Age-Related Differences in Auditory Statistical Learning of Complex Speech Streams: Evidence From ERPs

Ana Paula Soares (✉ asoares@psi.uminho.pt)

University of Minho

Francisco-Javier Gutiérrez-Domínguez

University of Minho

Alexandrina Lages

University of Minho

Helena M. Oliveira

University of Minho

Margarida Vasconcelos

University of Minho

Luis Jiménez

University of Santiago de Compostela

Research Article

Keywords: statistical learning, developmental changes, 'words' predictability, implicit learning, explicit learning, electrophysiological correlates

Posted Date: August 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-737920/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Statistical learning (SL), the ability to pick up regularities in the sensory environment, is a fundamental skill that allows us to structure the world in a regular and predictable way. Although extensive evidence has been gathered from children and adults, the changes that SL might undergo throughout development remain contentious, particularly in the auditory modality (aSL) with linguistic materials. Here, we collected Event-Related Potentials (ERPs) while 5 years-old children and young adults (university students) were exposed to a complex speech stream in which eight three-syllable nonsense words - four with high (1.0) and four with low (.50) transitional probabilities were embedded to further examine how aSL works under less predictable conditions and to enhance age-related differences in the neural correlates of aSL. Moreover, to ascertain how previous knowledge of the to-be-learned regularities might also affect the results, participants performed the aSL task, firstly, under implicit and, subsequently, under explicit conditions. Although behavioral signs of aSL were observed only for adult participants, ERP data showed evidence of aSL in both groups, as indexed by modulations in the N100 and N400 components. A detailed analysis of the neural responses suggests, however, that adults and children rely on different mechanisms to assist aSL.

Introduction

Although a large number of studies have shown that the ability to extract regularities from the sensory environment, an ability known as statistical learning (SL¹), is observed in young children [e.g.,¹⁻⁶] and adults [e.g.,⁷⁻¹⁰], little is known about how this ability, assumed to play a central role in human cognition [see¹¹], changes throughout development.

This occurs, at least in part, because early works in SL, as well as in the implicit learning (IL) related field [see¹² for a discussion], claimed that SL/IL is an early-maturing ability that remains quite stable across development as no differences in performance had been observed between children and adults [e.g.,^{8,13,14}]. However, a growing body of research conducted in the last decade has challenged this view by showing that SL/IL improves with age [see^{15,16} for reviews]. Moreover, recent works suggest that the developmental trajectory of this ability might not be the same across sensory modalities and types of stimuli [e.g.,^{5,6,17}]. For instance, Raviv and Arnon⁵, using auditory syllables and visual figures in visual and auditory SL tasks (vSL and aSL, respectively) modeled from Saffran et al's.¹ pioneering work, showed that while vSL improved from 5-to-12 years-old children, aSL did not, which could account for the mixed results in the literature showing age differences for vSL [e.g.,^{3,18}], but not for aSL [e.g.,⁸; see however, Emberson et al.⁴ for SL improvements in both modalities]. Nevertheless, in a subsequent work, Shufaniya and Arnon⁶ showed that the absence of age differences in the aSL was not due to the sensory modality *per se* but rather to the type of stimuli. Indeed, using non-linguistic auditory materials (i.e., familiar sounds) instead of auditory syllables, the authors found that vSL and aSL skills improved in the same age range. These findings suggest that SL is not age-invariant, as claimed by earlier works, except for auditory linguistic materials, which is also consistent with other works claiming that infancy is a

privileged time for language acquisition [e.g.,¹⁹⁻²²], as adults are not better than children at learning new languages due, probably, to competition with other mechanisms that appear later in development [see²³⁻²⁵].

It is worth noting, however, these findings were obtained using a two-alternative forced-choice (2-AFC) task, which has been subject to increasing criticism [see^{3,15,26-28} for an extended discussion]. Indeed, in this task participants are typically asked to identify which of two three-syllable nonsense words (e.g., 'tokibu' vs. 'kipopi') mostly resembles the ones heard in a previous familiarization phase in which, unbeknownst to them, syllables are grouped into triplets that always appear together in the stream (e.g., 'tokibutipolugopilatokibu'). If performance is greater than chance, SL is assumed to have occurred, as only the track of the statistical regularities embedded in the auditory stream presented during the familiarization phase allows a correct 'word' discrimination. Despite the widespread use of this task to test aSL, it is important to emphasize that it relies strongly on explicit judgments, which depend largely on other high-order cognitive skills (e.g., decision-making processes) that could be not fully developed in children of this age [see²⁹ for similar arguments]. Thus, further research using other tasks and/or methodologies is required to get a deeper understanding of how aSL with linguistic materials might change throughout development. Moreover, it is also important to note that most studies tested aSL using three-syllable nonsense words presenting the same level of predictability, i.e., transitional probabilities (TP) of 1.00, which might also contribute to limit not only our understanding of the conditions under which aSL works, but, importantly, the chances of age-related differences in aSL to be observed [see, however,^{9,30-33} for works using triplets with different TPs]. Note that a triplet with a TP of 1.00 means that a given stimulus (e.g., a syllable) is always followed by another stimulus (e.g., another syllable) in the stream, thus reducing the variance along with the aSL ability can be measured. Besides, as Soares et al.⁹ recently pointed out, removing all signs of uncertainty from the input does not mimic what occurs in natural languages, which also reduces the possibility of aSL results to be related with other linguistic outcomes. Indeed, in natural languages, syllables, as well as other linguistic units (e.g., phonemes, morphemes) do not follow each other with 100% of certainty, as the same syllable can occur in many different words at different syllable positions such as the syllable /*cur*/ in words as /*cur*.va.ture/, /in.*cur*.sion/ or /re.oc.*cur*/. Additionally, the 2-AFC task is an offline post-learning task, measuring the *result* of the learning that has occurred in the previous familiarization phase, and not the *processes* underlying that result [see^{9,28,34} for an extended discussion]. Thus, it is possible that even though children and adults might not differ in terms of aSL outcomes with linguistic materials, they might nevertheless differ in the mechanisms they recruit to reach those outcomes.

This work aimed to directly analyze these issues by examining the electrophysiological correlates (ERP) elicited during the familiarization phase of an aSL task with linguistic materials (syllables) while 5 years-old children and young adults (university students) were exposed to the repetition of eight three-syllable nonsense words - four with TPs of 1.00 (high-TP 'words') and four with TPs of .50 (low-TP 'words') - presented in two different blocks to further examine how ERPs might change as a function of words' predictability and length of exposure. Furthermore, to examine the role that previous knowledge of the to-

be-learned-regularities might play on aSL - an issue that is gaining increased interest given the proposal that previous linguistic knowledge might interfere with SL outcomes [see^{35,36}] - participants performed the aSL task, firstly, under implicit and, subsequently, under explicit conditions to minimize the role of individual differences in performance [see³² and also⁹ for a similar procedure]. Behavioral data were also collected from the standard 2-AFC task presented after the familiarization phase of each of the aSL tasks.

Previous studies analyzing the electrophysiological correlates of aSL during the exposure to the continuous stream showed the N100 component and particularly the N400 component as indexes of words' segmentation and of the emergence of a pre-lexical trace of 'words' in the brain [e.g.,^{9,37-43}]. It is relevant to note that the vast majority of previous ERP studies did not provide information regarding the changes that the neural correlates of aSL might undergo as the exposure to the input stream unfolds [for exceptions, see^{9,34,37,41}]. This is particularly important because children and adults might differ in the amount of exposure they need to unravel the statistical structure of the input, showing potential neurofunctional differences that might not be detected unless the length of exposure is taken into account. Therefore, if children and adults differ in the processes underlying aSL with linguistic materials, even though differences in the 2-AFC performance might not be observed as in previous studies [e.g.,^{5,6,8}], we expect modulations in the N100 and N400 ERP components to be observed across groups of participants for the high- and low-TP 'words' under implicit and explicit conditions as exposure to the speech stream unfolds.

Results

Behavioral

The mean percentages of correct responses in the implicit and explicit 2-AFC tasks for the high- and low-TP 'words' in each group of participants are presented in Table 1.

The results from the one-sample *t*-tests against chance level in the group of children showed that the 2-AFC performance did not differ from chance in either of the aSL tasks and type of 'words' (all *ps* > .115). In the adult group, the results showed that 2-AFC performance exceeded the chance level for the low-TP 'words', $t(20) = 2.264$, $p = .015$ in the implicit condition, and for the high-TP, $t(20) = 2.592$, $p = .017$, and low-TP 'words', $t(20) = 3.543$, $p = .002$ in the explicit condition. These findings indicated that, conversely to children, adults showed behavioral signs of SL in both aSL tasks and type of 'words' except for the high-TP 'words' in the implicit condition.

Moreover, the results obtained from the repeated measures ANOVA showed a main effect of group, $F(1, 39) = 4.791$, $p = .035$, $\eta_p^2 = .109$ indicating, unsurprisingly, that adults outperformed children (58.1% vs. 50.8%, respectively). A main effect of task was also observed, $F(1, 39) = 11.979$, $p = .001$, $\eta_p^2 = .235$. This effect indicated that participants showed better performance in the aSL task performed under explicit than implicit conditions (58.8% vs. 50.2%, respectively). Furthermore, the two-fold interaction group * type

of 'word' was marginally significant, $F(1, 39) = 4.023$, $p = .052$, $\eta_p^2 = .094$. This effect showed that while differences across the type of 'words' reached a marginal significance in the group of adults, which recognized significantly better low-TP than high-TP 'words' (62.0% vs. 54.3%, $p = .068$), in the group of children the difference across the type of 'word' failed to reach statistical significance (48.8% vs. 52.9%, $p = .338$). Moreover, the interaction also revealed that differences between children and adult participants only reached statistical significance for the low-TP 'words' with adults outperforming children (62.0% vs. 48.8%, $p = .004$). For the high-TP 'words' the difference failed to reach statistical significance (54.3% vs. 52.9%, $p = .765$).

ERP Data

N100

Children. In this ERP component, the ANOVA showed a main effect of block, maximal at the fronto-central ROI, $F(1,19) = 5.22$, $p = .034$, $\eta_p^2 = .215$ indicating that regardless of the aSL task and type of 'word', children showed a larger N100 amplitude in the second half than the first half of the aSL tasks (see Figure 3). No other main or interaction effects reached statistical significance.

Adults. Maximal effects were observed at the central ROI in this ERP component. The ANOVA showed a main effect of aSL task, $F(1,20) = 10.58$, $p = .004$, $\eta_p^2 = .346$, indicating an enhancement in the aSL task performed under explicit compared to the implicit condition. A main effect of block was also observed, $F(1,20) = 5.16$, $p = .034$, $\eta_p^2 = .205$, indicating, as in the case of children participants, a larger N100 amplitude in the second half than in the first half of the aSL tasks (Figure 3). In addition, the two-fold aSL task * type of 'word' interaction reached a marginally statistical significant level, $F(1,20) = 4.31$, $p = .051$, $\eta_p^2 = .177$. In this interaction, the effect of task was found for the high -TP 'words', showing larger N100 amplitudes in the aSL task performed under explicit than implicit condition ($p = .001$). In addition, the effect of type of 'word' was found under the explicit condition, with larger N100 amplitude for the high- than for the low-TP 'words' ($p = .039$). Figure 1 depicts that effect.

N400

Children. Maximal effects were observed at the central ROI in this ERP component. The ANOVA showed a main effect of aSL task, $F(1,19) = 8.23$, $p = .010$, $\eta_p^2 = .302$, indicating an enhancement in the aSL task performed under explicit than implicit condition (Figure 2). In addition, the three-fold type of 'word' * aSL task * block was also significant, $F(1,19) = 4.65$, $p = .044$, $\eta_p^2 = .197$. Figure 3 depicts this interaction.

Pairwise comparisons showed that the effect of aSL task resulted in an enhancement of the N400 component under explicit than implicit condition, observed for low-TP 'words' in the first half of the task ($p = .030$), while in the second half of the task that effect was observed for high-TP 'words' ($p = .027$). Moreover, a significant effect of type of 'word' was found in interaction with instructions and block effects, showing a larger amplitude for low-TP 'words' than high-TP 'words' in the first half of the explicit

task ($p = .041$). Finally, the effect of block reached significance for low-TP 'words' under explicit instructions, resulting in a larger N400 amplitude in the first half than in the second half ($p = .022$).

Adults. Adults showed a significant main effect of type of 'word' at central ROI, $F(1,20) = 6.88$, $p = .016$, $\eta_p^2 = .256$, showing a larger N400 for the high-TP than for the low-TP 'words' regardless of the aSL task (Figure 2). Moreover, a main effect of block was also observed, $F(1,20) = 8.15$, $p = .010$, $\eta_p^2 = .289$, indicating an enhancement in this component in the second than in the first half of the aSL tasks (Figure 2). No other main or interaction effect reached statistical significance.

Discussion

The present study aimed to examine age-related differences in the ERP correlates of aSL with linguistic materials. Five-years old children and young adults (college students) were exposed to a speech stream in which the statistical regularities of three-syllable nonsense words, with high- and low-predictability, had to be extracted through passive exposure (implicit condition) or after the nonsense words had been explicitly taught to the participants (explicit condition). With this design we aimed to further analyze if children and adults rely on the same core mechanisms to extract word-like units from the exposure to a more complex speech stream, and, ultimately, to shed light on the changes that SL might undergo throughout development since the few behavioral studies conducted so far have led to inconsistent results. Behavioral data were also collected from a standard 2-AFC task performed after the familiarization phase of each aSL task.

The behavioral data showed that the 2-AFC performance only differed from chance for adult participants in both aSL tasks and for both types of 'words', except for the high-TP 'words' in the aSL task performed under implicit conditions. In the group of children, performance did not differ significantly from chance regardless of words' predictability and the conditions under which they were presented. The absence of reliable signs of SL in the group of 5-years old children is not new. In Raviv and Arnon⁵, and also Shufaniya and Arnon's⁶ works, the authors did not find evidence of SL for children under 6 years of age, which led them to exclude these children from the analyses to avoid the age differences that simply reflected the move from a chance to an above-chance 2-AFC performance [see also¹⁷] for similar findings with the artificial grammar learning paradigm). In our case, the absence of reliable signs of aSL in the 2-AFC tasks performed by children might also stem from the complexity of the speech stream used, which entailed not only a higher number of 'words' (eight) than used in previous works (typically from four to six 'words'), but also 'words' that were more diverse on their composition. Indeed, against the vast majority of aSL studies conducted so far, which only used 'words' with high levels of predictability (TPs = 1.0), here we have also used 'words' with lower levels of predictability (TP = .50), which might have contributed to make the extraction of the statistical regularities embedded in the stream harder to achieve [see^{9,35}]. Moreover, the joint presentation of high- and low-TP 'words' in the same auditory stream might also have contributed to make the extraction of the regularities embedded in the input much more challenging as

‘words’ could be made up, or not, of unique syllables, which could have provided conflicting cues for ‘words’ segmentation.

Nevertheless, in the adults’ group, the behavioral results indicate a moderate, but reliable, level of learning in both aSL tasks although, interestingly, performance was better for the low- than for the high-TPs ‘words’. This unexpected result can be accounted for if we attend to an inevitable consequence that the TPs manipulation in our study, as well as in all other studies that used ‘words’ with different levels of predictability [e.g.,^{9,32,35}], entails. Indeed, even though high- and low-TP ‘words’ were presented exactly the same number of times during familiarization to control for ‘word’ frequency effects, the fact that high-TP ‘words’ comprise unique syllables - unlike low-TP ‘words’, whose syllables occurred in different ‘words’ in different syllable positions - made the learning of the low-TP ‘words’ to involve not only the encoding of a smaller number of syllables than the high-TP ‘words’ (4 vs. 12, respectively) but, importantly, syllables that occurred three times more frequently than the syllables of the high-TP ‘words’. The difference in the number of the different syllables and in the number of times each syllable appeared in the stream might have led participants when asked to decide which of two stimuli ‘sounded more familiar’ based on the stream presented before, to choose the low-TP ‘words’ against other possible candidates, as the low-TP ‘words’ were made of syllables that occurred more often in the stream, generating certainly a higher level of familiarity.

The absence of significant learning effects in the group of children, even when explicit instructions about the to-be-learned regularities were presented before exposure, converge with other works claiming that the 2-AFC task is not well-suited to test SL, particularly in children of this age [see^{15,29}]. As mentioned in the introduction section, the 2-AFC task requires participants to make explicit judgments about regularities that are expected to be acquired implicitly, which might be hard to accomplish by 5-years old children. Note that in each 2-AFC trial participants have to process two stimuli (a ‘word’ and a foil) that were very similar and to decide which one belongs to the artificial language they heard before. This requires working memory capacities that are still under development at this age [see⁴⁴] and that might also justify why adults, but not children, seem to take advantage of the previous knowledge of the to-be-learned regularities to boost 2-AFC performance [e.g.,^{3,45}].

The ERP data showed, however, modulations in the N100 and N400 components, taken as the neural signatures of SL in the brain [e.g.,^{37,40,46}] in both groups of participants. Specifically, the ERPs recorded during familiarization showed larger amplitudes in the N100 during the second half of the aSL tasks relative to the first, in both adults and children, suggesting this ERP component to be sensitive to the amount of exposure to the statistical structure embedded in the auditory stream. Previous research has considered the N100 a putative ‘marker’ of online segmentation in the brain [e.g.,^{37,46,47}], but the literature still presents divergent findings regarding how N100’s amplitude is modulated by specific factors [e.g.,^{9,38,40,48}]. Our findings are in line with previous research showing enhancements in the N100 in the last part of the familiarization phase [see³⁷], and suggest the N100 to index transient effects that change as learning progresses. More importantly, it also suggests that an early brain mechanism of SL is

already present in 5 years-old children for the decoding of linguistic input, which agrees with works claiming that SL is an early-maturing skill supporting language acquisition [e.g.,^{1,8}; see⁴⁹ for a review], even though behavioral evidence of SL might be only noticed at a later developmental stage, at least when using the 2-AFC task, which requires the maturation of other cognitive skills.

Additionally, in adults, we found evidence of a larger N100 when subjects were provided with prior knowledge about the 'words' they were exposed to during familiarization. Given the early, and sensory, nature of this neural component, this might indicate that explicit conditions facilitate the extraction of the regularities, particularly for those 'words' presenting high-TPs, triggering an enhancement in the this ERP component. This result is also consistent with other findings indicating that the development of executive functions could interfere with aSL results [e.g.,²³⁻²⁵]. Since, behaviorally, we found aSL to be increased when adults were provided with explicit instructions, we might as well consider the N100 as a highly sensitive neural proxy of aSL. In children, we did not find evidence that the amplitude of the N100 was modulated by previous knowledge about the 'words', only by mere exposure to the stream. However, ERP data showed that explicit instructions have influenced later stages, which could indicate that children's developing brains would use such information in further processing. In fact, the effects found on N400, related to the activity of specialized language circuits, seem to support this idea.

In the N400 component, we also found larger amplitudes for high- than low-TPs 'words' in the group of adult participants regardless of the aSL task performed and amount of input exposure. This result replicates the one of Soares et al.'s⁹ work and suggests that the N400 component can be taken as a reliable index of the emergence of a pre-lexical trace of 'words' in the brain, as modulations in this component reflect the degree to which representations of linguistic sequences are being integrated into memory and language-related brain circuits [see⁵⁰]. The fact that high-TPs 'words' elicited larger N400 amplitude indicates not only that this type of linguistic sequences is more easily extracted from the input [see also^{9,35} for similar conclusions], but it also corroborates the brain's ability to effectively decode the structure of continuous streams of syllables, distinguishing highly probable from less probable sequences, even when 'extra' (metalinguistic) information about the to-be-learned regularities was not provided.

In contrast, the N400 in children was found enhanced by the effect of explicit instructions. Under explicit conditions, the 'word' predictability effect showed that low-TPs 'words' elicited larger N400 amplitudes than high-TP 'words' in the first part of the aSL task, contrary to what was found for adult participants. This finding suggests that, unlike adults, children seem to guide the extraction of the statistical regularities embedded in the input by computing the frequency with which each syllable occurred in the stream rather than syllables' TPs. Indeed, the fact that high-TP 'words' entailed a higher number of syllables, that occurred three-times less often in the stream than low-TP 'words', might have make the immature cognitive system to use a more 'economic' strategy (syllable frequency instead of syllables' TPs) to predict the upcoming events, and to use that knowledge to facilitate the learning of lower frequency elements later on [see^{36,51,52}]. The fact that during the first half of the aSL task presented under

explicit conditions children showed larger N400 amplitudes for low-TP ‘words’, and during the second half of the same task larger N400 amplitudes for high-TP ‘words’ seem to be in accordance with this rationale, although future research should further explore this issue by manipulating the use of different statistics (conditional vs. distributive) in the input.

Conclusion

The present study is, to the best of our knowledge, the first reporting ERP evidence of age-related differences in the mechanisms used by children and adults to extract word-like units in the context of a much more complex and diverse speech stream that mimicked closely what occurs in ‘real’ environments. It highlights the usefulness of the ERP methodology to cope with the limitations of the post-learning behavioral methods to test SL (particularly the 2-AFC task with children under 6 years of age). It also sheds light on how the mechanisms underlying SL might change, not only in different sensory modalities, as previous authors have already highlighted [e.g.,^{4,6,8}], but importantly along different types of statistical predictability. Our results also support the idea we have been claiming in line with other authors [e.g.,³⁴] that online and offline/behavioral SL measures tap into distinct cognitive processes, the first related to how the brain computes statistical regularities embedded in the sensory input, and the second with how people retrieve patterns learned from memory. Future research should thus incorporate online measures to provide new insights on how the cognitive system segments complex speech streams and takes advantage of different regularities embedded in the sensory input at different developmental stages.

Method

Ethics Statement

The study was carried out in accordance with the guidelines of the Declaration of Helsinki and approved by the ethics committee of the local Ethics Committee (University of Minho, SECSH 028/2018). Written informed consent was obtained from each adult participant and from parents/legal representatives in the case of children participants.

Participants

Twenty-four children (13 female, $M_{\text{age}} = 5;7$; range 5;1 to 6;5) from Portuguese kindergarten institutions, and twenty-four students (22 female, $M_{\text{age}} = 20;3$; range 18;1 to 31;2) from the University of Minho participated in the study. All participants were native speakers of European Portuguese, with normal hearing and no reported history of learning or language disabilities and/or neurological problems. All were right-handed, as assessed by the Portuguese adaptation of the Edinburgh Handedness Inventory^{53,54}.

Stimuli

Thirty-two auditory European Portuguese CV syllables (e.g., *'tu'*, *'ga'*, *'ci'*, *'pa'*, *'be'*) evenly distributed over two syllabaries (syllabary A and syllabary B) were used from Soares et al.'s⁹ study to create the 16 nonsense words used in the implicit and explicit versions of the aSL tasks. Syllables were produced and recorded by a native speaker of European Portuguese with a duration of 300 ms each. Syllables in each syllabary were organized into eight three-syllable nonsense words: four high-TP 'words' (TP = 1.0), and four low-TP 'words' (TP = .50). For instance, the nonsense word *'tucida'* from syllabary A and the nonsense word *'todidu'* from syllabary B correspond to high-TP 'words' as the syllables they entail only appear in those 'words' and in that specific syllable positions, while the nonsense word *'dotige'* from syllabary A and the nonsense word *'pitegu'* from syllabary B correspond to low-TP 'words' as the syllables they entail appear in three different 'words' at different (initial, medial, final) syllable positions (*'tidomi'*, *'migedo'*, and *'tepime'*, *'megupi'*, respectively - see Table 2 for other examples).

The nonsense words were concatenated in a continuous 8.4 min stream with the Audacity® software (1999-2019) with a 50 ms interval between syllables. Each nonsense word was repeated 60 times in six blocks of 10 repetitions each (1.4 min per block - see Figure 4 ahead). In each block, the nonsense words were presented binaurally in random order with the restriction that the same nonsense word or the same syllable will never appear consecutively. TPs across 'word' boundaries were .14 for the high-TP 'words' as any of these 'words' could be followed by any other of the remaining seven 'words', and .17 for the low-TP 'words' as any of these 'words' could be only followed by any of six possible 'words' to prevent cases such as *'tidomi'* and *'migedo'* to occur successively in the stream. The speech stream was edited to include in ~10% of the syllables a superimposed chirp sound (a 0.1 s sawtooth wave sound from 450 to 1,450 Hz) to provide participants with a cover task (i.e., a chirp detection task) that ensured adequate attention to the auditory stimuli as in previous studies [e.g.,^{2,9}].

For the 2-AFC task, eight foils of three syllables each were also created for each syllabary (see Table 2). The foils were made up of the same syllables used in the 'words', presented with the same frequency and syllable positions as in the high- and low-TP 'words'. For example, the most frequent syllables used during familiarization from syllabary A (e.g., *'do'*, *'ti'*, *'mi'*, *'ge'*), which appeared three times in different low-TP 'words' (e.g., *'dotage'*, *'tidomi'*, *'migedo'*, *'gemiti'*), were also presented three times in the foils (e.g., *'dobage'*, *'tidemi'*, *'mipedo'*, *'geciti'*), whereas the less frequent syllables (e.g., *'tu'*, *'ci'*, *'da'*, *'bu'*, *'pe'*, *'po'*) which appeared only once in the high-TP 'words' (e.g., *'tucida'*, *'bupapo'*, *'modego'*, *'bibaca'*), were also presented once in the foils (e.g., *'tumica'*, *'bugego'*, *'modopo'*, *'bitida'*). However, conversely to the syllables in the high- and low-TP 'words', the syllables in the foils were never presented together during familiarization (TPs = 0). Note, however, that due to stimuli restrictions (number of syllables in each syllabary and the need to generate sequences of syllables never presented together before), the foils associated with the high-TP 'words' entailed two syllables from the high-TP 'words' and one-syllable from the low-TP 'words'. The same is observed for the foils associated with the low-TP 'words' that entailed two syllables from the low-TP 'words' and one syllable from the high-TP 'words'. Four lists of materials were created to counterbalance syllables across positions and stimuli in each syllabary. Participants in each group were randomly assigned to one list from syllabary A and one list from syllabary B to perform

the aSL under implicit and explicit conditions with the constraint that the same number of participants would complete a given list (six participants *per* list).

Procedure

Data were collected in a shielding cabin at the Psychological Neuroscience Lab (School of Psychology, University of Minho). Participants were firstly presented with the implicit version of the aSL task, and, subsequently, with the explicit version of an analogous aSL task (see Figure 4). In the implicit version, participants were instructed to pay attention to the auditory stream (sequences of syllables) presented at 60 dB SPL via binaural headphones, because occasionally a deviant sound (i.e., a click) would appear, and their task would be to detect it as soon and accurately as possible by pressing the spacebar from the computer keyboard (i.e., to perform a target detection task) as in previous works. Following familiarization, participants were asked to decide as accurately as possible which of two auditory stimuli (one 'word' and one foil) 'sounded more like' the stimuli presented before (i.e., to perform a 2-AFC task). The 2-AFC comprised 16 trials in which each of the 'words' were paired with two different foils.

After a brief interval, participants underwent the explicit version of the aSL task. Again, participants were exposed to a continuous stream made of other eight 'words' (four high-TP and four low-TP 'words') from the syllabary not used in the implicit version of the aSL task. The explicit version followed the same procedure, except that previously to the familiarization phase participants were presented with a training phase in which each of the eight new 'words' was presented individually and participants were asked to repeat each of them correctly before the familiarization phase began [see⁹ for details]. The procedure took about 90 min to be completed per participant. Figure 4 depicts a visual summary of the experimental design.

EEG Data Acquisition and Processing

Data collection was performed in an electric shielded, sound-attenuated room. Participants were seated in a comfortable chair, one meter away from a computer screen. During the familiarization phase, EEG data was recorded with a 64 channels BioSemi Active-Two system (BioSemi, Amsterdam, The Netherlands) according to the international 10-20 system and digitized at a sampling rate of 512 Hz. Electrode impedances were kept below 20 k Ω . EEG was re-referenced off-line to the algebraic average of mastoids. Data were filtered with a bandpass filter of 0.1 - 30 Hz (zero phase shift Butterworth). ERP epochs were time-locked to the nonsense words' onset, from -300 ms to 1,200 ms (baseline correction from -300 to 0 ms). Independent component analyses (ICA) were performed to remove stereotyped noise (mainly ocular movements and blinks) by subtracting the corresponding components. After that, epochs containing artifacts (i.e., with amplitudes exceeding ± 100 μ V) were excluded. EEG data processing was conducted with Brain Vision Analyzer, version 2.1.1. (Brain Products, Munich, Germany).

Data Analysis

Behavioral and ERP data analyses were performed using the IBM-SPSS® (Version 27.0). For behavioral data, the percentage (%) of correct responses was computed for each of the 2-AFC tasks and separately for the high-TP and low-TP 'words' in each group participants (coded as 1 for a correct and 0 for an incorrect response). One-sample *t*-tests against the chance level were conducted in each group of participants to determine whether performance in each aSL task and type of 'words' was significantly different from chance (50%). A mixed analysis of variance (ANOVA) using Group (children vs. adults) as a between-subject factor, and aSL task (implicit vs. explicit) and Type of 'word' (high-TP vs. low-TP) as within-subject factors was conducted to analyze if 2-AFC performance was significantly different across groups and experimental conditions.

Individual ERPs were averaged separately per condition. Grand averages waveforms were then calculated across individuals in each aSL task (implicit and explicit), Type of 'word' (high-TP vs. low-TP), and Block, i.e., considering the first half (block#1, block#2, block#3) and the second half (block#4, block#5, block#6) of the aSL task. Seven participants (four children and three adults) were excluded from the EEG analyses (and also from the behavioral analyses) due to artifact rejection. Based on previous literature, mean amplitudes were measured for the following time windows taken as the neural signatures of words' segmentation and the emergence of a pre-lexical trace of 'words' in the brain: 80-120 ms (N100 component) for both groups, 400-500 ms and 350-450 ms (N400 component) for the group of children and adults, respectively. We chose a slightly later time window for the children group since data inspection revealed a longer latency of the N400 component. This delay of the N400 component in children has already been described in the literature and considered a normative evolutionary phenomenon [see⁵⁵⁻⁵⁷]. To account for the topographical distribution of the abovementioned EEG deflections, mean amplitudes' values were obtained for the topographical regions where amplitudes were maximal: the fronto-central region of interest (ROI; F1, Fz, F2, FC1, FCz, FC2, C1, Cz, and C2) for N100 in children, and the central ROI (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, and CP2) for the rest of the cases.

Both for behavioral and ERP data, main or interaction effects that reached statistical or marginal significance levels in comparisons of interest are reported. The Greenhouse–Geisser correction for nonsphericity was used when appropriate. Post-hoc tests for multiple comparisons were adjusted with Bonferroni correction. Measures of effect size (Eta squared, η_p^2) and observed power (*pw*) for a single effect are reported in combination with the main effects of condition.

Declarations

Data Availability Statement

The datasets generated and analysed in the current study are available at Soares, Ana Paula; Gutiérrez Domínguez, Francisco Javier; Lages, Alexandrina; Oliveira, Helena Mendes; Vasconcelos, Margarida; Jiménez, Luis (2021): DATA: Developmental Differences in the Online Statistical Learning of Complex Speech Streams: An ERP Study with Children and Adults. figshare. Dataset.

<https://doi.org/10.6084/m9.figshare.14865870.v1>

Acknowledgements

This study was conducted at the Psychology Research Centre (PSI/01662), University of Minho, and supported by the Grant POCI-01-0145-FEDER-028212 from the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Science, Technology and Higher Education through national funds, and co-financed by FEDER through COMPETE2020 under the PT2020 Partnership Agreement.

Author contributions

A.P.S. and L.J. conceptualized the manuscript. F.G., A.L. and H.M.O. implemented the experiment and collected the data. F.G. analyzed the data. F.G., M.V., H.M.O. and A.P.S. interpreted the data. A.P.S. and M.V. wrote the first draft of the manuscript. A.P.S., L.J., F.G. and M.V. critically revised it. H.M.O. and A.L. prepared the revised versions. All authors reviewed and approved the final version submitted.

Additional information

The authors declare no competing interests.

References

1. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants., **274**, 1926–1928 <https://doi.org/10.1126/science.274.5294.1926> (1996).
2. Arciuli, J. & Simpson, I. C. Statistical learning is related to reading ability in children and adults. *Cogn. Sci.* **36**, 286–304. <https://doi.10.1111/j.1551-6709.2011.01200.x> (2012).
3. Bertels, J., Boursain, E., Destrebecqz, A. & Gaillard, V. Visual statistical learning in children and young adults: How implicit? *Front. Psychol.* **5**, 1541 <https://doi.org/doi:10.3389/fpsyg.2014.01541> (2015).
4. Emberson, L. L., Misyak, J. B., Shwade, J. A., Christiansen, M. H. & Goldstein, M. H. Comparing statistical learning across perceptual modalities in infancy: An investigation of underlying learning mechanism(s). *Dev. Sci.* **22**, e12847 <https://doi.org/10.1111/desc.12847> (2019).
5. Raviv, L. & Arnon, I. The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Dev. Sci.* **21**, e12593 <https://doi.org/10.1111/desc.12593> (2018).
6. Shufaniya, A. & Arnon, I. Statistical learning is not age-invariant during childhood: Performance improves with age across modality. *Cogn. Sci.* **42**, 3100–3115 <https://doi.org/10.1111/cogs.12692> (2018).
7. Fiser, J. & Aslin, R. N. Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psychol.: Learn., Mem. Cogn.* **28**, 458 <https://doi.org/10.1037/0278-7393.28.3.458> (2002).
8. Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A. & Barrueco, S. Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychol. Sci.* **8**, 101–105 <https://doi.org/10.1111/j.1467-9280.1997.tb00690.x> (1997).

9. Soares, A. P. *et al.* Not all words are equally acquired: Transitional probabilities and instructions affect the electrophysiological correlates of statistical learning. *Front. Hum. Neurosci*, **14**, 577991 <https://doi.org/10.3389/fnhum.2020.577991> (2020).
10. Turk-Browne, N., Jungé, J. & Scholl, B. The automaticity of visual statistical learning. *J. Exp. Psychol. Gen*, **134**, 552–564 <https://doi.org/10.1037/0096-3445.134.4.552> (2005).
11. Bogaerts, L., Frost, R. & Christiansen, M. H. Integrating statistical learning into cognitive science. *J. Mem. Lang*, **115**, 104167 <https://doi.org/10.1016/j.jml.2020.104167> (2020).
12. Christiansen, M. H. Implicit-statistical learning: A tale of two literatures. *Top. Cogn. Sci*, **11**, 468–481 <https://doi.org/10.1111/tops.12332> (2019).
13. Reber, A. S. Implicit learning and tacit knowledge. *J. Exp. Psychol. Gen*, **118**, 219 <https://doi.org/10.1037/0096-3445.118.3.219> (1989).
14. Reber, P. J. The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research., **51**, 2026–2042 <https://doi.org/10.1016/j.neuropsychologia.2013.06.019> (2013).
15. Arnon, I. Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behav. Res. Methods*, **52**, 68–81 <https://doi.org/10.3758/s13428-019-01205-5> (2020).
16. Zwart, F. S., Vissers, C. T. W., Kessels, R. P. & Maes, J. H. Procedural learning across the lifespan: A systematic review with implications for atypical development. *J. Neuropsychol*, **13**, 149–182 <https://doi.org/10.1111/jnp.12139> (2017).
17. Soares, A. P., Silva, R., Faria, F., Oliveira, H. M. & Jiménez, L. Literacy effects on artificial grammar learning (AGL) with letters and colors: Evidence from preschool and primary school children. *Lang. Cogn*, 1–28 <https://doi.org/https://www.10.1017/langcog.2021.12> (2021).
18. Arciuli, J. & Simpson, I. C. Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Dev. Sci*, **14**, 464–473 <https://doi.org/10.1111/j.1467-7687.2009.00937.x> (2011).
19. Kuhl, P. K. Brain mechanisms in early language acquisition., **67**, 713–727 <https://doi.org/10.1016/j.neuron.2010.08.038> (2010).
20. Lenneberg, E. H. The biological foundations of language. *Hosp. Pract*, **2**, 59–67 <https://doi.org/10.1080/21548331.1967.11707799> (1967).
21. Newport, E. L. Maturational constraints on language learning. *Cogn. Sci*, **14**, 11–28 https://doi.org/10.1207/s15516709cog1401_2 (1990).
22. Newport, E. L., Bavelier, D. & Neville, H. J.. Critical thinking about critical periods: Perspectives on a critical period for language acquisition. in *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler* (ed. Dupoux, E.)481–502. (MIT Press, 2001)
23. Ambrus, G. G. *et al.* When less is more: Enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. *J. Mem. Lang*, **114**, 104144 <https://doi.org/10.1016/j.jml.2020.104144> (2020).

24. Smalle, E. H., Panouilleres, M., Szmalec, A. & Möttönen, R. Language learning in the adult brain: Disrupting the dorsolateral prefrontal cortex facilitates word-form learning. *Sci. Rep*, **7**, Article 13966 <https://doi.org/10.1038/s41598-017-14547-x> (2017).
25. Thiessen, E. D., Girard, S. & Erickson, L. C. Statistical learning and the critical period: how a continuous learning mechanism can give rise to discontinuous learning. *Wiley Interdiscip. Rev. Cogn. Sci*, **7**, 276–288 <https://doi.org/10.1002/wcs.1394> (2016).
26. Erickson, L. C., Kaschak, M. P., Thiessen, E. D. & Berry, C. A. Individual differences in statistical learning: Conceptual and measurement issues. *Collabra*, **2**, 1–17 <https://doi.org/10.1525/collabra.41> (2016).
27. Franco, A., Eberlen, J., Destrebecqz, A., Cleeremans, A. & Bertels, J. Rapid serial auditory presentation. *Exp. Psychol*, <https://doi.org/10.1027/1618-3169/a000295> (2015).
28. Siegelman, N., Bogaerts, L. & Frost, R. Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behav. Res. Methods*, **49**, 418–432 <https://doi.org/10.3758/s13428-016-0719-z> (2017).
29. Lukics, K. S. & Lukács, Á. Tracking statistical learning online: Word segmentation in a target detection task. *Acta Psychol*, **215**, e103271 <https://doi.org/10.1016/j.actpsy.2021.103271> (2021).
30. Bogaerts, L., Siegelman, N. & Frost, R. Splitting the variance of statistical learning performance: A parametric investigation of exposure duration and transitional probabilities. *Psychol. Bull*, **23**, 1250–1256 <https://doi.org/10.3758/s13423-015-0996-z> (2016).
31. Koelsch, S., Busch, T., Jentschke, S. & Rohrmeier, M. Under the hood of statistical learning: a statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. *Sci. Rep*, **6**, 19741 <https://doi.org/10.1038/srep1974> (2016).
32. Siegelman, N., Bogaerts, L. & Frost, R. Measuring individual differences in statistical learning: current pitfalls and possible solutions. *Behav. Res. Methods*, **49**, 418–432 <https://doi.org/10.3758/s13428-016-0719-z> (2017).
33. Tsogli, V., Jentschke, S., Daikoku, T. & Koelsch, S. When the statistical MMN meets the physical MMN. *Sci. Rep*, **9**, 5563 <https://doi.org/10.1038/s41598-019-42066-4> (2019).
34. Batterink, L. J. & Paller, K. A. Online neural monitoring of statistical learning. *Cortex*, **90**, 31–45 <https://doi.org/10.1016/j.cortex.2017.02.004> (2017).
35. Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J. & Frost, R. Linguistic entrenchment: Prior knowledge impacts statistical learning performance., **177**, 198–213 <https://doi.org/10.1016/j.cognition.2018.04.011> (2018).
36. Palmer, S. D., Hutson, J., White, L. & Mattys, S. L. Lexical knowledge boosts statistically-driven speech segmentation. *J. Exp. Psychol.: Learn. Mem. Cogn*, **45**, 139–146 <https://doi.org/10.1037/xlm0000567> (2019).
37. Abia, D., Katahira, K. & Okanoya, K. On-line assessment of statistical learning by event-related potentials. *J. Cogn. Neurosci*, **20**, 952–964 <https://doi.org/10.1162/jocn.2008.20058> (2008).

38. Cunillera, T. *et al.* Time course and functional neuroanatomy of speech segmentation in adults., **48**, 541–553 <https://doi.org/10.1016/j.neuroimage.2009.06.069> (2009).
39. Cunillera, T., Toro, J. M., Sebastian-Galles, N. & Rodriguez-Fornells, A. The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Res*, **1123**, 168–178 <https://doi.org/10.1016/j.brainres.2006.09.046> (2006).
40. De Balaguer, D., Toro, R. & Rodriguez-Fornells, J. M. A. & Bachoud-Lévi, A. Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE*, **2**, e1175 <https://doi.org/10.1371/journal.pone.0001175> (2007).
41. François, C., Jaillet, F., Takerkart, S. & Schön, D. Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, **9**, e101340 <https://doi.org/10.1371/journal.pone.0101340> (2014).
42. Mandikal-Vasuki, P. R., Sharma, M., Ibrahim, R. & Arciuli, J. Musicians' online performance during auditory and visual statistical learning tasks. *Front. Hum. Neurosci*, **11**, 114 <https://doi.org/10.3389/fnhum.2017.00114> (2017).
43. Sanders, L. D., Ameral, V. & Sayles, K. Event-related potentials index segmentation of nonsense sounds. *Neuropsychologia* **47**, 1183–1186 <https://doi.org/10.1016/j.neuropsychologia.2008.11.005> (2009).
44. Gathercole, S. E., Pickering, S. J., Ambridge, B. & Wearing, H. The structure of working memory from 4 to 15 years of age. *Dev. Psychol*, **40**, 177 <https://doi.org/10.1037/0012-1649.40.2.177> (2004).
45. Batterink, L., Reber, P. J., Neville, H. & Paller, K. A. Implicit and explicit contributions to statistical learning. *J. Mem. Lang*, **83**, 62–78 <https://doi.org/10.1016/j.jml.2015.04.004> (2015).
46. Sanders, L. D., Newport, E. L. & Neville, H. J. Segmentation nonsense: An event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci*, **5**, 700–703 <https://doi.org/10.1038/nn873> (2002).
47. Sanders, L. D. & Neville, H. J. An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cogn. Brain Res*, **15**, 228–240 [https://doi.org/10.1016/S0926-6410\(02\)00195-7](https://doi.org/10.1016/S0926-6410(02)00195-7) (2003).
48. Buiatti, M., Peña, M. & Dehaene-Lambertz, G. Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses., **44**, 509–519 <https://doi.org/10.1016/j.neuroimage.2008.09.015> (2009).
49. Romberg, A. R. & Saffran, J. R. Statistical learning and language acquisition. *Wiley Interdiscip. Rev. Cogn. Sci*, **1**, 906–914 <https://doi.org/10.1002/wcs.78> (2010).
50. Lau, E. F., Phillips, C. & Poeppel, D. A cortical network for semantics: (de) constructing the N400. *Nat. Rev. Neurosci*, **9**, 920–933 <https://doi.org/10.1038/nrn2532> (2008).
51. Bortfeld, H., Morgan, J. L., Golinkoff, R. M. & Rathbun, K. Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychol. Sci*, **16**, 298–304 <https://doi.org/10.1111/j.0956-7976.2005.01531.x> (2005).
52. Lavi-Rotbain, O. & Arnon, I. Visual statistical learning is facilitated in zipfian distributions., **206**, 104492 <https://doi.org/10.31234/osf.io/xwgpk> (2021).

53. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory., **9**, 97–113 [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4) (1971).

54. Espirito-Santo, H. *et al.* Preliminary validation of the Portuguese Edinburgh Handedness Inventory in an adult sample. *Appl. Neuropsychol. Adult*, **24**, 275–287 <https://doi.org/10.1080/23279095.2017.1290636> (2017).

55. Cummings, A., Čeponienė, R., Dick, F., Saygin, A. P. & Townsend, J. A developmental ERP study of verbal and non-verbal semantic processing. *Brain Res*, **1208**, 137–149 <https://doi.org/10.1016/j.brainres.2008.02.015> (2008).

56. Juottonen, K., Revonsuo, A. & Lang, H. Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect. *Cogn. Brain Res*, **4**, 99–107 [https://doi.org/10.1016/0926-6410\(96\)00022-5](https://doi.org/10.1016/0926-6410(96)00022-5) (1996).

57. Hahne, A., Eckstein, K. & Friederici, A. D. Brain signatures of syntactic and semantic processes during children's language development. *J. Cogn. Neurosci*, **16**, 1302–1318 <https://doi.org/10.1162/0898929041920504> (2004).

Tables

Table 1. *Mean (SD) of the Number (%) of Correct Responses for the High- and Low-TP ‘Words’ in the Implicit and Explicit aSL Tasks per Group of Participants.*

‘Word’	aSL task			
	Implicit		Explicit	
Group	High-TP	Low-TP	High-TP	Low-TP
Children	50.1 (17.17)	45.1 (19.16)	55.7 (15.43)	52.5 (17.54)
Adults	47.1 (17.68)	58.4 (14.46)	61.4 (20.12)	65.5 (20.08)

Table 2. *Three-Syllable Nonsense Words and Foils From Syllabary A and Syllabary B*

Transitional Probabilities	Syllabary	Stimuli	
		'Words'	Foils
High-TP	A	<i>tucida</i>	<i>tumica</i>
		<i>bupepo</i>	<i>bugego</i>
		<i>modego</i>	<i>modopo</i>
		<i>bibaca</i>	<i>bitida</i>
	B	<i>todidu</i>	<i>tomeco</i>
		<i>cegita</i>	<i>cegube</i>
		<i>gapabe</i>	<i>gapita</i>
		<i>bomaco</i>	<i>botedu</i>
Low-TP	A	<i>dotige</i>	<i>dobage</i>
		<i>tidomi</i>	<i>tidemi</i>
		<i>migedo</i>	<i>mipedo</i>
		<i>gemiti</i>	<i>geciti</i>
	B	<i>pitegu</i>	<i>pimagu</i>
		<i>tepime</i>	<i>tepame</i>
		<i>megupi</i>	<i>megipi</i>
		<i>gumete</i>	<i>gudite</i>

Figures

Figure 1

Grand-Averaged Waveforms (Central ROI) and Topographic Maps for Adults and Children (First and Second Blocks Collapsed). Note: Black solid line = explicit, low-TP condition; Black dotted line = explicit, high-TP condition; Red solid line = implicit, low-TP condition; Red dotted line = implicit, high-TP condition. Values of the topographical images range from -0.5 to 1 μ V in adults and from -3 to 3 μ V in children. "IMP" stands for the aSL task performed under implicit conditions, whereas "EXP" for the aSL performed under explicit instructions. Grey shadowed rectangles indicate the analyzed time windows.

Figure 2

Block Effects in N100 and N400 Components Both in the Adult and Children Groups. Note: To assure the clarity of the graphical representation of the Block effects, the remaining conditions of interest were collapsed (i.e., learning condition and words' predictability). Black line = first block; Grey line = second block. Grand-averaged waveforms correspond to central ROI in adults and fronto-central ROI in children. Grey shadowed rectangles indicate the time windows in which the block effect was significant.

Figure 3

Graphical Representation of the N400 Triple Interaction Effect in the Children Group. Note: Grey shadowed rectangles indicate the N400 time window. A. Task effect in the Low-TP condition, in the first block. Black line = Explicit task; Red line: Implicit task. B. Task effect in the High-TP condition, in the second block. Black line = Explicit task; Red line: Implicit task. C. Type of 'word' effect under explicit instructions in the first block. Solid line = Low-TP conditions; Dotted line = High-TP condition. D. Effect of block in Low-TP 'words' under explicit instructions. Black line = first block; Grey line = second block.

Implicit aSL task

Figure 4

Visual Summary of the Experimental Design Note: Panels A to G illustrate the timeline of the experimental procedure in which the implicit and, subsequently, the explicit aSL task were administered. Each aSL task comprised three parts: instructions, familiarization phase, and test phase. Each task was initiated with specific instructions (Panels A and E) that determined the conditions under which the aSL task was performed: A. implicit instructions (i.e., without knowledge of the stimuli or the structure of the stream) or

E. explicit instructions (i.e., with explicit knowledge or pre-training on the 'words' presented in the stream). In the familiarization phase of both tasks (Panels B and F) during which EEG data were collected, participants were presented with a continuous auditory stream of four high-TP and four low-TP 'words', with chirp sounds (depicted as a speaker icon on the Figure) superimposed over specific syllables. The chirp sounds could emerge at any of the three syllable positions of the 'words', which precluded its use as a cue for stream segmentation. During this phase, participants had to perform a chirp detection task. Then, a test phase (Panels C and G) consisting on a 2-AFC task asked participants to indicate which of the two-syllable sequences (a 'word' and a foil) sounded more familiar considering the stream heard on the familiarization phase.