cortex 90 (2017) 31–45

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Online neural monitoring of statistical learning

Laura J. Batterink* and Ken A. Paller

Northwestern University, Department of Psychology, Evanston, IL, USA

ARTICLE INFO

Article history:
Received 20 June 2016
Reviewed 8 September 2016
Revised 17 October 2016
Accepted 3 February 2017
Action editor Ed Wilding
Published online 24 February 2017

Keywords:
Implicit learning
Intertrial coherence
Neural entrainment
Steady-state response
Word segmentation

ABSTRACT

The extraction of patterns in the environment plays a critical role in many types of human learning, from motor skills to language acquisition. This process is known as statistical learning. Here we propose that statistical learning has two dissociable components: (1) perceptual binding of individual stimulus units into integrated composites and (2) storing those integrated representations for later use. Statistical learning is typically assessed using post-learning tasks, such that the two components are conflated. Our goal was to characterize the online perceptual component of statistical learning. Participants were exposed to a structured stream of repeating trisyllabic nonsense words and a random syllable stream. Online learning was indexed by an EEG-based measure that quantified neural entrainment at the frequency of the repeating words relative to that of individual syllables. Statistical learning was subsequently assessed using conventional measures in an explicit rating task and a reaction-time task. In the structured stream, neural entrainment to trisyllabic words was higher than in the random stream, increased as a function of exposure to track the progression of learning, and predicted performance on the reaction time (RT) task. These results demonstrate that monitoring this critical component of learning via rhythmic EEG entrainment reveals a gradual acquisition of knowledge whereby novel stimulus sequences are transformed into familiar composites. This online perceptual transformation is a critical component of learning.

1. Introduction

Spoken words in an unknown foreign language can seem exceedingly rapid and incomprehensible compared to normal speech in one's native language, despite the fact that syllable rate across languages is relatively similar (Pellegrino, Coupe, & Marsico, 2011). Natural speech consists of a continuous stream of sound with no reliable pauses between words (Lehiste, 1960), and a major challenge for language learners is to discover word boundaries, a process known as speech segmentation. One reason that beginners learning a new language may perceive unfamiliar speech as unfolding quickly is because they are not yet capable of segmenting the speech stream, leading them to perceive a greater number of basic perceptual units (i.e., multiple individual syllables rather than multisyllabic words). Thus, one of the very first stages of word learning is a perceptual process, requiring a shift in the perception of smaller syllable-units (Bertoncini & Mehler, 1981; Mehler, Dommergues, Frauenfelder, & Segui, 1981) to that of larger word-units. Only after this perceptual shift has been accomplished can the extracted word forms, comprising

* *Corresponding author*. Northwestern University, Department of Psychology, 2029 Sheridan Road, Evanston, IL 60208, USA.
E-mail address: lbatterink@northwestern.edu (L.J. Batterink).

http://dx.doi.org/10.1016/j.cortex.2017.02.004

key building blocks of language, be stored in memory and undergo further processing (e.g., Graf-Estes, Evans, Alibali, & Saffran, 2007).

Statistical learning, the process of becoming sensitive to statistical structure in the environment, is thought to be a critical learning mechanism underlying speech segmentation (Saffran, 2003). Learners can discover word boundaries by computing the transitional probabilities between neighboring syllables, which are higher within words compared to across word boundaries (Saffran, Newport, & Aslin, 1996, Saffran, Aslin, Newport, 1996). In the first demonstration of statistical learning, infants were exposed to a continuous auditory stream of repeating trisyllabic nonsense words, in which transitional probabilities served as the only cue to word boundaries (Saffran, Aslin, et al., 1996). A subsequent test revealed that infants' visual fixation times differed between words from the stream and non-word foils made up of recombined syllables, demonstrating that infants were sensitive to the statistical properties of the input. Since this seminal study, a large literature has demonstrated that statistical learning operates across ages and sensory modalities (e.g., Bulf, Johnson, & Valenza, 2011; Conway & Christiansen, 2005; Fiser & Aslin, 2001; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997, Saffran, Johnson, Aslin, & Newport, 1999; Turk-Browne, Junge, & Scholl, 2005), and that it contributes to a wide range of cognitive functions in addition to speech segmentation (e.g., Creel, Newport, & Aslin, 2004; Fiser & Aslin, 2001; Goujon & Fagot, 2013; Hunt & Aslin, 2001; Saffran et al., 1999).

At a mechanistic level, statistical learning can be conceptualized as comprising at least two dissociable stages or components. As illustrated by the earlier example of foreign speech perception, the initial component is perceptual in nature, involving a transition from the perception and encoding of raw individual stimulus units (e.g., syllables) to that of larger integrated items (e.g., words). This perceptual process, which we will refer to as the *word identification component*, may be considered the central challenge of statistical learning. After words are encoded as units, the extracted representations can be stored as such in long-term memory. This *memory storage component* can in one sense be considered peripheral to the statistical learning process per se, but nonetheless critically influences performance on subsequent tests of statistical learning. Effective memory storage of integrated representations is also a prerequisite for further processing that occurs after initial segmentation, such as acquiring phonological patterns across words (e.g., Saffran & Thiessen, 2003) and mapping words to objects (Graf Estes et al., 2007; Mirman, Magnuson, Estes, & Dixon, 2008).

Although it is well-established that there are substantial individual differences in general long-term memory abilities (e.g., Bors & MacLeod, 1996), little is known about the extent to which the word identification component of statistical learning shows individual variability. These differences may be considerable, given that performance on post-exposure statistical learning tasks varies substantially, with at least one third of a sample often failing to perform the task above chance levels (Frost, Armstrong, Siegelman, & Christiansen, 2015; Siegelman & Frost, 2015). This variability in statistical learning performance could in principle be attributable to the word identification component of statistical learning, subsequent long-term memory storage, or both. One theoretical possibility is that the word identification component of statistical learning is relatively invariant across individuals with normal sensory processing, as some studies have suggested for implicit learning (e.g., Reber, Walkenfeld & Hernstadt, 1991; Reber, 1993). In this case, individual differences in performance on statistical learning tasks would be driven primarily by variability in long-term memory storage. Alternatively, there may also be substantial individual variability in the online perceptual and encoding processes contributing to word identification. If so, variability in the word identification component should at least partially account for observed individual differences in statistical learning task performance (Frost et al., 2015; Siegelman & Frost, 2015).

These two alternatives have not been previously explored, as previous statistical learning studies generally have not drawn a conceptual distinction between these two components of learning, nor attempted to disentangle them empirically. In large part, this may be due to the experimental approach that has been used to investigate statistical learning. With some notable exceptions (Cunillera, Toro, Sebastian-Galles, & Rodriguez-Fornells, 2006, Cunillera et al., 2009; Karuza et al., 2013; McNealy, Mazziotta, & Dapretto, 2006, 2010; Turk-Browne, Scholl, Chun, & Johnson, 2008), the vast majority of studies have followed the same general approach: an initial learning phase involving exposure to structured input, followed by an offline test. In infants, this test generally comprises assessing visual fixation times to test items (e.g., Aslin, Saffran, & Newport, 1998; Saffran et al., 1999, Saffran, Aslin, et al., 1996), and in adults, responses on a forced-choice recognition measure (Saffran et al., 1997, Saffran, Newport, et al., 1996; Turk-Browne et al., 2005). These tasks require learners to retrieve previously encoded knowledge after statistical learning is completed, focusing on the final outcome of learning rather than on the word identification component, the central process of online statistical learning. Poor performance on such tasks may be driven by a failure of the word identification component of statistical learning, or of long-term memory storage. Another limitation of such end-state outcome measures is that they are unable to assess the time course of learning.

According to our conceptualization of statistical learning, a shift in the perception and encoding of raw individual stimulus units towards that of larger integrated items is a critical component of statistical learning, and a prerequisite for above-chance performance on subsequent statistical learning tests. The goal of the present study was to characterize this word identification component of statistical learning, including its time course, its variability among individuals, and its relation to performance on post-exposure learning tasks. First, we hypothesized that a shift in perception of integrated items over individual units should increase as a function of exposure to structured input. Second, we hypothesized that individuals would show measurable differences in this word-identification component, with some individuals showing evidence of more effective segmentation than others. Third, we hypothesized that individual variability in the word-identification component would predict performance on post-exposure statistical learning tasks. Such a