Tap Tap Revolution: A study of rhythm reproduction from jazz solos with visual aids

Nithya Nadig Shikarpur
Human Computer Interaction: ECSE 542
McGill University

8th January, 2023

Abstract

This study observes design choices involved in building an interface to help users memorize the timing of note onsets (termed rhythm of the melody) from a jazz improvisation solo with respect to a backing track. During the task, the user is expected to listen to an improvisation snippet, memorize the note timings and then tap it out. In addition, this study looks at the roll of visualizations such as piano roll and musical notation in aiding the memorization of such rhythms. Through user tests, we are able to show that visualizations indeed help in better reproduction (or tapping) of the rhythm.

1 Introduction

Improvisation is the simultaneous act of composition and performance of a new work based on a traditionally established chordal framework [1]; it is an integral part of jazz performance [2]. As a beginner, a lot of practice goes into developing playing technique and expanding one's jazz vocabulary in order to aid better improvisation. [3] lists several important steps to follow while teaching jazz improvisation to beginners. Two of these steps are transcribing and copying. The former helps students build the connection between what they hear and the notes that they play on their instrument. Copying or mimicking musicians is essential to internalizing the nuances of the jazz language and also aids in expanding one's jazz vocabulary. In this accord, we study a rhythmic reproduction task where users have to listen to a snippet of melodic improvisation, decipher the note onset timings (rhythmic transcription) and reproduce it (copy) by tapping it out.

Studies have shown that humans are more easily able to capture temporal information in the auditory modality as compared to the visual modality [4, 5]. [5] studies the accuracy of the rhythmic reproduction task given auditory and audiovisual modality inputs. They found that even with multimodal input, the most appropriate modality (auditory for temporal tasks) appears to overpower the other modalities. In that sense, if one uses visual cues while listening to music, the visual modality shouldn't have too much of an effect on the ability of that person to capture or understand the music. However, it is common practice for musicians to write and read music (either as formal notation or any other individual-specific form) while listening to or playing a piece. Other than personal experience, one example that comes to mind is from the movie Whiplash [6], where the drummer, Carl Tanner, panics when he realizes his sheet music was misplaced just before his performance stating that he needed 'visual cues' to perform. This brings the hypothesis that in the context of music, perhaps the visual modality adds to the auditory modality to help musicians understand, learn, play and (in the case of Whiplash) perform music.

[7] was able to show through a corpus study that shorter licks (short melodic sequences in jazz improvisation) are indeed metrically dependent, i.e. a lick is more likely to occur on certain beats than others in a bar. This suggests a strong relationship between the timing of jazz improvisation and the rhythm section such as the bass and drums who are concerned with "keeping time" and thus constantly referencing certain beats of a measure [1]. Hence, through this study, I look at the rhythmic reproduction task in the context of the bass and drum tracks. In addition, since the act of improvisation is spontaneous, the position of the improvised notes can sometimes be unpredictable and difficult to grasp, thus making it a challenging and interesting rhythmic sequence to memorize. While the exact rhythmic reproduction of melodic improvisation lines might not be something a jazz performer would use directly in performance, it would be a useful exercise to help them incorporate the styles and phrasing of another musician in their vocabulary (which would fall under the 'copying' step in jazz pedagogy [3]).

With these arguments in mind, I study the role of visual aids for memorization of the rhythm of melodic sequences from jazz improvisation. Through this study, 2 questions are addressed:

(1) What are the considerations required to build an interface that can help users execute the rhythmic reproduction task, especially for a challenging task as described above? What kind of interactions need to be present?

(2) Does the visual modality help with the process of memorizing a rhythmic pattern in the context of music?

The rest of this paper is organised as follows: task definition, data and preprocessing, prototype evolution, experiment and results, conclusions and future work.

2 Task Definition

The goal of the user was to memorize the rhythmic pattern of a melodic sequence with respect to the bass and drums in the backing track. The melodic sequences were given as audio snippets with an average duration (difference between the timing of the last onset and timing of the first onset) of 4.86s, i.e. 1-2 bars. Each snippet had 2 bars of bass and drums prepended to the melodic improvisation to provide context to the user. This task is ideally intended for jazz students and performers who are trying to develop their skills and vocabulary for improvisation.

In order to carry out the task, each user was asked to perform 3 separate sub-tasks for each audio snippet provided.

1. Listening : In this sub-task the user was asked to listen to the snippet and memorize the note onsets (or the rhythm of the melody) with respect to the backing track (drums + bass). The user had 210 s to listen to the snippet that lasted for 12s (including the 2 bars prepended

Fig. 1. Piano roll visualization used in the interface

Fig. 2. Music notation visualization used in the interface

for context) and automatically played on loop. The duration of 210 s was decided empirically based on previous iterations of testing and feedback from the user. The user also had the option to change the tempo of the snippet in order to help with memorization. The user had access to the visualization if relevant in this sub-task.

2. Establishing the baseline: In this sub-task the user was asked to tap out the position of the notes while listening to the melodic improvisation and viewing the visualization if relevant. This is a partial reproduction task in the sense that the user still has access to the improvisation line (auditory input) while tapping out the onsets. This helps establish how well the user can reproduce the rhythm without expecting them to fully memorize it.

3. Testing: In this sub-task, the user could hear only the bass and the drums and was asked to tap out the positions of the notes. The users didn't have access to the melody of the improvisations and the visualizations (if relevant) here.

In order to study the role of visual inputs, users were given access to visualizations – piano roll and scores, hereafter referred as notation, for some of the snippets. Piano roll is a commonly used visualization method in most Digital Audio Workstations (DAWs) where each note is represented as a rectangle with the length of the rectangle proportional to the duration of the note and the y-position of the rectangle corresponding to the pitch of the note (see figure 1). This form of visualization is quite accessible to people regardless of the amount of their musical training. Notation is commonly used by musicians to read and write music. The vocabulary includes symbols to depict pitches, durations and expression of notes (see figure 2). To read a score comfortably, a student should have undergone some amount of training and practice, thus making this form less accessible to a general audience.

3 Data and Preprocessing

The jazz improvisation snippets were picked up from the Weimer Jazz Database [8]. Specifically, the audio snippets were extracted from the linked YouTube videos on their dataset website (see JazzTube) and the midi information was collected from the dataset.

The snippets were chosen from the Cool Jazz style of playing due to it's relative simplicity. This style of jazz emerged as a softer and more relaxed style of playing compared to it's predecessor bebop [9]. Songs of tempos ranging from 120-130 bpm were chosen in order to contain the difficulty of the task.

Once selected, the audio and midi files were aligned manually on a DAW. Source separation was performed using Demucs [10] on the audio file to extract the drums and bass from the song. The midi files were used to determine note onset timings from the solo. These were considered as the 'true' onsets to which the user's taps would be compared for the analysis. The extracted drums and bass were combined to form the backing track which would be used in the test sub-task.

In order to ensure the sequences were of comparable difficulty, the following statistics were kept in mind for each snippet:

a. Duration – Difference between the onset timing of the first note and onset timing of the last note in the snippet

b. Note Density – Number of note onsets per second in an audio snippet

c. Inter-onset Interval Mean – The mean of the inter-onset intervals of all the notes within a given snippet

d. Inter-onset Interval Std. Dev. – The standard deviation of the inter-onset intervals of all the notes within a given snippet

Keeping these statistics in mind and the author's judgement, 6 snippets were chosen across 2 songs - 'I Fall in Love Too Easily' by Chet Baker and 'Dancing in the Dark' by Zoot Sims, for user testing. Table 1 contains the statistics of the given snippets. They were randomly assigned to one of the 3 input modes (i.e. no visualization, piano roll or notation) resulting in 2 snippets for each input mode mode.

	Duration (s)	**Note Density (notes/s)**	**Inter-Onset Interval Mean (s)**	**Inter-Onset Interval Std. Dev. (s)**
Mean	4.86	2.55	0.44	0.29
Range	3.24	1.03	0.2	0.24
(min – max)	(3.1 – 6.3)	(1.9 – 2.9)	(0.37 – 0.57)	(0.16 – 0.4)

Table 1
Statistics on the audio snippets selected for the user tests.

4 Prototype Evolution

In order to develop an interface to (a) help the user memorize rhythmic sequences effectively with or without visual aids and (b) correctly record the tapping from the user remotely (as it wasn't feasible to conduct all tests in person at the time), 4 rounds of iterative development and testing were conducted. Listed below are the important interfaces and changes made based on testing and feedback:

4.1 Users

Before diving into the prototypes, here is a description of the users. Users for all rounds of testing were friends and acquaintances of the author. Although, not necessarily trained in jazz, they all had training in some form of music and had some music performance experience. The ideal target group for this task would be jazz musicians trying to learn or improve their improvisation, however due to time and geographic constraints we started with this group of users. Since this was intended to be a proof of concept for this task anyway, we decided this small user population shift should be okay.

4.2 Low Fidelity Prototype

To test the feasibility of building such an interface initially, a low fidelity prototype was created on Logic Pro X [11] (a DAW). Given a midi file, the software can automatically generate piano roll and notation visualizations. The user was able to see these representation while tapping out the rhythm. Figure 3 is a screenshot of the interface of the low-fidelity prototype.

4.2.1 Challenges

Based on this prototype a drawback brought to attention was the inability to slow down the audio, which is a common technique by musicians to learn new fast melodic sequences (based on personal experience and anecdotal evidence from the user of the low-fi prototype). In addition, the time for the listening sub-task in this iteration was only 60 s which appeared to be too little to memorize anything properly. This duration was increased in the next stages.

Notwithstanding some very hacky solutions to allow the user to view the visualizations and tap at the same time, this iteration seemed quite promising in terms of the type of task and interactions that could be possible for the user.

4.3 Computer Prototype

Based on feedback and interviews from the previous iteration, a computer prototype was built. This interface was built using magenta.js [12] for the visualizations and tone.js [13] for

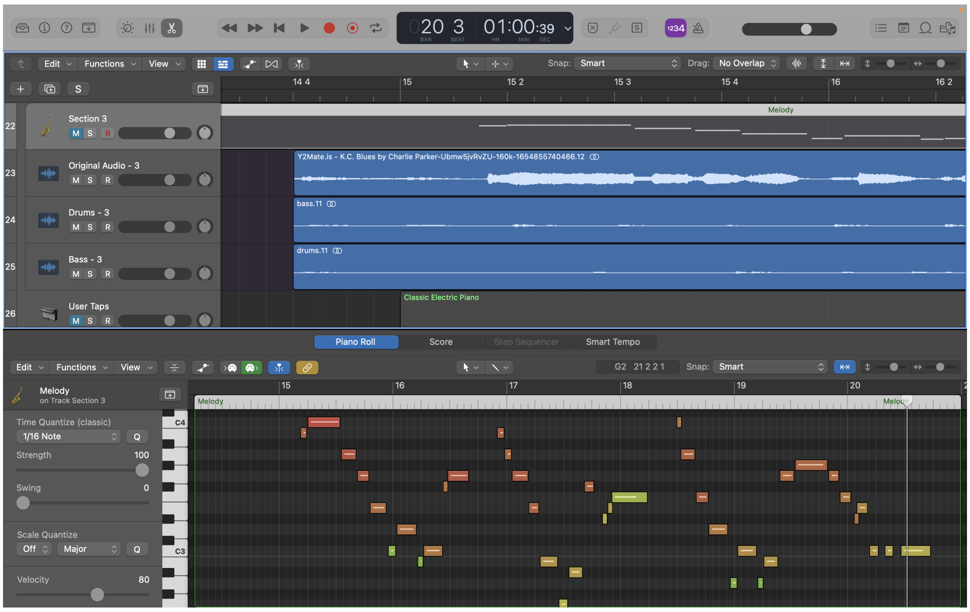

Fig. 3. Screenshot of the low-fidelity prototype on Logic Pro X. The piano roll visualization is available on the bottom half of the screen.

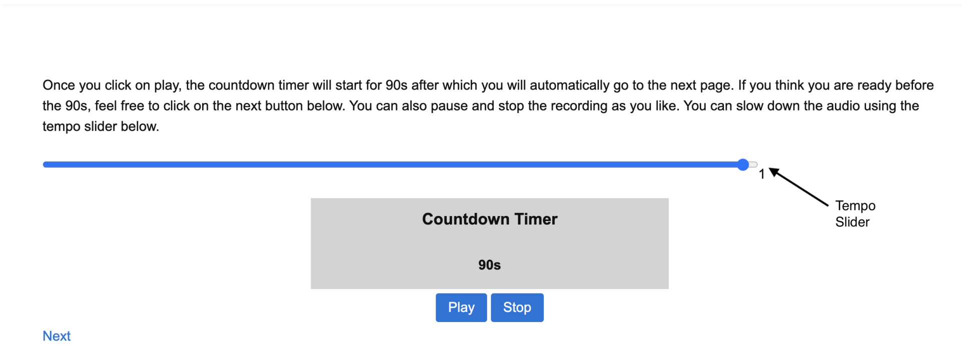

Fig. 4. Screenshot of the listening screen from the computer prototype

all audio related manipulations. The interface was accessed by users through the website [14]* while the author guided them through the interface on video call.

This prototype had 3 distinct screens for the 3 sub-tasks. The user had to enter the rhythm in sub-tasks 2 and 3 by tapping on the ‘b’ key. One of the biggest changes to this prototype was the addition of a tempo slider that allowed the user to slow down the tempo of the given audio snippet in the listening sub-task to aid the learning process. In addition, the user now had 90 s to listen to an audio snippet which lasts around 10 s. Each audio snippet included a count in of 4 metronome clicks. Figures 4, 5, 6 show screenshots of the screens for the 3 sub-tasks present in the computer prototype.

*This interface is still available online. If there are any questions or comments, please feel free to reach out to the author.

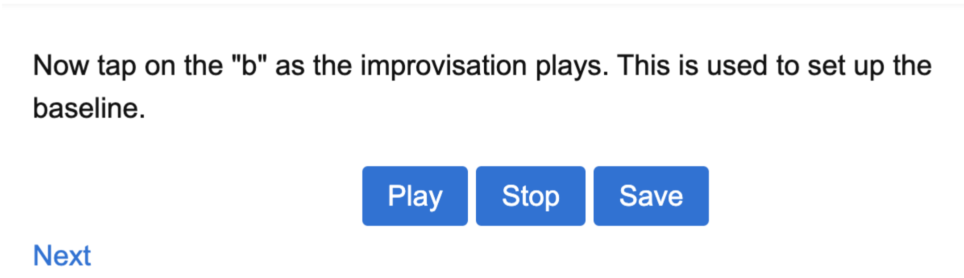

Fig. 5. Screenshot of the establishing baseline screen from the computer prototype

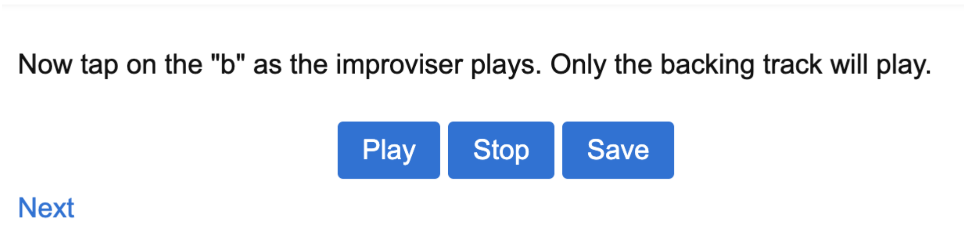

Fig. 6. Screenshot of the testing screen from the computer prototype

4.3.1 Challenges

Again, in this case the duration of 90 s felt insufficient to most users. Another remark received was that all the pieces presented felt like they were of different levels of complexity. This was probably because in this iteration, the complexity of an audio snippet was decided solely by the author's judgement.

A suggestion by one of the users to allow for a two key input versus a one key input also was very useful. The two keys allow users to more easily tap out fast sequences by alternating taps between the two fingers. In addition, users also claimed that they were finding it very difficult to find the starting of the melodic sequence. The problem was that the 4 click count in wasn't giving the user enough context of the song. Since the improvisations didn't necessarily start on beat, more context was required. Another problem that became apparent was that the piano was removed in the testing sub-task and hence users who took cues from the piano were thrown off.

These challenges were addressed in the next iteration as described below.

4.4 Final Prototype

The final prototype looked very similar to the computer prototype but had a few changes to the functionality and interaction of the interface. Based on previous suggestions, the duration of the listening sub-task was increased to 210 s. In addition, simple statistics were used to select audio snippets to ensure equal complexity as seen in Table 1.

The user could tap on either key 'b' or 'v' for the rhythm based on the suggestion listed in the previous section. In fact, this feature was found useful based on the post-test questionnaire by some participants who were common to the computer prototype and final prototype iterations. Also each audio snippet had 2 bars added prior to the improvisation to ensure that there was enough musical context for the user. Lastly, it was also explicitly stated that the piano will not be present in the test sub-task and thus users should refrain from taking cues from it.

5 Experiment and Results

The interface was accessed by the user remotely. They were then guided through the different sub-tasks: listening, establishing baseline and testing, and the various interactions possible through video call. The author was accessible throughout the test on video call in case of any questions or confusions. Each user was given 6 snippets – 2 with only audio input, 2 with audio and piano roll input and 2 with audio and notation input (in that order). Before these, the user also had access to a 'warmup' snippet which was intended to help users get accustomed to the interface and the sub-tasks.

The final prototype was used for analysis. The test was conducted on 3 users who had 7-10 years of music training. Prior to this, 5 other user tests were conducted with this interface but due to a bug in the code (that was found too late), this data became unusable. However I have used their post-test questionnaire answers to conduct a subjective analysis of the task and the interface.

5.1 Subjective Analysis

When asked to rate the easiness of memorization with either no visualization, piano roll (PR) or notation visualization, 6 of 8 participants gave PR a higher score than just audio (indicating PR is easier to memorize with than audio only input). Of the 2 participants who gave PR a lower score, one of them did so because they felt that the difficulty of the piano roll pieces was higher and not necessarily because of the mode of visual input.

The notation was computer generated directly from the midi file using magenta.js. Computers are notorious for automatically generating music notations that are unnecessarily hard to read. As a result, the notation was not very human-readable and was difficult to read by practicing musicians. This was pointed out by 2 participants during the study which resulted in one of them rating the notation visualization less than audio. The other participants (6 users) gave notation a higher rating than just audio despite being unfamiliar with music notation. In hindsight, their unfamiliarity with notation might have been the reason for the higher score for notation.

Except for one participant, who openly stated that he didn't look at the notation visualization because he hasn't studied music notation before, only 1 of the 7 gave notation a higher rating than PR. Everyone else seemed to PR to notation visualization. The scores have been plotted as boxplots for each mode of visual input in figure 7. The box contains values within the first quartile and third quartile and the whiskers contain the entire data.

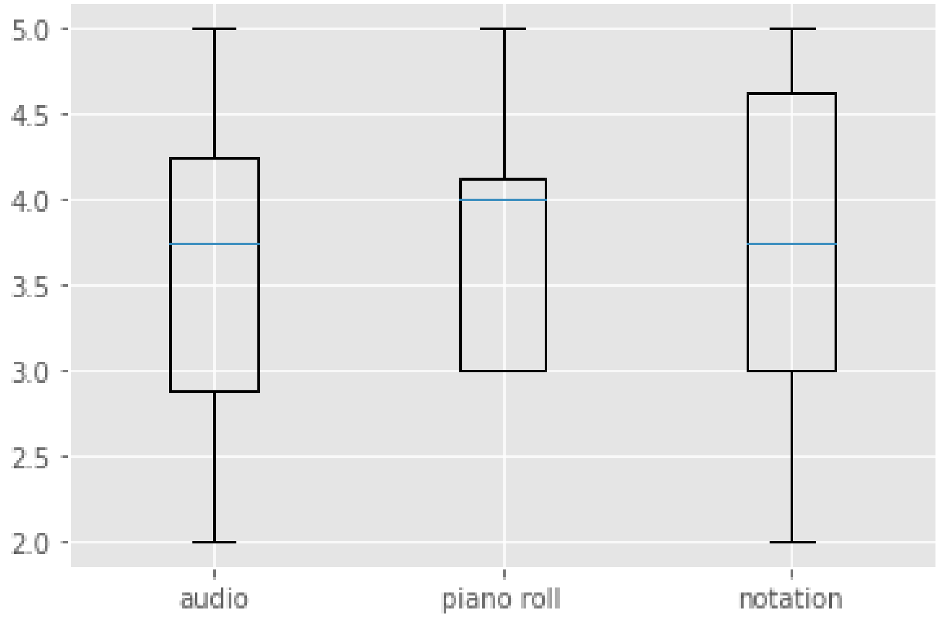

Fig. 7. Boxplots of the scores given by users rating the ease of memorization given 3 types of visual input: no visual input, piano roll, notation

Overall, the consensus seemed to be a preference for having visualization as opposed to not having one. One of the most common reasons stated for this was that it helped them clear confusions such as duration of the note, number of notes present etc. Within visualizations, the preferred form was the PR, although this could be heavily biased by the fact that 5 of the 8 participants weren't comfortable or didn't know how to read music notation.

5.2 Statistical Analysis

2 methods were incorporated to analyze the timing of the user's taps with respect to the true onset timing extracted from the midi files: F1 score and Absolute error.

5.2.1 F1 score

This method was adopted from [15]. Here, given the timing of the true onsets (t) and the timing of the user's tapped onsets (u), the following are calculated:
1. c – the number of correct taps in u
2. f^+ - the number of incorrectly detected taps in u
3. f^- - the number of taps missed in u

where the correctness of a user's tap is determined by whether it occurred within a window of +/- 0.1s of the closest true onset. The window size was relaxed from 0.07s in this study because of the complexity of the task.

To calculate the f1 score, we first calculate precision (p) and recall (r): $p = \frac{c}{(c+f^+)}$ and $r = \frac{c}{(c+f^-)}$.

$$f1\ score = \frac{2pr}{(p+r)}$$

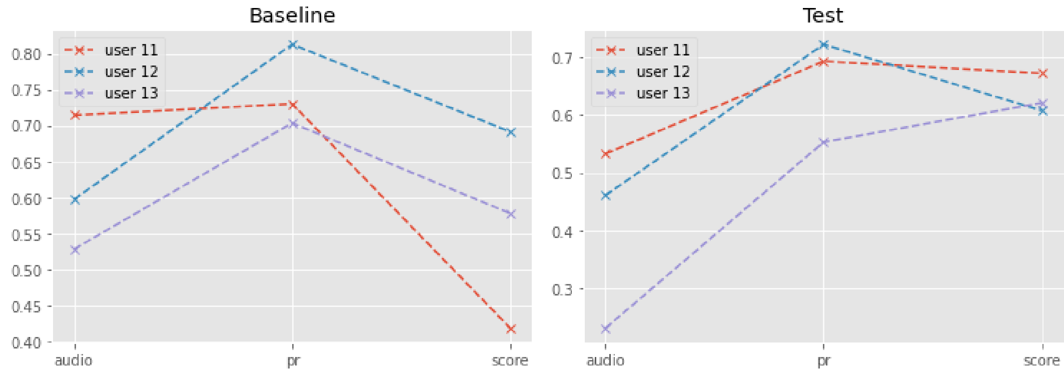

Fig. 8. Plots of the f1 score for each mode of input plotted for the user's baseline and test taps separately for each user.

Analysis

Fig 8 shows the average f1 score for each user with different modes of input for the user onsets from the baseline and test sub-tasks. In the baseline sub-task, it is clear that all 3 users, most of the time, appear to perform better (have a higher f1 score) with a visualization. Another point to note is that for two of the three users, notation has a higher f1 score than just audio. The drop in the notation measure for user 11 is due to a particularly low score in one of the notation tests. Testing with more users will help make certain if this is truly an outlier or not.

Next we look at the f1 score for the test sub-task. Here, it still holds true that tasks with visualization perform better. It is interesting to note that the performance with notation visualization appears to improve for user 13 compared to piano roll, despite that user being unfamiliar with music notation. Another point to note here is that for user 11, there is no significant drop in the scores from piano roll to notation visualization, thus bolstering to the theory that this user's data point for notation visualization in the baseline might be an outlier.

In order to compare the test sub-task results with respect to the baseline sub-task, we also plot the ratio of the test f1 score and baseline f1 score in figure 9. The baseline in some sense is capturing the user's ability to capture the rhythm of a melodic sequences and thus by taking the ratio, we are in some sense normalizing the user's test sub-task score with respect to their musical capabilities. Thus even with this normalization, the trend of having visualization appears to help with memorizing and correctly reproducing rhythmic sequences. 2 observations arise from here: (1) In most cases, the baseline score is greater than the test score (the ratios are less than 1). This is expected since in the baseline sub-task is relatively simpler since the user is able to listen to the melodic improvisation whereas they don't have access to the same in the test sub-task. (2) This ratio is higher with visualizations compared to no visualizations.

5.2.2 Absolute Error

This form of error was inspired from [16]. Each user's tap is matched to the closest true onset timing. In cases where 2 of the user's taps are mapped to the same true onset, only the mapping that has smaller error (absolute difference between timing of true onset and mapped

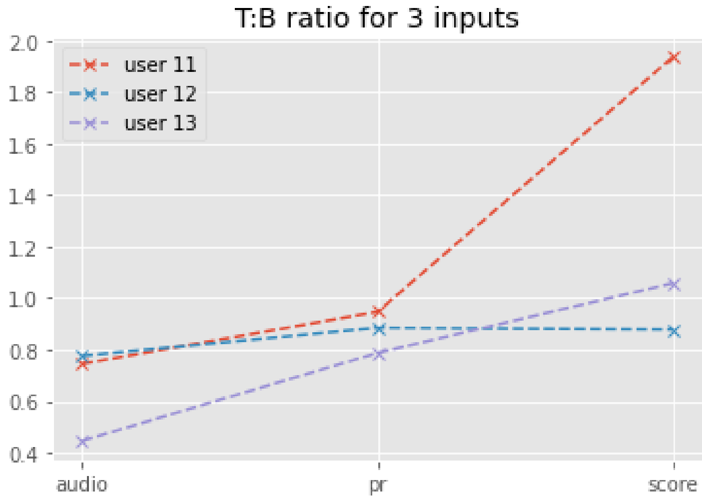

Fig. 9. Plots of the ratio of the f1 scores for the test and baseline sub-tasks for each mode of input plotted separately for each user.

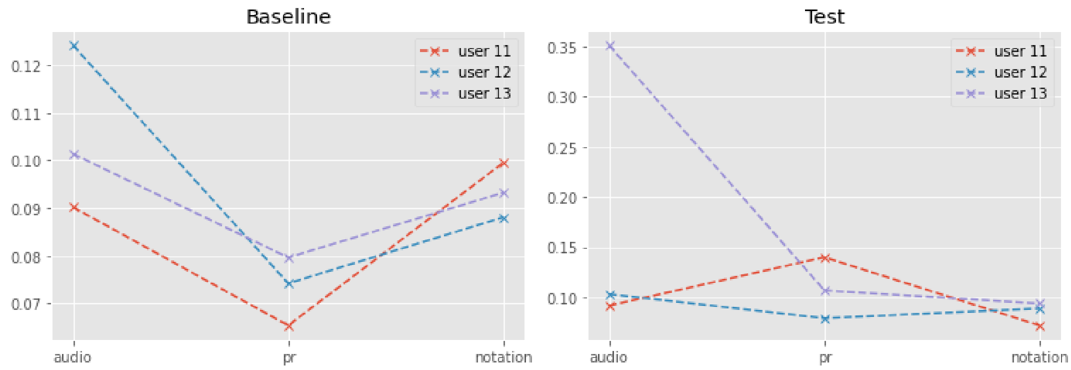

Fig. 10. Plots of the sum of absolute error from each of the users with different input modes for the baseline and test sub-tasks

user's onset) is retained. Once this mapping is established, we sum the absolute difference between the mapped pairs to generate error for the taps from each user for each song. In [16] they also involve a normalization factor which was relevant for the syncopated rhythm reproduction task that they were studying and is thus not relevant here.

Here we observe the errors from the comparison of baseline onsets to the true onsets and that of the test onsets to the true onsets as seen in figure 10. From the baseline figure, it is clear that the piano roll produced the least error, followed by notation which is closely followed by audio. However for the test sub-task, other than the fact that the error here is higher than the baseline sub-task, each of the 3 users seem to be performing differently for different input methods. Thus it is difficult to draw any meaningful conclusions from this.

6 Conclusion and Future Work

Through this paper we have (a) discussed the several considerations and features required in a remote interface built to help participants memorize rhythmic patterns with or without visual aids and (b) studied the effect of these visual inputs.

Since this task is complex, it is easy for the participant to get overwhelmed. Measures such as the tempo slider, clear play/pause/stop buttons and 2 key-input vs. 1 key input were added to make the experience more enjoyable and less frustrating. The audio snippets were also chosen such that they didn't have notes that were too fast or notes with too many off-beat occurrences. In fact, one of the participants admitted to getting overwhelmed due to the complexity of one of the audio snippets and said that that might have affected their performance in the next snippet as well. The goal was to minimize this sense of fluster as much as possible.

A study of the taps from 3 users clearly indicated that the errors in tapping decreases with the addition of visual aids such as piano roll and notation as measured by the f1 score and absolute error metrics. This was also confirmed by the post-test questionnaire conducted with each user. This is not in line with the observation in [4] that didn't find significant difference in the rhythm reproduction task between auditory and audiovisual input conditions. Here we do find an improvement in the audiovisual condition. However, in their study, [4] mentions that participants admit to mainly focusing on the audio in the audiovisual condition. This was also the case in the tests conducted here, participants claimed to focus mainly on audio but use visual cues to clear out confusions or confirm the rhythms that they had deciphered from the audio. [17] claims that music notation is intended to act merely as a memory prompt. Although in this study, the visual cues weren't present as a prompt in the test sub-task, we were able to show that they might also be a helpful prompt in the process of memorizing a rhythmic sequence as well.

There is a lot of room for future work which will be discussed in the following paragraphs. Starting with the drawbacks of this current study, testing with more people would help show statistically significant results as well as help remove outliers as we saw in the analysis. In addition, a common remark was that the complexity of pieces didn't feel equal despite the use of the statistical measures to choose the snippets. Upon further analysis, I realized that it is also important to consider the measure of how many of the notes appear on beat or off beat to measure the complexity of the piece. Another common suggestion was the addition of gridlines for the piano roll visualization. This is the common way piano rolls are represented in any DAW and also more clearly indicates note position and duration. This wasn't directly available in the magenta.js and thus couldn't be implemented for this test. Additionally, the notation displayed could also be altered to be more human readable.

2 of the 8 users who participated were primarily percussionists – one a drummer and the other a tabla player (Indian percussion instrument). In their post-test questionnaires, both

of them spoke about how this wasn't a task that they were necessarily trained to do. As the performer in charge of maintaining rhythm in a band, they are more concerned with constantly referencing the accented beats in the rhythm cycle at appropriate times rather than memorizing the micro timing of notes in jazz improvisation solos which wouldn't always be on beat. This pointed at the fact that different band members have a very different relationship to this task and it thus might not be relevant to every jazz musician. Due to time constraints, this test was conducted only on friends of the author who were all musically trained but not necessarily in the style of jazz. Since this kind of exercise would be aimed towards students of jazz, and perhaps particularly melodic soloists in jazz bands, conducting the next rounds of testing on this target audience would be useful to get meaningful results.

Another interesting point noticed was that none of the users consistently used the tempo slider. Few of them used it once just to check how the feature worked, but not more than that. The reasons for this as discussed in the post-test questionnaire was (a) the pieces were already slow enough that the users didn't require slowing down and (b) the time of 210 s was too little to listen to the slower version and then adapt that to the normal tempo in which the baseline and test sub-tasks would be conducted. This reminded me of the rule we often discussed in class - "users don't always know what they want". One last possible improvement could be the use of more advance beat tracking measures as discussed in [15]. Given the lack of data in this study, evaluation metrics were kept to their simplest forms to avoid over-analysis.

7 Acknowledgements

I would like to thank Prof. Jeremy Cooperstock for the wonderful HCI course and initial feedback for my idea. I would also like to thank Jonathan Lane-Smith and Max Henry for letting me bounce ideas off of them and giving me insights that helped shape this study. Most importantly I would like to thank all my friends and family who graciously agreed to take part in this study (sometimes multiple times!).

8 References

1. Tirro, F. Constructive Elements in Jazz Improvisation. *Journal of the American Musicological Society* **27,** 285–305 (1974).
2. Alperson, P. On Musical Improvisation. *The Journal of Aesthetics and Art Criticism* **43,** 17–29 (1984).
3. J. Richard Dunscomb, W. H. *Jazz Pedagogy: The Jazz Educator's Handbook and Resource Guide* (Alfred Music Publishing, 2002).
4. Barakat, B., Seitz, A. R. & Shams, L. Visual rhythm perception improves through auditory but not visual training. en. *Current Biology* **25,** R60–R61 (Jan. 2015).
5. Hildebrandt, A., Grießbach, E. & Cañal-Bruland, R. Auditory perception dominates in motor rhythm reproduction. en. *Perception* **51,** 403–416 (June 2022).
6. *Whiplash* 2015.

7. Cross, P. & Goldman, A. Interval patterns are dependent on metrical position in jazz solos. en. *Musicae Scientiae,* 102986492110339 (Aug. 2021).
8. *Inside the Jazzomat - New Perspectives for Jazz Research* (eds Pfleiderer, M., Frieler, K., Abeßer, J., Zaddach, W.-G. & Burkhart, B.) (Schott Campus, 2017).
9. Fripp, M. *What are the different styles and types of jazz music?* Last accessed on 2023-01-07. Oct. 2022.
10. Défossez, A. *Hybrid Spectrogram and Waveform Source Separation* in *Proceedings of the ISMIR 2021 Workshop on Music Source Separation* (2021).
11. (CA), A. *Logic Pro - Apple (CA)* Last accessed on 2023-01-07.
12. *@magenta/music - v1.23.1* https://magenta.github.io/magenta-js/music/. Last accessed on 2023-01-07.
13. *Tone.js* https://tonejs.github.io/. Last accessed on 2023-01-07.
14. Shikarpur, N. *Library* https://snnithya.github.io/rhythm-visualization/. Last accessed on 2023-01-07.
15. McKinney, M. F., Moelants, D., Davies, M. E. P. & Klapuri, A. Evaluation of Audio Beat Tracking and Music Tempo Extraction Algorithms. en. *Journal of New Music Research* **36,** 1–16 (Mar. 2007).
16. Fitch, W. T. & Rosenfeld, A. J. Perception and Production of Syncopated Rhythms. *Music Perception* **25,** 43–58. eprint: https://online.ucpress.edu/mp/article-pdf/25/1/43/564158/mp_2007_25_1_43.pdf (Sept. 2007).
17. Stenberg, A. & Cross, I. White spaces, music notation and the facilitation of sight-reading. *Scientific Reports* **9,** 5299 (Mar. 2019).