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 19 December 2015

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)