Rocznik Kognitywistyczny 8/2015, s. 91–98
doi: 10.4467/20843895RK.14.009.4258
www.ejournals.eu/Rocznik-Kognitywistyczny

KATARZYNA RĄCZY
ZUZANNA SKÓRA
MACIEJ J. SZUL
Jagiellonian University (Krakow, Poland)

The effect of multi-modal learning in Artificial Grammar Learning Task

Abstract: The aim of the following study was to answer the question whether multimodal grammar learning would improve classification accuracy as compared with a unimodal learning. To test this hypothesis, an experimental procedure was constructed based on the research conducted by Conway and Christiansen [2006]. Their study regarded modality-specific Artificial Grammar Learning task (AGL). The grammatical sequence that was used in the study presented here was based on an algorithm with a finite number of results. Two additional sets of ungrammatical sequences were generated in a random manner. One of them was used in the learning phase in the control group while the second one, in the classification phase, in both, control and experimental groups. The obtained results showed that participants performed classification task above the chance level. These findings supported the hypothesis, which stated that grammar learning would occur [Conway and Christiansen 2006; Reber 1989]. We did not observe any effect regarding the hypothesized accuracy enhancement in a multimodal learning condition.

Key words: implicit learning, artificial grammar learning, multimodal learning

Introduction

Perception of the world requires constant integration of the information from different sensory sources. Does this integration enhance our perception? Using the Artificial Grammar Learning Task (AGL) with visual and auditory stimuli [Conway and Christiansen 2006] could provide some answers. AGL is a paradigm designed to study implicit learning. In the basic design, the stimulus consists of strings of letters that follow a finite state grammar. Participants' task is to observe the consecutive strings during the learning task and then to classify new stimuli (that consist of grammatical and ungrammatical strings of letters). Accuracy in this task is usually above chance level [Reber 1989].

One of the main concerns regarding this paradigm is the nature of acquired representation of the knowledge. Reber [1989] assumed that participants gained abstract knowledge that is a kind of knowledge not directly tied to the superficial (sensory) features of a stimulus. Recently his claim has been challenged by proposing that learning in this task can be stimulus-specific, therefore can occur in parallel along separate perceptual dimensions [Conway and Christiansen 2006]. To test this hypothesis, Conway and Christiansen modified Reber's paradigm to study the learning process (learning of regularities within and between modalities. In this study, participants were able to classify a sequence as grammatical only when it was presented in the same modality (visual or auditory) as during the learning phase. Moreover, in a dual-grammar condition, when participants were simultaneously learning two different grammars through different modal channels there was no learning decrement observed. They have concluded that statistical learning observed in AGL task results in a knowledge that is stimulus-specific rather than abstract [Conway and Christiansen 2006].

More recently Johansson [2009] has conducted a partial replication of the study by Conway and Christiansen [2006]. He additionally manipulated the number of blocks during the learning phase and concluded that the extended period of learning results in strengthening the stimulus-specific knowledge rather than the abstract one [Johansson 2009].

After several decades of unimodal perceptual research, a great deal of interest was given to cross-modal interactions. Many objects in the external environment are represented in more than one sensory modality at a time. Studies in the motion detection paradigm revealed that the co-occurrence of both visual and auditory cues enhance performance [Alais and Burr 2004; Meyer, Wuerger, Röhrbein and Zetzsche 2005]. In both experiments the key point was the temporal co-occurrence of both cues and their relevance to the task. Similarly, the acquisition of a certain grammar is bind to a stimuli which can be represented not only in one modality. According to the studies up to now [Conway and Christiansen 2006], using AGL paradigm, letters are commonly used in such experiments. They, however, can be heard or seen when acquiring a grammar. Therefore, according to previous studies based on multimodal interaction there can be a difference in grammar acquisition performance due to the modality of the grammar. Based on the results obtained by Alais and Burr [2004], we assume that presenting a grammar in two modalities at the same time may improve learning.

In attempt of investigating the nature of representation of the knowledge acquired during AGL task and exploring the possible effects of multimodal learning a study partially based on the one by Conway and Christiansen [2006] was conducted. The focus of this study was on the influence of a simultaneous learning through two modal channels (visual and auditory) on a classification performance through either visual or auditory channel. If the information acquired through learning is abstract, then it is possible that multimodal learning could enhance the accuracy in the classification task when compared with unimodal learning. The main difference between the study conducted by Conway and Christiansen [2006] in which the authors replaced the real letters, commonly used in AGL research, with the colorful rectangels and various tones, was presenting the stimuli simultaneously. By using stimulus which have their representations either in visual or auditory system but detached from the experience of grammar acquisition in a daily life, we attempt to examine the acquisition of an artificial grammar based on the artificial stimuli learned in both modalities at the same time. Moreover, the previous authors [Conway and Christiansen 2006; Johansson 2009] used two grammars opposed to the research presented in this paper.

The first goal of the study is to assess whether it is possible to learn artificial grammar using colors and tones instead of strings of letters, therefore replicating the results of Conway and Christiansen [2006]. The main interest of this study, our hypothesis, states that learning a grammatical structure through both visual and auditory modality does improve performance as compared to learning only through one of these channels.

Method

Participants

80 participants with normal hearing and normal vision were recruited (58 women and 22 men). Mean age was 24 years old (SD = 8). Participants were not reimbursed.

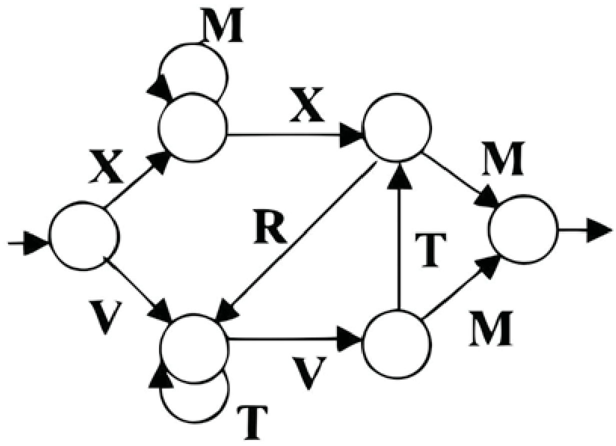
Materials and Procedure

Procedure included four groups divided into experimental and control conditions. In each group grammatical strings were made of the same finite state grammar as in Conway and Christiansen's paper [2006] but letters from the original paradigm were replaced by colorful rectangles and tones (see: Table 1 and 2). We used 8 grammatical sequences from the grammar in the training phase and 10 grammatical sequences from the grammar in the test phase; all sequences contained at least three and no more than seven elements.

V	R	T	X	M
R: 143 G: 207 B: 66	R: 64 G: 179 B: 163	R: 50 G: 168 B: 227	R: 81 G: 131 B: 196	R: 121 G: 52 B: 153
210 Hz	245 Hz	268 Hz	333 Hz	389 Hz

* Colors are represented by values in 8-bit digital RGB scale

Table 1: Letters used in grammar, matched with colors* and tones



STRINGS USED IN LEARNING PHASE	STRINGS USED IN CLASSIFICATION TASK
VVM	VTVM
XMXM	VVTM
XXRVM	XMMXM
VTVTM	VTTVM
XXRVTM	VTTTVM
XMMMXM	VVTRVM
XXRTTVM	VTTTVTM
VTVTRVM	XXRTVTM
	XMXRTVM
	XMMXRVM

Table 2: Depiction of finite state grammar used in research and generated strings (Conway and Christiansen 2006)

Two sets of ungrammatical strings were constructed randomly. One set was used in the learning phase of control groups (8 sequences). Second set was used in the classification phase (10 sequences) for both control and experimental groups. There were four groups, two of which were experimental groups whereas the another two were control ones. During the learning phase the experimental group (40 participants) was presented with grammatical strings in visual and auditory modality at the same time. The 8 training sequences in each block, for a total of six blocks. Thus, a total of 48 sequences were presented randomly, one at a time. During the test phase the experimental group was divided where one half was presented with the task in the auditory modality (20 participants) and the other half, in visual modality (20 participants) (Figure 1). The control groups, in the learning phase, were exposed either to a grammatical strings in auditory modality and ungrammatical strings in visual modality (20 participants) or ungrammatical strings in auditory modality and grammatical strings in visual modality (20 participants). The test phase was identical as in the experimental group. Participants were presented with the task in either auditory (20 participants) or visual modality (20 participants) (Figure 2) For the test phase, 20 sequences were used, 10 that were grammatical and other 10 that were ungrammatical.

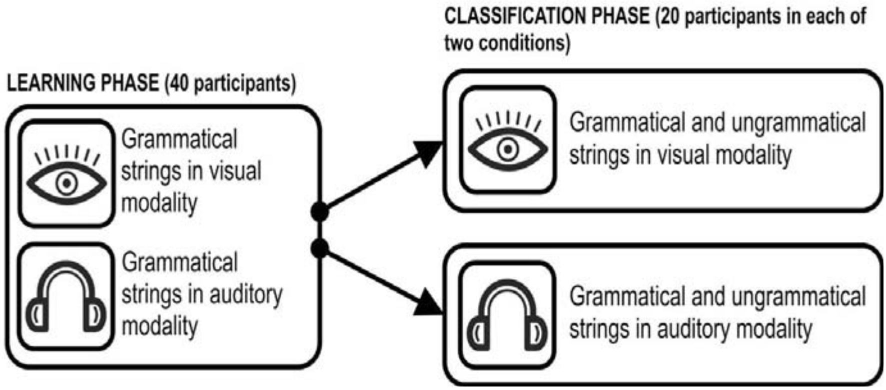


Figure 1: Scheme of experimental group tasks. Learning and classification

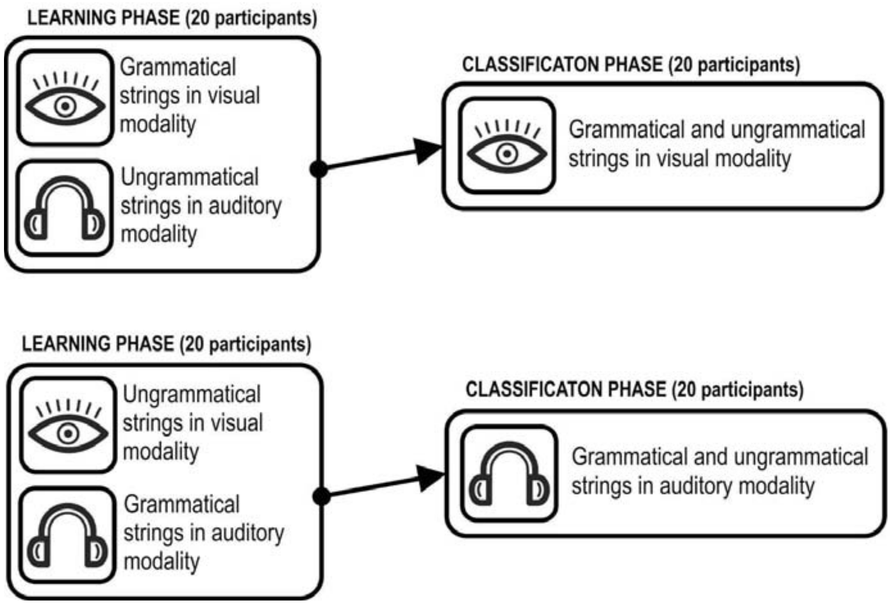


Figure 2: Scheme of control groups. Learning and classification

Results

The classification accuracy as expected, reached the level above 50% (all groups: M = .59, SD = .16, $t(79) = 3.02$, $p < .05$; visual experimental M = .59, SD = .16, $t(19) = 3.42$, $p < 0.05$; visual control M = .57, SD = .16, $t(19) = 2.93$, $p < 0.05$; auditory experimental M = .59, SD = .11, $t(19) = 2.39$, $p < .05$; auditory control M = .62, SD = .18, $t(19) = 2.10$, $p < .05$).

In order to examine whether the experimental group with grammatical strings in both modalities was associated with higher classification accuracy than the control group we have run independent samples t-test. There were no significant differences between these groups $t(78) = .21$, $p > .05$.

We have also tested the differences between the experimental and the control group within the same modality. Meaning, whether there were any differences between the groups when participants classified strings either through the auditory or the visual channel. There were no significant differences between the groups in any of the modalities. Visual classification: $t(38) = .28$, $p > .05$; and auditory classification: $t(38) = .61$, $p > .05$). However, the data presented here should be interpreted with caution and more research is needed to further verify the obtained results.

Discussion

The main goal of the study was to investigate whether the exposure to artificial grammar learning in two modalities in parallel will improve the classification accuracy. The first condition to achieve this goal was met, the learning of the artificial grammar in general occurred as in other studies administering the AGL task [Conway and Christiansen 2006; Reber 1989]. The hypothesis, regarding the differences between the groups, was not supported. Classification accuracy in the multimodal group did not differ from the classification accuracy in the unimodal group. The enhancing effect of double-modal learning observed by Meyer and colleagues [2005] was not obtained.

We have confirmed the results obtained by other researchers engaged in studying AGL task [Conway and Christiansen 2006; Reber 1989]. These scientific reports support the thesis that people have the ease of assimilating grammatical rules. This view is strongly rooted in linguistic research of Universal Grammar (UG), the theory usually credited by Noam Chomsky [Werry 2007] proposing that the ability to learn grammar is hard-wired into the brain. Although the results obtained in the study are not sufficient to confirm the existence of an internal grammar learning mechanism, certainly they show that artificial grammar learning occurs regardless of modality or stimulus used during the experiment.

The objective of the study was to investigate whether the exposure to artificial grammar learning in two different modalities (visual and auditory modality in our case) will result in the same level of classification accuracy as learning in just one modality (auditory or visual). The fact that we did not observe any enhancement in the classification accuracy after exposing the participant to both modalities speaks in favor of the stimulus-specific nature of the representation as in Conway and Christiansen's [2009] research.

It was found that inter-sensory interactions occur early in the primary sensory cortex [Okada, Venezia, Matchin, Saberi and Hickok 2013]. If the representation acquired in AGL were abstract the improvement in the classification accuracy would be observed in the double modal condition [Meyer et al. 2005]. On the other hand, if the representation were stimulus-specific the interaction between both channels to enhance performance would not be possible.

No difference between learning in the auditory modality and a visual learning, may be due to the fact that the visual stimulus was presented sequentially rather than spatially. The spatial presentation of a visual stimulus can produce better results in learning than the sequential one, which is not typical for visual perception [Conway and Christiansen 2009].

In contrast, no difference between the experimental group and the control one can be explained by separation of the perceptual pathways, and thus the lack of intermodal transfer. A system where the transfer could take place is the working memory [Baddeley 2003]. More specifically, the central executive system, which presumably unites the information flowing into the phonological loop and the visual-spatial sketchpad. Perruchet and Pacton [2006] proposed the use of heavily simplified grammar rules, in order not to overload the limited resources of working memory. Therefore, it is a promising direction for further studies.

The other possibility might be that the enhancing effect of a double-modal activation occurs only when the information conveys more meaning like in the McGurk effect (where a conflict between visual and auditory input modifies the perception in auditory input – Nahorna, Berthommier and Schwartz 2012).

According to Reber [1989] and Pothos [and Bailey 2000] there is a possibility that using the dual modality presentation during the learning phase in the AGL task would later support the accuracy in the classification phase. This would be due to the fact that the similarity between the strings both in the training phase and the classification phase would be significantly higher. According to the above-mentioned research it may be possible that learning in only one modality at a time and then classifying a string of letters in this same modality seems to be more similar in opposition to learning in one modality and classifying in another one. Although we did not check the effect of similarity between learning and classification phase this may be a promising direction for the future research.

Future studies should also look into the cortical areas activated during the classification phase (visual classification and auditory areas). If both areas are activated even though the task requires only one modality than it could still be possible that the representation is abstract in some amount (audiovisual integration area – Okada, Venezia, Matchin, Saberi and Hickok 2013).

Acknowledgments

Authors would like to thank Anna Anzulewicz for her invaluable advice and ideas throughout the process.

BIBLIOGRAPHY

Alais D., Burr D. (2004). *No direction-specific bimodal facilitation for audiovisual motion detection*. „Cognitive Brain Research", 19 (2), pp. 185–194.

Baddeley A. (2003). *Working memory: Looking back and looking forward*. „Nature Reviews Neuroscience", 4, pp. 829–839.

Conway C.M., Christiansen M.H. (2006). *Statistical learning within and between modalities pitting abstract against stimulus-specific representations*. „Psychological Science", 17 (10), pp. 905–912.

Conway, C. M., Christiansen, M. H. (2009*). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning*. „European Journal of Cognitive Psychology", 21, pp. 561-580.

Johansson T. (2009). *Strengthening the case for stimulus-specificity in Artificial Grammar Learning: No evidence for abstract representations with extended exposure*. „Experimental Psychology" (formerly „Zeitschrift für Experimentelle Psychologie"), 56 (3), pp. 188–197.

Meyer G.F., Wuerger S.M., Röhrbein F., Zetzsche C. (2005). *Low-level integration of auditory and visual motion signals requires spatial co-localisation*. „Experimental Brain Research", 166 (3–4), pp. 538–547.

Nahorna O., Berthommier F., Schwartz J.-L. (2012). *Binding and unbinding the auditory and visual streams in the McGurk effect*. „The Journal of the Acoustical Society of America", 132, p. 1061.

Okada K., Venezia J.H., Matchin W., Saberi K., Hickok G. (2013). *An fMRI study of audiovisual speech perception reveals multisensory interactions in auditory cortex*. „PLoS ONE", 8 (6), p. e68959.

Perruchet P., Pacton S. (2006). *Implicit learning and statistical learning: one phenomenon, two approaches*. „TRENDS in Cognitive Sciences", 10 (5), pp. 233–238.

Pothos E.M., Bailey T.M. (2000). *The role of similarity in artificial grammar learning*. „Journal of Experimental Psychology: Learning, Memory, and Cognition", 26 (4), p. 847.

Reber A.S. (1989). *Implicit learning and tacit knowledge*. „Journal of Experimental Psychology: General", 118 (3), p. 219.

Werry C. (2007). *Reflections on language: Chomsky, linguistic discourse and the value of rhetorical self-consciousness*. „Language Sciences", 29 (1), pp. 66–87, http://doi.org/10.1016/j.langsci.2006.01.001.