transient soundfield of a tone radiate approximately from the position of the key producing the corresponding note. To our knowledge, the most informative study on piano transients has been made by Askenfelt in 1993 [6]. In this study it was shown that the so-called touch precursor, which accounts for 20-25 ms of sound coming from the instrument *before* the hammer hits the string, contains in particular a distinct low-frequency “thump” radiating from the keybed. This noise originates when a key bottom bumps against the keybed wooden structure at the end of its fly. Such a bump can be roughly classified as a point-wise impact on a rigid part of the piano, hence potentially eligible for enabling the precedence effect. Yet, it is not safe to claim that a tone can be localized below the corresponding note key thanks to this process. In fact, the mechanical energy propagates much faster in rigid materials than on air; for this reason it must be reasonably assumed that pianists receive a relatively wide wavefront consequence of this bump instead of a well localized wave onset precurring the distributed emission of noise from the keybed.

2. MOTIVATIONS AND HYPOTHESIS

Fig. 1, previously published [7], seems to confirm the above assumption. It shows eight pressure signals measured in July 2012 inside a silent chamber at Viscount International SpA in Mondaino (RN), Italy, over the keyboard of a Seiler model 1849 playing a note C4, using a Bruel&Kjær model 4188 omnidirectional microphone array calibrated and made available by Angelo Farina’s research group. Considering that microphone channels 8 to 22 ranged from C2 to C6, hence covering four octaves, the plots in fact show too little relative delay among signals around microphone channel 14, which was closest to the key C4. Hence, this sound should not enable the precedence effect.

In the same publication it was shown that even a weaker hypothesis did not hold. In fact, pianists who were passively listening to a sound reproduction of the above eight channels scored scrambled versions of the same reproduction to be equivalently realistic in both sound and auditory scene quality [7]. Scrambling was obtained by re-wiring the input-output connections between channels into alternative configurations. Now, only the heaviest reconfigurations (e.g., random re-wiring or swapping of channel pairs between the left and right quartet) were scored significantly lower.

2.1 Active listening

So, why digital piano tonmeisters find consensus by customers (including skilled musicians) when designing piano tones which render definite localization cues? We approached the question by hypothesizing that — partially quoting [7] — independently of the existence of a precedence effect the localization of a note during playing could be locked to the corresponding key position by somatosensory cues of hand proprioception. Holding this “somatosensory precedence”, then the same position could be robustly recalled during listening: this previously learnt process may overwhelm or even suppress the auditory localization of the same note via lateralization cues, in particular resolving any potential incongruence between proprioceptive and auditory information.

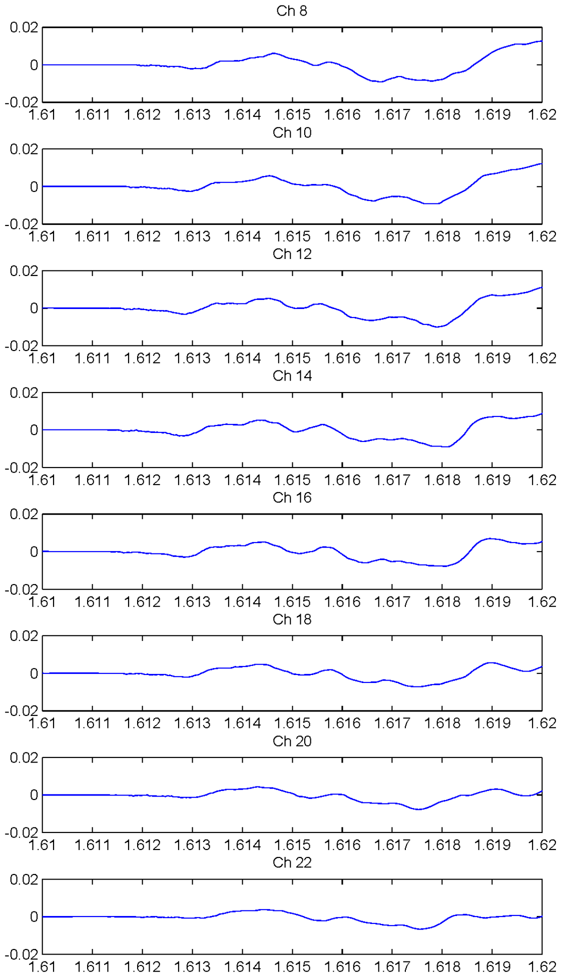

Figure 1. Temporal plots (10 ms) containing the attack of a C4 note on eight selected channels of a microphone array spanning four octaves (C2 to C6) of a Seiler 1849 piano keyboard. Microphone on channel 14 is closest to the key C4 [7].

While a role of the proprioceptive memory in support of the localization of passively listened piano tones is not hypothesized in the following, we experimented on the influence of active listening on the same process. In a previous experiment, whose outcome has been not yet systematically analyzed and for this reason it is still waiting to be published [8], pianists were asked to listen to the multichannel piano tones seen above. The only difference in the stimuli was that we used fourteen- instead of eight-channel reproduction of those tones, hence achieving a higher quality. Concerning the listening conditions, pianists attended also a session in which they actively played the notes instead of just passively listening to the corresponding tones. At each trial, every pianist rated the realism of the tone’s spatial presentation. The result, summarized by the histograms in Fig. 2 which are still waiting for an analysis of their exact significance, tells that pianists preferred presentations where the channels were not scrambled, in both

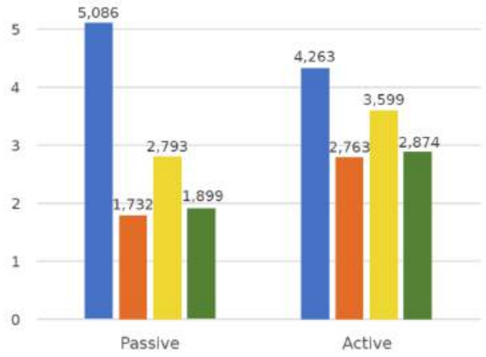

Figure 2. Histograms showing the perceived realism of tone presentations for passively (left) vs. actively (right) listening performers, respectively for unscrambled (blue), pan-reversed (red), randomly scrambled (yellow), and monophonic (green) array listening.

passive and active listening conditions. Furthermore, active listening supported the spatial realism of scrambled presentations.

In spite of its potential insignificance, this result is interesting as it suggests a logical cross-modal process at the base of spatial perception in pianists: if tones were perceived to sound realistic in space, then the auditory process was not supported by the kinesthetic action of playing the keys; if tones did *not* sound realistic in space, then the kinesthetic action conversely supported the same process. So, at least preliminarily, this experiment could be in favor of a key position-dependent auditory localization hypothesis, supported by somatosensory cues if the auditory process becomes confusing.

2.2 Current hypothesis

We investigated whether listeners localized tones on the corresponding keys. Binaural listening was enabled through headphones. Furthermore, on top of somatosensory feedback we tested also the role of visual feedback, by including in the experiment a condition in which participants while listening to a tone could see the corresponding piano key to move down. The test was targeted to individuals who were not pianists, to minimize the effects of repeated exposition to auditory feedback cues of piano tones following by long-practiced kinesthetic actions of key pression, and by the related somatosensory and visual perceptions.

3. SETUP

3.1 Recordings

As a preliminary step to the creation of suitable experimental stimuli, we made an extensive campaign of binaural recordings of piano notes. A grand piano Yamaha Disklavier (model DC3 M4) was used, equipped with sensors and actuators on keys and pedals which can be accessed via MIDI. The piano was placed in a laboratory space, therefore recordings contained also the room response: we do not consider this to be a limitation of the study, as we were interested in investigating ecological listening conditions. In order to avoid strong reflections from the walls, the piano was placed at the center of the room.

The recording setup is shown in Fig. 3 (left) and employed a KEMAR mannequin [9] with microphones placed at the entrance of the ear canal. As shown in Fig. 3 (right) the KEMAR mannequin was placed at the position of the pianist. The setup was moved to the center of the room. The height of the head was chosen to match that of one of the experimenters when sitting at the piano stool, resulting in stool height of about 45 cm from the ground. Additional binaural recordings were taken for different placements of the mannequin, although only frontal recordings taken at the pianist position are used in the remainder of this paper.

The two signals transduced by the mannequin were amplified first by the KEMAR 12AL preamps and then further amplified and sampled at 96 kHz and 24 bits through an RME Fireface 800 audio interface, which in turn was connected to a laptop. A SuperCollider patch [1] running on the laptop performed two tasks: (1) triggering the production of single tones on the Disklavier (through MIDI events), and (2) recording the binaural signals and storing them with appropriate names.

All the 88 keys were recorded. For each key, 10 key velocity values were recorded, from 12 to 111 with 11-step increments. These values were chosen for consistency with a previous study [10] where we collected vibration signals at all the piano keys for the same velocity values: the long-term goal is to develop a multimodal dataset of stimuli containing both binaural recordings and vibrotactile stimuli. While earlier recordings of vibration signals lasted 16 s [10], in this case it was found that this duration is not always sufficient, as the decays for the lowest keys at the highest velocities last much longer (up to 30 s). Therefore, variable recording times were chosen empirically depending of the key and the velocity value, with the goal of reducing the total time of a recording session (which lasted as long as six hours). Recording times for single keys range from 30 s for A0 at velocity 111, to 12 s for C8 at velocity 12.

Recording sessions were performed automatically at night time. This choice had a twofold motivation: first, it was a necessity in order to avoid clashes with other laboratory activities carried out during the day; second, it served the purpose of minimizing environmental noise coming from the exterior (the lab is in a pedestrian-only university area, which is noisy during the day and extremely silent during the night). To this end, we used the launchd service-management framework to temporize the execution of the SuperCollider patch.

3.2 Playback

The setup for the playback of recorded binaural samples employed a laptop, connected on one side to a RME Fireface 800 audio interface and to a pair of Sennheiser HD600 headphones, and on the other side to the Yamaha Disklavier. Fig. 4 depicts the scheme of a SuperCollider patch that

---

[1] http://supercollider.github.io/

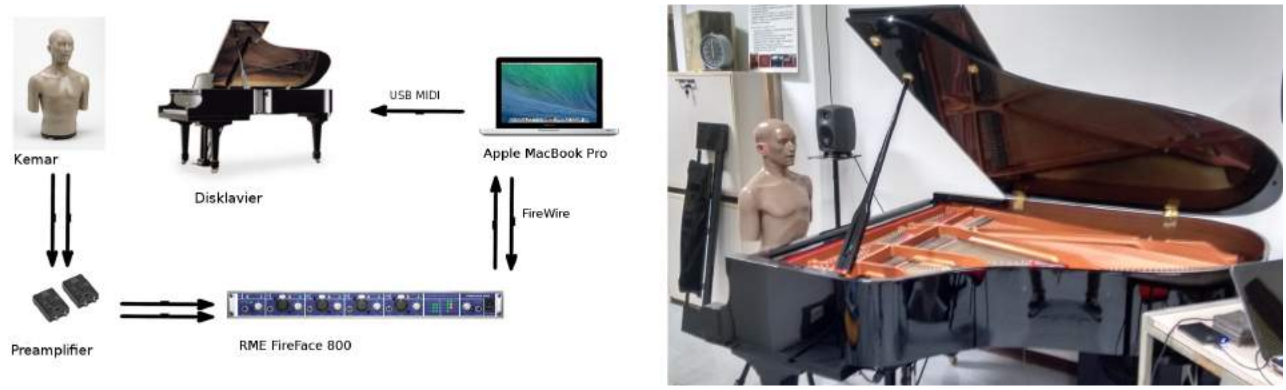

Figure 3. The recording setup. Left: block scheme of the acquisition chain. Right: a picture of the KEMAR mannequin positioned at the pianist location. The setup was moved to the center of the room for the recording session.

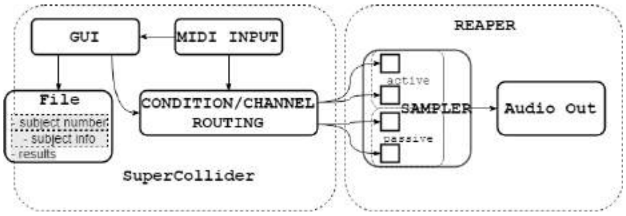

Figure 4. Scheme of the software employed for passive and active playback of binaural samples.

controls the playback conditions. Specifically, in a "passive" condition a binaural sample was played back through headphones, and the corresponding key on the Disklavier was simultaneously depressed by the actuator through a MIDI event. Conversely, in an "active" playback condition the patch received MIDI messages from piano keys depressed by the pianist, and played back the corresponding binaural sample through headphones. In all cases the playback of samples was delegated to a VST sampler hosted by Reaper [2] which in turn was controlled by SuperCollider.

Note that Disklavier pianos can be switched from normal operation to a "silent mode", which disconnects the string board while leaving the MIDI features active. In this modality the hammers do not hit the strings, and therefore the instrument does not sound, while the remaining mechanical operations of the keyboard are left unaltered. Both active and passive playback conditions employed the Disklavier in silent mode, in such a way that no acoustic piano sound was ever superimposed to the recorded samples.

The Sennheiser HD600 are open headphones. We chose to use open headphones to provide subjects with ecological listening conditions where environmental sound was not masked. Moreover, these headphones have a very flat frequency response, which does not disrupt spatial information contained in the binaural signals. This was verified by comparing some of the recorded samples with the same samples played back through headphones and again recorded by the KEMAR wearing headphones. Pairwise comparisons showed very low spectral distortion. In light of this, it was decided that no equalization filter was needed to compensate for the headphone frequency response.

A calibration procedure was devised to match the loudness of the samples played back through headphones to that of the true Disklavier sound at the pianist position. Again, the samples were played back through headphones and recorded by the KEMAR wearing headphones. Comparing the loudness of these re-recordings to that of the original ones led to the insertion of a $+18$ dB gain. Such a high gain had the drawback that the background noise became perceptible in the late stage of the note decay. However this problem was mitigated by the fact that experimental stimuli lasted only 5 s (see Sec. 4 below), and by the use of open headphones which let environmental sound go through.

4. EXPERIMENT

In order to investigate the role of the auditory, visual and somatosensory feedback in piano tone localization, we devised an experiment in which these three types of information were combined in different conditions.

4.1 Stimuli

Three listening conditions were used. In all of them the subject was sitting on the piano stool at the pianist's position in front of the Disklavier; the lid of the piano was fully raised as during the recordings. The first condition employed passive listening of binaural samples with the fallboard closed ("passive closed" condition hereafter). The second one again employed passive listening of samples, but this time with the fallboard open in such a way that the subject could see the movement of the corresponding key ("passive open" condition hereafter). Finally, the third condition employed active listening of samples, in which the subject actively pressed a key on the piano and heard as a consequence the corresponding sample ("active" condition hereafter).

A second factor was channel routing, for which two settings were used. In the first one ("standard" setting hereafter), the two channels of the frontal binaural recordings were routed to the correct headphone channels, while in

---

[2] http://www.reaper.fm

the second one ("reverse" setting hereafter), they were reversed so that the left channel was routed to the right headphone channel and viceversa.

Experimental stimuli were then constructed using only a subset of the A tones, namely the five tones A0, A1, A3, A6, A7. This choice was driven by the need of keeping the total duration of a single experimental session to a reasonable time.

The duration of the stimuli was limited to 5 s, for the same reason. Additionally, 5 s long stimuli minimized the perceptibility of background noise in the late decay stage of the samples, as discussed in Sec. 3 above. In the active condition, the 5 s stimulus duration was enforced by showing a timer to the subjects and instructing them to release the key after this duration; if the key was not released after this duration, a note-off event was automatically sent.

The total number of trials was 120 (i.e., 5 samples × 2 routing settings × 4 repetitions × 3 conditions), thus 40 for each of the three conditions. In the passive closed and passive open conditions the key velocity was set to 84 in all samples. In the active condition the key velocity was set by the playing subject, controlling the corresponding sample play back.

4.2 Procedure

Subjects with an even ID attended the active, passive closed and passive open condition in sequence; those with an odd ID attended the passive closed, passive open, and active condition instead. Stimuli were randomized within each condition.

After each trial, the subject had to rate the stimulus along two dimensions: the overall "sound quality", rated on a 7-points Likert scale (from 1 to 7), and the perceived "lateral direction" of the sound, again on a 7-points Likert scale (from -3 to $+3$). Subjective ratings were recorded through a graphical interface on a laptop placed at the piano music stand. The interface was composed of a vertical slider for the sound quality ratings and a horizontal slider for the lateral direction ratings. The same interface guided subjects through the experiment. For the active condition, the actual (MIDI) note values and velocity values of the keys played by subjects were recorded as well.

Experimental sessions started with a brief verbal introduction to the subjects, who were told that the goal of the experiment was to evaluate the "quality" of piano sound, with no mention of localization or spatial features of the sound. Then the various phases of the experiment were explained. In particular, subjects were instructed to keep their head still during each trial and to maintain the visual gaze on a single location in front of them. This ensured that, during the playback of the binaural stimuli, the reference frame of the subjects' heads remained coherent with that of the Kemar head (with which the same binaural stimuli were acquired).

Then, a pre-experimental interview was carried out to collect age, self-reporting of hearing problems, degree of familiarity with the acoustic piano as an instrument. Three alternative answers to the last question were given: performer, listener, no familiarity (meaning that the subject had never experienced the sound of an acoustic piano in live conditions). During the session, the experimenter annotated relevant reactions or behaviors by the subjects. Finally after the session an informal post-experimental interview was carried out with questions related to musical competencies of the subject and her/his impressions of the experiment.

Experimental sessions lasted about $35-40$ minutes. A total of 13 subjects took part in the experiment (age $17-26$ years, average 22.2). None of them was a pianist. Interviews revealed that 8 of them were not familiar with the acoustic piano, 8 were listeners, and 1 was a performer although not at a professional level. Moreover, 6 subjects had no music training, 2 had practiced a musical instrument as amateurs for a short amount of time, and 5 received music training and played a musical instrument although not at a professional level.

4.3 Results

We analyze and discuss results regarding the second dimension rated by subjects, i.e. perceived "lateral direction" of the sound. One-way ANOVAs with factor Condition were performed separately for each note, and for the two routing settings independently. Such a choice is preliminary at this stage of the research, since clearly it prevents from bringing to surface the interactions among factors. Even without analyzing such interactions, the results we found using individual ANOVAs are sufficiently rich to lead to an articulate discussion.

Fig. 5 shows boxplots for the standard routing setting. Physical positions of the five A keys have been normalized in the range [-3,3], see Sec. 4.2. Furthermore, perceived note position means are reported for each of the three listening conditions (passive closed, passive open, active). The result is mixed: lower tones were localized on positions having clear relationship with the keys; localization, though, ceased to depend on key positions for the higher notes, which in fact were localized slightly leftward from the center.

In the reverse routing setting, similar boxplots were found. Fig. 6 in fact shows that participants located the tones assimilably on reversed positions, again displacing the higher notes to the wrong half of the keyboard.

For each note a one-way ANOVA was conducted among the three listening conditions, separately for the standard and reverse settings. The ANOVA showed significant differences between conditions for note A1 with the standard setting ($F(2, 140) = 6.15$, $p < 0.01$), and for note A7 with the reverse setting ($F(2, 125) = 5.25$, $p < 0.01$). Pairwise post-hoc t-tests for paired samples were ran, using Holm-Bonferroni correction on p-values. The post-hoc analysis revealed that note A1 was localized in significantly different positions in active vs. passive open conditions ($t(89) = 3.52$, $p < 0.01$), and in active vs. passive closed conditions ($t(89) = 3.17$, $p < 0.01$). The same analysis showed that note A7 was localized in significantly different positions in active vs. passive closed conditions ($t(74) = 3.09$, $p < 0.01$), and in passive open vs. passive closed conditions ($t(102) = 2.34$, $p < 0.05$).

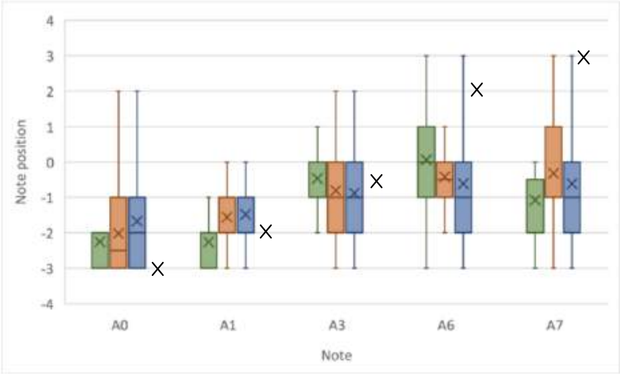

Figure 5. Key vs. perceived note position for the standard routing setting. Right to left boxplots for each note: perceived position in the passive closed condition; perceived position in the passive open condition; perceived position in the active condition. Note positions are normalized in the range [-3,3] and marked with 'X': A0=-3; A1=-2; A3=-0.5; A6=2; A7=3.

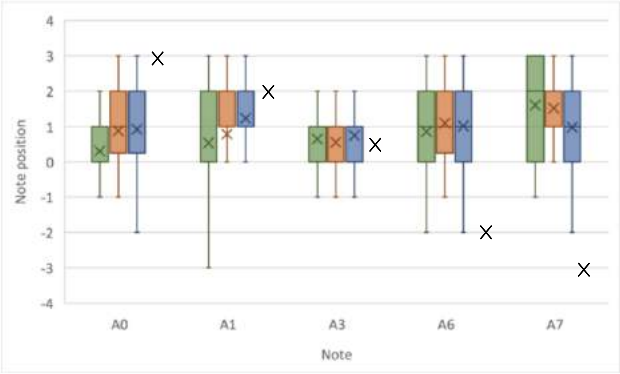

Figure 6. Key vs. perceived note position for the reverse routing setting. Right to left boxplots for each note: perceived position in the passive closed condition; perceived position in the passive open condition; perceived position in the active condition. Note positions are normalized in the range [-3,3] and marked with 'X': A0=3; A1=2; A3=0.5; A6=-2; A7=-3.

A further series of one-way ANOVA was conducted for each condition and routing setting, this time to uncover significant differences among notes. These ANOVA showed that all notes were localized on significantly different positions and routing setting, except in the case of passive closed condition with reverse setting. Pairwise post-hoc t-tests for paired samples were ran, using Holm-Bonferroni correction on p-values. The post-hoc analysis revealed a number of significantly different pairs. For brevity, individual t-statistics are omitted and only the significant p-values are listed in Table 1.

5. DISCUSSION

First of all, with standard routing setting listeners did localize all tones without significant support (except for note A0) from the visual or somatosensory feedback. Notes A0, A1 and A3 were localized around the corresponding key positions: in this case the visual and somatosensory feedback in general helped refine the localization, meanwhile with apparently stronger role of the somatosensory than visual feedback in locking the sound source in correspondence of the respective key. The data are also less dispersed when somatosensory perception was enabled. On the contrary, notes A6 and A7 were localized off key position. In this case the contribution of the visual and somatosensory feedback to the auditory localization is not clear: for note A6, the two sensory modalities seemed to shift the sound source position toward the key; for note A7, they probably wandered around the sound source position where it was localized through listening.

	standard			reverse		
	A	C	O	A	C	O
A0-A1						
A0-A3	●	○	●			
A0-A6	●	●	●			
A0-A7		●	●			
A1-A3	●		●			
A1-A6	●	○	●			
A1-A7	●	○	●			
A3-A6						○
A3-A7				●		●
A6-A7						

Table 1. Significantly different note pairs under different listening conditions and routing settings. A: active; C: passive closed; O: passive open. '○': $p < 0.05$; '●: $p < 0.01$'.

In the reverse routing setting, listeners appear more confused. Notes are largely localized in the normalized portion [0,2] of the keyboard range. Once again there was no significant support from the visual or somatosensory feedback, except for note A7 which, however, suffers from the most inaccurate localization. In general the somatosensory data are more dispersed, suggesting that active playing did not help reduce confusion in subjects when listening to reverse routing.

In the limit of its significance, the ANOVA among conditions suggests the existence of a progressively supportive role of the visual and then somatosensory feedback in localizing a tone over the corresponding key when the auditory system did not contradict this localization process. Conversely, such two sensory modalities would have no decisive role when the auditory localization failed to match the sound source with the corresponding key position.

The same analysis was repeated for the amateur musicians subgroup (5 subjects), and for subjects with an even and odd ID who were respectively exposed to a different order of the experimental conditions—refer to the beginning of Sec. 4.2. Results were analyzed informally, due to insufficient population in those subgroups. Amateur musicians provided answers that are less dispersed around the note position, whereas they did not perform differently from non-musicians under reverse routing conditions. The same effect seems to affect subjects with an even ID, i.e., those who attended the active condition first. A previous knowledge of music playing, then, could have supported subjects in making more correct decisions in normal routing conditions. Similar support could have helped those subjects who first played the notes during the test.

The ANOVA among note positions did not contradict the previous analysis among conditions, and suggests further observations. First, in the reverse routing setting there was certainly more confusion in localizing the tones on significantly different positions of the keyboard. Visual feedback was relatively more decisive in this sense, not only with the reverse but also with the standard routing setting. In this case tone localization was more reliable: Table 1 shows a generally increasing role of the somatosensory (A), auditory (C) and visual (O) modality in supporting the localization of the notes on different keyboard positions. One would be tempted to sort the rows in that table in key distance order, and expect to find an increasingly different perceived position of the tones based on this order. Unfortunately this is true only to a limited extent: pairs A3-A6 and A3-A7, for instance, which fall about on the middle rows if the table is reorganized in key distance order, did not give rise to distinct perceived positions in the keyboard under any type of feedback.

A comparison between our analysis and the data in Fig. 2 does not permit to align this discussion to one possible result suggested by that figure, i.e., that somatosensory feedback may have supported tone localization on the corresponding keys when the auditory feedback was unreliable; in fact, our experiment mainly suggests that visual and somatosensory localization worked better if not contradicting the auditory localization. This misalignment indeed could be expected, since we switched from loudspeaker to headphone listening and furthermore we experimented with participants who were not pianists.

6. CONCLUSIONS

An experiment has been made to investigate the role of the auditory, visual and somatosensory feedback in piano tone localization. We inherited previous literature about the characteristic transients precurring a steady tone in the piano, along with yet partially unstable findings about spatial realism of piano tones using auditory and somatosensory cues. Moving from this literature, we hypothesized that somatosensory and visual feedback supported the auditory localization of piano tones on the corresponding keys. Overall the results suggest that this happens in the cases when the auditory feedback is coherent with the somatosensory and visual feedback, in the sense of the hypothesized specific localization process. However, these results are difficult to be compared with the previous findings.

Given also the different hypothesis as well as conditions existing between the previous tests and the experiment proposed here, we remain noncommittal about the existence of an auditory precedence effect leading the localization of piano notes, whose key positions are well-known by pianists in any case through their visual and somatosensory experience. Rather, we have collected results under a broad set of conditions which may prove precious to build future experiments. Such experiments, first of all, should include a systematic search for robust acoustic precursors that may be responsible for the activation of the precedence effect.

7. REFERENCES

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

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

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

[4] S. Koseki, R. Mantani, N. Sugiyama, and T. Tamaki, “Method for making electronic tones close to acoustic tones, recording system and tone generating system,” EU Patent EP1 357 538A3, Apr., 2003.

[5] C. Tsakostas and J. Blauert, “Some new experiments on the precedence effect,” *FORTSCHRITTE DER AKUSTIK*, vol. 27, pp. 486–487, 2001.

[6] 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.

[7] 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.

[8] F. Fontana, S. Zambon, and Y. D. Pra, “Designing on subjective tolerance to approximated piano reproductions,” in *Proc. of the Third Vienna Talk on Music Acoustics*, Vienna, Austria, Sep. 16-19 2015, pp. 197–204, invited paper.

[9] W. G. Gardner and K. D. Martin, “HRTF measurements of a KEMAR,” *J. Acoust. Soc. Am.*, vol. 97, no. 6, p. 3907–3908, Jun. 1995.

[10] F. Fontana, F. Avanzini, H. Järveläinen, S. Papetti, F. Zanini, and V. Zanini, “Perception of interactive vibrotactile cues on the acoustic grand and upright piano,” in *Proc. Joint ICMC/SMC Conf.*, Athens, Greece, 2014, pp. 948–953.