

Research report

Modality effects in implicit artificial grammar learning: An EEG study

Susana Silva^{a,b,*}, Vasiliki Folia^c, Filomena Inácio^b, São Luís Castro^a, Karl Magnus Petersson^{b,c}^a Center for Psychology, University of Porto, Porto, Portugal^b Center for Biomedical Research, University of Algarve, Faro, Portugal^c Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

ARTICLE INFO

Article history:

Received 9 October 2017

Received in revised form 26 December 2017

Accepted 15 February 2018

Available online 21 February 2018

Keywords:

Modality

Implicit learning

Artificial Grammar Learning

EEG

ERD/ERS

ABSTRACT

Recently, it has been proposed that sequence learning engages a combination of modality-specific operating networks and modality-independent computational principles. In the present study, we compared the behavioural and EEG outcomes of implicit artificial grammar learning in the visual vs. auditory modality. We controlled for the influence of surface characteristics of sequences (Associative Chunk Strength), thus focusing on the strictly structural aspects of sequence learning, and we adapted the paradigms to compensate for known frailties of the visual modality compared to audition (temporal presentation, fast presentation rate). The behavioural outcomes were similar across modalities. Favouring the idea of modality-specificity, ERPs in response to grammar violations differed in topography and latency (earlier and more anterior component in the visual modality), and ERPs in response to surface features emerged only in the auditory modality. In favour of modality-independence, we observed three common functional properties in the late ERPs of the two grammars: both were free of interactions between structural and surface influences, both were more extended in a grammaticality classification test than in a preference classification test, and both correlated positively and strongly with theta event-related-synchronization during baseline testing. Our findings support the idea of modality-specificity combined with modality-independence, and suggest that memory for visual vs. auditory sequences may largely contribute to cross-modal differences.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Humans are capable of apprehending the regularities of sequential input after being exposed to it, even without any intention to do so. Two different research traditions have focused on this human ability: research on implicit learning dealt mainly with the learning of structural, all-or-none, rule-based regularities as it is probed in Artificial Grammar Learning (AGL) paradigms (Pothos, 2007; Reber, 1967, 1989), while research on statistical learning dealt mainly with distributional, frequency-related properties of the input, such as transitional probabilities subtending word segmentation (Saffran et al., 1996). Although different in some aspects, both traditions emphasized the acquisition of implicit knowledge (i.e., learning without awareness) from passive exposure, and how this affords managing the complexities of the world. Therefore, they are often viewed as two sides of the same coin (Batterink et al., 2015; Perruchet and Pacton, 2006), and they raise common questions.

A prominent question in both implicit learning and statistical learning – which we will refer to as *sequence learning* from now on – is whether learning is modality-specific, modality-independent, or both. Sequence learning – at least statistical learning – has been approached as a domain-general (hence modality-independent) ability (Baker et al., 2004; Goujon and Fagot, 2013; Saffran et al., 1996), but evidence against a purely amodal learning system is already abundant. First, it is known that learning is tied to perception: sequence learning unfolds in parallel in different perceptual dimensions (Conway and Christiansen, 2006), and stimulus-specific knowledge increases with prolonged exposure (Johansson, 2009). Second, stimulus modality seems to impact the efficiency of the learning process. Initial studies found a quantitative advantage of the auditory modality in sequence learning (Conway and Christiansen, 2005; Conway et al., 2009). Recent studies found that it is not necessarily so, but modalities are still subjected to different constraints, such as the presentation rate of sequence items (i.e., syllables, letters), or the temporal vs. spatial format of sequence presentation (item-by-item vs. whole sequence, see Conway and Christiansen, 2009; Emberson et al., 2011). In an effort to make sense of these results, Frost and colleagues (Frost et al., 2015) proposed a distinction between the

* Corresponding author at: Center for Psychology at the University of Porto, Rua Alfredo Allen, s/n, 4200-135 Porto, Portugal.

E-mail address: susanamsilva@fpce.up.pt (S. Silva).

mechanisms for learning, representing and processing the input, and the neural networks in which these principles would be instantiated. The former would be modality-independent, the latter subjected to modality constraints. While statistical learning is clearly based on an assumption of modality-independence (Frost et al., 2015), this is not so well-established within implicit learning paradigms, and thus the hypothesis of modality-specificity *plus* independence is lacking empirical evidence.

In the present study, we investigated the hypothesis of modality-independent mechanisms coupled with modality-specific networks in the context of implicit Artificial Grammar Learning (AGL), where participants are exposed to examples of a grammar and tested for grammar knowledge before and after exposure. AGL paradigms capture the structural, rule-like aspects engaged in natural language learning, namely those of natural syntax (Christiansen et al., 2012; Zimmerer et al., 2014). This explains why an analogue of the natural-syntax-related P600 component (Hagoort et al., 1993; Osterhout and Holcomb, 1992) has been found in EEG AGL studies in response to violations of the artificial grammar, in humans (Christiansen et al., 2012; Lelekov-Boissard and Dominey, 2002; Tabullo et al., 2013) and even in primates (Attaheri et al., 2015). We saw that component in a previous AGL study of ours (Silva et al., 2017), where we examined the behavioural and EEG outcomes of learning an artificial auditory grammar. In the present study, we did a direct comparison of our previous results on the auditory grammar (Silva et al., 2017) with those from a new experiment, using a matched sample and an analogous visual grammar. Thus, we compared behavioural and EEG learning indices across modalities in order to examine the similarities and differences indicating modality-independence vs. modality-specificity.

Concerning the EEG, we predicted that learning indices would differ in morphology (i.e., latency and/or topography) across modalities, but they would show similar underlying computational principles and thus the ERP components would be functionally equivalent. We tested for the functional equivalence of ERPs across modalities by focusing on the late (post 400 ms) component, and we considered three levels of comparison. First, we checked whether the visual ERP was free from the influence of the surface characteristics of sequences (free from Associative Chunk Strength, ACS influences, see Knowlton and Squire, 1996; Meulemans and Van der Linden, 1997), just as it was in the auditory modality: in Silva et al. (2017), we analysed both the effects of grammatical status (grammatical vs. non-grammatical sequences) and of the surface characteristics of sequences (low- vs. high ACS) on participants' post-exposure EEG responses, and we found that the late (500–700 ms) posterior component only depended on grammatical status, no effects from, or interactions with ACS being observed. Here, we examined whether the same happened in the visual modality. Second, we analysed ERP differences between two post-exposure tests of grammar knowledge (preference and grammaticality classification) in both modalities, and checked whether the difference profile was similar: in the auditory modality, we saw more expanded latencies for the late positive component in grammaticality classification (500–900 ms) than in preference classification (500–700 ms), and we wanted to see if the same went for the visual grammar. Our third approach to the functional equivalence of ERPs across modalities was to correlate subject-level learning-related ERPs (i.e., increase in ERP amplitude from pre-exposure to post-exposure testing) with a subject-level measure of general engagement with sequential input. We chose to measure event-related synchronization or desynchronization (ERS/ERD, see Pfurtscheller and Lopes da Silva, 1999) of low-frequency (1–12 Hz) brain oscillations time-locked to the onset of pre-exposure sequences, assuming that this could provide either an index of entrainment to the presentation rate of sequence items,

leading to improved sequence learning (Riecke et al., 2015; Selchenkova et al., 2014a,b), or an index of more general brain processes (e.g., attention) triggered by sequence presentation and leading to the same result. Entrainment would be indicated by the dominance of brain oscillations matching the presentation rate of sequence items (1.72 Hz for visual, 3.03 Hz for auditory), but we were uncertain about having this kind of evidence with short stimulation periods (see Nozaradan et al., 2015; Nozaradan, 2014; Nozaradan et al., 2011). Even if we did not find evidence of entrainment, we expected to find indices of processes such as working memory maintenance or sustained attention following the onset of sequence presentation, which have been associated with synchronization in the theta (4–8 Hz) band (working memory: Gevins et al., 1997, 1998; Gevins and Smith, 2000; Grunwald et al., 1999; Jensen and Tesche, 2002; Klimesch, 1999; Mecklinger et al., 1992; attention: Aftanas and Golocheikine, 2001; Asada et al., 1999; Griesmayr et al., 2010; Kubota et al., 2001). We reasoned that, if the two modalities showed similar subject-level correlations between learning-related ERPs and engagement with sequence input (entrainment and/or theta ERS), this would support the idea that the respective learning-related ERPs are functionally equivalent.

Concerning behavioural indices of learning, we predicted that these would be similar, provided that we control for known constraints, such as the deleterious effects of temporal (vs. spatial) presentation on the visual modality, particularly when the presentation rate is fast (Conway and Christiansen, 2009; Emberson et al., 2011). Therefore, we chose to use a temporal presentation format for both the visual and the auditory grammar, but we used a slower presentation rate (1.72 Hz) in the visual grammar than in the auditory one (3.03 Hz), and we decreased the memory load in the visual exposure task, by asking participants to do serial recall (retype) of only one sequence, instead of doing same-different judgements on two sequences, as we had done in the auditory version. Although less loading in terms of memory, the serial recall task is potentially more akin to the use of memorization strategies, and this may enhance explicit learning when there is a predisposition to do so. In order to grant that the serial recall task we used in the visual modality did not facilitate participants' attempts to learn explicitly, we complemented our questionnaire measures of awareness with the results of a control study. In this complementary study, we provided instructions for explicit learning (Witt et al., 2013) in a matched sample (visual/serial recall vs. auditory/same-different). If there were no differences in awareness between groups, this would mean that the visual/serial recall task did not potentiate any predisposition for explicit learning, we would be dealing with similar learning processes across the two modalities, and thus differences in EEG indices could not be imputed to different learning processes concerning awareness.

2. Results

2.1. Behavioural results

Accuracy was at chance levels on TEST1 (*Mean ± SD*: 47% ± 8%; $t(21) = -2.05, p = .053$) and rose above chance in TEST2 (*Mean ± SD*: 58% ± 12%; $t(21) = 3.15, p = .005$) and TEST3 (*Mean ± SD*: 74% ± 14%; $t(21) = 8.16, p < .001$). Results did not differ from the auditory version (Silva et al., 2017) in any of the three tests ($ps > .27$).

The analysis of endorsement rates (Fig. 1) for TEST1 vs. TEST2 showed a significant interaction between Test and Gram ($F(1,21) = 17.83, p < .001, \eta^2 p = .459$): on TEST1, G and NG sequences were equally endorsed ($F(1,21) = 4.24, p = .052, \eta^2 p = .168$). On TEST2, the endorsement rates for G sequences were higher than for NG ($F(1,21) = 9.94, p = .005, \eta^2 p = .321$). Comparisons with the audi-

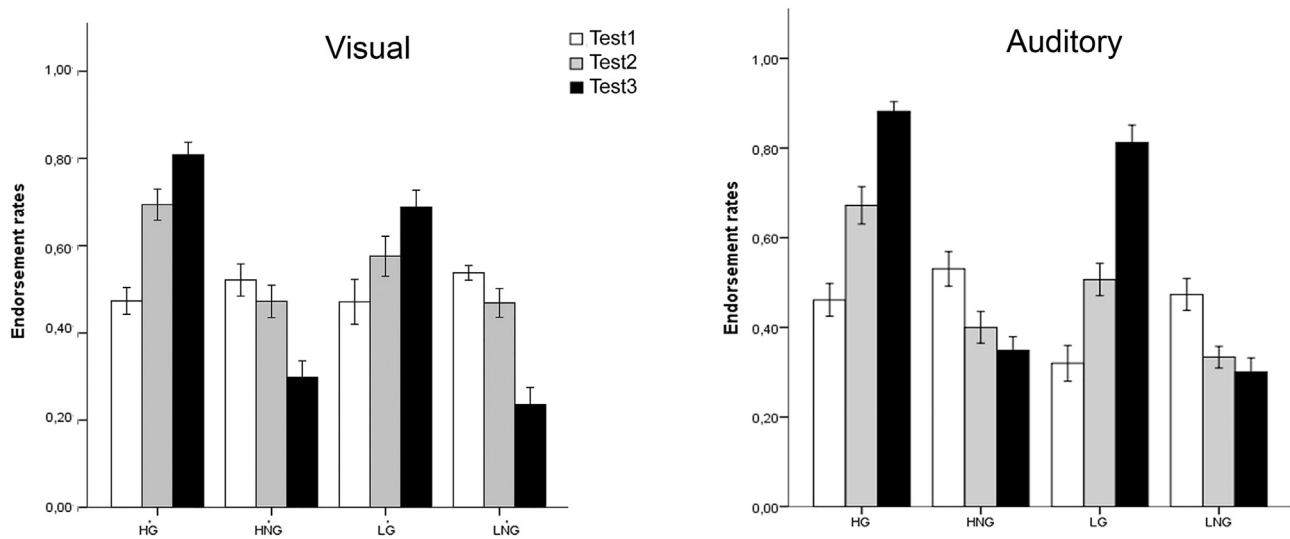


Fig. 1. Left: Endorsement rates as a function of grammatical status \times ACS (HG = High-ACS grammatical, HNG = High-ACS non-grammatical, LG = Low-ACS grammatical, LNG = Low-ACS non-grammatical). Right: Results for auditory stimuli (Silva et al., 2017). Error bars represent the standard error of the mean.

tory modality (Silva et al., 2017) showed no differences (Gram \times Modality, TEST1: $p = .32$; TEST2: $p = .28$). Both the Test \times ACS ($p = .27$) and the Test \times Gram \times ACS interactions ($p = .11$) were non-significant. Neither of these interactions differed from the ones previously obtained for auditory stimuli (comparison with Silva et al., 2017; Test \times Gram \times Modality: $p = .11$; Test \times ACS \times Modality: $p = .49$; Test \times Gram \times ACS \times Modality: $p = .29$). On TEST3, G sequences were significantly more endorsed than NG ($F(1,21) = 66.59$, $p < .001$, $\eta^2 p = .760$), paralleling the results for the auditory modality ($p > .64$). Neither ACS effects nor ACS \times Gram interactions differed across modalities (ACS \times Modality, $p = .196$; ACS \times Gram \times Modality, $p = .599$). Comparisons between TEST2 and TEST3 showed a significant Test \times Gram interaction ($F(1,21) = 19.04$, $p < .001$, $\eta^2 p = .476$) and non-significant Test \times ACS ($p > .37$) or Test \times Gram \times ACS interactions ($p > .31$). The results were similar to those obtained for the auditory modality (Test \times Gram \times Modality: $p > .71$; Test \times ACS \times Modality: $p > .11$; Test \times Gram \times ACS \times Modality: $p > .73$).

Consistent with the endorsement rates analysis, the d -prime increased from TEST1 over TEST2 ($t(21) = -4.16$; $p < .001$) and from TEST2 over TEST3 ($t(21) = -4.34$; $p < .001$), similar to auditory stimuli ($ps > .11$). The ACS d -prime did not change significantly across tests ($ps > .27$). Changes did not differ from the ones observed for auditory stimuli ($ps > .11$).

2.2. Explicit knowledge: Questionnaires and control study

Questionnaires showed no evidence of explicit knowledge. Since these did not provide a quantitative measure of explicit knowledge, but rather a qualitative one, we conducted a control study to make sure that the two experiments were equivalent in this respect. Specifically, the goal of this control study was to see whether the serial recall (retype) task, which was used for acquisition in the current visual study, fostered explicit learning, when compared to the same-different-judgement task of our previous auditory experiment (Silva et al., 2017).

In the control study, 44 participants did an explicit analogue of the visual and auditory implicit AGL experiments. In a single session, they did one acquisition task (80 Grammatical items) under the explicit orientation to learn the grammar, followed by one grammaticality test (96 items, 48 grammatical + 48 non-grammatical). Twenty-two participants were assigned to the

visual/serial recall group. As in the current study, they were asked to rewrite the sequences after viewing them during acquisition. The other 22 were assigned to the auditory/same-different group, and they were asked to judge whether the sequences were equal or different, as in Silva et al. (2017). Before acquisition, all participants were asked to pay attention to the structure of the sequences, namely which letters/syllables were at the beginning, which ones were at the end, and how the letters/syllables followed one another. After acquisition, they did the grammaticality classification test and, finally, they were given a questionnaire. They were asked three open questions (which letters/syllables at the beginning, at the end, and which could take part in adjacent repetitions), and they were given eight sequencing rules to classify as true or false.

We analysed questionnaire data to test whether explicit knowledge was higher in the visual/serial recall group compared to the auditory/same-different group. Correct responses to all questions would lead to a total maximum score of 17, and this score would index the amount of explicit knowledge. Errors in open questions (e.g., providing a letter/syllable as initial when it was not) were discounted. Comparisons between the two groups showed no differences in explicit knowledge ($t(44) = 1.38$, $p > .17$, Fig. 2), providing no evidence that explicit rule-acquisition is favoured under visual/serial recall settings compared to auditory/same-different settings.

In addition, we saw that accuracy did not differ between groups ($t(42) = .315$, $p > .55$, Fig. 2) and it was below chance levels for both groups ($t(21) = -2.69$, $p = .014$; auditory: $t(21) = -4.18$, $p < .001$). This indicates that an explicit orientation to learn is not enough to grant above-chance levels of performance, at least with a one-day acquisition session.

2.3. Event-related potentials

2.3.1. Effects of grammatical status

2.3.1.1. No effects between 100 and 300 ms. Only TEST2 showed Gram effects, which were seen at the central region ($F(1,21) = 6.92$, $p = .016$, $\eta^2 p = .248$). Neither TEST1 nor TEST3 showed Gram effects at any region ($ps > .31$). There were no significant Gram \times ACS interactions in any test ($ps > .06$, Fig. 3). Despite the Gram effects in TEST2, direct cross-test comparisons (Fig. 4A) showed no significant changes in Gram effects from TEST1 over TEST2 (Te

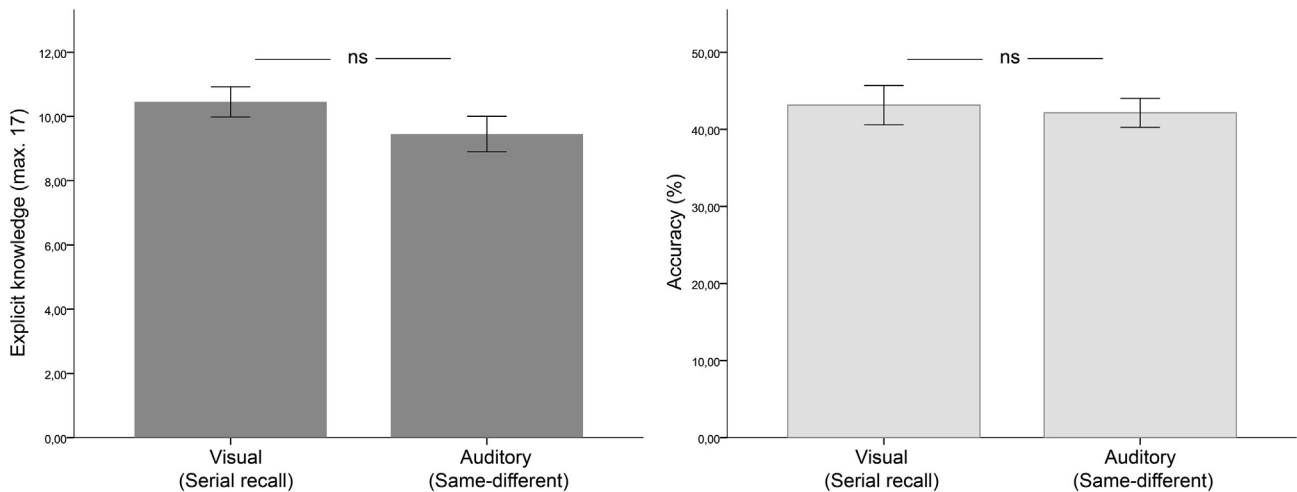


Fig. 2. Control study: explicit knowledge (scores from 0 to 17) and accuracy in grammaticality classification after a 1-day exposure with explicit instructions to learn.

st \times Gram: $ps > .09$) or TEST3 ($ps > .63$), neither from TEST2 over TEST3 ($ps > .24$). Test \times Gram \times ACS interactions were not significant ($ps > .10$). Thus, according to a strict proper-learning approach (Pettersson et al., 1999), there were no grammatical-status-related changes in ERPs in this time window.

2.3.1.2. No effects between 300 and 400 ms. There were no gram effects (TEST1: $ps > .52$; TEST2: $ps > .19$; TEST3: $ps > .39$) and no Gram \times ACS interactions (TEST1: $ps > .06$, TEST2: $ps > .11$; TEST3: $ps > .60$, Fig. 3). Cross-test comparisons (Fig. 4A) showed no significant Test \times Gram interactions (TEST1&2: $ps > .18$; TEST1&3: $ps > .30$; TEST2&3: $ps > .73$) or Test \times Gram \times ACS interactions (TEST1&2: $ps > .17$; TEST1&3: $ps > .35$; TEST2&3: $ps > .40$). Again, no grammatical-status-related changes in ERPs were seen, leaving no early ERP markers of learning.

2.3.1.3. Effects between 400 and 600 ms for TEST2 and TEST3. TEST2 and TEST3 showed Gram effects at anterior (TEST2: $F(1,21) = 4.57$, $p = .045$, $\eta^2p = .179$; TEST3: $F(1,21) = 9.97$, $p = .005$, $\eta^2p = .322$) and central (TEST2: $F(1,21) = 4.82$, $p = .040$, $\eta^2p = .187$; TEST3: $F(1,21) = 8.94$, $p = .007$, $\eta^2p = .299$) regions. TEST1 showed no Gram effects ($ps > .23$). There were no Gram \times ACS interactions at any test (TEST1: $ps > .08$; TEST2: $ps > .33$; TEST3: $ps > .13$). Cross-test comparisons (Fig. 4A) showed changes in Gram effects from TEST1 over TEST2 at central ROIs ($F(1,21) = 6.52$, $p = .019$, $\eta^2p = .237$), but not at anterior ($p > .11$) or posterior regions ($p > .34$). Differently, changes in Gram effects from TEST1 over TEST3 extended to all regions (anterior: $F(1,21) = 6.92$, $p = .016$, $\eta^2p = .248$; central: $F(1,21) = 9.61$, $p = .005$, $\eta^2p = .314$; posterior: $F(1,21) = 5.36$, $p = .031$, $\eta^2p = .203$). Changes from TEST2 over TEST3 were non-significant (anterior: $p > .06$; central: $p > .08$; posterior: $p > .18$). Test \times Gram \times ACS interactions were non-significant (TEST1&2: $p > .23$; TEST1&3: $p > .07$; TEST2&3: $ps > .26$). Thus, both TEST2 and TEST3 showed learning-related changes, which were most obvious in anterior-central regions. This topography differs from the one we saw in Silva et al. (2017), which was clearly posterior (Fig. 4 B).

2.3.1.4. Effects between 600 and 900 ms for TEST3. There were no Gram effects at TEST1 ($ps > .29$) or TEST2 ($ps > .07$), but TEST3 showed GRAM effects at all regions (anterior: $F(1,21) = 6.61$, $p = .018$, $\eta^2p = .239$; central: $F(1,21) = 10.60$, $p = .004$, $\eta^2p = .335$; posterior: $F(1,21) = 8.04$, $p = .010$, $\eta^2p = .277$). Gram \times ACS interactions were non-significant in all tests (TEST1: $ps > .08$, TEST 2: $ps > .17$, TEST3: $ps > .24$, Fig. 3). Consistent with this, there were no changes in GRAM effects from TEST1 over TEST2 (Test \times Gram:

$ps > .20$, Fig. 4A). TEST3 showed increased Gram effects compared to TEST1 ($F(1,21) = 7.81$, $p = .011$, $\eta^2p = .271$) and TEST2 ($F(1,21) = 5.12$, $p = .034$, $\eta^2p = .196$) at central regions. There were no Test \times Gram \times ACS interactions for any cross-test comparison (TEST1: $ps > .26$, TEST2: $ps > .16$, TEST3: $ps > .08$). Thus, TEST3 but not TEST2 showed an extended late component, located at the central region.

2.3.2. Effects of ACS

Although we found ACS effects in the test-by-test analysis, cross-test comparisons showed no differences in any time window (Fig. 4A, 100–300 ms: Test \times ACS: $ps > .24$; 100–400 ms: Test \times ACS: $ps > .24$; 400–600 ms: Test \times ACS: $ps > .51$; 400–600 ms: Test \times ACS: $ps > .51$; 600–900 ms: Test \times ACS: $ps > .73$). This contrasts with the auditory experiment (Silva et al., 2017), where we saw a decrease in ACS effects from baseline to grammaticality classification, showing some sensitivity to surface characteristics (although not the expected one, which would be an increase in ACS effects, rather than a decrease). In sum, the auditory modality seemed more prone to surface influences.

2.4. ERD/ERS at baseline preference

We did a first analysis over fine-grained frequency bins (steps of 0.25 Hz) in order to look for possible indices of entrainment to the presentation rates of visual (1.72 Hz) and auditory (3.03 Hz) sequences. We saw no evidence that these target frequencies were enhanced relative to the remaining spectrum and, therefore, we did not find evidence of entrainment. This might be due to the short time window used in the analysis.

We then averaged frequency bins into the delta (1–4 Hz), theta (4–8 Hz) and alpha (8–12) bands. We found Event-Related Synchronization (positive values) for the two modalities in all frequency bands. In the visual grammar, delta and theta bands showed increased ERS values compared to the auditory grammar (delta: $t(42) = -2.18$, $p = .034$; theta: $t(42) = -5.84$, $p < .001$), but the alpha band did not ($p > .24$).

Correlations between learning-related changes from TEST1 over TEST3 and ERS at TEST1 showed strong, positive and significant values for the theta band (Table 1). Theta ERS correlated with ERP changes at central and anterior sites in the visual modality, and with changes at posterior sites in the auditory one. These were the sites where ERP changes related to structural (grammatical-status based) learning were significant. Thus, increased baseline

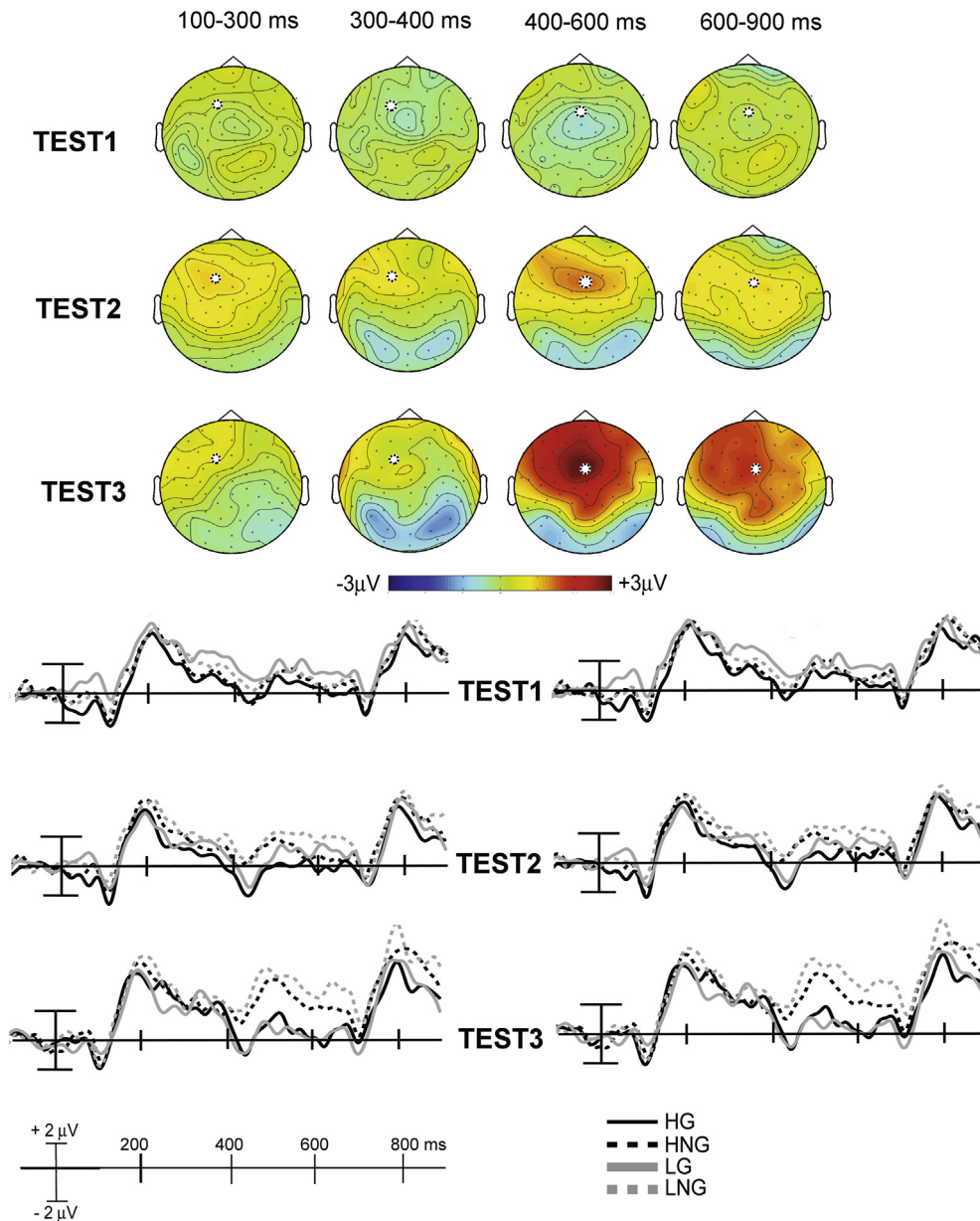


Fig. 3. Topographic maps (top) of early (100–400 ms) effects of grammatical status (NG-G) in each classification test (TEST1 = baseline preference; TEST2 = final preference; TEST3 = grammaticality classification) and illustrative ERP waveforms (down) for each level of grammatical status \times ACS effects (HG, high-ACS grammatical; HNG, high-ACS non-grammatical; LG, low-ACS grammatical; LNG, low-ACS non-grammatical). The waveforms show the signal at the marked electrodes.

theta-ERS seems to relate to the EEG indices of learning an artificial grammar, whether visual or auditory.

3. Discussion

We tested the hypothesis that modality-specific networks coexist with modality-independent processing mechanisms in implicit artificial grammar learning (AGL). We examined learning-related changes (post- vs. pre-exposure) in behavioural and ERP responses to visual and auditory grammars, and we found evidence in support of the hypothesis in question.

Specifically, we found similar behavioural responses in both grammars after compensating for known modality constraints on behavioural responses (demands of presentation format and presentation rate). This suggests that these modality-specific influences were cancelled, and thus that they may be real.

Concerning ERP responses to grammar violations, we saw differences across modalities in both the topography and the latency of the late component (earlier and more frontal in the visual modality), indicating modality-specific networks. These differences could not be attributed to different levels of explicit learning, as both our questionnaires and our control study suggested. Critical to the hypothesis being tested, indices of modality-specific networks concurred with modality-independent (shared) features: both the visual and the auditory late ERPs were free from ACS influences, both were more extended in the grammaticality classification test (400–900 ms visual, 500–900 ms auditory) than it is in the final preference test (400–600 ms visual, 500–700 ms auditory), and, in both, learning-related changes correlated with theta event-related-synchronization (ERS) during pre-exposure.

Finding the *meaning* of the correlation between learning-related ERP changes and ERS at the pre-exposure phase was not central to

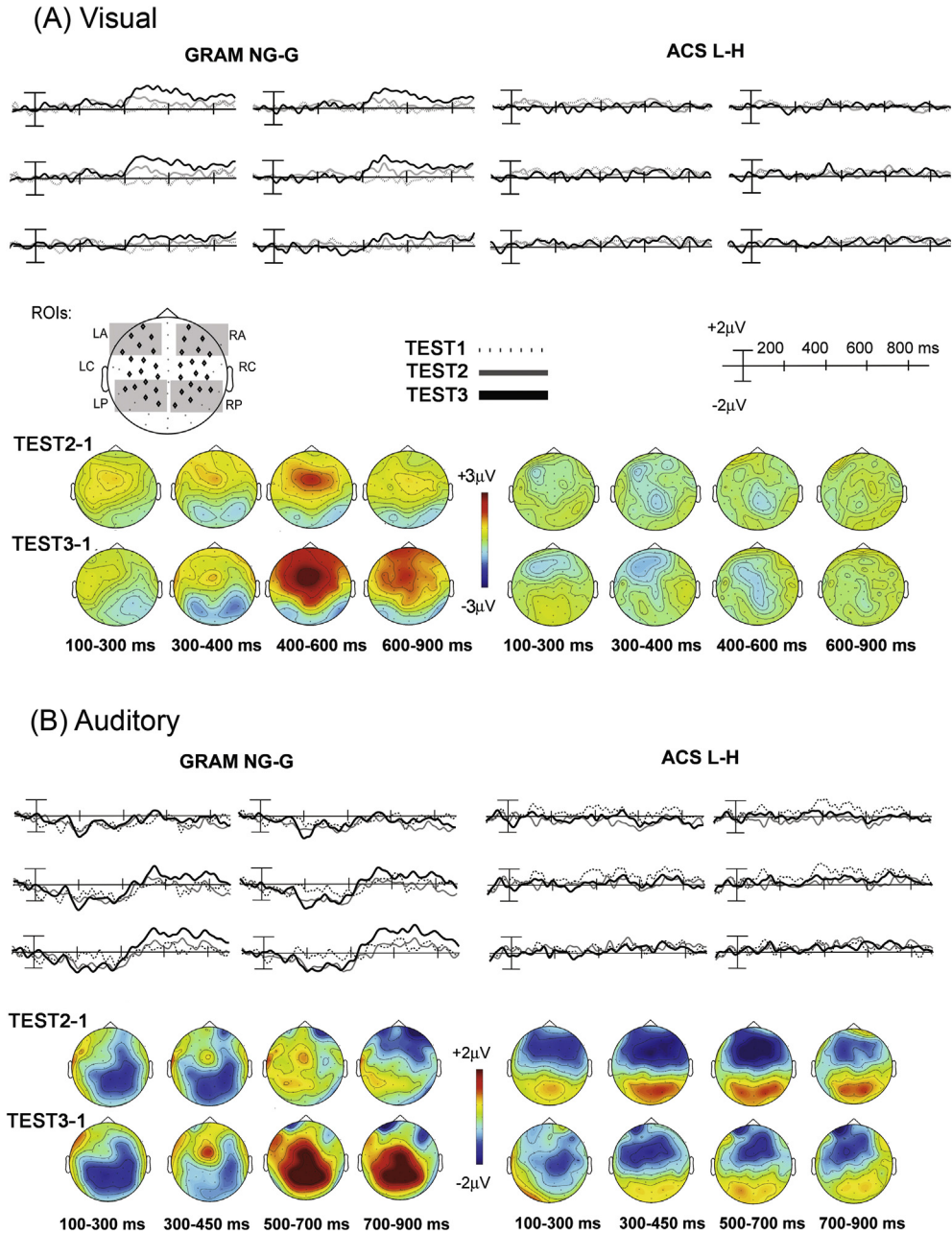


Fig. 4. Learning-related ERPs in visual and auditory (Silva et al., 2017) grammars. Top A and Top B: Average difference waves at the six regions of interest (ROIs) for grammatical status (grammatical status) and ACS effects across the three tests (NG, non-grammatical; G, grammatical; L, low ACS; H, high-ACS; LA, left anterior; LC, left central; LP, left posterior; RA, right anterior; RC, right central; RP, right posterior). Bottom A and bottom B: topographic maps comparing grammatical status/ACS effects across TEST1 and TEST2 and across TEST1 and TEST3.

Table 1

Correlation between learning-related ERP changes and ERS (Delta, Theta, Alpha) at the onset of baseline (TEST1) sequences (Ns = non-significant).

	Visual ERP 400–600 ms			Auditory ERP 500–700 ms		
	Anterior	Central	Posterior	Anterior	Central	Posterior
ERPs TEST 1–3 ^a	Yes	Yes	Yes	No	No	Yes
Delta with ERPs	Ns	Ns	Ns	Ns	Ns	Ns
Theta with ERPs	$r = .640$ $p = .001$ (.003) ^b	$r = .653$ $p = .001$ (.003) ^b	Ns	Ns	Ns	$r = .514$ $p = .014$ (.042) ^b
Alpha with ERPs	Ns	Ns	Ns	Ns	Ns	Ns

^a Yes = significant changes in late ERPs from TEST1 over TEST3; No = non-significant changes.

^b Significance of the correlation when correcting for multiple comparisons (3 frequency bands).

our goal, since our focus was on whether the ERPs of both modalities would correlate with the same variable. In any case, our findings support our initial hypothesis and suggest that individual levels of working memory maintenance and/or attention towards experimental sequences increase EEG indices of learning. The posterior topography of theta ERS in our study favours a working memory interpretation, since attention has been more related to frontal theta (Deiber et al., 2007). The fact that we saw increased theta synchronization in the visual modality may indicate increased working memory demands for visual temporal compared to auditory presentation, strengthening the idea of non-overlapping working memory networks subtending structural learning in the two modalities.

Adding to the latency and topography differences we saw between the late components of visual vs. auditory grammars (earlier and more anterior component for visual), ERP waveforms showed two other modality effects. First, the visual modality did not show any early posterior negativity (100–450 ms), analogous to the one we found in the auditory experiment (Silva et al., 2017). This early negativity was then affected by interactions between grammatical status and ACS, and we raised the hypothesis that it could be a marker of non-structural processes in AGL, preceding but not competing with structural processes indexed by the P600 (see Bahlmann et al., 2006; Friederici et al., 2002). The fact that we did not see this early negativity in response to the learning of the visual grammar suggests that this ERP is not an essential component of the response to structural violations, thus strengthening our previous interpretation. Second, we did not see any cross-test ACS effects in the ERP waveforms (see Fig. 4A). In the auditory experiment, we saw baseline ACS effects that were cancelled after exposure. This was an atypical response to ACS features, since effects should come after exposure, but it nevertheless indicated some sensitivity to surface features. The fact that we did not see any change in ACS effects in the visual modality is possibly related to the heavier memory constraints that temporal presentation formats pose on the visual modality, impairing memory for bigrams and trigrams. This would be consistent with the increased theta synchronization we saw in the visual modality, possibly signalling increased working memory demands.

Overall, neither the absence of an early negativity, nor the absence of cross-test ACS effects in the visual modality threaten the core finding of structural learning of visual grammars. Therefore, these additional differences do not seem to indicate the presence of different computational principles for implicit structured-sequence learning across modalities. Rather, these differences seem to reflect more superficial influences, and thus they reinforce the idea of modality-specific networks that instantiate shared computational principles.

Our main goal was to characterize the domain of sequence learning in terms of modality effects, and this is why we focused on the subdomain of implicit structured-sequence learning. Given that implicit structured-sequence learning captures syntax-learning mechanisms, abstracted from the phonological, lexico-semantic or pragmatic components of natural language, our findings may also have implications for psycholinguistic approaches. First, they indicate that syntactic processing *per se* engages mechanisms that are shared across modalities, thus strengthening previous findings on natural language (Hagoort and Brown, 2000). Second, we did not observe an analogue of the left anterior negativity found in natural-syntax experiments (LAN: Hagoort and Brown, 2000; Kutas et al., 2006, ELAN: Friederici et al., 1996) in any of the two modalities. This raises the possibility that this negative ERP reflects processes related to the interaction between syntax and other linguistic dimensions. In alternative, it is possible that it requires a high level of syntactic proficiency, or the explicit processing of syntactic materials (see Silva et al., 2017).

4. Conclusion

The coexistence of modality-independence with modality-specificity has been widely noted in the field of statistical learning. The findings of our implicit AGL study showed that this coexistence is valid for structural, rule-based learning. In addition, the present results indicate that differences in memory systems supporting structural learning may be key to explain modality-specific outcomes.

5. Methods and materials

5.1. Participants

Twenty-two healthy, right-handed participants (9 male; mean age \pm SD = 22 + 3.6; mean years of education \pm SD = 15 \pm 2) took part in the experiment. They were age- and schooling-matched with the 22 participants of Silva et al.'s (2017) experiment, and had normal or corrected-to-normal vision. None was taking medication. All participants signed informed consent according to the declaration of Helsinki.

5.2. Stimuli

We used the same right-linear artificial grammar as in our auditory experiment (Silva et al., 2017), with 5 letters (Fig. 5) replacing the 5 CV syllables. Although using printed CV syllables in the visual modality could seem like a better solution to grant the equivalence of items across modalities, each visual CV syllable might be processed as a group of two items, and we did not want that to happen. Since our visual letters vs. auditory syllables were functionally equivalent – they represent fundamental units in the linguistic domain, and the combinations made with each item type would lead to meaningless patterns – this seemed like the optimal choice. From the structural viewpoint, our grammar generated right-linear dependencies and allowed for contiguous repetition of some items (e.g., S and R in the visual format, see Fig. 5). We controlled for the amount of repetitions across Grammaticality (Grammatical vs. Non-grammatical) and ACS (High- vs- low) levels when generating test sequences (see below). Details on this control procedure are presented in Silva et al. (2017).

Sequences generated by the grammar ranged between 5 and 12 items in length. For the exposure phase, we generated 100 (Grammatical, G) sequences. Test sequences ($n = 360$) were manipulated for both grammatical status (grammatical vs. non-grammatical)

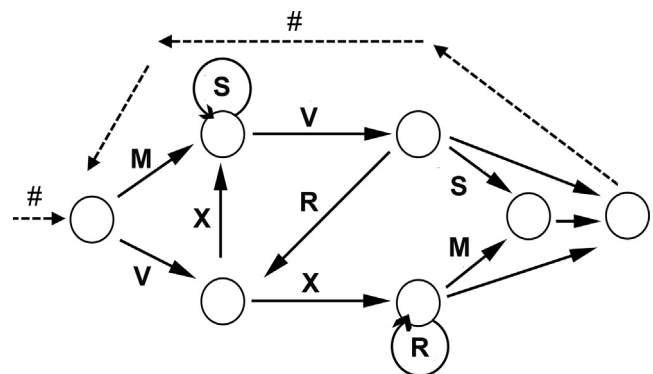


Fig. 5. The five-letter (M, R, X, V, S) grammar used in the visual experiment. MSVRXR is a grammatical sequence, which can be obtained by following the arrows. MSVMXR is a non-grammatical sequence, since there is no path creating this particular sequence. The letters within circular arrows can be repeated, such as S and R: MSSVRXR would be grammatical too.

and Associative Chunk Strength (ACS). The ACS of a test sequence is the average frequency of its composing bi- and trigrams in the exposure stimulus set, and it indicates how often participants were confronted with the surface characteristics of a given sequence (different from structural, rule-related characteristics) during exposure. We started with grammatical sequences: 60×3 G sequences were generated for the 3 test phases. Half these sequences had low ACS, and the other half had high ACS. Using these 3 sets of 60 sequences, we generated 60×3 Non-Grammatical sequences (NG), by replacing 2 letters with a violating letter. We first did a random replacement, and we later checked whether there were actual violations at those points. Overall, violations were inserted in all positions except terminal (i.e., from second to eleventh position, balanced across Grammaticality and ACS levels). We ended up with 3 sets for the 3 test phases (baseline preference, final preference, grammaticality classification), each set comprising 30 grammatical, high ACS sequences (HG), 30 grammatical, low ACS (LG), 30 non-grammatical, high ACS (HNG) and 30 non-grammatical, low ACS (LNG), with a total of 120×3 (360) test sequences. Sequence length ranged between 5 and 12 letters. For further details on the stimulus design, see [Silva et al. \(2017\)](#).

5.3. Procedure

Participants were exposed to the grammar for 5 days (one session a day). They were asked to look at the sequences, try to memorize them and then retype them using the computer keyboard. The task was self-paced. On day 1, right before exposure, they did the baseline preference test (TEST1), containing both G and NG sequences, and where they were asked to indicate whether they liked or disliked each sequence. On day 5, after the last exposure session, they did another preference test (final preference, TEST2) with a different stimulus set. After they completed the preference test, they were informed that all exposure sequences obeyed the same rules, and were asked to classify a new set of G + NG sequences according to grammatical status (grammaticality classification: 'correct'/grammatical vs. 'incorrect'/non-grammatical sequences, TEST3). In the 3 tests, response time was limited to 4 s, and participants were encouraged to respond as fast as they could. In all tasks, letters were presented one-by-one (visuo-temporal presentation) with an inter-onset-interval of 580 ms (290 ms presentation time + 290 ms inter-letter interval). This presentation rate was made slower than the one used in the auditory modality (330 ms) so as to cancel the advantage of the auditory modality in fast presentation rates (see introduction). The EEG was recorded during the three tests.

5.4. EEG recording and processing

The EEG was recorded at 512 Hz, with a 64-channel Brainvision system (Brain Products UK, London). Electrodes were mounted on an elastic cap with an equidistant triangular arrangement. An electrode was placed under the left eye, to measure vertical electro-oculographic (EOG) activity from a bipolar derivation between this and a left prefrontal channel. Bipolar derivations between two bilateral temporal channels provided the horizontal EOG, and muscular activity was monitored with two bilateral posterior channels. The EEG was referenced to the average of the two mastoids. Impedances were kept below 5 K Ω . We processed the EEG data in MatLab with the FieldTrip toolbox ([Oostenveld et al., 2011](#)), both for ERPs and frequency analysis.

The ERPs for the three tests (TEST1, baseline preference, TEST2, final preference, TEST3, grammaticality classification) were computed in relation to the onset of the first violation letter and its non-grammatical counterpart. Epochs ranging from – 300 ms to 900 ms around the trigger point were extracted and scanned for

ocular or muscular artefacts. Trials with artefacts were rejected. Baseline correction was performed using the 100 ms interval preceding the trigger point (violating letter), and data were band-pass filtered from 0.016 Hz to 30 Hz. For further details, see [Silva et al. \(2017\)](#).

For frequency analysis, we considered only TEST1 (baseline) and we locked the EEG to the onset of each sequence. We defined new epochs – from –500 to 1650 ms around this trigger point – because we needed longer baselines for analysing low frequencies, and because we were interested in capturing the activity related to the whole sequence. Our shortest sequence was 5-items long, and it was shorter in the auditory modality (330 ms \times 5 = 1650 ms), so 1650 ms was the largest time frame for trial segmentation we could use. Artefact rejection was done again for these longer epochs, and frequency analysis was done trial-by-trial. We analysed the frequency spectrum of each trial from 1 to 12 Hz, in steps of 0.25 Hz and 50 ms. We used Morlet wavelets for the analysis, since these provide the necessary increased temporal resolution for higher frequencies ([Demiralp et al., 1999](#)). The power spectrum was averaged across all trials of each subject, regardless of condition. We were thus left with one power spectrum per subject (44 frequency bins \times 61 channels \times 33 post-baseline time bins), characterizing his/her ERD/ERS locked to sequence onset during TEST1. We first looked for evidence of entrainment following the approach of Saupe and colleagues ([Saupe et al., 2009](#)), who looked for steady-state responses matching the target frequency in induced (single-trial) activity. Since we found no evidence of entrainment, we investigated ERS/ERD in wider frequency bands. To that end, frequency bins were averaged such that we remained with 3 frequency bands: 1–4 Hz (delta) band, 4–8 Hz (theta) and 8–12 Hz (alpha). All channels were averaged, and so were time bins from 0 to 1650 ms. We were left with 3 ERD/ERS values per subject, characterizing delta, theta and alpha activity time-locked to sequence presentation.

5.5. Statistical analysis

We analysed behavioural data for accuracy, endorsement rates and d-prime. Accuracy per test was compared to chance levels (50%) with one-sample t-tests. Endorsement rates – the amount of sequences classified as liked (in TESTS 1 and 2) or correct (in TEST3) – were analysed with repeated measures ANOVAs with *Test* (TEST1, TEST2, TEST3), *grammatical status* (Gram: G vs. NG sequences) and *ACS* as within-subject factors. The key comparison was made between TEST1 and TEST2, where we looked for significant *Test* \times *Gram* interactions (structure-based learning) as well as possible *Test* \times *ACS* interactions (surface-based learning) or *Test* \times *Gram* \times *ACS* interactions (structure-based learning modulated by surface characteristics). In case of interactions, we analysed *Gram* and *ACS* effects per test. For TEST3, we started by analysing *Gram* and *ACS* effects, and then examined differences between TEST2 and TEST3. Finally, we computed the grammatical status *d*-prime (discrimination between G and NG). For each analysis of behavioural data, we did direct comparisons with the results from our previous auditory AGL experiment ([Silva et al., 2017](#)). Thus, we examined the effects of *Modality* (visual vs. auditory) on accuracy, endorsement rates and *d*-prime using independent-samples t-tests (visual vs. auditory) as well as mixed ANOVAs with *Modality* as between-subjects factor.

For the analysis of event-related potentials, we did not rely on direct comparisons between auditory and visual grammars, given that waveforms differed considerably across modalities, and the relevant time windows did not overlap. This already suggested some level of modality-specificity. Thus, mean voltages per subject were computed at time windows 100–300, 300–400, 400–600, and 600–900 ms, instead of 100–300, 300–450, 500–700, and 700–900

ms of Silva et al. (2017). As in the auditory experiment, we ran repeated-measures ANOVAs, with comparisons of interest based on the factors grammatical status (two levels, G vs. NG) and local subsequence familiarity (ACS, two levels, H and L). We used the same six regions of interest as in Silva et al. (2017, Fig. 4): caudality (Caud) entered the analysis with three levels (Anterior, Central, Posterior) and laterality (Lat) with two (Left, Right). We analysed the main effects of grammatical status, ACS, and grammatical status \times ACS interactions one test at a time (TEST1, TEST2, TEST3). Additional comparisons (Test \times grammatical status, Test \times ACS, Test \times grammatical status \times ACS) were done across tests. Greenhouse-Geisser corrections were applied in case of non-sphericity. Unless otherwise specified, a significance level of .05 was adopted.

For both modalities, subject-level ERD/ERS delta, theta and alpha values on TEST1 were correlated with the learning-related changes in the late (post 400 ms) ERP components, with different topographic areas being considered (correlation with anterior, central and posterior). In addition, we compared delta, theta and alpha ERD/ERS values across modalities. Bonferroni corrections were applied to all multiple comparisons.

Acknowledgements

This work was supported by Max Planck Institute for Psycholinguistics, Donders Institute for Brain, Cognition and Behaviour, Fundação para a Ciência e Tecnologia (PTDC/PSI-PCO/110734/2009; UID/BIM/04773/2013, CBMR 1334; PEst-OE/EQB/LA0023/2013 and UID/PSI/00050/2013). We thank Christian Forkstam for help with the experimental set-up.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.brainres.2018.02.020>.

References

- Aftanas, L.I., Golosheikine, S.A., 2001. Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation. *Neurosci. Lett.* 310 (1), 57–60. [https://doi.org/10.1016/S0304-3940\(01\)02094-8](https://doi.org/10.1016/S0304-3940(01)02094-8).
- Asada, H., Fukuda, Y., Tsunoda, S., Yamaguchi, M., Tonoike, M., 1999. Frontal midline theta rhythms reflect alternative activation of prefrontal cortex and anterior cingulate cortex in humans. *Neurosci. Lett.* 274 (1), 29–32. [https://doi.org/10.1016/S0304-3940\(99\)00679-5](https://doi.org/10.1016/S0304-3940(99)00679-5).
- Attaheri, A., Kikuchi, Y., Milne, A.E., Wilson, B., Alter, K., Petkov, C.I., 2015. EEG potentials associated with artificial grammar learning in the primate brain. *Brain Lang.* 148 (Supplement C), 74–80. <https://doi.org/10.1016/j.bandl.2014.11.006>.
- Bahlmann, J., Gunter, T.C., Friederici, A.D., 2006. Hierarchical and linear sequence processing: an electrophysiological exploration of two different grammar types. *J. Cognitive Neurosci.* 18 (11), 1829–1842. <https://doi.org/10.1162/jocn.2006.18.11.1829>.
- Baker, C.I., Olson, C.R., Behrmann, M., 2004. Role of attention and perceptual grouping in visual statistical learning. *Psycholog. Sci.* 15 (7), 460–466. <https://doi.org/10.1111/j.0956-7976.2004.00702.x>.
- Batterink, L.J., Reber, P.J., Neville, H.J., Paller, K.A., 2015. Implicit and explicit contributions to statistical learning. *J. Memory Lang.* 83, 62–78. <https://doi.org/10.1016/j.jml.2015.04.004>.
- Christiansen, M.H., Conway, C.M., Onnis, L., 2012. Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Lang. Cognitive Process.* 27 (2), 231–256. <https://doi.org/10.1080/01690965.2011.606666>.
- Conway, C.M., Christiansen, M.H., 2005. Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J. Exp. Psychol.: Learn., Memory, Cognition* 31 (1), 24–39. <https://doi.org/10.1037/0278-7393.31.1.24>.
- Conway, C.M., Christiansen, M.H., 2006. Statistical learning within and between modalities pitting abstract against stimulus-specific representations. *Psycholog. Sci.* 17 (10), 905–912. <https://doi.org/10.1111/j.1467-9280.2006.01801.x>.
- Conway, C.M., Christiansen, M.H., 2009. Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *Eur. J. Cognitive Psychol.* 21 (4), 561–580. <https://doi.org/10.1080/09541440802097951>.
- Conway, C.M., Pisoni, D.B., Kronenberger, W.G., 2009. The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Curr. Directions Psychol. Sci.* 18 (5), 275–279. <https://doi.org/10.1111/j.1467-8721.2009.01651.x>.
- Deiber, M.-P., Missonnier, P., Bertrand, O., Gold, G., Fazio-Costa, L., Ibañez, V., Giannakopoulos, P., 2007. Distinction between perceptual and attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics. *J. Cognitive Neurosci.* 19 (1), 158–172. <https://doi.org/10.1162/jocn.2007.19.1.158>.
- Demiralp, T., Yordanova, J., Kolev, V., Ademoglu, A., Devrim, M., Samar, V.J., 1999. Time-frequency analysis of single-sweep event-related potentials by means of fast wavelet transform. *Brain Lang.* 66 (1), 129–145. <https://doi.org/10.1006/brln.1998.2028>.
- Emberson, L.L., Conway, C.M., Christiansen, M.H., 2011. Timing is everything: changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Q. J. Exp. Psychol.* 64 (5), 1021–1040. <https://doi.org/10.1080/10.1080/17470218.2010.538972>.
- Friederici, A.D., Hahne, A., Mecklinger, A., 1996. Temporal structure of syntactic parsing: early and late event-related brain potential effects. *J. Exp. Psychol. Learn. Memory Cognition* 22 (5), 1219–1248.
- Friederici, A.D., Steinhauer, K., Pfeifer, E., 2002. Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proc. National Acad. Sci.* 99 (1), 529–534. <https://doi.org/10.1073/pnas.012611199>.
- Frost, R., Armstrong, B.C., Siegelman, N., Christiansen, M.H., 2015. Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cognitive Sci.* 19 (3), 117–125. <https://doi.org/10.1016/j.tics.2014.12.010>.
- Gevins, A., Smith, M.E., 2000. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex* 10 (9), 829–839. <https://doi.org/10.1093/cercor/10.9.829>.
- Gevins, A., Smith, M.E., Leong, H., McEvoy, L., Whitfield, S., Du, R., Rush, G., 1998. Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors* 40 (1), 79–91. <https://doi.org/10.1518/001872098779480578>.
- Gevins, A., Smith, M.E., McEvoy, L., Yu, D., 1997. High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex* 7 (4), 374–385. <https://doi.org/10.1093/cercor/7.4.374>.
- Goujon, A., Fagot, J., 2013. Learning of spatial statistics in nonhuman primates: contextual cueing in baboons (*Papio papio*). *Behav. Brain Res.* 247, 101–109. <https://doi.org/10.1016/j.bbr.2013.03.004>.
- Griesmayr, B., Gruber, W.R., Klimesch, W., Sauseng, P., 2010. Human frontal midline theta and its synchronization to gamma during a verbal delayed match to sample task. *Neurobiol. Learn. Memory* 93 (2), 208–215. <https://doi.org/10.1016/j.nlm.2009.09.013>.
- Grunwald, M., Weiss, T., Krause, W., Beyer, L., Rost, R., Gutberlet, I., Gertz, H.-J., 1999. Power of theta waves in the EEG of human subjects increases during recall of haptic information. *Neurosci. Lett.* 260 (3), 189–192. [https://doi.org/10.1016/S0304-3940\(98\)00990-2](https://doi.org/10.1016/S0304-3940(98)00990-2).
- Hagoort, P., Brown, C.M., 2000. ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia* 38 (11), 1531–1549. [https://doi.org/10.1016/S0028-3932\(00\)00053-1](https://doi.org/10.1016/S0028-3932(00)00053-1).
- Hagoort, P., Brown, C., Groothusen, J., 1993. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Lang. Cognitive Process.* 8 (4), 439–483. <https://doi.org/10.1080/01690969308407585>.
- Jensen, O., Tesche, C.D., 2002. Frontal theta activity in humans increases with memory load in a working memory task. *Eur. J. Neurosci.* 15 (8), 1395–1399. <https://doi.org/10.1046/j.1460-9568.2002.01975.x>.
- Johansson, T., 2009. Strengthening the case for stimulus-specificity in artificial grammar learning. *Exp. Psychol.* 56 (3), 188–197. <https://doi.org/10.1027/1618-3169.56.3.188>.
- Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29 (2–3), 169–195. [https://doi.org/10.1016/S0165-0173\(98\)00056-3](https://doi.org/10.1016/S0165-0173(98)00056-3).
- Knowlton, B.J., Squire, L.R., 1996. Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *J. Exp. Psychol.: Learn., Memory, Cognition* 22 (1), 169–181. <https://doi.org/10.1037/0278-7393.22.1.169>.
- Kubota, Y., Sato, W., Toichi, M., Murai, T., Okada, T., Hayashi, A., Sengoku, A., 2001. Frontal midline theta rhythm is correlated with cardiac autonomic activities during the performance of an attention demanding meditation procedure. *Cognitive Brain Res.* 11 (2), 281–287. [https://doi.org/10.1016/S0926-6410\(00\)00086-0](https://doi.org/10.1016/S0926-6410(00)00086-0).
- Kutas, M., Van Petten, C., Kluender, R., 2006. Psycholinguistics electrified II (1994–2005). In: Traxler, M., Gernsbacher, M.A. (Eds.), *Handbook of psycholinguistics*, pp. 659–724.
- Lelekov-Boissard, T., Dominey, P.F., 2002. Human brain potentials reveal similar processing of non-linguistic abstract structure and linguistic syntactic structure. *Neurophysiologie Clinique/Clin. Neurophysiol.* 1 (32), 72–84.
- Mecklinger, A., Kramer, A.F., Strayer, D.L., 1992. Event related potentials and eeg components in a semantic memory search task. *Psychophysiology* 29 (1), 104–119. <https://doi.org/10.1111/j.1469-8986.1992.tb02021.x>.

- Meulemans, T., Van der Linden, M., 1997. Associative chunk strength in artificial grammar learning. *J. Exp. Psychol.: Learn., Memory, Cognition* 23 (4), 1007–1028. <https://doi.org/10.1037/0278-7393.23.4.1007>.
- Nozaradan, S., 2014. Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Phil. Trans. R. Soc. B* 369 (1658), 20130393. <https://doi.org/10.1098/rstb.2013.0393>.
- Nozaradan, S., Peretz, I., Missal, M., Mouraux, A., 2011. Tagging the neuronal entrainment to beat and meter. *J. Neurosci.* 31 (28), 10234–10240. <https://doi.org/10.1523/JNEUROSCI.0411-11.2011>.
- Nozaradan, S., Zerouali, Y., Peretz, I., Mouraux, A., 2015. Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cerebral Cortex* 25 (3), 736–747. <https://doi.org/10.1093/cercor/bht261>.
- Oostenfeld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, e156869. <https://doi.org/10.1155/2011/156869>.
- Osterhout, L., Holcomb, P.J., 1992. Event-related brain potentials elicited by syntactic anomaly. *J. Memory Lang.* 31 (6), 785–806. [https://doi.org/10.1016/0749-596X\(92\)90039-Z](https://doi.org/10.1016/0749-596X(92)90039-Z).
- Perruchet, P., Pacton, S., 2006. Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cognitive Sci.* 10 (5), 233–238. <https://doi.org/10.1016/j.tics.2006.03.006>.
- Petersson, K.M., Elfgren, C., Ingvar, M., 1999. Learning-related effects and functional neuroimaging. *Human Brain Mapping* 7, 234–243. [https://doi.org/10.1002/\(SICI\)1097-0193\(1999\)7:4<textless234::AID-HBM2_textgreater3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0193(1999)7:4<textless234::AID-HBM2_textgreater3.0.CO;2-O).
- Pfurtscheller, G., Lopes da Silva, F.H., 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110 (11), 1842–1857. [https://doi.org/10.1016/S1388-2457\(99\)00141-8](https://doi.org/10.1016/S1388-2457(99)00141-8).
- Pothos, E.M., 2007. Theories of artificial grammar learning. *Psycholog. Bull.* 133 (2), 227–244. <https://doi.org/10.1037/0033-2909.133.2.227>.
- Reber, A.S., 1967. Implicit learning of artificial grammars. *J. Verbal Learn. Verbal Behav.* 6 (6), 855–863. [https://doi.org/10.1016/S0022-5371\(67\)80149-X](https://doi.org/10.1016/S0022-5371(67)80149-X).
- Reber, A.S., 1989. Implicit learning and tacit knowledge. *J. Exp. Psychol.: Gen.* 118 (3), 219–235. <https://doi.org/10.1037/0096-3445.118.3.219>.
- Riecke, L., Sack, A.T., Schroeder, C.E., 2015. Endogenous delta/theta sound-brain phase entrainment accelerates the buildup of auditory streaming. *Curr. Biol.* 25 (24), 3196–3201. <https://doi.org/10.1016/j.cub.2015.10.045>.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Saupe, K., Schröger, E., Andersen, S.K., Müller, M.M., 2009. Neural mechanisms of intermodal sustained selective attention with concurrently presented auditory and visual stimuli. *Front. Human Neurosci.* 3. <https://doi.org/10.3389/neuro.09.058.2009>.
- Selchenkova, T., François, C., Schön, D., Corneillie, A., Perrin, F., Tillmann, B., 2014a. Metrical presentation boosts implicit learning of artificial grammar. *Plos One* 9 (11), e112233. <https://doi.org/10.1371/journal.pone.0112233>.
- Selchenkova, T., Jones, M.R., Tillmann, B., 2014b. The influence of temporal regularities on the implicit learning of pitch structures. *Q. J. Exp. Psychol.* 67 (12), 2360–2380. <https://doi.org/10.1080/17470218.2014.929155>.
- Silva, S., Folia, V., Hagoort, P., Petersson, K.M., 2017. The P600 in implicit artificial grammar learning. *Cognitive Sci.* 41 (1), 137–157. <https://doi.org/10.1111/cogs.12343>.
- Tabullo, Á., Sevilla, Y., Segura, E., Zanutto, S., Wainelboim, A., 2013. An ERP study of structural anomalies in native and semantic free artificial grammar: evidence for shared processing mechanisms. *Brain Res.* 1527 (Supplement C), 149–160. <https://doi.org/10.1016/j.brainres.2013.05.022>.
- Witt, A., Puspitawati, I., Vinter, A., 2013. How explicit and implicit test instructions in an implicit learning task affect performance. *PLOS ONE* 8 (1), e53296. <https://doi.org/10.1371/journal.pone.0053296>.
- Zimmerer, V.C., Cowell, P.E., Varley, R.A., 2014. Artificial grammar learning in individuals with severe aphasia. *Neuropsychologia* 53 (Supplement C), 25–38. <https://doi.org/10.1016/j.neuropsychologia.2013.10.014>.