Acta Psychologica 164 (2016) 27–33

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Multimodal sequence learning

Ferenc Kemény [a,b,c,*], Beat Meier [c,d]

[a] Institute of Psychology, University of Graz, Austria
[b] Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary
[c] Institute of Psychology, University of Bern, Bern, Switzerland
[d] Center for Cognition, Learning and Memory, University of Bern, Bern, Switzerland

ARTICLE INFO

Article history:
Received 30 April 2015
Received in revised form 27 October 2015
Accepted 29 October 2015
Available online xxxx

Keywords:
Task Sequence Learning
Modality independence
Correlated sequences
Implicit sequence learning

ABSTRACT

While sequence learning research models complex phenomena, previous studies have mostly focused on unimodal sequences. The goal of the current experiment is to put implicit sequence learning into a multimodal context: to test whether it can operate across different modalities. We used the Task Sequence Learning paradigm to test whether sequence learning varies across modalities, and whether participants are able to learn multimodal sequences. Our results show that implicit sequence learning is very similar regardless of the source modality. However, the presence of correlated task and response sequences was required for learning to take place. The experiment provides new evidence for implicit sequence learning of abstract conceptual representations. In general, the results suggest that correlated sequences are necessary for implicit sequence learning to occur. Moreover, they show that elements from different modalities can be automatically integrated into one unitary multimodal sequence.

In everyday life we are required to respond to sequentially organized stimuli and our daily routines involve ordered sequences of tasks and actions. The ability to acquire and use knowledge involving structured sequences of events and actions is fundamental to adaptive behavior. Sequence learning is involved in tasks such as speaking and writing, driving, preparing meals, performing sports and music, and far more. These activities typically involve the integration of information from different modalities such as visual and auditory. Although such learning is usually goal-driven and perfected through deliberate practice, it can happen incidentally and unintentionally. Sometimes we are not even aware that learning has taken place. In the laboratory, this kind of learning is termed *implicit* and is typically assessed using a serial reaction time task (SRTT; Nissen & Bullemer, 1987). In this paradigm, a visual stimulus is presented at one of several horizontally aligned locations on a computer monitor, and participants respond by pressing keys that correspond directly to the locations. Unbeknownst to them, the order of locations (and thereby the order of required key press responses) is determined by a repeating pattern, or sequence. With practice, response times decrease. However, when the sequence is replaced by a random order, response times increase again substantially. This increase in response times is taken as indirect evidence of implicit sequence learning. Subsequent assessment of sequence awareness often reveals that knowledge of the sequence is implicit rather than explicit. The purpose of the present study is to investigate the role of auditory as well as visual stimuli in *implicit sequence learning* and the potential integration of information in the different modalities. To this end, we employed a Task Sequence Learning paradigm, as described below.

There is ample evidence that different surface features can form the basis of learning in the SRTT (such as effector-based information Deroost, Zeeuws, & Soetens, 2006; perceptual information Remillard, 2003; or response-based information Willingham, Wells, Farrell, & Stemwedel, 2000; for a detailed summary see Kemény & Lukács, 2011). There is, however, less evidence for the learning of abstract sequences. Goschke and Bolte (2007) tested participants in an object naming task, in which the underlying semantic categories were sequenced. Results showed faster reaction times with sequenced as opposed to random organization in the categories. On the other hand, neither Dominey and colleagues (Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998), nor Pacton and colleagues (Pacton, Perruchet, Fayol, & Cleeremans, 2001) found evidence of sequence learning at an abstract level (see Abrahamse, Jiménez, Verwey, & Clegg, 2010 for a detailed review).

Experiments focusing on the different sources of sequenced information shed light on the fact that the contribution of these types of information is difficult to contrast. As a possible solution, a *task sequence learning* (TSL) paradigm was introduced (Heuer, Schmidtke, & Kleinsorge, 2001; Koch, 2001). In the standard SRTT, the different streams of information are necessarily correlated (i.e., visual–spatial stimulus positions, eye-movements, motor responses), but in the TSL these streams can be uncoupled and manipulated separately (cf. Cock & Meier, 2007; Meier & Cock, 2010). In the TSL, participants are asked to respond to a series of different intermixed tasks: in the Animals

* Corresponding author at: Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Egry József u. 1, Hungary.
E-mail address: fkemeny@cogsci.bme.hu (F. Kemény).

http://dx.doi.org/10.1016/j.actpsy.2015.10.009
0001-6918/

Task they have to decide whether the presented animal is a mammal or a bird, in the Plants Task they have to decide between trees and flowers, and in the Implements Task, between kitchen utensils and musical instruments (Meier et al., 2013; Meier & Cock, 2010; Weiermann & Meier, 2012b). On each trial, participants are required to respond by pressing one of two keys (the same two keys being used for all three tasks). Hence, the design enables the selective manipulation of the order of responses (1 out of 2 possible responses, organized in a sequenced or pseudo-random order) as well as the order of the tasks (1 out of 3 possible tasks, organized in a sequenced or pseudo random order). Most important for the purpose of the present study, it is also possible to vary the *modality* in which the stimuli are presented (i.e., visual pictures or auditory words), which allows a) a comparison of sequence learning in different modalities as well as b) the testing of multimodal integration of sequences.

Selectively manipulating different streams of repeated sequences in the TSL paradigm has revealed that sequence learning only takes place if there are at least two correlated sequences present (Correlated Sequences Approach by Meier & Cock, 2010; Weiermann, Cock, & Meier, 2010). In fact, previous studies have shown that the kind of information within a sequence did not seem to matter as long as two correlated sequences were present, for example, correlated sequences of tasks and responses, tasks and response mappings, tasks and stimulus locations, stimulus locations and responses, or tasks and task cues (Cock & Meier, 2007, 2013; Meier & Cock, 2012; Meier, Weiermann, & Cock, 2012; Weiermann et al., 2010; Weiermann & Meier, 2012a, 2012b; also note, that perceptual and response sequences are always present in the case of the classical SRTT. For an overview of correlated streams in the SRT task, see Meier & Cock, 2010). Removing either of the sequences led to an increase in reaction times, suggesting that sequence learning took place. However, if only one sequence was present, removal of that sequence did not lead to an RT increase, suggesting no sequence learning in this case (Cock & Meier, 2007; Meier & Cock, 2010; Weiermann et al., 2010; Weiermann & Meier, 2012b).

All the previously mentioned studies have used unimodal stimuli, comprising either a visual or an auditory sequence. In contrast, the current study uses two modalities. We were motivated by the fact that in the real world, we are often exposed to several sequences at once, with each in a different modality. A simple example would be listening to and watching a televised song and dance routine, or following a cookery demonstration, or attending to the visual and auditory patterns of someone speaking a foreign language. In such cases, sequence learning may occur, particularly if the activity is repeated, but it need not be intentional and the person may have no idea that anything has been retained. It would be useful to know whether sequences presented in different modalities are learned in much the same way and to the same degree.

This issue has been partially addressed by previous studies of sequence learning. In the case of simple repeating patterns, Saffran and colleagues showed that infants use similar statistical learning mechanisms across modalities. Similar statistical learning was observed with auditorily presented linguistic (Saffran, Aslin, & Newport, 1996) and non-linguistic stimuli (Saffran, Johnson, Aslin, & Newport, 1999), and with visually presented spatial stimuli (Fiser & Aslin, 2002). On the other hand, Marcus and colleagues (2007) showed that infants only extract simple ABA rules from linguistic and not non-linguistic stimuli. Another study by Saffran and colleagues, however, showed learning in an identical non-linguistic visual setting (Saffran, Pollak, Seibel, & Shkolnik, 2007).

Apart from infant studies, previous results from Artificial Grammar Learning (AGL) showed that adult participants perform better in the case of auditory than in the case of visual or tactile stimuli (Conway & Christiansen, 2005). The difference, however, was not only quantitative, but also qualitative. Another study using probabilistic category learning found no modality-based difference (Kemény & Lukács, 2013). Hence results are not conclusive either in infants or in adults.

So far, only one study has tested *task sequence learning* with not visual, but auditory stimuli (Weiermann & Meier, 2012a). The results showed that implicit sequence learning took place, but only when the order of tasks and the order of left vs. right key press responses were correlated (i.e. when the streams of information could be integrated). In the case of a single sequence being present (either task-based or response-based, with the other order being random and hence uncorrelated), no sequence learning occurred. A comparison with previously published visual data (Meier & Cock, 2010) showed a very similar pattern across experiments, and, importantly, there was no statistical difference between sequence learning in the two modalities (p. 472, Weiermann & Meier, 2012a). This evidence was indirect however. The current study is an extension of Weiermann and Meier (2012a) as it tests sequence learning with auditory stimuli and provides a direct comparison to a visual task with picture stimuli.

The current study also tests the learning of multimodal sequences. Learning multimodal sequences has already been addressed by previous studies using the SRTT and Statistical Learning paradigms. In both paradigms, novel theoretical contributions suggest that learning mechanisms typically take place within modality or dimension boundaries, as independent modality-based learning mechanisms may exist for the different modalities (Frost, Armstrong, Siegelman, & Christiansen, 2015; Goschke & Bolte, 2012). To test whether elements from different modalities can be integrated into a single sequential representation, we added a set of conditions in which the modality of stimulus presentation varied randomly (between visual and auditory items). If sequence learning takes place within modality boundaries, we expect no integration of multimodal stimuli, hence no implicit sequence learning under these circumstances. On the other hand, if multimodal implicit sequence learning were to be found here, then we might be able to conclude that task sequence learning of this kind can indeed take place across modalities.

As stimuli from different modalities tap on the same concepts, learning on the multimodal conditions require abstraction of the stimuli. The question as to whether exposure to sequential information can give rise to the integration of abstract as well as visuo-spatial and motor knowledge may have theoretical implications (Abrahamse et al., 2010; Altmann, Dienes, & Goode, 1995; Dienes & Altmann, 1997; Gomez & Gerken, 2000; Pacton et al., 2001). The main aim of the present study is to address the role of abstract and modality-based information in TSL and to ascertain whether implicit sequence learning is integrated across modalities.

Furthermore, it has been suggested that adding a random stream of information may interfere with sequence learning (Keele, Ivry, Mayr, Hazeltine, & Heuer, 2003) — an explanation that may apply to the lack of single stream learning in previous TSL studies. In the current experiment, modalities change randomly in the multimodal condition. If processing a random stream of information alongside a sequenced stream of information impedes implicit learning of the sequence, we would expect no or reduced sequence learning in the multimodal conditions. Throughout the experiments, we employed a particular TSL paradigm with three different tasks that has been used successfully in previous work (Meier et al., 2013, 2012; Meier & Cock, 2010; Weiermann & Meier, 2012a, 2012b).

1. Method

1.1. Participants & design

A total of 324 people participated in the Experiment (230 female, 94 male, mean age = 24.2 years, $SD = 5.14$, range: 18–41). Participants with known neurological or cognitive deficits were not included in the study. All participants had normal or corrected to normal vision, and all had Hungarian as their native language. They were randomly assigned to one of twelve experimental conditions. The conditions differed in Modality condition (Auditory versus Visual versus Multimodal) and Ordering (Task + Response, Task, Response and Random), resulting in a 3×4 design. Stimuli were color pictures in the Visual Condition, auditorily presented words in the Auditory Condition and randomly changing color pictures and auditory words in the Multimodal Condition. For each Modality condition there were four ordering types. In the Task + Response condition, both tasks and responses followed a predefined 6-element sequence. Only a task sequence was present in the Task condition, only a response sequence was present in the Response condition, and no sequences were present in the Random condition.

1.2. Materials

There were three different tasks: the *Animals* task, the *Implements* task, and the *Plants* task. In each task there were two types of stimuli. For *Animals* there were *Birds* and *Mammals*, for *Implements* there were *Musical instruments* and *Kitchen utensils*, and for *Plants* there were *Vegetables* and *Fruits*. Each category had 15 members, with an auditory (sound file) and a visual stimulus (color picture) for each item. That is, within *Birds* among 14 others there was a picture of an eagle, as well as – among 14 others – the sound clip with the word 'eagle'.

Visual stimuli were 500×400 pixel color photographs appearing in the center of a 640×480 screen. Auditory stimuli were the Hungarian names of the pictures used in the Visual conditions. Duration of each word was between 700 and 800 ms. Words were recorded by a male voice. The volume of all sound files was matched. All sound files were 16 bit mono '.wav' files, with a sample rate of 11 kHz. Stimulus presentation and data collection were done using PSTNet's E-prime (Psychology Software Tools, Pittsburgh, PA).

Each experimental condition consisted of 11 blocks of 90 items. Blocks 1–7, Block 9 and Block 11 served as training blocks. If a task sequence was present (Task + Response and Task conditions), it was replaced with a random order of tasks in Block 8, whereas the response sequence that was present (Task + Response and Response conditions) was broken in Block 10. Table 1 provides the Block design of the Experiment, Fig. 1 illustrates the items.

Each sequence was composed of six elements. Thus, a block with a sequential organization consisted of 15 presentations of that sequence, that is, no stimuli were repeated within a block (there were 15 pictures and words in each category). The computer-implemented experiment randomly assigned one out of two possible sequences to each participant, both for tasks and for responses. Specifically, the sequences *Plant–Animal–Implement–Animal–Plant–Implement* and *Implement–Plant–Animal–Plant–Implement–Animal* were used as the task sequences, and the sequences *R–L–R–R–L–L* and *L–R–R–L–R–L* were used as the response sequences.

Table 1
Conditions and their compositions across blocks regarding the ordering of stimulus elements.

Ordering type	Blocks 1–7	Block 8	Block 9	Block 10	Block 11
Task + Response					
Task	Sequence	Random	Sequence	Sequence	Sequence
Response	Sequence	Sequence	Sequence	Random	Sequence
Task					
Task	Sequence	Random	Sequence	Sequence	Sequence
Response	Random	Random	Random	Random	Random
Response					
Task	Random	Random	Random	Random	Random
Response	Sequence	Sequence	Sequence	Random	Sequence
Random					
Task	Random	Random	Random	Random	Random
Response	Random	Random	Random	Random	Random

1.3. Procedure

Participants were tested individually. After giving consent, they were informed that they were to carry out a computer implemented task about perception that involved making binary decisions such as whether a word or picture was an animal or a plant and so forth. They were instructed to respond by pressing one of two response buttons and were told that there would be several blocks of trials. All participants were required to put on headphones regardless of condition. For the *Animals* task they were asked to respond with the *Left* button to *Mammals* and with the *Right* button to *Birds*. For the *Plants* task, the *Left* button was to be pressed for *Vegetables*, and the *Right* for *Fruits*. In the case of the *Implements* task, participants were asked to respond with the *Left* button for *Kitchen utensils*, and with the *Right* button for *Musical instruments*. The left button was the 'S', and the right button was the 'L' on the keyboard.[1] Participants were instructed to work as fast as possible, and were informed that they would have a short self-paced break between blocks. No other instructions were given and participants were not told about the possible existence of any kind of sequence in the materials or responses. A reminder of the response mapping was present throughout the experiment to guarantee that participants do not forget or swap the response mappings. A short 600-msec-long sound followed each erroneous response in order to keep response accuracy high.

At the end of the experiment, a post-test interview was conducted to test for explicit knowledge of the sequences. First, participants were asked whether they were aware of the presence of the task sequence, then to try to reproduce the sequence, that is, to try to tell how the different tasks followed each other. Second, they were asked whether they were aware of the presence of a response sequence, and then they were asked to reproduce it either from memory or simply by guessing.

1.4. Data analysis

Since we were primarily interested in *implicit learning*, we excluded those participants, who were able to correctly reproduce the actual six-element-sequences to which they had been exposed during the training phase.[2] Data of 25 participants (out of 53) in the Auditory Task + Response, 21 (out of 22) in the Auditory Task, 16 (out of 24) in the Auditory Response, 18 (exclusion not possible) in the Auditory Random, 19 (out of 39) in the Visual Task + Response, 20 (out of 20) in the Visual Task, 17 (out of 19) in the Visual Response, 20 (exclusion not possible) in the Visual Random, 25 (out of 44) of the Multimodal Task + Response, 20 (out of 20) in the Multimodal Task, 17 (out of 23) in the Multimodal Response, and 21 (exclusion not possible) in the Multimodal Random remained for further analysis. Random Participants were also asked to try to reproduce the sequence despite the fact that there was no sequence present. One participant in the Auditory Random, two in the Visual Random, and two in the Multimodal Random conditions provided an answer that was identical to one of the sequences used in the sequenced condition. That is, it is likely that some participants produced a high score on the explicit test just by chance. No participants were excluded from the Random conditions, as these conditions did not include a sequence.

A Hierarchical Loglinear Analysis was conducted to test whether the inclusion rates (labeled Implicit versus Explicit) differ along modalities (Auditory versus Visual versus Multimodal), the presence of the

[1] Note that both keys are located in the center row of Hungarian QWERTZ keyboards. None of the tasks or categories started with the letter 'S' or 'L'.

[2] Based on the posttest we calculated an explicit knowledge score, which was the length of the longest fragment that was reproduced. Due to the nature of the posttest, participants randomly providing a six-element sequence necessarily reach a score of at least 4 (except for six identical elements). E.g. if a participant provides *L–R–L–L–L–L* as a response sequence, there is a four element overlap with both the *R–L–R–R–L–L* and *R–L–L–R–R–L* sequences: *L–L–R–L* are four overlapping consecutive elements with the first, and *L–R–L–L* are four overlapping consecutive elements with the second sequence.

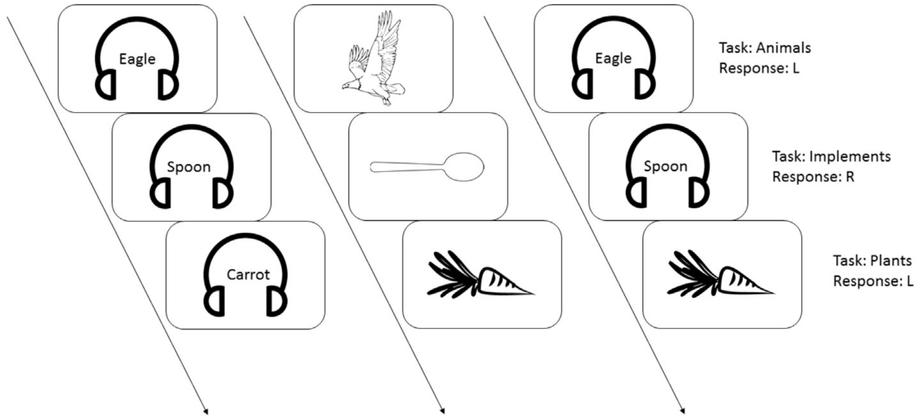

Fig. 1. The procedure of the Experiment. The auditory condition is shown on the left, the Visual in the middle, and the Multimodal on the right.

response sequence (present in Response and Task + Response orderings, absent in Random and Task orderings), and the presence of the task sequence (Task and Task + Response versus Random and Response). The Hierarchical Loglinear Analysis revealed that no three-way or higher order effects were significant, χ^2 (9) = 4.865, p = 0.846. On the other hand, two-way effects were significant, χ^2 (9) = 94.022, $p < 0.001$. Partial associations showed that the distribution of implicit versus explicit learners differed by the presence or absence of a response sequence, χ^2 (1) = 76.871, $p < 0.001$, and by the presence or absence of a task sequence, χ^2 (1) = 12.191, $p < 0.001$. No other two-way interactions were significant (all *ps* > 0.266).

We calculated the median Reaction Times for correct items for each participant in each block. As the task sequence was removed in Block 8 in the Task + Response and Task conditions, and the response sequence was removed in Block 10 in the Task + Response and Response conditions, we calculated two *disruption scores*, one for task sequence and one for response sequence. The disruption scores were computed as the median reaction time of the critical block, minus the average of the median Reaction Times of the surrounding two blocks. Conditions with no sequence served as control conditions. For both *disruption scores*, first separate ANOVAs with Modality condition (Auditory vs Visual vs Multimodal) and Type of Ordering (Task + Response vs Task vs Response vs Random) as between subject variables were conducted. In addition we compared the *disruption scores* in each condition to zero using one-sample T-tests. The calculation and analysis of *disruption scores* is based on previously published studies (e.g. Meier et al., 2012).

2. Results

The overall accuracy of the participants was 96.3%. As these scores are close to ceiling we did not analyze them further. As seen below, results showed that removing either of the sequences resulted in an increased reaction time, but only for the Task + Response conditions, not the single sequence conditions. Reaction times by Block, by Modality condition and by Ordering are presented in the left column of Fig. 2, while the right column provides *disruption scores*.

Block 8 (removal of task sequence if present) and Block 10 (removal of response sequence if present) disruption scores were analyzed separately using a 3 × 4 ANOVA for each *disruption score* with Modality condition (Visual versus Auditory versus Multimodal) and Ordering (Task + Response versus Response versus Task versus Random) as between subject variables.

For Block 8 *disruption scores* (removal of the Task sequence if present), the ANOVA revealed a significant main effect of Ordering, $F(3, 228) = 15.559$, $p < 0.001$, $\eta_p^2 = 0.170$. Neither the main effect of Modality condition, nor the Modality condition x Ordering interaction were significant, both *ps* > 0.626. Post hoc pairwise comparisons revealed that *disruption scores* of the Task + Response condition were significantly higher than *disruption scores* of all other conditions (all *ps* < 0.001), and no other differences were significant (all *ps* > 0.297). The above analysis compared the *disruption scores* across groups, but did not test whether the *disruption scores* are different from zero. Thus we used one-sample T-tests separately for each Modality condition and Ordering, to test whether learning took place at all. In all three modalities, the Block 8 *disruption score* was only significant in the Task + Response condition, $t(24) = 3.717$, $p < 0.01$ for the Auditory, $t(18) = 3.125$, $p < 0.01$ for the Visual, and $t(24) = 3.452$, $p < 0.01$ for the Multimodal modality condition. No other t-tests were significant, all *ps* > 0.145 (> 0.145 in the Multimodal, > 0.223 in the Visual, and >0.737 in the Auditory modality condition).

Disruption scores for Block 10 (removal of the Response sequence) showed a similar pattern as described for Block 8. The ANOVA revealed a significant main effect of Ordering, $F(3, 228) = 9.826$, $p < 0.001$, $\eta_p^2 = 0.114$. Neither the main effect of Modality condition, nor the Modality condition x Ordering interaction were significant, both *ps* > 0.414. Post hoc pairwise comparisons revealed that *disruption scores* of the Task + Response condition were significantly higher than *disruption scores* of all other conditions (all *ps* < 0.001), and no other differences were significant (all *ps* > 0.530). *Disruption scores* by Ordering are presented in the right column of Fig. 2.

One sample T-tests showed that in all three modality conditions, learning only occurred in the Task + Response condition, $t(24) = 2.891$, $p < 0.01$ for the Auditory, $t(18) = 5.496$, $p < 0.001$ for the Visual, and $t(24) = 3.685$, $p < 0.01$ for the Multimodal condition. All other *ps* > 0.163 (>0.163 in the Auditory, >0.251 in the Visual, and >0.552 in the Multimodal condition).

3. Discussion

The current study tested *implicit sequence learning* across modalities with visually presented pictures and auditorily presented words. By the introduction of the Multimodal conditions we also tested the integration of stimuli originating from different modalities. In all cases, sequence learning only took place when a correlated sequence of tasks and responses was present. A single task sequence or a single response sequence did not lead to learning. Patterns were identical across modalities, and are in line with the Correlated Sequences Approach, that is, learning requires at least two correlated streams of sequential information (Cock & Meier, 2007; Meier & Cock, 2010; Weiermann et al., 2010). These results suggest that *implicit task sequence learning* is modality invariant, and takes place in the absence of repeating modality-based information.

Previous studies addressed the possibility that the introduction of a stream of random information can interfere with implicit sequence learning (Keele et al., 2003).The current results provide evidence that *sequence learning* can take place even with concurrent changes in

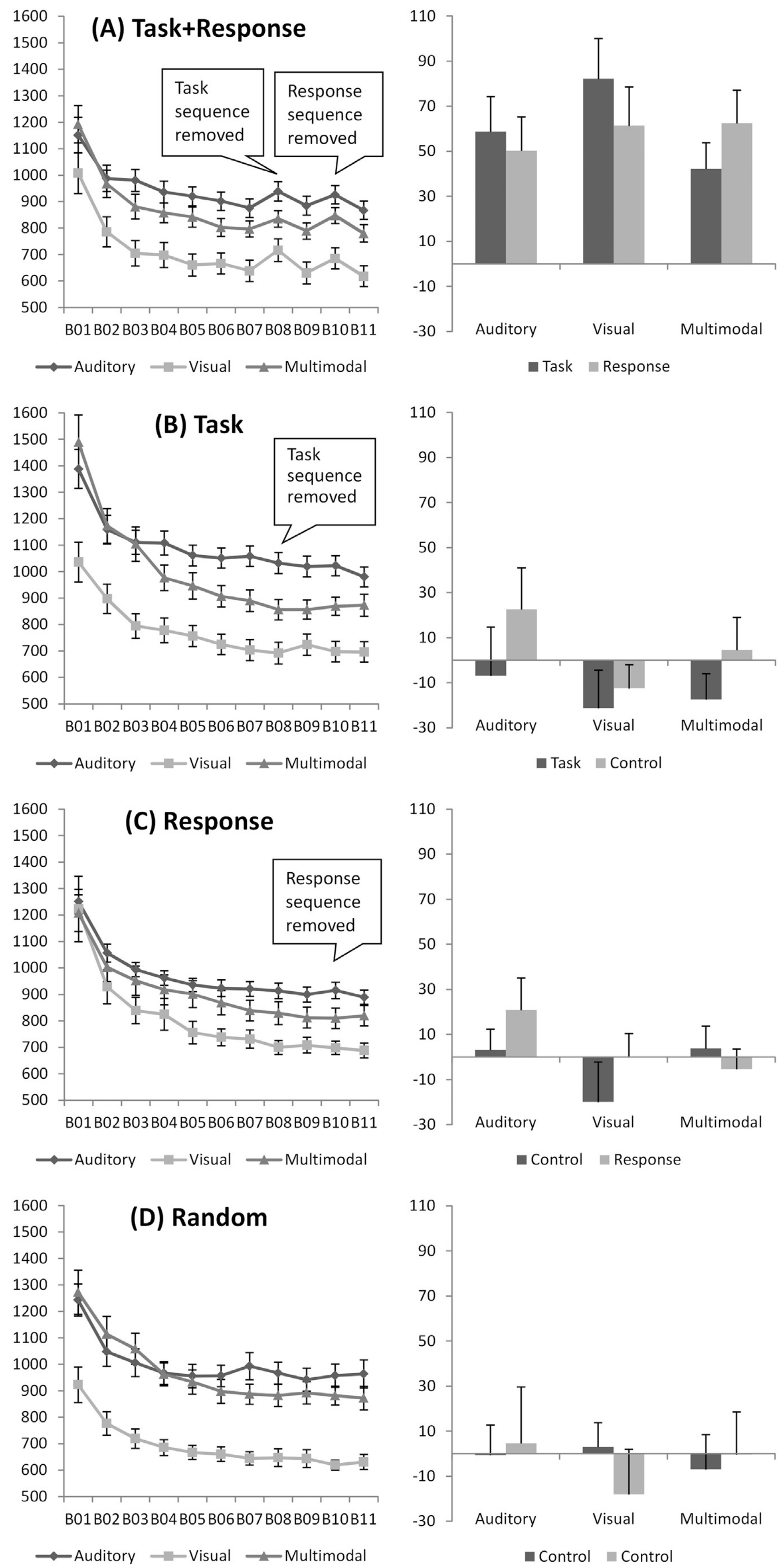

Fig. 2. Reaction times and disruption scores. The left column illustrates reaction times by Block and by Modality condition in the four Types of Ordering. The right column illustrates *disruption scores*. Error bars indicate SEM.

modality happening in random order, but only if correlated task and response sequences are still present. On the other hand, it is a question how it is related to the non-learning of unitary sequences. The lack of no learning in the single sequence conditions could be due to the increased amount of random information. However it is unlikely, since a visual or auditory condition with a single stream of sequenced information probably has the same amount of random factor, as a multimodal condition with correlated task and response sequences. But it is also possible that the relative and not the absolute amount of random information is important. The current study has no clear implications on that. Still, the Correlated Sequences Approach implies that the multiple correlated sequences are required to filter the incoming amount of structured information. That is, the cognitive system is flooded by repeating patterns, but not all of these are processed: only the ones that are correlated. It is a question though, whether correlated sequences can be considered a single integrated sequence with a high amount of predictive information, or whether the cognitive system processes them as separate sequences highlighted by each other.

The central finding of the current experiment is that elements from different modalities can be integrated into a multimodal sequence. While this is the first study on implicit multimodal sequence learning, there are some theoretical approaches to implicit multimodal cognition. A theory that focuses specifically on sequence learning is the “cognitive and neural architectures of sequence representation” (Keele et al., 2003). According to this model there are two parallel pathways underlying the acquisition of sequences. The dorsal pathway works on unimodal uninterpreted stimuli, and results in implicit learning, while the ventral pathway operates on categorized stimuli, can operate across different modalities and domains, and results in implicit or explicit learning. The former is composed of parietal and supplementary motor areas, whereas the latter involves temporal and lateral prefrontal cortical areas.

While the dual pathways hypothesis has a clear implication on the existence of implicit multimodal sequence learning, to our knowledge no previous studies provided empirical support. We have now provided evidence that multimodal implicit sequence learning can take place. It is important to note that sequence elements were abstract, interpreted stimuli (Multimodal Task + Response condition), and previous studies have already shown the existence of abstract sequence learning, but only within modality boundaries (Abrahamse et al., 2010; Goschke & Bolte, 2007). As the characteristics of the sequence fit well with the focus of the ventral stream, this experimental design is a good candidate for future neuroimaging studies.

The current study provides further support to the correlated sequences hypothesis (Meier & Cock, 2010). Conditions with auditory, visual and randomly changing stimuli (Multimodal conditions) showed that sequence learning only occurred in the case of correlated sequences, that is, where a response and a task sequence were present. Unitary task and response sequences led to no learning.

3.1. Modality-dependent statistical learning

The central aim of the current study was to test modality-based differences in sequence learning. This issue has mostly been neglected in the SRTT and TSL, but there have been earlier studies using AGL. While differences in methodology do not allow direct comparisons, it is important to integrate findings across methods. Previous studies showed that infants at the age of 8 months are able to extract and utilize statistical information on how consecutive syllables follow each other (Saffran et al., 1996). The same results were borne out with non-verbal sounds (Saffran et al., 1999), and with spatial configurations (Fiser & Aslin, 2002). A later study of AGL compared visual, auditory and tactile learning in adult participants (Conway & Christiansen, 2005). Results showed severe differences between modalities. There was an overall auditory advantage over the other two modalities, and there were also strategic differences. Participants of the auditory condition showed a strong sensitivity to sequence-final chunks, while no such bias was observed in the other modalities.

In sum, previous studies argue for similar learning across modalities in infants, but a general advantage of audition in adults. Current results, however, showed that TSL performance is independent of the source modality. To integrate these results we have to take into account that AGL and TSL differ in at least two major points: motor relatedness and abstraction. On the one hand, AGL is a motor-free task relying on choice preference methodology, while TSL is a motor-related task that tests reaction times. On the other hand, the AGL studies cited above do not require abstract sequential representations, as sequences of surface elements can describe variance, while in the case of TSL, focus on surface elements on their own is simply not enough for solving the task. To integrate results, it is possible that sequence learning is only independent of modalities, if an abstraction to conceptual level takes place. If abstraction is not required, we can expect modality-based (and stimulus-based) differences. This idea is further supported by a study using a different, non-sequential statistical learning paradigm, the Weather Prediction task, which found that performance is independent of the source modality (Kemény & Lukács, 2013). In this task participants faced different geometric shapes, and had to decide whether the outcome would be sunshine or rain. While focusing on surface elements could be sufficient for solving the task, previous clinical as well as experimental studies showed that high performance requires abstraction and generalization (Gluck, Shohamy, & Myers, 2002; Reber, Knowlton, & Squire, 1996). That is, in accordance with the above assumption, our results suggest that implicit sequence learning can operate on abstract as well as perceptual representations, thereby giving rise to modality-independent learning.

Acknowledgment

During the publication process of this work, the authors were supported by an SNF International Short Visit 157925 award, entitled ‘Multi-modal sequence learning’, made to Beat Meier and Ferenc Kemény. The postdoctoral fellowship of Ferenc Kemény is funded by SciEx NMS-CH fellowship 14.120 ‘MUST — Multimodal Sequences in Task Sequence Learning’. We are grateful for the helpful suggestions of Josephine Cock.

References

Abrahamse, E.L., Jiménez, L., Verwey, W.B., & Clegg, B.A. (2010). Representing serial action and perception. *Psychonomic Bulletin & Review*, *17*(5), 603–623. http://dx.doi.org/10.3758/PBR.17.5.603.

Altmann, G.T.M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 899–912.

Cock, J., & Meier, B. (2007). Incidental task sequence learning: Perceptual rather than conceptual? *Psychological Research*, *71*(2), 140–151. http://dx.doi.org/10.1007/s00426-005-0005-7.

Cock, J., & Meier, B. (2013). Correlation and response relevance in sequence learning. *Psychological Research*, *77*(4), 449–462. http://dx.doi.org/10.1007/s00426-012-0444-x.

Conway, C.M., & Christiansen, M.H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 24–39.

Deroost, N., Zeeuws, I., & Soetens, E. (2006). Effector-dependent and response location learning of probabilistic sequences in serial reaction time tasks. *Experimental Brain Research*, *171*, 469–480.

Dienes, Z., & Altmann, G.T.M. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract? *How implicit is implicit learning?* (pp. 107–123). Oxford, UK: Oxford University Press.

Dominey, P.F., Lelekov, T., Ventre-Dominey, J., & Jeannerod, M. (1998). Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences. *Journal of Cognitive Neuroscience*, *10*(6), 734–751.

Fiser, J., & Aslin, R.N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 15822–15826.

Frost, R., Armstrong, B.C., Siegelman, N., & Christiansen, M.H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. http://dx.doi.org/10.1016/j.tics.2014.12.010.

Gluck, M.A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory*, *9*, 408–418.

Gomez, R.L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178–186.

Goschke, T., & Bolte, A. (2007). Implicit learning of semantic category sequences: Response-independent acquisition of abstract sequential regularities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 394–406.

Goschke, T., & Bolte, A. (2012). On the modularity of implicit sequence learning: independent acquisition of spatial, symbolic, and manual sequences. *Cognitive Psychology*, *65*(2), 284–320. http://dx.doi.org/10.1016/j.cogpsych.2012.04.002.

Heuer, H., Schmidtke, V., & Kleinsorge, T. (2001). Implicit learning of sequences of tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 967–983. http://dx.doi.org/10.1037/0278-7393.27.4.967.

Keele, S.W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, *110*(2), 316–339. http://dx.doi.org/10.1037/0033-295X.110.2.316.

Kemény, F., & Lukács, Á. (2011). Perceptual effect on motor learning in the serial reaction time task. *Journal of General Psychology*, *138*, 110–126.

Kemény, F., & Lukács, Á. (2013). Stimulus dependence in probabilistic category learning. *Acta Psychologica*, *143*(1), 58–64. http://dx.doi.org/10.1016/j.actpsy.2013.02.008.

Koch, I. (2001). Automatic and intentional activation of task sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1474–1486. http://dx.doi.org/10.1037/0278-7393.27.6.1474.

Marcus, G.F., Fernandes, K.J., & Johnson, S.P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, *18*(5), 387–391. http://dx.doi.org/10.1111/j.1467-9280.2007.01910.x.

Meier, B., & Cock, J. (2010). Are correlated streams of information necessary for implicit sequence learning? *Acta Psychologica*, *133*(1), 17–27. http://dx.doi.org/10.1016/j.actpsy.2009.08.001.

Meier, B., & Cock, J. (2012). The role of cues and stimulus valency in implicit task sequence learning — A task sequence is not enough. In A.L. Magnusson, & D.J. Lindberg (Eds.), *Psychology of performance and defeat* (pp. 155–166). Hauppauge, NY: Nova Science Publisher.

Meier, B., Weiermann, B., & Cock, J. (2012). Only correlated sequences that are actively processed contribute to implicit sequence learning. *Acta Psychologica*, *141*(1), 86–95. http://dx.doi.org/10.1016/j.actpsy.2012.06.009.

Meier, B., Weiermann, B., Gutbrod, K., Stephan, M.A., Cock, J., Müri, R.M., & Kaelin-Lang, A. (2013). Implicit task sequence learning in patients with Parkinson's disease, frontal lesions and amnesia: The critical role of fronto-striatal loops. *Neuropsychologia*, *51*(14), 3014–3024. http://dx.doi.org/10.1016/j.neuropsychologia.2013.10.009.

Nissen, M.J., & Bullemer, P. (1987). Attentional requirements of learning — Evidence from performance-measures. *Cognitive Psychology*, *19*, 1–32.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: the case of orthographic regularities. *Journal of Experimental Psychology. General*, *130*(3), 401–426.

Reber, P.J., Knowlton, B.J., & Squire, L.R. (1996). Dissociable properties of memory systems: Differences in the flexibility of declarative and nondeclarative knowledge. *Behavioral Neuroscience*, *110*, 861–871.

Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 581–597.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.

Saffran, J.R., Pollak, S.D., Seibel, R.L., & Shkolnik, A. (2007). Dog is a dog is a dog: infant rule learning is not specific to language. *Cognition*, *105*, 669–680.

Weiermann, B., Cock, J., & Meier, B. (2010). What matters in implicit task sequence learning: perceptual stimulus features, task sets, or correlated streams of information? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(6), 1492–1509. http://dx.doi.org/10.1037/a0021038.

Weiermann, B., & Meier, B. (2012a). Implicit task sequence learning with auditory stimuli. *Journal of Cognitive Psychology*, *24*(4), 468–475. http://dx.doi.org/10.1080/20445911.2011.653339.

Weiermann, B., & Meier, B. (2012b). Incidental sequence learning across the lifespan. *Cognition*, *123*(3), 380–391. http://dx.doi.org/10.1016/j.cognition.2012.02.010.

Willingham, D.B., Wells, L.A., Farrell, J.M., & Stemwedel, M.E. (2000). Implicit motor sequence learning is represented in response locations. *Memory & Cognition*, *28*, 366–375.