The role of metacognition in recognition of the content of statistical learning

Mikhail Ordin[1,2,&,*], Leona Polyanskaya[1,&],

[&] equal contribution

[1]BCBL – Basque Centre on Cognition, Brain and Language, Mikeletegi 69, 20009, San Sebastián, Spain

[2]Ikerbasque – Basque Foundation for Science, Maria Diaz de Hro 3, 48013, Bilbao, Spain

[*]corresponding author, address for correspondence: Basque Centre on Cognition, Brain and Language, Paseo Mikeletegi 69, San Sebastian, 20009 Spain. E-mail: mordin@bcbl.eu

ABSTRACT

Despite theoretical debate on the extent to which statistical learning is incidental or modulated by explicit instructions and conscious awareness of the content of statistical learning, no study has ever looked into the metacognition of statistical learning. We used an artificial language learning paradigm and a segmentation task that required splitting a continuous stream of syllables into discrete recurrent constituents. During this task, statistical learning potentially produces knowledge of discrete constituents as well as about statistical regularities that are embodied in familiarization input. We measured metacognitive sensitivity and efficiency (using hierarchical Bayesian modelling to estimate metacognitive sensitivity and efficiency) to probe the role of conscious awareness in recognition of constituents extracted from the familiarization input and recognition of novel constituents embodying the same statistical regularities as these extracted constituents. Novel constituents are conceptualized to represent recognition of statistical structure rather than recognition of items retrieved from memory as whole constituents. We found that participants are equally sensitive to both types of learning products, yet subject them to varying degrees of conscious processing during the post-familiarization recognition test. The data point to the contribution of conscious awareness to at least some types of statistical learning content.

INTRODUCTION

Statistical learning is a process for extracting statistical regularities from the environment that enables efficient processing of continuous sensory inputs. One of the tasks that relies on statistical learning is segmenting continuous inputs into discrete constituents (Baldwin et al., 2008; Gómez, Bion, & Mehler, 2011; Hard et al., 2018; Siegelman, Bogaerts, et al., 2019; Siegelman, 2019. It is generally assumed that statistical learning is incidental and happens without awareness and across modalities (Arciuli, Torkildsen, Stevens & Simpson, 2014; Aslin & Newport, 2012; Dienes, Broadbent, & Berry, 1991). However, some empirical evidence suggests that performance can be modulated by attention (Fernandes, Kolinsky, & Ventura, 2010; Toro, Sinnett & Soto-Faraco, 2005), and that conscious focus on a task improves performance (Alamia & Zenon, 2016; Reber Kassin, Lewis, & Cantor, 1980). Surprisingly, despite theoretical tension in regard to how conscious statistical learning may be, studies on metacognition in statistical learning are exceptionally rare.

Metacognition is cognition about cognition. It helps humans to evaluate past decisions, make better future decisions, and monitor their own cognitive processes and cognitive states (e.g., Flavell, 1979; Nelson, 1996). Flavel (1979) and Schraw (1998) suggested that metacognition includes two aspects: knowledge about cognition and regulation of cognition. These components are served by different cognitive processes, referred to as *metacognitive monitoring* and *metacognitive control* (Nelson & L. Narens, 1990; Dunlosky et al., 2007). Monitoring processes track decisions, cognitive states and behaviour in uncertain situations and estimate retrospective confidence associated with cognitive states and past decisions (Kepecs, Uchida, Zariwala, & Mainen, 2008). Control processes guide future behaviour, taking current cognitive states and available evidence about the current environment and past outcomes into account. In this study, we will focus on metacognitive monitoring in statistical learning tasks by exploring the retrospective confidence that humans assigned their decisions in an artificial language learning paradigm.

Efficient metacognitive monitoring is associated with accurate estimation of the degree of uncertainty associated with past and future decisions (Kepecs et al., 2008). It manifests as an individual's ability to estimate the likelihood of error for each conscious decision they make. This estimate is reflected by assigning higher confidence ratings to decisions when the estimated likelihood of making an error is lower, and lower confidence ratings when the estimated likelihood of making an error is higher. Individuals, whose confidence ratings accurately discriminate between correct and incorrect responses are considered to exhibit higher *metacognitive sensitivity*. The same logic applies to different types of decisions (e.g., in different domains, modalities, tasks) made by the same individual. If, for example, a person's confidence ratings better discriminate between correct and incorrect decisions in the visual than the auditory modality, we can say that their metacognitive sensitivity is higher in the visual modality.

In this research, we explore statistical learning mechanisms that operate on a subtype of statistical relations, that is, conditional statistics. Conditional statistics measure the predictive relationships between consecutive events (Harris, 1955; Saffran et al., 1996). The strength of a predictive relationship is measured as the likelihood that an element or event B will occur given that element or event A has just happened (i.e., the transitional probability – TP – between A and B). Frequently, this kind of statistical learning mechanism is explored in the context of splitting a sequence of syllables into word-like units (Saffran, 2001): TPs between syllables within a discrete word are higher than between syllables straddling word boundaries. However, imagine a sequence of syllables, in which syllable triplets ROSENU and PASETI are embedded and recurrently experienced by an individual. Statistical regularities predict SE given either RO or PA with equal probability, and the transition from SE to either TI or NU also happens with equal probability. Thus, while triplets ROSETI and PASENU have not been embedded as whole units in the familiarization input, they could nevertheless be endorsed as eligible constituents because they are statistically congruent with the collective exemplars from the learning input (Endress & Langus, 2017; Endress & Mehler, 2009; Nosofsky & Zaki, 2002; Ordin, Polyanskaya, & Soto, 2020a; Ordin, Polyanskaya, Soto, & Molinaro, 2020; Roediger & McDermott, 1995). The ability to extrapolate the eligibility of these units from learned input would be useful; it would allow the system to learn regularities from a limited set of exemplars, then generalize these to previously unencountered items, or to novel situations that exhibit the previously encountered statistical features. Here, we try to understand

the role of awareness and metacognitive processes in the recognition of old items (encountered and learnt during familiarization) as well as novel items that are statistically congruent with the old ones.

METHODS

Material, participants, and experimental procedure

Here, we use the same dataset as Ordin, Polyanskaya, Soto, & Molinaro (2020) where the participant sample, experimental procedure and materials are detailed. Below, we provide only the information essential for this empirical report.

Participants listened to a stream of 12 recurrent, randomly concatenated triplets (ROSENU, PASETI, etc.). TPs were set to 0.5 between syllables within triplets and to 0.16 between syllables at the boundaries of triplets. Participants were familiarized with this stream for 18 minutes and explicitly instructed to detect and memorize the "words" of this "alien language". The syllables in the embedded triplets (*words*) were also used to construct 12 novel triplets (*phantoms,* e.g., ROSETI, PASENU) that embodied the same TPs as the words. After familiarization, participants performed a dual forced-choice test, in which they listened to a pair of possible word candidates and had to indicate whether the first or the second candidate was a word from the alien language they had been exposed to. Participants were not informed that the words were composed according to rules. After choosing one of the two candidates, participants were next asked to indicate – on a 4-point scale – how sure they were in their response. In the test pairs, we pitted *words* against *phantoms*, *words* against *non-words* (concatenations of syllables that had never co-occurred sequentially in the familiarization stream, thus violating the statistical regularities embedded in both words and phantoms), and *phantoms* against *non-words* (the order of token types was counter-balanced in the test pairs). In the original study, EEG recording was made throughout the experiment. However, here we analyze only behavioral data, and therefore also include responses from participants who were discarded from the original analysis due to technical issues related to the EEG signal. This brings the total number of participants to 38.

Analytical approach

To estimate metacognition, we used a signal detection (SDT) approach (Galvin, Podd, Drga, & Whitmore, 2003; Maniscalco and Lau, 2012). Correct responses which also receive high confidence ratings are considered to be metacognitive (type-2) *hits*; incorrect responses given with high confidence are conceptualized as meta-*false alarms*; incorrect responses with low confidence rating are analyzed as meta-*correct rejections*; and correct responses with low confidence are taken as meta-*misses.* This logic is applied to binary confidence choices. The ability to estimate the likelihood of making an error and thus to discriminate between correct and incorrect responses by assigning different confidence ratings on a wider scale can be quantified by Type-2 ROC analysis (Galvin et al., 2003). Maniscalco and Lau (2012) proposed a modeling approach to quantify Type-2 ROC area in units of Type-1 d'. The basic idea behind this approach is to estimate the pseudo d' measure that would perfectly fit confidence ratings, not type-1 decisions. This estimated value – which accounts for confidence ratings instead of cognitive decisions – provides a measure of subject-specific metacognitive sensitivity, i.e., *meta-d*. Meta-d' estimates the reliability of confidence ratings, and may theoretically reflect fluctuations in confidence between groups or conditions even when type-1 performances do not differ. Importantly, meta-d' is independent on metacognitive bias, i.e., individual tendency to assign higher or lower rating overall. However, meta-d' also reflects the quality of the type-1 information which is subject to metacognitive processing, and in practice may scale with performance (Fleming and Lau, 2014). Thus it sometimes makes sense to estimate metacognitive efficiency relative to type-1 performance. This is represented by the meta-d' to d' ratio (*M-ratio*). Meta-d' reflects the presence or absence of metacognition in a particular task. However, the M-ratio is more informative when the goal is to compare metacognition across groups or conditions, especially if type-1 performances differ; it is even more useful when type-1 differences across conditions and groups are caused by: unequal variance in the number of participants and trials per group/condition; by potentially different neural and cognitive mechanisms underlying metacognitive judgments across different tasks; or by differences (including purposefully modulated differences) in the perceptual salience of the signal, individually-adjusted perceptual or performance thresholds, individual differences in encoding information at different

presentation rates or different modalities, etc. Meta-d' can easily be compared to d' (because they are measured in the same units on the same scale) but comparing across conditions/domains/tasks/experiments may be challenging. Thus, M-ratio - the measure of metacognitive efficiency, or a subject-specific level of metacognitive sensitivity which takes individual level of type-1 performance into account – can be more informative.

Fleming (2017) developed a method for hierarchical Bayesian estimation of metacognitive sensitivity and efficiency, which we have adopted for our analysis here. Bayesian estimation is superior to maximum likelihood estimation (MLE) and sum-of-squared error (SSE) methods (Maniscalco & Lau, 2012; Barret, Dienes, & Seth, 2013) for several reasons: it does not require data "padding" when participants give zero responses with a particular (usually the maximum and minimum) confidence level; it naturally accounts for situations when the number of trials for the levels of confidence differs; it reduces the influence of a single outlying participant on group results witout requiring that this outlier be corrected or removed. It thus allows for all data to be included in the analysis and decreases subjectivity in sampling. Finally – although this is not applicable to the current analysis – Bayesian estimation allows groups with different number of participants and different number of trials per participant/condition to be compared. This can be important in studies comparing metacognition in adults and children at different ages (it is not always possible to collect equal number of trials from young children and adult participants), or comparing patients and healthy individuals (patient samples often depend on "convenience" while healthy adult participants are easier to recruit). For hierarchical Bayesian modeling, we used the code developed by Fleming (2017) available at https://github.com/metacoglab/Hmeta-d

A limitation of this method is that data need to be amended to accommodate a two-choice SDT framework, that is, a 2(stimulus)*2(response)*N(confidence rating) matrix. We adopted the algorithm applied by Ordin, Polyanskaya, Soto (2020b) for this purpose (essential details are given below):

Word vs. non-word pairs (i.e., test trials in which words were paired with non-words, words happened equal number of times in the first and in the second positions of the test pairs): When the word appeared in position 1 and the non-word in position 2, the response was considered a *hit* if the participant responded "1" (i.e., chose a word over a non-word) but a *miss* if the participant responded "2" (i.e., missing a word in position 1). If the word occurred in position 2, the response was defined as a *correct rejection* if the participant responded "2" (i.e., rejected the non-word in position 1) and a *false alarm if* the participant responded "1" (falsely identified position 1 as containing the target).

Phantoms vs. non-words: A phantom was operationalized as a signal to be detected, so choosing a phantom over a non-word was considered a correct response. When a phantom appeared in the position 1 and the non-word in position 2, the response was considered a *hit* if the participant responded "1" but a *miss* if the participant responded "2" (missing the phantom in position 1). When a phantom appeared in position 2, the response was defined as a *correct rejection* if the participant responded "2" (i.e., rejecting the option that the phantom was in position 1) and a *false alarm* if the participant responded "1" (falsely identified position 1 as containing the target).

Words vs. phantoms: In a similar fashion, when a word appeared in position 1 and a phantom in position 2, the trial was considered a *hit* if the participant responded "1", but a *miss*, if the participant responses "2" (i.e., missing a target in position 1) When a word appeared in position 2, the trial was considered a *false alarm* if the participant responded "1" (falsely identifying position 1 as containing the target), and a *correct rejection*, if the participant responses "2" (i.e., correctly rejecting the possibility of the target in position 1).

We estimated type-1 sensitivity and bias, metacognitive sensitivity, efficiency and bias separately for these three pair types to explore metacognitive processes when people select old vs. statistically congruent novel items (*words* vs. *phantoms*), novel statistically congruent items vs. novel statistically incongruent items (*phantoms* vs. *non-words*), and old vs. novel statistically incongruent items (*words* vs. *non-words*).

RESULTS

We calculated bias (a tendency to select the first or the second item in the test pair), d', meta-d', M-ratio (Figure 1a) and metacognitive bias as mean confidence (Figure 1b) for the three pair-types. We applied a repeated measures ANOVA to each measure to find whether the scores differed significantly between pair-types. In all cases, we used Maulchy's test to check the sphericity assumption. Where necessary we corrected DFs and p-values using the Greenhouse-Geiser epsilon, reporting uncorrected DFs and corrected p-values. Each significant test was then split into a series of pairwise comparisons using paired 2-tailed t-tests, applying Bonferroni correction and reporting corrected p-values.
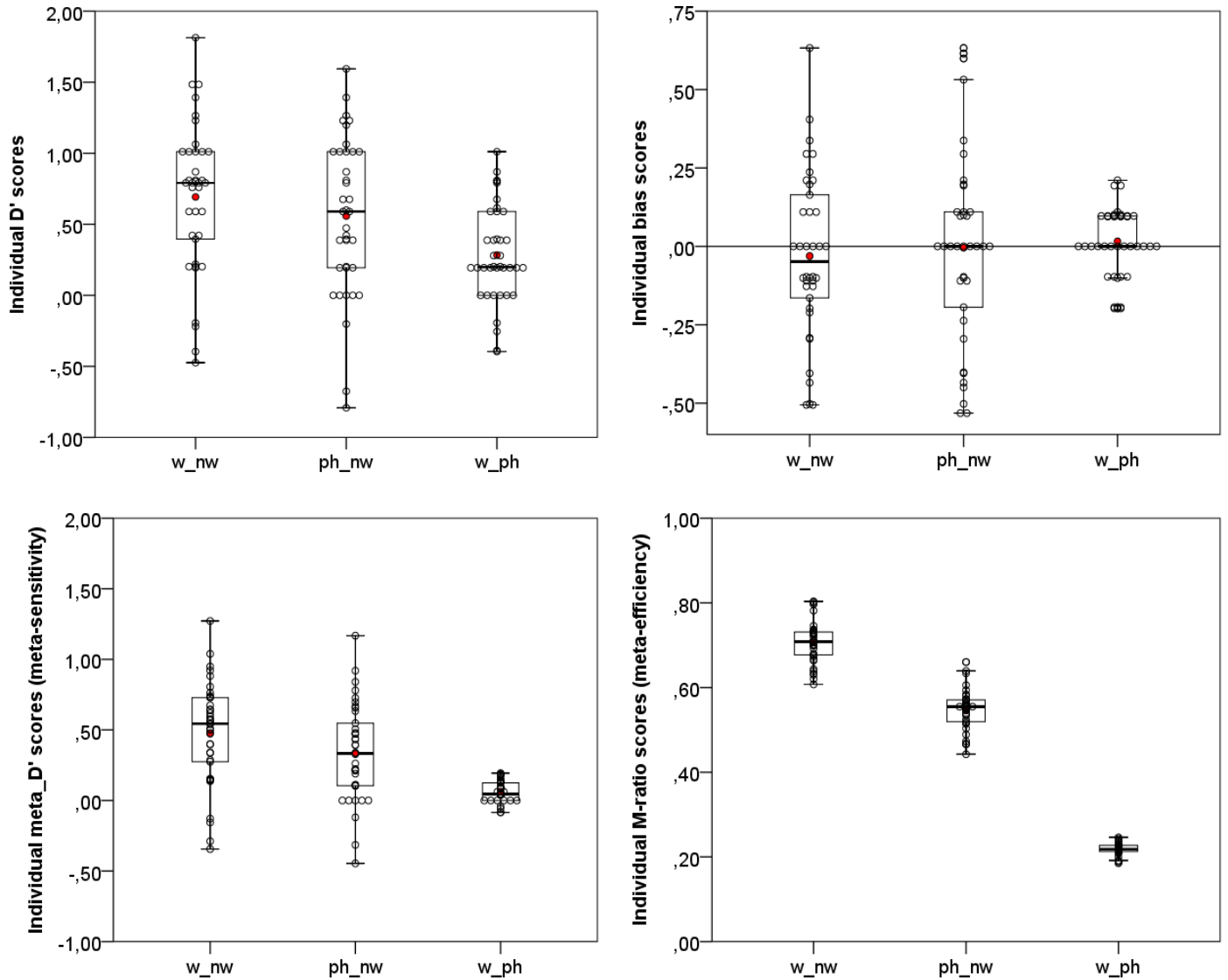

Figure 1a. Type-1 (cognitive) and type-2 (metacognitive) performance in the test trials pairing words against non-words (w_nw), phantoms against non-words (ph_nw) and words against phantoms (w_ph). Type-1 performance is represented on the upper figures (perceptual-cognitive sensitivity on the left and bias on the right, negative sensitivity shows below chance individual performance, positive bias scores mean the overall individual tendency to prefer the first of the two options in the alternative forced-choice test, and negative values show individual tendency to choose the second token). Type-2 performance is represented on the lower figures (metacognitive sensitivity on the left and metacognitive efficiency on the right). On this and the following figure, the dots represent the scores of individual participants, the boxes contain 50% of all data (second and third quartiles), whiskers represent upper and lower (the first and the fourth) quartiles. The horizontal lines inside the boxes represent medians in each sample, and filled dots (red dots in online version) show the means.
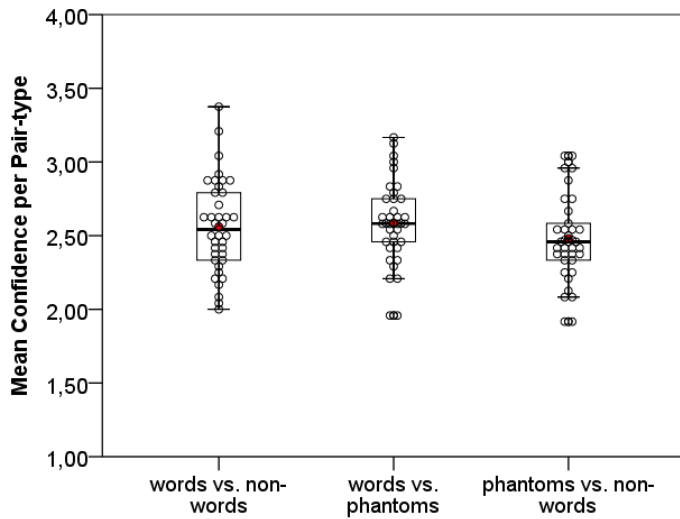
Figure 1b. Metacognitive bias expressed as mean confidence rating assigned to the trials pairing words versus non-words, words versus phantoms and phantoms versus non-words.

D-prime

The analysis showed a significant effect of *pair-type*, $F(2, 74)=9.328$, $p<.0005$, $\eta_p^2=.201$. Pairwise comparisons revealed that sensitivity was significantly higher for words over non-words than for words over phantoms, M=.479 (.1), [.278:.681], $t(37)=4.817$, $p<.0005$. The difference in sensitivity between phantoms over non-words and words over phantoms was also significant, M=.329 (.126), [.073:.586], $t(37)=2.607$, $p=.039$. We did not observe that sensitivity to words over non-words was significantly different from sensitivity to phantoms over non-words, M=.15 (SE of the difference .113), 95%CI of the difference [-.08:.38], $t(37)=1.325$, , $p=.579$. To estimate the evidence for absence of difference in the latter contrast, we calculated the Bayes factor using Bayesian t-tests (2-tailed, with the prior that both outcomes are equally likely) for the latter contrast. The result, $BF_{10}=.039$, shows that, given the data, the null hypothesis (no difference) is 2.55 times more likely than the alternative hypothesis, which provides some – although weak – evidence for absence of significant difference.

Bias (C')

The difference in bias between conditions was not significant, $F(2, 74)=2.07$, $p=.133$, $\eta_p^2=.053$, suggesting that people did not have different preferences across pair-types in endorsing the first or the second item.

Meta-d-prime

The analysis showed a significant effect of *pair-type*, $F(2, 74)=27.52$, $p<.0005$, $\eta_p^2=.427$. Pairwise comparisons revealed that sensitivity to words over non-words was significantly different from sensitivity to phantoms over non-words, M=.208 (.07), [.067:.349], $t(37)=2.99$, $p=.015$. Meta-d' was significantly higher on words over non-words than on words over phantoms, M=.482 (.07), [.342:.623], $t(37)=6.95$, $p<.0005$. Finally, the difference in meta-sensitivity to phantoms over non-words and words over phantoms was also significant, M=.275 (.06), [.161:.388], $t(37)=4.92$, $p<.0005$.

Overall, this shows that even though people select both words and phantoms over non-words with equal sensitivity, their confidence ratings better discriminate between correct and incorrect responses in word vs. non-word test trials. This suggests higher metacognitive sensitivity to old items, which can potentially be retrieved from memory and thus benefit from memory representations during the recognition test.

M-ratio

The analysis showed a significant effect of pair-type, $F(2, 74)=1079.415$, $p<.0005$, $\eta_p^2=.967$. Pairwise comparisons revealed that sensitivity to words over non-words was significantly different from sensitivity to phantoms over non-words, M=.156 (.11), [.133:.179], $t(37)=13.85$, $p<.0005$. M-ratio was significantly higher on words over non-words than on words over phantoms, M=.491 (.008), [.474:.508], $t(37)=58.56$, $p<.0005$. Finally, the difference in M-ratios

on phantoms over non-words and words over phantoms was also significant, M=.334 (.12), [.31:.361], t(37)=27.09, p<.0005.

Metacognitive bias

Overall, mean confidence did not differ across pair-types, F(2, 74)=2.653, p=.085, η_p^2=.067. This suggests that participants used similar confidence references across all conditions and reveals no tendency to assign higher or lower confidence on any of the pair-types.

Importantly, differences in discrimination performance across the three pair-types could have influenced mean confidence, indicating that this commonly used measure of metacognitive bias would not be bias-free in our experiment. To verify that people had no tendency to assign overall higher or lower confidence ratings across conditions, we explored the relations between type-1 sensitivity and mean confidence ratings using by-subject correlations. Lower correlations would signify overconfidence (low performance and high confidence). Comparison of correlations could then be used to compare the bias across conditions. We compared correlations between sensitivity and mean confidence in word versus non-word trials (r=.197) and word versus phantom trials (r=-.052), controlling for the correlations in phantom versus non-word trials (r=-.05). The correlations were not significantly different, z=1.026, p=.152. Difference in correlations between type-1 sensitivity and mean confidence in word versus phantom and phantom versus non-word pairs was not significant either, z=-.009, p=.496 (controlling for the correlations in the third pair-type). Finally, difference in correlations for word versus non-word and phantom versus non-word trials were not significant, z=1.017, p=.155. This analysis converges with the results of ANOVA on mean confidence ratings. No evidence was detected that participants tended to assign higher or lower confidence ratings to any pair-type.

DISCUSSION

Overall, the results show that people easily discriminate between tokens that violate statistical regularities and tokens that conform to statistical regularities, both when statistically eligible tokens are old (i.e., were embedded in the familiarization stream) or novel. Sensitivity to words over non-words and phantoms over non-words appears to be similar, although direct evidence for absence of difference is weak (given the data, the no difference scenario is only 2.6 times more likely than the alternative). Discrimination between old and novel statistically congruent items is more challenging and less accurate, yet still possible (people preferred words to phantoms, when these two types of tokens are presented as word candidates during the recognition test). This pattern of results suggests that when people make a choice between two word candidates, deviations from learnt structures are more salient than the potential to retrieve the token from memory as a whole constituent. However, relying on memory representations may still occur when people need to choose between two candidates that are both congruent with regularities. Memory representations of recurrent constituents extracted from continuous sensory input during familiarization do modulate confidence judgments: metacognitive sensitivity is higher in trials in which words are paired with non-words than trials in which phantoms are paired with non-words. Despite the insignificant difference in d' scores between these conditions, M-ratio is higher when words are paired with non-words, pointing to a facilitatory effect of memory representations for old items.

Although metacognitive processing is indeed possible without conscious awareness (Jachs, Blanco, Grantham-Hill, & Soto, 2015; Kentridge & Heywood, 2000), in this experiment participants were explicitly told to detect and memorize the words of "an alien language" during the familiarization stage. During the recognition test, participants also made conscious choices. Thus we assume that metacognitive processing in our experiment is strongly coupled with conscious awareness of knowledge or conscious assessment of the feeling of familiarity with respect to the content of statistical learning (Kunimoto, Miller, & Pashler, 2001; Persaud, Davidson, Maniscalco, Mobbs, Passingham, Coway, & Lau, 2011; Persaud, McLeod, & Cowey, 2007; Shimamura, 2008). If participants were not consciously aware of the content of their knowledge, their confidence ratings would not be informative with respect to the correctness of their responses, even if recognition performance were above chance. Following this logic (Ko & Lau, 2012; Maniscalco & Lau, 2012; Persaud et al., 2007; Shimamura, 2008), we believe that – at least in the current

experimental setup – post-decision confidence judgments objectively measure conscious awareness in the choice of words over non-words and phantoms, and phantoms over non-words. Thus, our results suggest that people are more aware of statistical regularities (and violations of such regularities) than memorized triplets, although they can still discriminate these old triplets from novel items that conform to acquired regularities. This means that the statistical learning process produces both discrete constituents and statistical rules (e.g., TPs), yet these are subjected to varying degrees of conscious processing. In this regard, it is important to emphasize that explicit instructions equipped participants with information that the familiarization stream contained discrete constituents (aka were words in an "alien language"), but the existence of statistical rules was not mentioned to the participants. We propose that, over the course of learning, attention was re-shifted from constituents to structures, subjecting them to more conscious awareness. This could support the extraction of rules from a small number of cases, and generalization of these extracted rules to previously unencountered situations. It would also be interesting to explore how conscious awareness is modulated by drawing attention – via explicit instructions – not only to the presence of discrete constituents in the sensory input, but also to the regularities embedded in that input. It would also be interesting to explore metacognitive processes elicited by an identification task (yes/no recognition task, where the response requires deciding whether a presented token is a word constituent from a sensory input one has been exposed to) as opposed to a discrimination task.

Our data leaves many questions unanswered. We do not know whether the observed differences in metacognitive processing were based on different strategies for assigning confidence ratings for the pair-types, or on reduced error detection (people might not have perceived rejected phantoms as errors in word-phantom pairs, which would prevent error detection mechanisms from functioning). Furthermore, it is important to note that d' provides a bias-free measure of perceptual and cognitive sensitivity to the signal (in this case, presence of structure, presence of memory representations, or both), without making any conclusions in regard to the processing architecture, and meta-d' provides an estimate of metacognitive sensitivity, also without pointing to a particular type of processing mechanism. Both measures can be driven by multiple, sometimes confounding factors. Studies regarding type-1 sensitivity in statistical learning have already spawned a rich body of literature. Siegelman, Bogaerts, Christiansen, and Frost (2017) and Frost, Armstrong, Siegelman, and Christiansen (2015) have summarized that individual differences in statistical learning efficiency can be modulated by a) computational ability (i.e., how well an individual registers statistical regularities and employs them to detect boundaries between discrete constituents); b) individual differences related to experience with a particular domain, task, or type of material (more expertise makes an individual better at dealing with statistical regularities with higher degrees of complexity); and c) psychophysics differences related to individual ability to encode information in a particular modality at a particular rate with a particular just-noticeable detection or discrimination threshold. Each of these factors can also affect metacognitive sensitivity and feelings of confidence. Additionally, metacognitive sensitivity can be influenced by (a) information – usually a subjective feeling for which the participant has no logical explanation because the information is not subject to conscious processing - that is taken into account when making metacognitive judgments but not cognitive decisions (Jachs et al., 2015; Scott, Dienes, Barret, Bor, & Seth, 2014); (b) further processing of the stimulus information after the decision is already taken, leading to meta-d' scores being higher than d' scores (Rabbit & Vyas, 1981); or (c) self-evaluation of one's feeling of confidence (Fleming & Daw, 2017). Further research is needed to explore these multiple drivers of metacognitive sensitivity, efficiency, and bias in statistical learning tasks, which are reflecting in meta-d' and M-ratio measures.

OPEN PRACTICES STATEMENT

We used the same dataset reported in Ordin, Polyanskaya, Soto, & Molinaro (2020). The raw data is available from the corresponding author on request (mordin@bcbl.eu), and is deposited online on Figshare (DOI: 10.6084/m9.figshare.9989000 and DOI: 10.6084/m9.figshare.9988610). The analysis scripts for hierarchical Bayesian modelling of metacognitive sensitivity are developed by Prof. Stephen Fleming, who made them publicly available at https://github.com/metacoglab/Hmeta-d (last accessed on 25/04/2020).

REFERENCES

1. Alamia, A., & Zenon, A. (2016). Statistical Regularities Attract Attention when Task-Relevant. *Frontiers in Human Neuroscience 10*: 42.
2. Arciuli J, Torkildsen JvK, Stevens DJ and Simpson IC (2014) Statistical learning under incidental versus intentional conditions. *Front. Psychol. 5:747*. doi: 10.3389/fpsyg.2014.00747
3. Aslin, R. N., and Newport, E. L. (2012). Statistical learning: from acquiring specific items to forming general rules. *Current Directions in Psychological Science 21*, 170–176. doi: 10.1177/0963721412436806
4. Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. Cognition 106, 1382–1407.
5. Barrett, A.B., Dienes, Z., & Seth, A.K. (2013). Measures of metacognition on signal-detection theoretic models. *Psychological Methods 18*, 535–552.
6. Dienes, Z., Broadbent, D., and Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar. *J. Exp. Psychol. Learn. Mem. Cogn. 17*, 875–887.
7. Dunlosky, J., Serra, M. J., & Baker, J. M. C. (2007).*Metamemory.*In F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandowsky, & T. J. Perfect (Eds.). *Handbook of applied cognition* (137–160). John Wiley & Sons.
8. Endress, A.D. & Langus, A. (2017). Transitional probabilities count more than frequency, but might not be used for memorization. *Cognitive Psychology 92*, 37-64.
9. Endress, A.D. & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language 60(3)*, 351-367.
10. Fernandes, T., Kolinsky, R. & Ventura, P. (2010). The impact of attention load on the use of statistical information and co-articulation as speech segmentation cues. *Attention, Perception, & Psychophysics 72*, 1522-1532.
11. Flavell, J.H. (1979). "Metacognition and cognitive monitoring. A new area of cognitive-development inquiry".*American Psychologist 34(10)*, 906–911.
12. Fleming, S. & Lau, H. (2014). How to measure metacognition. *Frontiers in Human Neuroscience 8:443*. doi: 10.3389/fnhum.2014.00443
13. Fleming, S. (2017). HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness 1*: mix007.
14. Fleming, S., & Daw, N.D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review 124(1)*, 91-114.
15. Frost, R., Armstrong, B., Siegelman, N., & Christiansen, M. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Science 19(3)*, 117-125.
16. Galvin, S.J., Podd, J.V., Drga, V., & Whitmore (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review 10*, 843–876.
17. Gómez, D.M., Bion, R., & Mehler, J. (2011). The word segmentation process as revealed by click detection. Language and Cognitive Processes 26(2), 212-223.
18. Hard, B.M., Meyer, M. & Baldwin, D. Attention reorganizes as structure is detected in dynamic action. Mem Cogn 47, 17–32 (2019).
19. Harris, Z. (1955). From phoneme to morpheme. *Language 31,* 190-222.
20. Jachs, B., Blanco, M., Grantham-Hill, S., & Soto, D. (2015). On the independence of visual awareness and metacognition: A signal detection theoretic analysis. *Journal of Experimental Psychology: Human Perception and Performance 41(2),* 269-276.
21. Kentridge, R.W. & Heywood, C.A. (2000). Metacognition and Awareness. *Consciousness and Cognition 9,* 308-312.
22. Kepecs, A., Uchida, N., Zariwala, H. & Mainen, Z. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature 455*, 227–231.
23. Ko Y, & Lau H. (2012). A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. *Philosophical Transactions of the Royal Society B: Biological Sciences 367*, 1401–1411.
24. Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition 10(3)*, 294 –340.
25. Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition 21(1)*, 422-430.
26. Nelson, T. (1996). Consciousness and Metacognition. *American Psychologist 51*, 102-116.
27. Nelson, T., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation 26*, 125-173.

28. Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning Memory and Cognition 28(5)*, 924-940.

29. Ordin, M., Polyanskaya, L., & Soto, D. (2020a). Neural bases of learning and recognition of statistical regularities. *Annals of the New York Academy of Sciences 1467*, 60-76.

30. Ordin, M., Polyanskaya, L., & Soto, D. (2020b). Metacognitive Processing in Language Learning Tasks Is Affected by Bilingualism. *Journal of Experimental Psychology: Learning Memory and Cogn*ition 46(3), 529-538.

31. Ordin, M., Polyanskaya, L., Soto, D., & Molinaro, N. (2020). Electrophysiology of statistical learning: Exploring the online learning process and offline learning product. *European Journal of Neuroscience 51(9)*, 2008-2022.

32. Persaud, N., Davidson, M., Maniscalco, B., Mobbs, D., Passingham, R. E., Cowey, A., & Lau, H. (2011). Awareness-related activity in prefrontal and parietal cortices in blindsight reflects more than superior visual performance. *NeuroImage 58*, 605– 611.

33. Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience 10*, 257– 261.

34. Rabbit, P. & Vyas, Sh. (1981). Processing a display even after you make a response to it. How perceptual errors can be corrected. The Quarterly *Journal of Experimental Psychology A 33(3)*, 223-239.

35. Reber, A. S., Kassin, S. M., Lewis, S., and Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *J. Exp. Psychol. Hum. Learn. Mem. 6, 492–502*.

36. Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition 21*, 803–814.

37. Saffran, J. (2001). Words in a sea of sounds: the output of infant statistical learning. *Cognition 81(2),* 149-169.

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

39. Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science* 26, 113–125.

40. Scott, R., Dienes, Z., Barret, A., Bor, D., & Seth, A. (2015). Blind insight: Metacognitive discrimination despite chance task performance. *Psychological Science 25(12)*, 2199-2208.

41. Shimamura, A. P. (2008). A neurocognitive approach to metacognitive monitoring and control. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 373–390). New York, NY: Psychology Press.

42. Siegelman, N. (2019). Statistical learning abilities and their relation to language. *Language and Linguistics Compass.*

43. Siegelman, N., Bogaerts, L., Armstrong, B. C., & Frost, R. (2019). What exactly is learned in visual statistical learning? Insights from Bayesian modeling. *Cognition 192:* 104002.

44. Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 372*(1711).

45. Toro, J.M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition 97*, B25-B34.