Influence of metrical structure on learning of positional regularities in movement sequences

Talieh Kazemi Esfeh1 · Javad Hatami1 · Masoud Gholamali Lavasani1

Abstract
Sequential stimuli are usually perceived to have hierarchical temporal structures. However, some of these structures are only investigated in one type of sequence, regardless of the existing evidence, showing the domain-generality of the representation of these structures. Here, we assess whether the hierarchical representation of regularly segmented action sequences resembles the perceived metrical patterns that organize the representation of events hierarchically in temporally regular sequences. In all our experiments, we presented the participants with sequences of human movements and tested the perception of metrical pattern by segmenting the movement streams into temporally equal groups containing four movements. In Experiment 1, we found that a movement sequence with temporally equal groupings improves the learning of positional regularities inherent within each group of movements. To further clarify the degree to which this learning mechanism is affected by the perceived metrical patterns, we conducted Experiments 2a and 2b, in which the relative saliencies of the first and last positions in the movement groups, respectively, were studied. The results showed that, although in the learning of positional regularities, the rule-conforming first positions are as effective as when both first and last positions are legal, the last positions are not as influential. Based on these findings we conclude that, in grouped sequences, learning of positional regularities may be modulated by the metrical saliency patterns that are imposed by the temporal regularity of the sequential grouping pattern.

Introduction
We are surrounded by a world full of stimuli which are extended through time and from which we extract meaning by representing them in hierarchically organized structures (Fitch, & Martins, 2014; Zacks, & Swallow, 2007). The common feature of the sequential stimuli, namely, their unfolding through time, makes them prone to possible influences of hierarchical temporal structures present in the stimuli. Accordingly, it has been shown that the hierarchical representation of various kinds of sequences is processed in the same region of the brain (Fiebach, & Schubotz, 2006). However, some of these temporal hierarchical structures in perceptual and representational levels, despite being well investigated in one discipline (e.g. music), remain unknown in others (e.g. action perception). Correspondingly, the aim of the current paper is to investigate whether the temporal structure of movement sequences provides a perceptual input that forms a metrical representation of these kinds of sequences.

The first step for creating hierarchical representations is to segment the perceived stimuli into groups and this could be done based on different subtle structures and features of a sequence. Human action is shown to be grouped through two different mechanisms. First, according to the principles of the event horizon model, people make predictions about the next movements based on the previous segmented groups of the stream of human activity (i.e. event models) and when their expectations turn out to be false, they start to update their current event model into a new one, by perceiving an event boundary that separates the previous group of events from the new one (Kurby, & Zacks, 2012; Radvansky, 2012; Swallow, Zacks, & Abrams, 2009).

The second mechanism for segmentation of action streams is based on its statistical structure. This specific structure is first studied to clarify the way in which children are able to segment the fluent speech signal into separate words. Such a segmentation mechanism, in the face of language streams, has been shown to be dependent on transitional probabilities. That is, the number of times two (non-) adjacent syllables co-occur, divided by the times when only...
the first one is present in the stream (Saffran, Newport, & Aslin, 1996). Likewise, a number of studies have indicated that this mechanism is not language specific and could also be exploited for the segmentation of human movements, as well (e.g. Endress, & Wood, 2011). In other words, transitional probabilities also facilitate the segmentation of human movement streams into fine-grained units.

Furthermore, according to the More than One Mechanism (MOM) hypothesis (Endress, & Bonatti, 2016), while the statistical learning mechanism helps people segment the movement sequence, another mechanism becomes active simultaneously only when a sequence is already segmented. This second mechanism has first been revealed in studies of artificial language learning. These studies have brought to light the fact that the learning of statistical structures is not sensitive to the silences or prosodic indicators of word-boundaries. That is, this kind of learning neglects all the boundaries and works as if they do not exist. It has been argued that such a learning mechanism falls short of extracting the grammar-like regularities that usually depend on the position of an event within the boundaries of a word-like unit. Therefore, Endress and Bonatti (2007) have introduced a second learning mechanism, which, by taking the explicit boundaries of a given unit into account, is more sensitive to the positions of events within segmented groups and, consequently, extracts and generalizes the grammar-like positional regularities (i.e. positional rules). This separate mechanism for rule-extraction has also been investigated in the framework of movement streams. Endress and Wood (2011), in their second experiment, showed that participants can learn that different classes of movements tend to appear at different positions regularly, and once these class-rules are extracted, the participants choose the rule-compatible movement-groups over the statistical ones as correct groups. However, the presence of this rule-abstraction mechanism is restricted to conditions in which the stream of movements is already segmented by cues. Under these circumstances, people seem to encode the positions of the movements within the cued groups. Nevertheless, there still exists no evidence showing how these positional codes form the final hierarchy of action representations.

Here, we suggest that the metrical structure of a movement sequence facilitates the organization of movement representations in a hierarchical structure. Generally, auditory exposure seems to be an important factor in the formation of sequential learning abilities (Conway, Pisoni, Anaya, Karpicke, & Henning, 2011); therefore, well-known temporal structures in the auditory stimuli, such as metre, may also play a role in domain-general sequence-learning abilities. A metrical structure is a subjective experience in which the stream of temporal events is perceived to have a hierarchically organized pattern of accentuations (Lerdahl, & Jackendoff, 1981) that may or may not be based on objective features of the event sequence (Bolton, 1894; Woodrow, 1909). It has been indicated that once a sequence has temporally equal groups (e.g. Hoch, Tyler, & Tillmann, 2012) and a pattern of regularly occurring accents (i.e. phenomenal accents) a subjective sense of metre emerges (Ellis, & Jones, 2009; Povel, 1984). Both of these criteria are present in the familiarization movie used in Endress and Wood’s (2011) paradigm for measuring learning of positional regularities. In their second experiment, they presented their participants with a familiarization movie made up of an actor performing a sequence of movements. After performing each of three movements, the actor remained in a neutral standing position to help the participants group the movements. The presented sequence of movements not only had statistical structure (i.e. movements usually followed each other in a similar fashion) but also followed positional rules—that is, the first and last movements of each movement-triplet were always presented in their fixed position during the familiarization. After watching the familiarization movie, the participants had to choose between movement-triplets that conformed to the positional rules and movement-triplets that conformed to the statistical structure. Their results indicated a preference for positional rule-conforming triplets. Here, we argue that their familiarization movie could have given rise to a sense of metre in their participants by fulfilling the aforementioned two criteria for the emergence of a subjective sense of metre. First, in the familiarization phase of their experiments, these researchers used movement-triplets, fulfilling the first criteria by yielding equal lengths for each movement group, and, regarding the second criteria, we hypothesize that the neutral positions, in which the actor remained motionless to group the sequence, could have been a type of accentuation. The possible presence of both criteria raises the possibility of perception of a metrical structure which, as a consequence, could have affected participants’ learning of positional regularities and their performance in the test phase. Thus, the metrical patterns are shown to affect the representation and perception of other hierarchical structures that already exist in the sequential stimuli (Schmidt-Kassow, & Kotz, 2008). For instance, learning grammatical structures in linguistic streams are shown to be influenced by perceived metrical structures (Selchenkova, Francois, et al., 2014; Selchenkova, Jones, & Tillmann, 2014). However, no previous studies have investigated the capability of metrical structure existing in movement sequences (or any other type of sequence) to influence the learning of positional regularities.

If our aforementioned argument, that is, the emergence of a perceived metrical structure in regularly segmented movement sequences, turns out to be true, then the pattern of positional saliency at each position of the movement groups should also correspond with the metrical accentuation patterns. In metrically perceived groups, the first event of each group is always perceived as the most prominent (Brochard,
Abecasis, Potter, Ragot, & Drake, 2003), while the last is usually expected to have low prominence compared to the other events in the group in double, triple, and quadruple metrical patterns. However, many studies, investigating the saliency patterns of continuous action and event streams, have reported an equally high saliency for both the first and last positions of perceptual groups. For instance, as stated in Event Segmentation Theory (Zacks, Speer, Swallow, Braver, & Reynolds, 2007), participants are more likely to detect changes at the event boundaries (Baker, & Levin, 2015) as a result of enhanced prediction errors. However, it was shown that these predictions are not a default process in the perception of events, but are resource intensive and need attentional resources; hence, there remains an open question of why such predictions take place at edges and not at the other positions of a sequence (see Hymel, Levin, & Baker, 2016).

Moreover, regarding the statistical organization of a sequence, Hochmann et al. (2016) showed that, for young infants, the perceptual edges and more specifically, the first two and last two syllables are equally the most statistically informative positions of a linguistic sequence. Additionally, testing learning of positional regularities of movement sequences, Endress and Wood (2011) used equal-length groups of movements and tested whether their participants learned that specific classes of movements occur at both unit edges (i.e. the first and last positions of groups) or not. In line with previous experiments, they found that compared to a condition in which the to-be-learned classes were in the unit-middles during the familiarization phase, the legal edge-movements, which were occurring at the edge positions and belonged to the previously learned classes of movements for those positions, were recognized more often as correct answers, and, therefore, they concluded that both edges are of equal importance in rule-learning tasks. This is even though they tested both edge positions simultaneously and, based on these results, could not decide if the saliency of the first and last positions was equal or not. The present Experiments 2a and 2b aim to clarify this issue.

By and large, in the current experiments, we hypothesized that the presence of a temporal regularity in the sequences of movements affects the learning of positional rules by inducing a metrically organized representation of the sequence. Experiment 1 aimed to investigate the effect of temporal regularity on positional learning, and to do so, we employed a task similar to that of Hoch et al.’s (2012) and Endress and Wood’s (2011) experiments, and tested whether similar lengths of the movement groups could enhance learning of positional regularities as a rule-learning mechanism. Subsequently, based on the results of the first experiment, we hypothesized that in movement streams that induce a metrical structure, the importance of the first and last positions would conform to metrical saliency patterns. These issues concerning the first and last positions are addressed in Experiments 2a and 2b, respectively. In addition, in all present experiments, learning of positional regularities was measured compared to learning of transitional probabilities between adjacent movements.

**Experiment 1**

In this experiment, we utilized Endress and Wood’s (2011) task for measurement of learning of positional regularities and assessed whether or not this kind of learning is modulated by the temporal regularity of the stimuli. Based on Hoch et al’s (2012) experiment, we assumed that the temporal regularity of movement group lengths could lead to the perception of a metrical structure. According to Endress and Wood (2011), the edges of the movement groups are the most important parts of the groups. Based on this argument, we predicted that the perception of temporally equal groups can produce a pattern of perceptual metrical saliencies that could lead to a higher preference for the positional rules embedded in the sequence.

**Materials and methods**

**Participants**

26 employees of the Railways of the Islamic Republic of Iran (15 females, mean age 28.13, SD 7.50) voluntarily participated in the experiment. Each participant was assigned to one of two experimental conditions in which movement-groups with equal and unequal lengths were presented. All participants were Persian native speakers and had normal or corrected-to-normal vision.

**Apparatus**

An animated sequence of actual recorded movements performed by an actor was made with Poser Pro 2014 software (Smith Micro Software, Inc., Aliso Viejo, CA, USA) and was presented using a custom program written in C# with a screen size of 640 × 480 pixels and a framerate of 29.97 fps. Each movement contained 24 frames that began and ended with a neutral position shown in Fig. 1. The largest deviation from the neutral position was placed in the 12th frame of each movement. Three body parts were involved in the movement-groups, namely, hands, feet, and the rest of body. Each group of movements could contain three or four movements depending on the participant’s experimental condition. The movements were combined into groups as movement sequences without any explicit boundaries between the within-group movements.
Design

Participants were assigned randomly to one of the two experimental conditions. In the equal-length condition, the familiarization movie was solely made up of movement-quadruplets and, thus, had a temporal pattern assumed to induce the perception of metre. The unequal-length condition, on the other hand, had a familiarization movie made up of movement groups that could contain either three or four movements. In this experimental condition, the order of movement-triplets and quadruplets were pseudo-random, so that participants could not perceive a metrical structure due to lack of any temporal regularities in the familiarization movie. Participants in both experimental conditions, then, completed a structurally similar test phase. The comparison between the results of the equal-length with the unequal-length conditions would show the effect of a perceptual metrical structure created through equal length of the movement-groups on the learning of positional regularities.

Procedure

In the familiarization phase, participants were asked to watch a person during an exercise, after which they would have to answer some questions about it. Then, as in Endress and Wood’s (2011) experiments, a sequence of movements was presented for them, in which the actor remained in a neutral position between the groups for 1000 ms, to cue the grouping of the movement sequence. In the equal-length experimental condition, each group contained four movements and had a structure of $A_iXYB_j$. Three movements were possible for each of the four positions in the groups (see Fig. 2). Thereupon, three action frames with the form of $A_i...B_j$ (i.e. $A_1...B_1$, $A_2...B_2$, $A_3...B_3$) could have been produced. The subscripts of A and B movements show the frame to which they belonged; the occurrence of an $A_1$ movement predicted the later appearance of a $B_1$ movement with a 100% probability. Within the $A_i...B_j$ frames, the X and Y movements were randomly chosen from the three possible movements for each. Thus, given that the groups could share the middle movements (i.e. X and Y movements), the combination of the middle movements and the $A_i...B_j$ frames produced a sum of 27 distinct groups. Each movement-frame was repeated 60 times during the movie, making a sequence of 180 groups. Each group lasted 3030 ms, making a whole duration of 12 min and 16 s.

The familiarization movie for the unequal-length experimental condition consisted of groups with either three or four movements with a form of $A_iXB_j$ and $A_iXYB_j$, respectively. The construction of the movement-triplets was the same as the movement-quadruplets, having one of the three possible frames. To balance the number of groups between the equal-length and the unequal-length conditions, in movement-quadruplets, each X movement could be followed by
only two of the Y movements, making a total of 27 possible groups.

In the unequal-length condition, as in the equal one, each movement frame was presented 60 times, 30 times of which the groups had three movements. The duration of each movement-triplet was 2.10 s, and the quadruplets had a length of 3.03 s, making a total length of 11 min 08 s 20 ms for the familiarization movie. The movement groups were ordered pseudo-randomly with two constraints: no more than three actions in a row could share the same length, and the consecutive groups could not share a movement. The transitional probabilities between A and B movements were 1, while each B could be followed by only two other A movements (TPs = 0.5). Both A-X and A-Y TPs were 0.33. The time-course of both familiarization movies and the formation of the movement groups are depicted in Fig. 3.

In the test phase, participants were told that they were going to watch pairs of movement-quadruplets, and after watching each pair they would have to choose an action that is more likely a part of the exercise they saw during the familiarization movie; however, they were kept ignorant to the structural differences between the two movement sequences presented as a pair. The test was made of a series of two-alternative forced choice (2AFC) tasks, in which participants had to choose between a pair of movement sequences (see Fig. 4). Each pair consisted of two different kinds of movement-groups, namely, class-groups and part-groups. Class-groups were made of four consequent movements the first and the last of which had a form of $A_1...B_j$. The non-adjacent transitional probability between A and B was zero, but the movements had appeared in these positions during the familiarization movie, and hence, conformed to the positional regularities existing within each movement group in the familiarization phase. Thus, choosing class-groups, in each pair, as the most plausible sequence to be a part of the familiarization movie would show that the participants had learned the positional regularities presented earlier. Part-groups, on the other hand, were consistent with the statistical pattern in the movie but their movements were not legal for their positions (e.g. $YB_1A_jX$, see Fig. 4) in these movement groups, the ordering of the movements was identical to what the participants had seen in the familiarization phase and, hence, the statistical relationships remained intact. However, the position of each movement regarding the edges of group was new to participants. In each pair of movement groups, class-quadruplets and part-quadruplets were separated by 1000 ms of actor’s neutral standing position.

The structure of the test phase for both experimental conditions was identical. That is, for each of the equal- and unequal-length experimental conditions, 12 test pairs were presented, half of which started with part-groups and another half with the class ones, to enable the comparison between the learning of statistical and positional regularities within each condition. The sole difference between conditions in this phase was due to the difference between their movement group-lengths during the familiarization phase. In other words, in the test phase of the equal-length condition, which had a familiarization movie made up of only movement-quadruplets, both class- and part-groups contained four movements. The part-quadruplets could have one of the forms of $XYB_1A_j$, $YB_1A_jX$, or $B_1A_jXY$. Each of these forms was used four times during the test phase. The class-quadruplets in the equal-length condition had a form of...
of movement-groups for both experimental conditions. The written times on the time axes mark the onset of the first frame of the movements. Each movement was given a specific title (i.e. A$_1$, A$_2$, etc.), and the positions of the movements followed a regular pattern in both familiarization movies. Movement frames of the groups were identical across both experimental conditions—that is, the first and last movements of each group were always bound together and had one of the forms of A$_1$…B$_1$, A$_2$…B$_2$, or A$_3$…B$_3$ (the movements in each movement frame are shown with a different color). In the equal-length condition, there were two movements (i.e. X and Y) in the middle of the movement frames. Each of the middle movements was chosen from X and Y classes of movements, respectively, to make the movement-quadruplets with a duration of 3030 ms. In the unequal-length condition, on the other hand, groups could be either quadruplets or triplets, in a random manner. In the movement-triplets, there was only one movement in between, randomly chosen from the class of X movements. Thus, these groups had a duration of 2272.5 ms. The movement groups, in both conditions, were randomly ordered. Hence, in the unequal-length condition, after each movement group, the possibility of the next group to be a triplet or a quadruplet was 50%. The actor remained in a neutral standing state between the groups for 1000 ms. The regular duration of the movement groups in the equal-length condition was hypothesized to induce a sense of metre in participants, and, consequently, improve their learning of positional regularities.

In the unequal-length condition, movement-groups in each of the test pairs were identical in their length. That is, they both had either three or four movements. The part-quadruplets were the same as those in the equal-length condition, and the part-triplets had one of the forms of XB$_i$A$_j$ or B$_i$A$_j$X. Each of these forms was used three times in the test phase. The class-quadruplets had a form of A$_i$X’Y’B$_j$, where X’ had been either A or B during the familiarization phase. Each of the six possible frames was presented twice: one time in a movement-quadruplet and another time in a movement-triplet.

**Results and discussion**

In the equal-length experimental condition, each preference for class-quadruplets over part-quadruplets was considered as a correct answer. Running a two-tailed $t$ test with a chance
level of 50% and a significance threshold of 0.05, participants showed a significant preference for class-quadruplets over part-quadruplets (M 0.66, SD 0.18), \( t (14) = 3.47, p < .01 \), Cohen’s \( d = 0.89 \). Thus, familiarized with movement-groups of similar length, they learned the positional regularities of the movements in each group.

In the unequal-length experimental condition, on the other hand, there were no significant preferences for the class-groups or the part-groups (M 0.46, SD 0.16), \( t (14) = -0.73, p > .05 \), Cohen’s \( d = -0.18 \), ns. The performance of the participants in each experimental condition is demonstrated in Fig. 5. These results show that a segmented movie of human actions with unequal-lengths disrupts both of the class-groups did not appear in the middle positions in the familiarization and were actually As and Bs in the original sequence. Part-groups, on the other hand, measuring the statistical learning of the familiarization movie, were, in fact, part of the original sequence and, thus, maintained the statistical structure of the movie. Faded movements, depicted above, illustrate the original familiarization sequence from which the part-groups are extracted. The positions of the movements in these groups, however, were all wrong. That is, they were never located in these positions within the groups during the familiarization, and the edge movements did not conform to the positional regularities of the original sequence. The test phase of the two experimental conditions differed with regards to the length of the groups in the test pairs. In other words, the test pairs of the equal-length condition always consisted of two movement-quadruplets, while the pairs of the unequal-length condition could consist either of two movement-triplets or movement-quadruplets.
interaction between the two independent variables, \( F(1, 28) = 7.98, p < .01 \) (see Fig. 6). That is, the equality of the movement-groups in the familiarization sequence caused a significant difference between the preference for the class-groups, compared to the part-groups. Nonetheless, neither the equality of groups, \( F(1, 28) = 0.17, p > .05, \) ns, nor the group types, \( F(1, 28) = 3.77, p > .05, \) ns, had a significant main effect. Thus, the equality of the groups in the familiarization phase did not affect the overall performance of the participants, and, interestingly, the group types in the test phase did not yield a significant influence on the performance of the participants. These findings are in contrast with that of Endress and Wood (2011).

Additionally, given the variable group-lengths used in the second condition of the experiment, it was possible to compare the influence of the group-lengths with the group types. Therefore, in the unequal-length condition, a two-way repeated measures ANOVA with group-lengths (movement-triplets vs. movement-quadruplets) and group types (class-groups vs. part-groups) as the independent variables showed a significant main effect of group-lengths \( F(1, 14) = 4.41, p < .05 \). That is, the movement-triplets, compared to the movement-quadruplets, cast a stronger effect on the within-group positional saliencies. Nevertheless, there has been no significant main effect of the group types, \( F(1, 14) = 0.06, p > .05, \) ns, and the interaction between group-lengths and group types was also non-significant, \( F(1, 14) = 2.73, p > .05, \) ns. Thus, with a familiarization sequence of variable group lengths, the type of the test groups had no influence on the participants’ preference, and, in the test phase, the movement-triplets and the movement-quadruplets did not influence the participants’ performance.

The findings of Experiment 1 suggest that while movement-groups of equal length result in the learning of the class-rules, with unequal group-lengths, the participants failed to learn the class-rules as well as the statistical structure. Our results also show that the temporal regularity of the sequence causes a perceptual preference for the items occurring at the edge-positions. Nonetheless, based on these data, we could not decide whether both edges of a group are of equal importance in the learning of positional regularities. Considering the assumption that the perception of an underlying metrical structure is the reason for the effect, we hypothesized that, the saliencies of different positions within the sequential groups are in accordance with the corresponding metrical pattern of the sequence, and thereby produce a relatively different positional learning at each position. That is, in the movement-quadruplets presented in the familiarization phase, based on the binary metrical saliency patterns, the movement presented at the first position would be more salient than the one in the fourth (i.e. last) position. Accordingly, in Experiments 2a and 2b, we tested the recognition of movement-quadruplets in which either their first or their last movement conformed to the class-rules presented in the familiarization movie. In Experiment 2a, the preference for the statistical structure over groups with legal first movements was assessed.
Experiment 2a

In this experiment, based on the results of the previous experiment, we proposed that the temporal regularity of the equal-length movement groups could also have given rise to the perception of a binary metrical structure. To test this hypothesis, in accordance with the saliency patterns of binary metrical structures (i.e. strong–weak–strong–weak), we tested whether the movement-quadruplets with only a legal first movement are enough for participants to recognize them as rule-conforming answers, due to the strong saliency of the first positions of each movement-quadruplet resulted from a perceived metrical structure. That being the case, there should not be a significant difference between the results of this experiment and the equal-length condition of Experiment 1, in which participants chose movement groups with both legal first and last movements. However, if the last movement was as important as the first one, participants would not choose these kinds of groups, being only half-correct, as correct answers.

Materials and methods

Participants

15 new employees of the Railways of the Islamic Republic of Iran (9 females, mean age 27.93, SD 9.49) voluntarily took part in the experiment. All participants were Persian native speakers and had normal or corrected-to-normal vision.

Apparatus and procedure

The familiarization and test phase were identical to those of the equal-length condition in Experiment 1, except that in the test phase, instead of class-quadruplets, which tested for the positional learning of both first and last movements of a group, the quadruplets used here had a semi-class formation with a structure of A,X’Y’B’. Unlike the class-quadruplets, tested for the positional learning of only first (or last, in Experiment 2b) movements—that is, only the first movement of each group conformed to the positional rules established during the familiarization. None of the X, Y, and B movements used in the test pairs were presented at these positions during the familiarization phase and were in fact chosen from one of the remaining three movement categories, and hence, the only legal movement in these actions was their first movement. Each of the three possible semi-class-quadruplets was presented four times during the test phase.

Results and discussion

A one-sample t test revealed that participants preferred the semi-class-quadruplets over the part ones at more than a chance level of 50% (M 0.57, SD 0.13), \( t(14) = 2.07, p < .05 \), Cohen’s \( d = 0.53 \). Thus, participants were able to recognize positional regularities, even when only the first movement conformed to these regularities, while the last one violated them. The results demonstrate the importance of the first movement in abstracting the positional rules of a movement sequence.

An independent samples t-test between these findings (M 0.57, SD 0.13) and that of the equal-length condition in Experiment 1 (M 0.66, SD 0.18) showed no significant difference, \( t(26) = 1.49, p > .05 \), Cohen’s \( d = 0.58 \). Levene’s test indicated unequal variances between the experiments, \( F = 1.22, p > .05 \), so degrees of freedom were adjusted from 28 to 26. The performance of each participant compared with the results of other experiments is shown in Fig. 7. These results suggest that the first movements were sufficient for the participants to choose semi-class-quadruplets over the part ones, which implies the importance

![Fig. 7 Comparison between the results of Experiments 1, 2a, and 2b. The average performance of each participant is shown as a dot and the diamonds indicate the average of the experimental conditions compared to the chance level of 50% (i.e. the dotted line). In all the indicated experimental conditions, participants were presented with the same familiarization movie. The results show that while the movement groups with first legal movement were equally chosen as those with both first and last legal movements, the movement groups with only a final legal movement were chosen less, compared to the transitional probabilities (TPs), and thus, participants in Experiment 2b preferred the statistical units, instead](image-url)
of the first movement in learning positional regularities. In what follows, to maintain the number of questions in the test phase, we executed a third experiment with an identical familiarization movie, in which the only difference was in the test phase. Experiment 2b tested the preference for movement-quadruplets with legal last movements over the statistical groups.

**Experiment 2b**

**Materials and methods**

**Participants**

15 new volunteers from employees of the Railways of the Islamic Republic of Iran (7 females, mean age 29.20, SD 10.64) took part in the experiment. All participants were Persian native speakers and had normal or corrected-to-normal vision.

**Apparatus and procedure**

In this experiment, both familiarization and test phases were identical to Experiment 2a, with an exception that the semi-class-quadruplets in the test phase had the form of $A'X'Y'B_1$.

**Results and discussion**

The selection of the semi-class-quadruplets in each pair was counted as a correct answer. The findings indicated that the participants chose part-quadruplets over semi-class-quadruplets at more than the chance level of 50% (M 0.41, SD 0.12), $t (14) = −2.77, p < .01$, Cohen’s $d = −0.71$. Accordingly, the results reflect that, in contrast with the equal-length experimental condition of Experiment 1 and with Experiment 2a, the semi-class-groups in which only their last movements conformed to the regularities in the familiarization phase were not identified by participants as correct answers. Thereby, we conclude that the last position in the movement-quadruplets is less important in learning positional regularities. This outcome contrasts with the recency effect proposed in Gervain and Endress (2017), which showed that the last event of a sequence (a sequence of artificial language in their case) was more important than the other parts of the sequence.

Comparing this experiment and Experiment 2a also showed that the participants chose the rule-conforming first positions (M 0.57, SD 0.13) more than the last positions (M 0.41, SD 0.12) in the movement-quadruplets, $t (27) = 3.38$, $p < .01$, Cohen’s $d = 1.29$ (see Fig. 7). Since the variances between the two experiments were unequal according to Levene’s test ($F = 0.34, p > .05$), degrees of freedom were adjusted from 28 to 27. The difference between these two experiments implies different saliencies between the first and last positions in learning the regularities. In the general discussion, we further explain these findings and the possible explanations for the contradictory recency effect reported by Gervain and Endress (2017).

**General discussion**

In the current paper, we tried to clarify the possible influences of the perceptual saliencies resulting from the perception of a metrical structure on learning structural regularities within a sequence of human movements. Experiment 1 revealed that the rapid learning of class-rules seen in Endress and Wood’s (2011) experiments might have been for the regularities of the movement group lengths, as we showed that the group types (i.e. class-groups and part-groups) alone do not induce this kind of learning, rather, they only affect performance when the familiarization sequence comprises equal groups, and such learning is significantly reduced when the movement-units have variable lengths within the sequence. However, this reduction did not result in a substantial preference for the statistical units. Thus, we concluded that the perceived metrical structure produced by the equality of movement group lengths could have influenced the hierarchical representation of the movements, and hence, increased the positional learning of class-rules. To further investigate this possible influence, Experiments 2a and 2b examined whether the saliency of the first and last movements conformed to the metrical patterns. Given the fact that each movement-group consisted of four movements, the pattern of metrical saliencies within each group would have been like strong–weak–strong–weak, in accordance with the binary metric patterns. Thus, we expected first movements to be more salient than the last ones. Accordingly, our results indicated a significant difference between recognition of the movement groups with first and last legal movements. It was shown that compared to a condition in which both edges conform to the class-rules, groups with a first legal movement did not make any difference, whereas those with a last legal movement significantly reduced the positional learning.

Nonetheless, as Endress and Bonatti (2007) have shown, with longer familiarization phases, the statistical structure of the sequence tends to overshadow learning of the class-rules. Similarly, the results seen in the unequal-length condition in Experiment 1 indicated a significant main effect of group-lengths on the learning of class-rules, with a strengthening effect of the shorter groups (i.e. movement-triplets) on learning. However, it should be noted that this length-effect did not differ across the statistical and positional units (i.e. part-groups and class-groups, respectively), illustrating that
the group-lengths affected both the statistical and positional learnings to the same degree, which rules out the possibility of a longer duration to produce a preference for the statistical structure. Accordingly, in the equal-length condition, a longer familiarization phase was not able to reduce learning of positional regularities in favour of the TPs. In addition, we indicated that in Experiment 2b, with an identical familiarization phase, participants chose the statistical units significantly more. These results show that the length of the movement sequence was enough for learning of the statistical structure. Nevertheless, there remains an open question about possible changes in the results with longer familiarization phases.

In the current paper, we have suggested that the temporal regularity of a movement sequence creates a sense of metre by having the essential elements of a metrical structure. That is, the regular occurrence of the movements could potentially be perceived as beats and the equal durations between the movements working as the periodicity element of the metre. Moreover, as suggested by Lerdahl and Jackendoff (1981), the metrical accents could be inferred from the accentuation patterns present in a sequence. Here it is implied that in the sequences used in learning positional rules, which are usually segmented explicitly, the segmentation signals, in which the actor remains in a neutral position in the current study, could hypothetically count as a type of accentuation put on the first movement of each group, which in turn would cause a metrical representation of the sequence. In the presented experiments, we tried to support this implication by showing that in the unequal-length condition of Experiment 1, the unpredictability of the sequence with regards to its accentuation patterns (i.e. the placement of the neutral positions between the groups of varying lengths) results in a worsened preference for the positional rule-based groups of movements that could theoretically be the consequence of the ambiguity of the temporal structure in the sequence. Given the difference between the grouping of a sequence and its metrical structure (Lerdahl, & Jackendoff, 1981), and the results of the current experiments showing the influence of the metrical structure, it would be necessary to investigate the possible interactions between the grouping of the movement sequences and their metrical structure. Notably, it would be relevant to study the two aforementioned structures being in-phase or out-of-phase with each other in the sequences of human movements.

An alternative underlying mechanism for the different representational saliencies of the sequential positions found in the current study could be the entrainment of the internal oscillators to the temporal regularity of the sequence. According to the Dynamic Attending Theory (DAT; Large, & Jones, 1999), sequential stimuli with a regular temporal pattern are capable of creating attentional cycles through the entrainment, resulting in recurring attentional peaks synchronized with the onsets of the external stimuli. Considering the findings of Schmidt-Kassow and Kotz (2008), which show that when faced with a grammatical sequence, participants process syntax only in attentional peaks instead of the most predictive points of the sequence, it would be possible to attribute the enhanced learning of the class-rules at the first positions to the synchronization of the attentional peaks to these specific positions. Such synchronization would lead to an expectancy for the moments in time, in which the first movement of each group would occur, and the difference between the magnitude of expectancies for each position would bring about the metrical representation of a sequence (Palmer, & Krumhansl, 1990). It could be argued, however, that the effect of the representation of metrical structure found here is not necessarily due to the entrainment of internal oscillators. Studies addressing internal oscillator entrainment mechanisms usually have the participants familiarized with sequences of events with regular short inter-onset intervals (IOIs). Then, their internal entrainment is probed by presenting events that perturb the already established stimulus onset patterns (e.g. Barnes, & Jones, 2000; Large, & Jones, 1999). Although the experiments presented here do not directly address the entrainment mechanisms as in studies like that of Barnes and Jones (2000), we suggest that the observed influences could have resulted from the entrainment mechanisms activated through the perceived metrical structure of the stimuli. Moreover, regarding the duration of inter-onset intervals (IOI), taking the first frame of each movement as an event, the IOIs used in all of our experiments were 757.5 ms which is near the 600–700 ms IOI that is the range of the most comfortable IOI for reproducing intervals (i.e. ‘spontaneous tempo’; Tinker, Fraisse, & Leith, 1965), is perceived to be neither too long nor too short (i.e. ‘indifference interval’; Wundt, 1911), and, thus, is the range of IOI in which people optimally perceive a given rhythmic structure. This optimality of the IOIs in our stimuli raises the possibility of an entrainment mechanism involved in the perception of the rhythm in the movements. Nevertheless, owing to the findings that have indicated the optimal tempo for rhythm perception to be less than 600 ms (e.g. 500 ms as proposed by Moelants, 2002), the results of the current study are not enough to support such explanations, and more psychological and neuroscientific studies are needed to assess the entrainment mechanism proposed here more directly and to study whether or not the presented results have roots in entrainment.

Considering that learning of positional regularities has its roots in serial short-term memory (Endress, & Bonatti, 2016; Endress, & Mehler, 2009b), studies of this kind of memory provide more evidence for the perception of a metrical structure through grouping of sequences. It has been indicated that grouping of serial stimuli requires encoding of time to prevent confusion of one’s memories.
(Brown, Neath, & Chater, 2007; Farrell, Wise, & Lelièvre, 2011), which in turn might also increase the influence of perceived temporal structures, such as metrical ones, on perception of these stimuli. The unique saliency of each position within the metrical groups could also provide a reason for the findings that grouping reduces transposition errors (i.e. recall of a correct item in a wrong position) during recall of a sequence, which is in line with findings that rhythmic temporal grouping reduces order errors in auditory sequences (Ryan, 1969), and raises the argument of a perceived metrical structure in grouped sequences. Additionally, grouping a sequence also increases interposition errors (i.e. recall of an item in its correct position in a wrong group), which might show that a separate mechanism is responsible for encoding of within-group positions when the sequence is segmented into regular groups.

What we found in Experiments 2a and 2b is also compatible with studies about the Hebb effect, which is the enhancement of recall by performing a task in separate sessions (Hebb, 1961). Schwartz and Bryden (1971) have shown that the Hebb effect disappears by elimination of the first two items of a list, while change or elimination of the last two items does not reduce the effect. These findings show that even in memorizing lists in different sessions, grouping of the sequence could result in the perception of a metrical structure that might determine the saliency patterns of items in a list. This argument becomes more probable knowing that a change in temporal grouping of items in a list would reduce or eliminate learning (Hitch, Flude, & Burgess, 2009). Therefore, based on our results, we suggest that regular grouping of a sequence creates a perceived metrical structure which gives each position a unique degree of saliency.

The results of Experiment 2b contrast with the findings of Endress and Wood’s (2011) fourth experiment which showed an equal saliency for both edges of the movement groups. Nevertheless, this contradiction becomes clear knowing that these researchers measured the saliency of both positions simultaneously; hence, it might have been the first position upon which the participants decided whether a group was rule-compliant. Furthermore, inasmuch as there is an expectation of an unimportant event occurring at the final group position in stimuli with metrical structure, we argue that the observed significance of these positions in various studies (e.g. Gervain, & Endress, 2017) might have been due to the unexpectedness of the methodological accentuations on these positions in their experimental paradigms. That is, the repetitions and lengthening of the stimuli at the final positions might have been perceived as accentuations presented at the least expected position of the sequences, and therefore, they might have attracted the attention of participants. This phenomenon is called syncopation in musical sequences and is defined as the accentuation of a less metrically salient position in a sequence (Scholes, & Ward, 1970).

Additionally, as explained in the introduction, studies about the grouping mechanisms of movements and events have also shown the role of the saliencies of group-bounaries on the segmentation of sequences. For example, Baker and Levin (2015) showed that predictions about the next movements do not occur during the whole stream of events; rather these predictions only take place at unit edges. Henceforth, it is conceivable that the perception of a metrical structure also plays a role in the segmentation of the event streams. As can be seen in the statistical segmentation of artificial linguistic sequences, temporal regularities of the unit lengths affect on the segmentation process (Hoch et al., 2012). Thus, based on our results, we propose that the temporal regularity of movement streams not only facilitates learning of positional regularities and the segmentation of the sequences, but could also be capable of creating metrical perceptual saliences.

### Conclusion

In our experiments, we have tried to understand how positional encodings could help the hierarchical representation of human movements, and, by suggesting the role of the perceived metrical structure in this kind of representation, we seem to be closer to the solution of some questions about the learning of positional regularities. First, regarding the necessity of grouping cues (i.e. pauses) for the learning of positional rules, we proposed that by grouping the continuous stream, these cues bring about the perception of a metrical structure which has a unique saliency pattern differentiating the within-group positions from each other. We propose that this function of the perceived metrical structure could also provide a reason for the shorter time course required for the positional learning compared to the statistical one, for it has been shown that the perception of a metrical structure enhances processing of stimuli (Bausenhart, Rolke, & Ulrich, 2007). This finding could explain the faster learning of positional information in exposure to temporally grouped sequences. Together, our findings introduce a new explanation for the hierarchical representation of regularly segmented action sequences, by demonstrating the effects of the perceived metrical saliencies on the learning of positional regularities.

**Acknowledgements** We would like to acknowledge II’ya Suhov for his help in programming the software of the experiments.
Compliance with ethical standards

Conflict of interest All authors declare that they have no conflicts of interest. We agree to make raw data available if requested.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References


