Reflected Intuition: Statistical and Musical Implications of a Listening Experiment in Post-Tonal Music

Ana Rebrina[1] and Thomas Wozonig[2]

[1, 2] *University of Music and Performing Arts Graz, Austria*

[1]a.rebrina.du@gmail.com, [2]twozonig@gmail.com

Abstract

This study explores listener perception of post-tonal music through a two-step listening experiment conducted as part of the research project *Points of Discontinuity*. In the experiment, subjects annotated aurally "important" musical events during initial real-time listening (step 1) and refined their annotations upon re-listening (step 2). Annotational markers from both steps were analyzed using DBSCAN cluster analysis. Comparing the analyses of the two steps revealed five distinct tendencies in how clusters evolved: (1) clusters became more precise, (2) new clusters emerged, (3) noise (non-matching) areas transformed into clusters, (4) some noise areas remained unchanged, and (5) clusters dissolved into noise areas.

Results show that repeated listening often refined initial impressions, with clusters becoming denser and more precise. However, in processual and layered textures, areas of scattered annotations without clusters persisted, highlighting the complexity of post-tonal structures. These findings underscore the individualized nature of listener engagement and highlight the analytical importance of both clustered and scattered markers in understanding post-tonal music.

Keywords: post-tonal music, perception, first listening impression, repeated listening

Introduction

Understanding how listeners engage with post-tonal musical structures remains a challenge within contemporary music research. The Points of Discontinuity (PoD) research project contributes to this ongoing inquiry by building on key insights from psychological and theoretical studies of listening strategies in post-tonal contexts (e.g., Hasty, 1981; Deliège et al., 1989; Clarke & Krumhansl, 1990; Lock, 2020; Utz, 2023). A particularly influential foundation is Irène Deliège's theory of cue abstraction, which posits that salient auditory features—whether discrete events or extended processes—are selectively registered during listening, forming a cognitive thread that shapes the listener's perception of form (Deliège & Mélen, 1997).

The method of repeated listening and its influence on music perception were of particular interest for our project. Repeated listening has been a common feature of earlier experiments investigating musical segmentation, regardless of the repertoire (Deliège, 1989; Clarke & Krumhansl, 1990; Deliège et al., 1996; Hartmann et al., 2014). Although some research has identified differences in brain network engagement between expert and non-expert listeners during real-time segmentation of music (Burunat et al., 2024), the resulting segmentations may be broadly comparable, although experts often provide more detailed analyses (Phillips et al., 2020). This applies to both initial and repeated listening. In an experiment based on Franz Schubert's *Valse sentimentale*, D. 779, op. 50, no. 6, Deliège et al. (1996) demonstrated that first and second listening resulted in similar identification of landmarks and their perceived importance (however, it should be noted that the piece used was very short, lasting less than one minute). Hartmann et al. (2014), employing excerpts from a variety of primarily tonal Western musical styles , found that musicians identified more segmentation boundaries during repeated listening, presumably due to increased familiarity with the material.

Although this body of research has significantly advanced understanding of listening, many empirical studies on post-tonal music have largely focused on case studies (e.g., Deliège, 1989; Addessi & Caterina, 2005; Phillips et al., 2020). The PoD project seeks to expand this scope by investigating a broader spectrum of post-tonal styles. To this end, we conducted a listening experiment featuring a diverse range of instrumental and electronic post-tonal music. From an initial analysis of 100 works, 23 were selected for the experiment (seven of which were presented as excerpts, see Table 1). These selections reflect considerable diversity in historical era, style, genre, duration, ensemble size (ranging from solo to full orchestra and electronics), instrumentation (and therefore timbre), and degrees of structural complexity.[1] Although not exhaustive, the selection offers a representative overview of the diversity within post-tonal music.

Aims

The experiment aimed to explore how subjects aurally structure post-tonal music during listening and to examine differences in listener perception between initial exposure and repeated listening. Specifically, the study examined the perception of distinct musical "events," a term that was used in a broad sense for various phenomena, ranging from discrete and distinct moments—such as sharp onsets, endings, or abrupt changes—to more gradual and ambiguous processes that evolve and transform over more extended periods of time.

Methods

Each piece in the experiment was listened to by 40 subjects (25 non-experts and 15 experts).[2] The experiment consisted of two steps. In step 1, subjects were asked to identify and rate (on a scale of 1–3) "important events" during an uninterrupted listening session by tapping on a keyboard in real-time. The instructions emphasized that an "event" could be either an isolated moment or an extended process. Subjects were informed that there were no "right" or "wrong" responses and that the perceived significance of events might vary between individuals. In step 2, subjects re-listened the same pieces with the option to pause, rewind, or fast-forward to specific sections. During this step, they could revise their initial annotation by moving, deleting, adding, or re-rating markers on a timescale displayed above the waveform visualization of the piece. No scores or other graphical representations of the music were provided, nor was any information given about the pieces or composers. Additionally, subjects were required to provide verbal descriptors to further explain the placement of each marker.[3] The outputs from the two steps included the timestamps of the markers, their assigned importance ratings, and the verbal descriptors from step 2. Data recorded from both steps enabled tracking changes, including whether markers were deleted, relocated, re-rated, or added in step 2.

The evaluation consisted of a quantitative and qualitative comparison between the markers from step 2 and the model analysis (i.e., the research team's hypotheses about outcomes)[4], as well as a comparison between the markers identified in steps 1 and 2—the primary focus of this paper.

Tracking changes between steps 1 and 2 involved comparing the visual representations of all markers in Sonic Visualiser[5] and conducting a cluster analysis for both steps. The cluster analysis was performed using DBSCAN (Density-Based Spatial Clustering of Applications with Noise) (Ester et al., 1996), an unsupervised clustering algorithm that incorporates noise. DBSCAN requires two parameters to be set: the minimum number of points to form a cluster (minPts) and the distance between points (ε). For this study, minPts was consistently set to 10, representing 25% of subjects—a proportion deemed sufficient to indicate significance, while ε was set in several different settings (0.1–1 second for short and medium-length pieces, and 0.1–2 seconds for longer pieces).[6] As a result of ε setting, a cluster may not correspond to a single point in time but could span several seconds, as long as the density of markers is sufficiently high. Conversely, noise areas denote regions of high activity where

[1] Complexity refers to the difficulty of the task in the experiment and is defined by the research team's consensual assessment on a scale of 1–5. This assessment considers factors such as duration, genre, and texture of the piece.

[2] In the experimental design, an "expert" was defined as someone who has pursued or completed higher education in music (practical or theoretical). The number of non-experts was set higher because their results were expected to show greater variability compared to those of the experts.

[3] For details on experiment design, see "PoD_web_application.pdf" and "PoD_web_application_video.mp4" (https://doi.org/10.5281/zenodo.13981069).

[4] The model analyses are primarily based on listening, in order to mirror the experimental procedure, and are subsequently refined using the score to establish clear textural reference points for later evaluation. The scores including model analyses are available in the respective datasets on the Zenodo community Perception of Contemporary Music – Empirical Research.

[5] Sonic Visualiser is an open-source application for visualization and analysis of music (Cannam et al., 2010).

[6] The different distance settings enable the tracking of changes in cluster formation across different conditions. For instance, if a cluster consistently appears across multiple settings, it indicates robustness. In the following examples, only one setting will be used.

markers are too scattered to form a cluster. These areas can be assessed based on their density, determined by the number of markers within a specific area. Comparing clusters and noise areas across both steps enabled us to interpret how subjects responded to the music during their initial and repeated listening experiences.

Results

Across all 23 pieces, the results of expert and non-expert listeners did not differ pronouncedly. However, in every case, the experts' segmentations aligned more closely with the model analysis (Rebrina & Utz, 2025, pp. 121–22). Although previous research has identified differences in brain network engagement between experts and non-experts during real-time music segmentation (Burunat et al., 2024), the resulting segmentations may still be similar, with experts offering greater detail (Phillips et al., 2020). Across all 23 PoD experiment pieces, listeners provided 27,948 descriptors.[7] The most frequently chosen descriptor was *increase of tension* (11.85%), followed by *change of loudness* (6.71%), *change of timbre* (5.97%), *local beginning* (5.93%), *change of density or texture* (5.03%), and *change of pitch* (4.95%). These aspects and parameters are consistent with findings from previous research on the factors underlying segmentation in post-tonal music (Deliège, 1989, Addessi & Caterina, 2005; Phillips et al., 2020).

The comparison between step 1 and step 2 showed that many subjects reassessed the "importance" of events after hearing them within the broader context of the entire piece in step 2. Some events were deemed less significant, leading subjects to delete markers, while others were recognized as (more) noteworthy only during repeated listening, resulting in the addition of new markers.

In 13 of the 23 pieces analyzed in the PoD experiment, the total number of markers decreased in step 2, while it increased in 10 pieces due to the addition of new markers. Table 1 illustrates the list of pieces and the marker counts in the two steps. The quantitative results can be summarized as follows: In some cases, subjects marked events extensively in step 1 and deleted markers in step 2, while in others, they marked sparingly in step 1 and added new markers in step 2. Notably, these variations did not correlate with duration, complexity, or genre but instead appeared to depend on the unique characteristics of each individual piece.

Following earlier research on musical segmentation (see above), it was hypothesized that clusters identified in step 1 would primarily be "refined" in step 2, becoming denser and more precise in pinpointing the specific events that elicited them. While this pattern was observed in many cases, other trends also emerged. Overall, five distinct tendencies were identified:

1. clusters formed in step 1 tend to become more stable and denser in step 2;
2. new clusters emerged in step 2;
3. noise areas in step 1 transformed into clusters in step 2;
4. some noise areas remained noise areas in step 2, even though subjects adjusted their markers;
5. clusters identified in step 1 dissolved and became noise areas in step 2.

To further illustrate these five tendencies, the following section presents examples drawn from specific musical case studies.

(1) The first tendency can be observed in Salvatore Sciarrino's *Sei capricci* no. 1 (1976)[8], consisting of twelve phrases and an echo-like gesture at the end. The piece is characterized by a constant rhythmic pulse in 32nd and 64th notes. Each phrase is shaped by an increase of tension, followed by a decrease and a caesura. In the second half particularly, most clusters formed in step 1 persisted into step 2. Figure 1 illustrates the clusters formed in both steps ($\varepsilon = 0.3$).[9] Even with the tight setting of ε, 6 out of 7 clusters identified in step 1 persisted in step 2, albeit with slight adjustments, indicating that listeners primarily refined and

[7] See "PoD_global_data.xlsx," table "descriptor population ALL" (https://doi.org/10.5281/zenodo.13981069). The criteria for selecting descriptors are detailed in Rebrina & Utz (2025, p. 120).

[8] Recording used in the experiment (CD Salvatore Sciarrino – Marco Fusi – Complete Works for Violin and for Viola, Stradivarius STR 37057, ℗ 2017 Milano Dischi, Track 2). The complete dataset produced in the experiment for this piece is available at https://doi.org/10.5281/zenodo.13985428.

[9] The higher the ε value, the more likely it is that clusters will span longer durations. Since Sciarrino's piece is brief and features short phrases, a low ε setting was chosen to illustrate the clusters. For the remaining examples, the ε setting was adjusted to correspond to the duration of the pieces: for *Ge Xu (Antiphony)* (7:47), $\varepsilon = 0.7$; for *Quartetto no. 4* (13:52), and for *Lichtbogen* (19:24), $\varepsilon = 1.5$.

Table 1. The list of 23 pieces of the experiment sorted by the absolute percentage difference in the number of markers between the two listening steps (last column).

year	duration	composer	work title	complexity (1–5)	number of markers			
					step 1	step 2	diff.	diff. perc.
1976	01:09	Salvatore Sciarrino	*Sei capricci per violino*, no. 1: Vivace	2	395	483	+88	**+20.05%**
1913	01:17	Alban Berg	*Vier Stücke für Klarinette und Klavier* Op. 5, no. 1: Mäßig	3	278	339	+61	**+19.77%**
1957	01:05	Pierre Boulez	*Le Marteau sans Maître*, no. 7	3	289	351	+62	**+19.38%**
1997	02:30	Liza Lim	*The Heart's Ear* (excerpt)	3	354	301	-53	**-16.18%**
1992	02:09	Mathias Spahlinger	*furioso* für Ensemble (exc.)	4	480	419	-61	**-13.57%**
1998	12:46	Beat Furrer	*spur* für Klavier und Streichquartett	2	1,193	1,064	-129	**-11.43%**
2010	02:10	Natasha Barrett	*Animalcules* (exc.)	3	318	285	-33	**-10.95%**
1957	02:18	Karlheinz Stockhausen	*Gruppen* for 3 orchestras (exc.)	5	471	428	-43	**-9.57%**
1964	13:52	Giacinto Scelsi	*Quartetto no. 4*	5	1,008	917	-91	**-9.45%**
1970	07:08	Helmut Lachenmann	*Pression* for one cellist	3	659	611	-48	**-7.56%**
1994	07:47	Chen Yi	*Ge Xu* (*Antiphony*)	2	1,128	1,046	-82	**-7.54%**
1914	01:54	Igor Stravinsky	*Trois pièces pour quatuor à cordes*, no. 2	1	572	606	+34	**+5.77%**
1959	01:39	György Kurtág	*Quartetto per archi* op. 1, iv	2	518	545	+27	**+5.08%**
1982	01:51	Jonty Harrison	*Klang* (exc.)	2	413	432	+19	**+4.50%**
1931	02:08	Edgard Varèse	*Ionisation* (exc.)	1	344	331	-13	**-3.85%**
1909	01:57	Arnold Schönberg	*Fünf Orchesterstücke* op. 16, no. 1: *Vorgefühle*	2	616	631	+15	**+2.41%**
1977	09:44	Brian Ferneyhough	*Time and Motion Study* I	5	1,221	1,245	+24	**+1.95%**
1986	19:24	Kaija Saariaho	*Lichtbogen*	5	1,185	1,164	-21	**-1.79%**
1999	06:17	Curtis Roads	*Half-life*	3	848	836	-12	**-1.43%**
1981	06:32	Nicolaus A. Huber	*6 Bagatellen*, vi	3	815	806	-9	**-1.11%**
1956	09:38	Iannis Xenakis	*Pithoprakta*	4	903	910	+7	**+0.77%**
1993	02:13	Annette Vande Gorne	*Bois* (exc.)	1	308	310	+2	**+0.65%**
1978	31:07	Morton Feldman	*Why Patterns?*	5	1,492	1,490	-2	**-0.13%**

reinforced their initial impressions. The boundaries of the clusters shifted because subjects moved their markers in step 2, often to mark an event more precisely, mostly aligning with the loudest parts of each phrase (i.e., the "peaks" of the waveform visualization). Out of the 933 descriptors provided, the most frequently used were *increase of tension* (12.33%), *change of loudness* (9.43%), and *climax* (7.29%), which corresponds to the locations of the markers on the "peaks" of phrases.

(2) Some clusters, on the other hand, emerged exclusively in step 2. Particularly in the first half of Sciarrino's piece, these new clusters mostly formed at the beginnings and endings of phrases, as well as at the loudest parts of phrases, due to markers being added or shifted to these positions in step 2. This indicates that, upon reflecting and engaging further in step 2, subjects marked the piece with greater detail. The high number of markers added in step 2 (88, see Table 1) further supports the observation that subjects paid closer attention to new details during this phase.

(3) The first two tendencies—which point to step 2 as a more "sorted version" of step 1—also occur in longer pieces. The difference is that the noise areas therein might span over several minutes.

Even those long noise areas tend to transform into clusters to some degree (tendency 3). In Chen Yi's *Ge Xu (Antiphony)* for chamber orchestra (1994)[10], a piece characterized by clearly perceivable yet extended transformative sections—reflected in the three predominant of the 2,011 descriptors chosen by participants: *increase of tension* (15.07%), *change of timbre* (8.4%), and *local beginning* (6.46%)—a striking example occurs at the 4:30 mark, where the percussion section begins a long buildup. Gradually joined by other instruments, a layered and intense texture is created. In the experiment, this buildup yielded a 73-second noise area containing 115 scattered markers, with a high density of 1.58 markers per second. Subjects distributed their markers across numerous local impulses, such as rhythmic and pitch changes, occurring throughout the buildup. In step 2, after subjects had gained a more profound impression of the piece, this area became partially organized, forming at least two clusters around events where new timbres were introduced. These clusters demonstrated agreement among 21 subjects (52.5%) in the first case and 18 subjects (45%) in the second (Figure 2). However, even after this reorganization, the same area still contained 76 non-clustered markers, underscoring the inherent ambiguity of this section.

(4) Given the experimental setup in which subjects reacted in real-time in step 1, the three discussed tendencies—that most clusters formed in step 1 persist, that new clusters emerge, and that

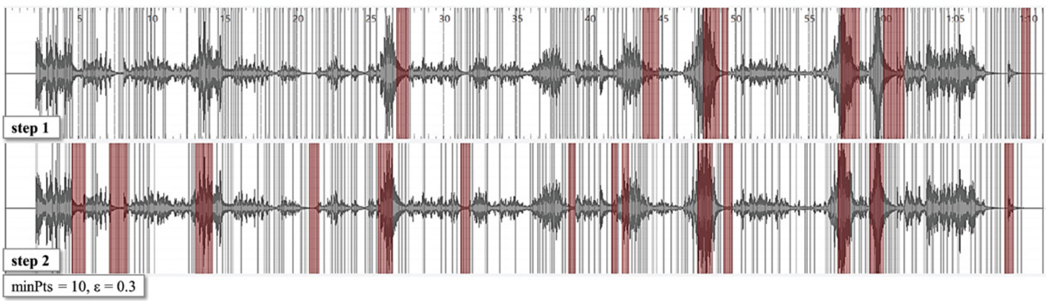

Figure 1. Salvatore Sciarrino, *Sei capricci* no. 1, subject markers (vertical lines) against a waveform visualization of the presented recording. Top: subject markers in step 1, bottom: subject markers in step 2. Red areas indicate the formation of clusters. In step 1, a total of 395 subject markers were recorded (131 clustered in 7 clusters, and 264 classified as noise). In step 2, 483 markers were recorded (239 clustered in 14 clusters, 244 classified as noise).

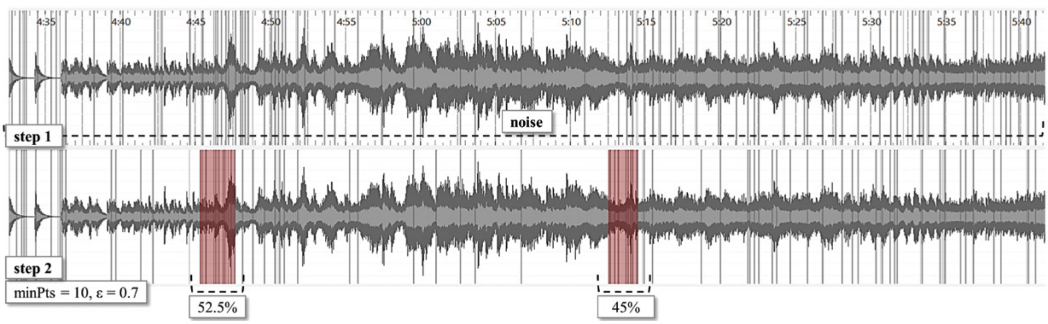

Figure 2. Chen Yi, *Ge Xu (Antiphony)*, 4:33–5:41, top: subject markers in step 1, bottom: subject markers in step 2. Red areas indicate the formation of clusters.

[10] Recording: CD The Music of Chen Yi, Chen Yi – The Women's Philharmonic, JoAnn Falletta, NA 090 CD, ℗ 1996 New Albion Records, Track 3, Complete dataset: https://doi.org/10.5281/zenodo.13982412.

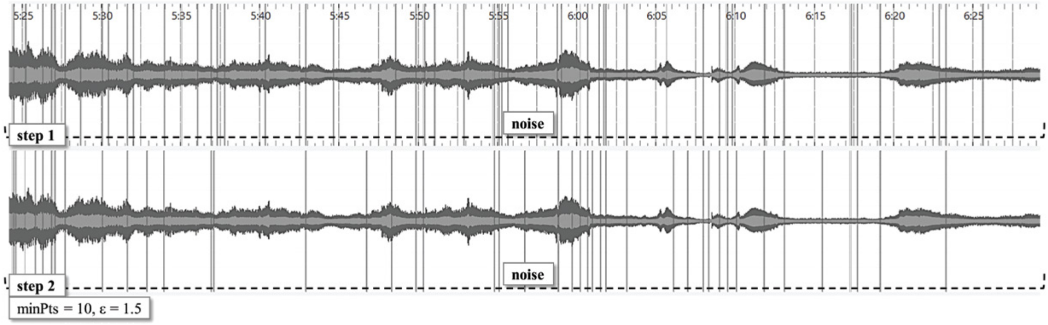

Figure 3. Giacinto Scelsi, *Quartetto No. 4*, 5:24–6:29, top: subject markers in step 1, bottom: subject markers in step 2.

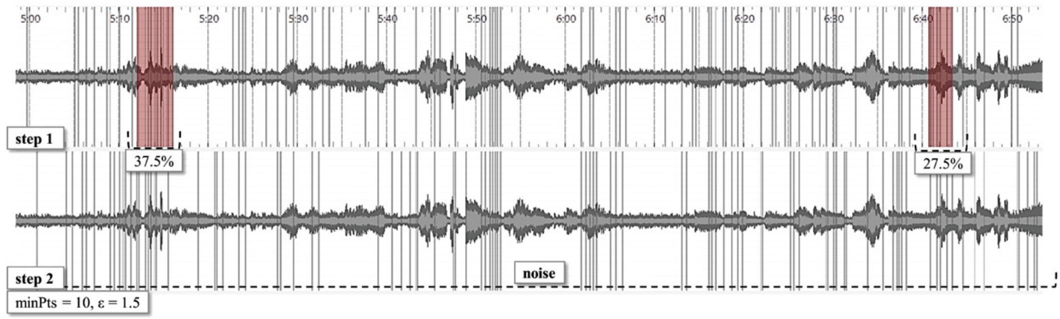

Figure 4. Kaija Saariaho, *Lichtbogen*, 5:00–6:51, top: subject markers in step 1, bottom: subject markers in step 2. Red areas indicate the formation of clusters.

noise areas evolve into clusters—were expected to a certain degree. However, in more complex pieces, additional tendencies were observed. In Giacinto Scelsi's *Quartetto No. 4* (1964)[11], a low-hierarchical piece with a processual texture featuring numerous local cues, the most frequently placed markers (total number: 1,721) were *increase in tension* (16.68%), *change in loudness* (10.75%), *change in pitch* (8.25%), and *change in timbre* (7.09%). In a 68-second-long noise area starting at 5:25, 63 markers were placed in step 1 (density 0.92). After step 2, while subjects moved or removed some markers, the area remained classified as noise, with 48 markers and a reduced density of 0.71 (Figure 3). Despite adjustments, no consensus emerged regarding specific events, and no clusters formed in step 2 in this area.

(5) In the final example, Kaija Saariaho's *Lichtbogen* for nine musicians and live electronics (1985–86)[12], ambiguity in step 2 was even more pronounced. Some areas that contained clusters in step 1 dissolved into noise areas in step 2. For instance, two clusters formed in step 1 in a 114-second long section (starting at 4:57), with agreement levels of 37.5% and 27.5%, no longer existed in step 2. Instead, subjects' adjustments led to the creation of a long noise area containing 98 markers, resulting in a density of 0.86 (Figure 4). This dissolution was probably driven by the presence of numerous local cues—once again, *increase in tension* was the most frequently used descriptor (16.43% of 1,936 descriptors), followed by *change in timbre* (6.92%) and *change in pitch* (5.37%). Additionally, subjects often marked different phases of ongoing processes, resulting in a scattering of markers even when they appeared to refer to the same underlying phenomenon.

11 Recording: CD Giacinto Scelsi – Klangforum Wien, Hans Zender – Streichquartett Nr. 4 / Elohim / Duo / Anagamin / Maknongan / Natura Renovatur, 0012162KAI, ℗ 2001 Kairos Production, Track 1. Complete dataset: https://doi.org/10.5281/zenodo.13985403.

12 Recording: CD Endymion Ensemble – Lindberg. Kaipainen. Hämeenniemi. Saariaho, FACD 361, ℗ 1989 Finlandia Records, Track 6. Complete dataset: https://doi.org/10.5281/zenodo.13985390.

Discussion

The five tendencies of cluster formation, or their absence, are closely linked to the textures of the individual pieces and the specific events around which the clusters are formed.[13] These events can be broadly categorized into two types: (A) distinct events that can be pinpointed to a specific moment, such as sudden changes, caesuras, and interruptions; and (B) ambiguous events, such as transitions and transformations, which are processes rather than discrete points.

The first two tendencies—clusters present in step 1 becoming more stable (denser and more precise) in step 2, and new clusters emerging in step 2—are evident in both textures with clear phrasing and in processual textures. However, clusters associated with distinct events tend to become more compact in step 2. Repeated listening not only reinforces the initial impressions but also enables subjects to identify and mark additional details, resulting in a more refined and nuanced interpretation. In this sense, the first annotation functions as a preliminary "sketch" that becomes increasingly precise as the ambiguity and complexity of post-tonal structures are reduced through repeated listening, thereby making them more comprehensible—an effect most likely connected to the well- investigated fact that music becomes more predictable with repeated listening (Huron, 2006, p. 241).

In contrast, tendencies (3), (4), and (5) are more pronounced in pieces with complex, layered, and transformative textures. In such cases, noise areas identified in step 1 are often a result of ambiguous textures. As demonstrated in the example of Chen Yi's *Ge Xu*, these noise areas can become somewhat more robustly organized in step 2, with clusters forming around certain events. However, in more intricate and layered textures, such as in Scelsi's *Quartetto No. 4*, noise areas tend to persist. Even after repeated listening, the uncertainty about marker placement remains unresolved, reflecting a higher degree of ambiguity regarding the determination of structurally "important" cues within the musical flow. The most extreme case of this phenomenon occurs in pieces such as Saariaho's *Lichtbogen*, in which clusters identified in step 1 dissolve in step 2, leaving behind larger noise areas. This dissolution suggests that, for some textures, the ambiguity of the music in fact increases with repeated listening.

Despite the lack of consensus of identifiable "events" represented in clusters in these cases, areas with scattered markers still provide valuable insights for analysis. A high number of scattered markers within a noise area indicates that subjects collectively perceive a particular "region of significance," even if they do not agree on its precise location. For instance, the dissolution of clusters observed in *Lichtbogen* highlights the analytical importance of processual textures. These regions, despite their lack of defined clusters, offer a rich field for exploring subjects' understanding and navigation of complex post-tonal structures.

Further investigation into individual listening strategies could shed light on different ways individuals perceive and orient themselves within complex musical textures. Additionally, the available dataset, which includes detailed information about the reasoning behind marker placement, offers an opportunity for future research. This might involve analysing the descriptors, potentially allowing for a more parameter-sensitive analysis, and further differentiating types of changes in cluster formations between the two steps, an area that remains a promising avenue for further study.

Conclusion

While the differences in subjects' annotational behavior observed between the two listening steps of our PoD experiment are highly context-dependent, two primary tendencies emerge. First, relatively clear segmentation points tend to become more defined and clusters around them more robust as a result of the subject's repeated listening experiences. Second, ambiguous local processes may either come to highlight certain events within the musical flow more strongly after repeated listening or, conversely, remain ambiguous, as evidenced by the lack of or even decrease of cluster formations.

It has to be noted that in all examples, noise areas without clusters are by no means considered "unimportant." A relatively high number and density of markers in such areas might suggest that they are still somewhat "important" to listeners, even though no clear "moment" can unanimously be identified. These areas are meaningful because they might reflect the perception of multiple micro-events occurring in close proximity to one another, pointing to particularly intriguing processual textures. These findings highlight the individualized nature of subjects' interpretations, underscoring how listeners may conceptualize the temporal flow and structure within post-tonal music.

[13] For detailed discussion on cluster formations in various repertoires, see Utz (forthcoming), Wozonig (forthcoming), and Rebrina (forthcoming).

Acknowledgements. The article was written as part of the research project *Points of Discontinuity: Theory, Categorization, and Perception of Cadences and Openings in Post-tonal Music* (Austrian Science Fund, FWF P 34097-G, 01/03/2021–30/04/2024, https://pod.kug.ac.at). Ana Rebrina, Christian Utz, and Luca Danieli created the initial model analyses and developed the concept of the web application for the listening experiment. Rebrina, Utz, Thomas Wozonig, Daniel Ambrosch, and Marián Štúň contributed to the finalisation of the model analyses, Rebrina, Wozonig, and Utz evaluated the results of the listening experiment. Robert Höldrich, Oscar Bandtlow, Bruno Gingras, Markus Neuwirth, Antares Boyle, Robert Hasegawa, and Daniel Meyer provided important suggestions for the project's methodology, and Zvonimir Kovač programmed scripts for data analysis. The research data collected during the project are published in the Zenodo repository https://zenodo.org/communities/music-perception.

References

Addessi, A. R., & Caterina, R. (2005). Analysis and perception in post-tonal music: An example from Kurtág's String Quartet Op. 1. *Psychology of Music*, 33(1), 94–116. https://doi.org/10.1177/0305735605048015

Burunat, I., Levitin, D. J., & Toiviainen, P. (2024). Breaking (musical) boundaries by investigating brain dynamics of event segmentation during real-life music-listening. *Proceedings of the National Academy of Sciences*, 121(36), e2319459121. https://doi.org/10.1073/pnas.2319459121

Cannam, C., Landone, C., & Sandler, M. (2010). Sonic Visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the 18th ACM International Conference on Multimedia* (pp. 1467–1468). Association for Computing Machinery. https://doi.org/10.1145/1873951.1874248

Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving musical time. *Music Perception: An Interdisciplinary Journal,* 7(3), 213–251. https://doi.org/10.2307/40285462

Deliège, I. (1989). A perceptual approach to contemporary musical forms. *Contemporary Music Review, 4*(1), 213–230.

Deliège, I., Mélen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real-time listening to a piece of music. *Music Perception, 14*(2), 117–159.

Deliège, I., & Mélen, M. (1997). Cue abstraction in the representation of musical form. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 387–412). Psychology Press.

Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (pp. 226–231).

Hartmann, M., Toiviainen, P., & Lartillot, O. (2014). Perception of segment boundaries in musicians and non-musicians. In M. K. Song (Ed.), *Proceedings of the ICMPC-APSCOM 2014 Joint Conference* (pp. 165–170). Yonsei University.

Hasty, C. (1981). Segmentation and process in post-tonal music. *Music Theory Spectrum*, 3, 54–73. https://doi.org/10.2307/746134

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.

Lock, G. (2020). Salienz, Narrativität und die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung von post-tonaler Orchestermusik. *Zeitschrift der Gesellschaft für Musiktheorie,* 17(2), 311–349. https://doi.org/10.31751/1074

Phillips, M., Stewart, A. J., Wilcoxson, J. M., Jones, L. A., Howard, E., Willcox, P., du Sautoy, M., & De Roure, D. (2020). What determines the perception of segmentation in contemporary music? *Frontiers in Psychology, 11*, 1001. https://doi.org/10.3389/fpsyg.2020.01001

Rebrina, A. (forthcoming). Listening to complex forms of post-tonal music: Insights into listeners' temporal orientation in the works of Brian Ferneyhough, Iannis Xenakis, Beat Furrer, and Kaija Saariaho. Manuscript submitted for publication.

Rebrina, A., & Utz, C. (2025). Orientierungen im Klang: Empirische Studien zum hörenden Erfassen temporaler Strukturen in posttonaler Musik. *Die Musikforschung*, 78(2), 116–150.

Utz, C. (2023). *Unerhörte Klänge: Zur performativen Analyse und Wahrnehmung posttonaler Musik und ihren historischen Voraussetzungen*. Georg Olms Verlag.

Utz, C. (forthcoming). Towards a performative listening of new music: Empirical studies on the auditory perception of temporality in post-tonal sound structures. Manuscript submitted for publication.

Wozonig, T. (forthcoming). Objects, processes, and airplanes: Listening approaches to electroacoustic works by Jonty Harrison, Annette Vande Gorne, Curtis Roads, and Natasha Barrett. Manuscript submitted for publication.

https://doi.org/10.17234/9789533793085.04