The title

Atsuko Tominaga[1], Günther Knoblich[1], & Natalie Sebanz[1]

[1] Department of Cognitive Science, Central European University

Author Note

Correspondence concerning this article should be addressed to Atsuko Tominaga, Quellenstraße 51, 1100 Vienna, Austria. E-mail: tominaga_atsuko@phd.ceu.edu

Abstract

AAA

Keywords: keywords

Word count: X

The title

Introduction

Methods

Participants

We recruited 20 participants who already had a degree (above bachelor's or equivalent) in piano performance/teaching or were studying advanced piano performance at a music school. Most participants were right-handed (left: 2, ambidextrous: 2). The mean age of the participants was 28.25 years ($SD = 10.95$). They had 21.55 years of practice on average ($SD = 11.59$). 17 participants had teaching experience in piano ($M = 7$ years, $SD = 6.68$). All participants were recruited through an online participant platform (SONA system, https://www.sona-systems.com). The study (No. 2020_05) was approved by the Psychological Research Ethics Board (PREBO) CEU PU in Austria.

Apparatus and stimuli

The experiment was programmed in Max/MSP (8.1.11; https://cycling74.com/products/max) on a Mac Book Pro with Mac OS X Catalina 10.15.7. A weighted Yamaha MIDI digital piano was used to record participants' performances. The pitch, onset and offset time of each note, and key velocity profiles were obtained from MIDI data using Max/MSP patchers. All auditory feedback was given to participants through headphones (Audio-Technica ATH-M50X). Sheet music was displayed on a computer monitor in front of the participants.

As stimuli, one piece of music, which was used in our previous experiment, was selected (for details, Tominaga et al., 2022). The piece was taken from Clementi's Sonatina Op.36 (No.3) in C major. The first 12 measures of the original piece were used and modified so that the piece had an almost equal number of data points for each dependent variable. The modified piece consisted of a 12-measure isochronous melody notated in 4/4 meter to be

played with the right hand only. Two expressive notations (i.e., articulation and dynamics) were added as *Fig 1.*

We created artificial recordings of students to manipulate their skill levels. In this experiment, we made the recordings by changing two factors of articulation (present or absent) and dynamics (present or absent). Therefore, there were four types of recordings. We aimed to create a quarter of the recordings where both articulation and dynamics were implemented. The second quarter included the recordings where only articulation was implemented whereas dynamics was missing. The third quarter included the recordings where only dynamics was implemented whereas articulation was missing. The last quarter included the recordings where neither articulation nor dynamics was implemented. We generated 4 instances for each type, therefore there were 16 stimuli (i.e., recordings) in total. How we generated the stimuli is written in *Supplementary Material.*

Procedure

Prior to the experiment, participants were required to memorise the piece so that they had enough time to practise and perform it without pitch errors while implementing notated expressions in the experiment.

First, we recorded participants' baseline performance by asking them to perform the piece without listening to any students' performance. A leading metronome (100 quarter beats per minute, 8 beats) was given before participants started performing the piece. Sheet music (*Fig 1*) was displayed in front of the participants. Also, they were told to perform the piece expressively with their interpretation and do their best as a performer. This instruction was given to make sure that they paid attention to expressive aspects of the performance. Each participant performed the piece twice.

After we recorded the participants' baseline performance, participants were informed that they were going to listen to a number of recordings from 16 different students, who were

learning musical expressive techniques. Participants were required to listen to each student's recording first, and then to perform the same piece to teach musical expressive techniques to the student (i.e., recording). In total, there were 16 trials and participants played the piece for each recording only once. The order of the recordings was randomised for each participant. A leading metronome (100 quarter beats per minute, 8 beats) was given before participants started performing the piece.

After participants completed the 16 trials, they were asked to perform the piece in the same situation as we recorded their baseline performance. At the end of the experiment, participants filled in a questionnaire asking about their demographic information and experience in piano performance/teaching.

Data analysis

The dependent variables were computed from MIDI data for data analysis. Interonset intervals (IOIs) are the intervals between onsets of adjacent notes and provide a measure of tempo. Key-overlap time (KOT) is the difference between the offset time of the current tone (i.e., key release time) and the onset time of the ensuing tone and is a measure for the smoothness of musical sequences. A positive value indicates smooth legato styles due to overlap between the current and ensuing tone whereas a negative value indicates sharp staccato styles due to separation between the current and ensuing note. Tone intensity is assessed by key velocity (KV) and measures the loudness of a musical note. A higher value indicates forte styles whereas a lower value indicates piano styles. The value of KV in MIDI varies between 0 (minimum) and 127 (maximum). Also, KV difference was calculated by subtracting the KV value of the current note from that of the following note. We particularly focused on specific points where each subcomponent changed from one to the other (i.e., forte to piano or piano to forte) to measure dynamics contrast between forte and piano.

Data processing and statistical analysis were performed in R version 4.0.5. For

statistical analysis, we included 8th notes with expressive notations only. Pitch errors were identified by comparing the sequence of musical notes produced by a participant with the sequence of musical notes according to the sheet music. Pitch errors included either, extra, missing or substituted tones and were manually removed by using the *editData* R package. For note onsets, 31.87 % of the trials contained at least one pitch error (extra notes: 5.94 %, missing notes: 24.38 %, substituted notes: 1.56 %). For note offsets, 35.31 % of the trials contained at least one pitch error (extra notes: 5.94 %, missing notes: 24.38 %, substituted notes: 5 %). We found that some participants did not precisely follow the sheet music (e.g., they held some notes longer than notated), therefore the order of offsets did not correspond to that of onsets. We considered these as errors and removed the erroneous notes even if the order of onsets was correct. As a result, less than 1 % of total responses were corrected. After removing pitch errors, we removed outliers for IOIs, KOT, KV and KV Difference, defined as values more than 3 standard deviations from the mean of each dependent variable. For each dependent variable, this resulted in less than 2 % of overall responses being removed as outliers.

We performed a 2 x 2 repeated-measures analysis of variance (ANOVA) with the factors Articulation (present vs. absent) and Dynamics (present or absent) for each dependent variable (i.e., IOIs, KOT, KV, KV Difference). The *aov_ez* function in the *afex* R package was used for a repeated-measures ANOVA. For post-doc comparisons on the estimated marginal means, we used the *emmeans* R package.

Results

All effects are reported as significant at $p < .05$. For KOT, KV and KV Difference, we performed two-way ANOVAs separately for each subcomponent (i.e., legato, staccato, forte, piano).

IOIs

Neither main effect of Articulation ($F(1, 19) = 2.79$, $p = 0.111$, $\eta_G^2 = 0.010$) or Dynamics ($F(1, 19) = 3.27$, $p = 0.086$, $\eta_G^2 = 0.003$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 1.48$, $p = 0.24$, $\eta_G^2 = 0.002$). Therefore, participants did not change the tempo depending on the types of the recordings (*Fig 2*).

KOT

Legato. There was a significant main effect of Articulation ($F(1, 19) = 5.59$, $p = 0.029$, $\eta_G^2 = 0.002$). However, there was no significant main effect of Dynamics ($F(1, 19) = 0.06$, $p = 0.81$, $\eta_G^2 = 0.000$) or interaction between Articulation and Dynamics ($F(1, 19) = 1.89$, $p = 0.19$, $\eta_G^2 = 0.002$). Participants produced longer legato when listening to the recordings where articulation was not implemented (*Fig 3*, left).

Staccato. There was a significant main effect of Articulation ($F(1, 19) = 4.88$, $p = 0.040$, $\eta_G^2 = 0.009$). However, there was no significant main effect of Dynamics ($F(1, 19) = 0.68$, $p = 0.42$, $\eta_G^2 = 0.000$) or interaction between Articulation and Dynamics ($F(1, 19) = 3.13$, $p = 0.093$, $\eta_G^2 = 0.001$). Participants produced shorter staccato when listening to the recordings where articulation was not implemented (*Fig 3*, right).

KV

Forte. Neither main effect of Articulation ($F(1, 19) = 1.66$, $p = 0.213$, $\eta_G^2 = 0.00$) or Dynamics ($F(1, 19) = 0.33$, $p = 0.57$, $\eta_G^2 = 0.000$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 0.00$, $p = 0.96$, $\eta_G^2 = 0.000$). Participants did not change performances in terms of dynamics (loudness) depending on the types of the recordings (*Fig 4*, left).

Piano. There was a significant main effect of Articulation ($F(1, 19) = 7.18$, $p = 0.01$, $\eta_G^2 = 0.060$). However, there was no significant main effect of Dynamics ($F(1, 19) = 1.80$, $p = 0.20$, $\eta_G^2 = 0.003$) or interaction between Articulation and Dynamics ($F(1, 19) = 0.04$, $p =$

0.85, $\eta_G^2 = 0.000$). Participants produced softer piano when listening to the recordings where articulation was implemented (*Fig 4*, right).

KV Difference

Forte to Piano.   Neither main effect of Articulation ($F(1, 19) = 3.40$, $p = 0.081$, $\eta_G^2 = 0.012$) or Dynamics ($F(1, 19) = 1.58$, $p = 0.22$, $\eta_G^2 = 0.002$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 0.97$, $p = 0.336$, $\eta_G^2 = 0.001$).

Piano to Forte.   Neither main effect of Articulation ($F(1, 19) = 2.25$, $p = 0.150$, $\eta_G^2 = 0.005$) or Dynamics ($F(1, 19) = 3.76$, $p = 0.07$, $\eta_G^2 = 0.004$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 1.08$, $p = 0.31$, $\eta_G^2 = 0.001$).

These results indicated that participants did not change performances in terms of dynamics contrast depending on the types of the recording (*Fig 5*).

Comparison with baseline performance

In order to investigate how their performances during the experiment differed from their baseline performances (i.e., performances that participants produced before listening to any of the recordings), we performed one-way repeated-measures ANOVAs. For KOT, KV and KV Difference, we performed ANOVAs separately for each subcomponent. To recall the instruction for the baseline performance, we asked participants to perform the piece expressively with their interpretation and do their best as a performer. By comparing with the baseline performance, we can investigate how didactic intentions influence participants' performances.

IOIs

We compared the baseline performance with the performances where either both articulation and dynamics were implemented or neither of them was implemented to examine how participants performed differently for these extreme examples of students' skill levels.

We categorised performances into three groups (baseline, both, none) and treated them as a factor Category.

There was no significant main effect of Category ($F(1.11, 21.18) = 0.58$, $p = 0.47$, $\eta_G^2 = 0.012$; Greenhouse-Geisser corrected), suggesting that participants kept the same tempo regardless of whether they had an intention to teach (*Fig 6*).

KOT

We compared the baseline performance with the performances where dynamics was implemented to examine how participants performed differently depending on whether articulation was implemented or not on each recording. We categorised performances into three groups (baseline, both (i.e., articulation-present, dynamics-present), dynamics-only (i.e., articulation-absent, dynamics-present)) and treated them as a factor Category.

Legato. There was a significant main effect of Category ($F(1.45, 27.53) = 11.82$, $p < 0.001$, $\eta_G^2 = 0.060$; Greenhouse-Geisser corrected). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that there were differences between baseline and the other two categories (baseline and both: $p < .001$, baseline and dynamics-only: p .008), suggesting that participants produced longer legato when they had an intention to teach (*Fig 7*, left).

Staccato. There was no significant main effect of Category ($F(1.12, 21.28) = 0.75$, $p = 0.41$, $\eta_G^2 = 0.011$; Greenhouse-Geisser corrected), suggesting that participants did not play staccato differently depending on whether they had an intention to teach (*Fig 7*, right).

KV

We compared the baseline performance with the performances where articulation was implemented to examine how participants performed differently depending on whether dynamics was implemented or not on each recording. We categorised performances into three groups (baseline, both (i.e., articulation-present, dynamics-present), articulation-only

(i.e., articulation-present, dynamics-absent)) and treated them as a factor Category.

Forte. There was no significant main effect of Category ($F(1.59, 30.18) = 1.43$, $p = 0.25$, $\eta_G^2 = 0.004$; Greenhouse-Geisser corrected), suggesting that participants did not play forte differently depending on whether they had an intention to teach (*Fig 8*, left).

Piano. There was a significant main effect of Category ($F(1.79, 34.09) = 5.68$, $p = 0.009$, $\eta_G^2 = 0.072$; Greenhouse-Geisser corrected). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that there were differences between baseline and the other two categories (baseline and both: p .05 , baseline and articulation-only: p .03), suggesting that participants produced softer piano when they had an intention to teach (*Fig 8*, right).

KV Difference

We compared the baseline performance with the performances where articulation was implemented to examine how participants performed differently depending on whether dynamics was implemented or not on each recording. We categorised performances into three groups (baseline, both (i.e., articulation-present, dynamics-present), articulation-only (i.e., articulation-present, dynamics-absent)) and treated them as a factor Category.

Forte to Piano. There was no significant main effect of Category ($F(1.53, 29.15) = 1.73$, $p = 0.20$, $\eta_G^2 = 0.014$; Greenhouse-Geisser corrected).

Piano to Forte. There was no significant main effect of Category ($F(1.28, 24.39) = 2.27$, $p = 0.14$, $\eta_G^2 = 0.013$; Greenhouse-Geisser corrected).

These results indicated that participants did not make dynamics contrast between forte and piano differently depending on whether they had an intention to teach (*Fig 9*).

Discussion

The present study investigated whether and how expert pianists adapt their performance depending on the skill levels of students. We manipulated the implementation

of articulation and dynamics in the piece to create four types of artificial recordings. These recordings were supposed to be the reflection of students' skill levels. In one quarter of the recordings, both articulation and dynamics were implemented [SUMMARY]. We found that Limitation
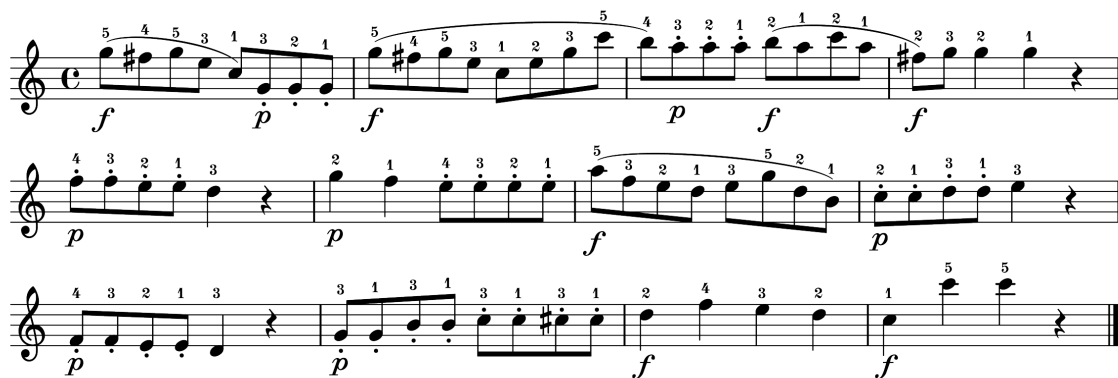
References

Acknowledgement

Figure 1. Sheet music. Articulation. For articulation notation, the curved line (slur) indicates legato and the dots indicate staccato. For dynamics notation, the symbol 'f' denotes forte and the symbol 'p' denotes piano. For data analysis, only the 8th notes were included.
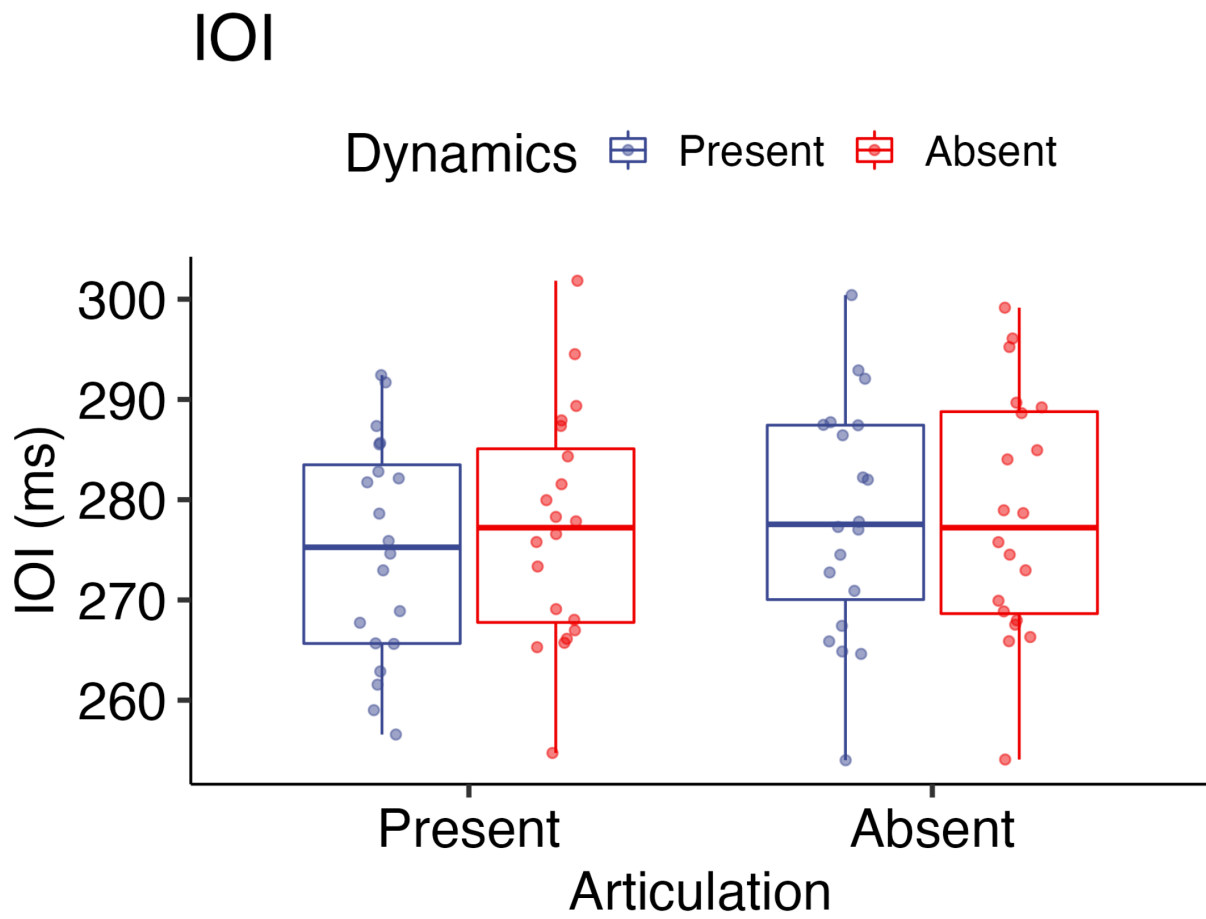
Figure 2. IOIs (ms). Each box indicates the IQR with the median, and whiskers extend to a maximum of $1.5 \times$ IQR beyond the box.
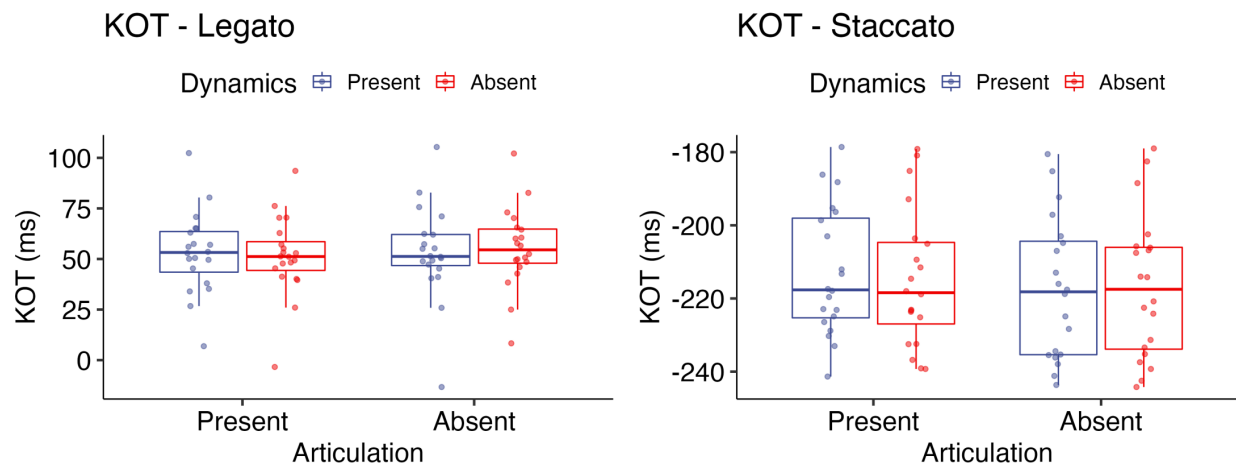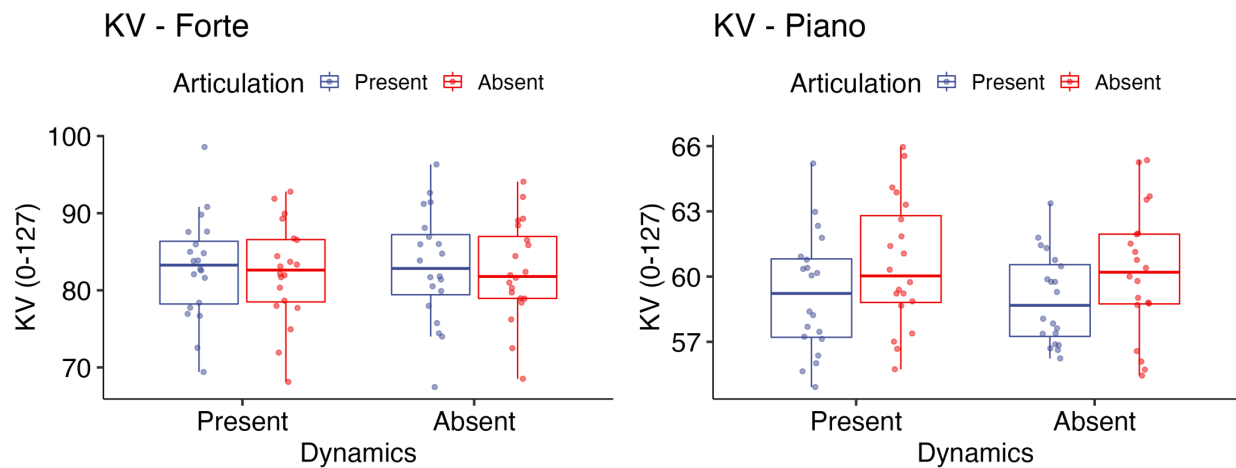
Figure 3. KOT(ms) for each subcomponent; legato (left) and staccato (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of $1.5 \times$ IQR beyond the box.
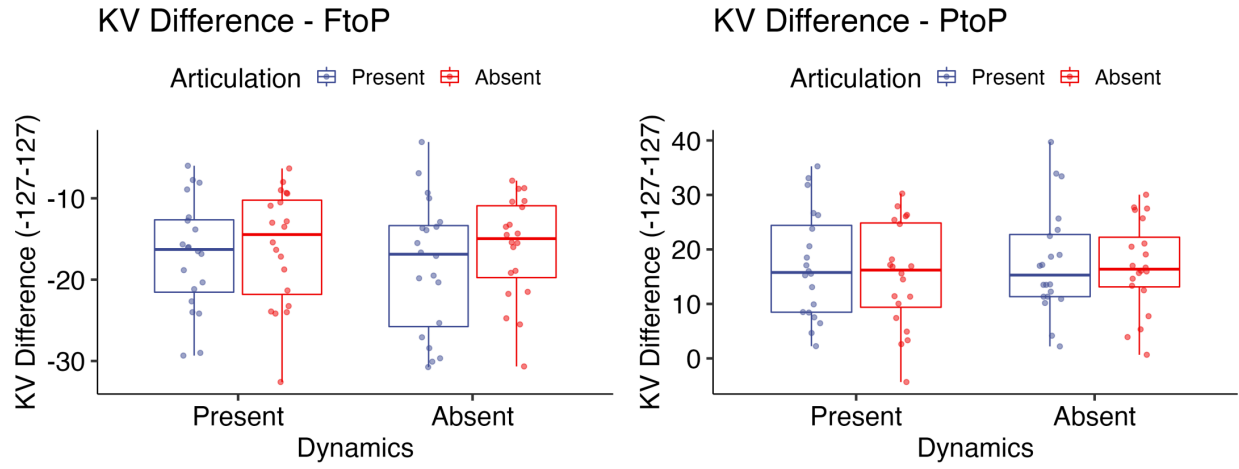

Figure 4. KV (0-127) for each subcomponent; forte (left) and piano (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of $1.5 \times$ IQR beyond the box.

Figure 5. KV Difference (-127-127) for each subcomponent; forte to piano (left) and piano to forte (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of $1.5 \times$ IQR beyond the box.
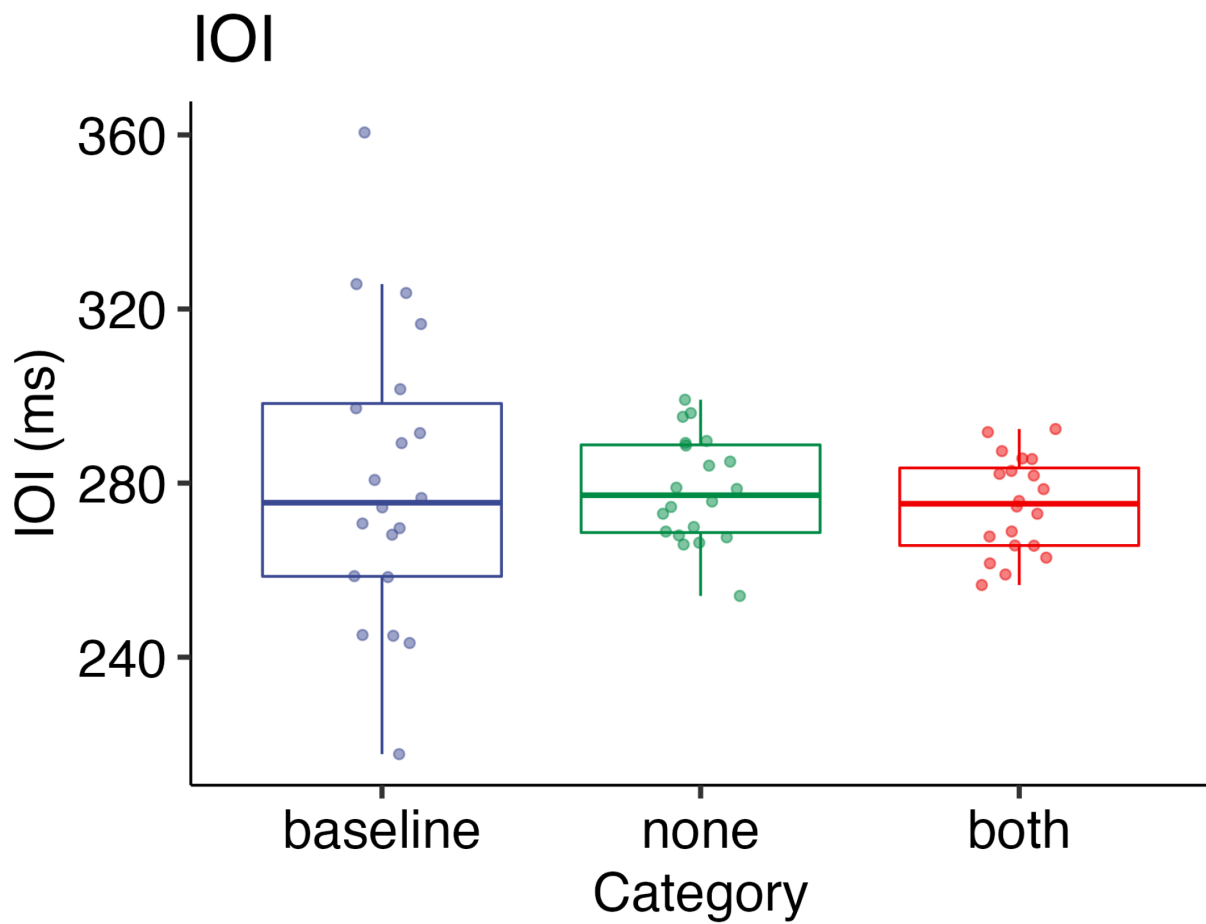
Figure 6. Comparison with the baseline performance in terms of IOIs (ms). Each box indicates the IQR with the median, and whiskers extend to a maximum of $1.5 \times$ IQR beyond the box.
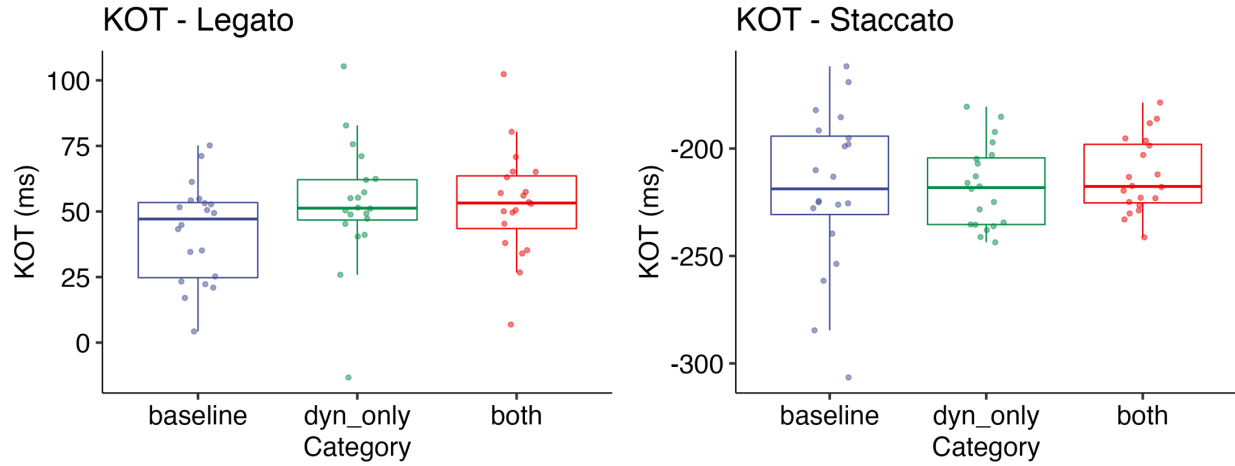
Figure 7. Comparison with the baseline performance in terms of KOT(ms) for each subcomponent; legato (left) and staccato (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of $1.5 \times$ IQR beyond the box.
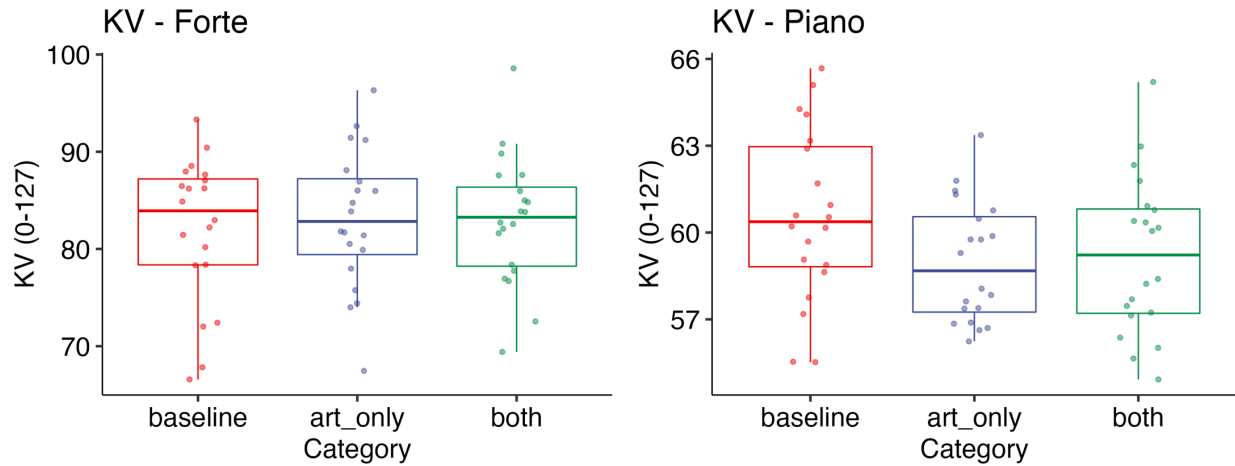

Figure 8. Comparison with the baseline performance in terms of KV (0-127) for each subcomponent; forte (left) and piano (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of $1.5 \times$ IQR beyond the box.
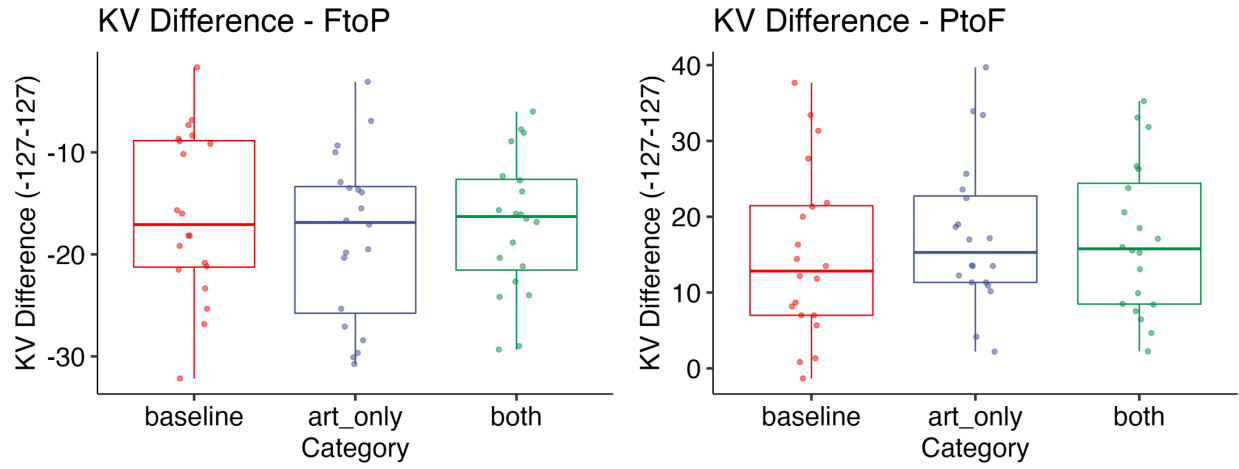
Figure 9. Comparison with the baseline performance in terms of KV Difference (-127-127) for each subcomponent; forte to piano (left) and piano to forte (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.