Journal of Memory and Language 81 (2015) 105–120

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Statistical learning as an individual ability: Theoretical perspectives and empirical evidence

Noam Siegelman [a,*], Ram Frost [a,b,c]

[a] Department of Psychology, The Hebrew University, Israel
[b] Haskins Laboratories, New Haven, United States
[c] The Basque Center for Brain and Language, Spain

ARTICLE INFO

Article history:
Received 1 June 2014
revision received 2 February 2015
Available online 11 March 2015

Keywords:
Statistical learning
Individual differences
Predicting linguistic abilities

ABSTRACT

Although the power of statistical learning (SL) in explaining a wide range of linguistic functions is gaining increasing support, relatively little research has focused on this theoretical construct from the perspective of individual differences. However, to be able to reliably link individual differences in a given ability such as language learning to individual differences in SL, three critical theoretical questions should be posed: Is SL a componential or unified ability? Is it nested within other general cognitive abilities? Is it a stable capacity of an individual? Following an initial mapping sentence outlining the possible dimensions of SL, we employed a battery of SL tasks in the visual and auditory modalities, using verbal and non-verbal stimuli, with adjacent and non-adjacent contingencies. SL tasks were administered along with general cognitive tasks in a within-subject design at two time points to explore our theoretical questions. We found that SL, as measured by some tasks, is a stable and reliable capacity of an individual. Moreover, we found SL to be independent of general cognitive abilities such as intelligence or working memory. However, SL is not a unified capacity, so that individual sensitivity to conditional probabilities is not uniform across modalities and stimuli.

Introduction

The detection of regularities and quasi-regularities in the environment is a necessary prerequisite for making sense of the infinitely rich stimulation provided to the brain. It underlies categorization and segmentation of continuous information, allows prediction of upcoming events, and shapes basic mechanisms of perception and action. Originally labeled “Artificial Grammar Learning” (AGL), “Implicit Learning” (IL) (Reber, 1967, 1993), and then “Statistical Learning” (SL) (Saffran, Aslin, & Newport, 1996), the ability to pick up regularities in the world is taken as a domain-general central mechanism by which cognitive systems discover the underlying structural properties of any input for the purpose of generating expectations. A large body of research has focused, therefore, on tracing the extent of this ability, showing a remarkable sensitivity of subjects to even strikingly low correlations within large sets of stimuli (e.g., Kareev, Fiedler, & Avrahami, 2009), detected automatically (Fiser & Aslin, 2001), even without overt attention (Evans, Saffran, & Robe-Torres, 2009), and from a very early age (as young as 1–3 day old newborns, Bulf, Johnson, & Valenza, 2011).

The term SL was coined specifically in the domain of language, and its explanatory power is gaining increasing recognition. The ability to extract repeated patterns of regularities and transitional probabilities from sequential and continuous auditory or visual inputs has proved useful in

---

* Corresponding author at: Department of Psychology, The Hebrew University, Jerusalem 91905, Israel.

E-mail address: noam.siegelman@gmail.com (N. Siegelman).

http://dx.doi.org/10.1016/j.jml.2015.02.001
0749-596X/© 2015 Elsevier Inc. All rights reserved.

explaining basic linguistic phenomena such as syntax acquisition (e.g., Saffran & Wilson, 2003), segmentation of speech and detection of word boundaries (e.g., Onnis, Waterfall, & Edelman, 2008), picking up on orthographic (e.g., Pacton, Perruchet, Fayol, & Cleeremans, 2001) or phonotactic regularities (e.g., Chambers, Onishi, & Fisher, 2003), detecting the internal structure of words for adjacent (e.g., Endress & Mehler, 2009) or nonadjacent (e.g., Newport & Aslin, 2004) components, even with long distance dependencies (e.g., Gómez, 2002). This form of learning has been demonstrated to be automatic and exceedingly fast (sometimes just two minutes of exposure suffice; e.g., Saffran et al., 1996), as well as incidental (e.g., Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). Setting aside the complicated and loaded question of whether linguistic knowledge requires domain-specific abstract rule-like generalizations (e.g., Marcus, Vijayan, Rao, & Vishton, 1999), it is well-established that the domain-general capacity of SL provides the cognitive system with reliable cues regarding the structural properties of printed or spoken words, thereby serving language acquisition. Recent studies have also suggested that similar neural correlates underlie sequential learning and language processing (see Christiansen, Conway, & Onnis, 2012, for a discussion).

Although the power of statistical learning in explaining a wide range of linguistic functions is gaining increasing support, relatively little research has focused on this theoretical construct from the perspective of individual differences. Some studies have suggested that individual variance in implicit learning is significantly smaller than in explicit learning (e.g., Reber, 1993). In the same vein, it has been argued that SL displays developmental invariance, so that performance is mostly unaffected by age or mental health (e.g., Thiessen, Kronstein, & Hufnagle, 2013; but see Arciuli & Simpson, 2011; Karuza et al., 2013; Saffran, 2001 for counter arguments and evidence). Therefore, most studies on statistical and/or implicit learning have typically focused on the *average* success rate in a given experimental paradigm aiming to map the extents and limits of this ability across participants. From this perspective, SL is often studied as a unified theoretical construct, and researchers design experimental tasks that implicate some form of embedded transitional probabilities to explore it. Success in the task beyond chance level in the sampled population is taken to suggest that SL has occurred. Although theoretical distinctions have been made between SL and IL (see Perruchet & Pacton, 2006, for a comprehensive review), both are taken to tap in one way or another similar domain-general mechanism of learning the structural properties of the input.

Considering SL from the perspective of individual differences has one main theoretical motivation: to examine whether individual differences in picking up embedded correlations reliably predict performance in a variety of cognitive and language-related tasks or even personality traits. For example, in a recent study, Kaufman et al. (2010) launched an extensive investigation of implicit learning as an individual ability using the Serial Reaction Time (SRT) task (e.g., Schvaneveldt & Gomez, 1998). In general, the findings demonstrated a relatively weak correlation of RTs in the task with psychometric intelligence and with working memory (see Reber, Walkenfeld, & Hernstadt, 1991, for similar conclusions). Interestingly, individual differences in the SRT task were associated with academic performance in two foreign language exams. Performance in the SRT paradigm was also found to strongly correlate with syntax acquisition in children, as measured by a syntactic priming task (Kidd, 2012), thus confirming the link between IL and language acquisition. Similar conclusions were suggested by Misyak and Christiansen (2012), who tested subjects in two tasks of AGL with adjacent and nonadjacent dependencies, measuring in parallel the subjects' sentence comprehension. They reported that performance in the two AGL tasks significantly predicted sentence comprehension. Moreover, performance in a nonadjacent combined AGL-SRT task was found to specifically predict individual differences in processing sentences with relative clauses that involve long-distance dependencies (Misyak, Christiansen, & Tomblin, 2010). Implicit learning as measured by the AGL paradigm was also found to correlate with speech perception abilities, even when controlling for general cognitive measures such as memory or IQ (Conway, Bauernschmidt, Huang, & Pisoni, 2010; Conway, Karpicke, & Pisoni, 2007). In the same vein, Arciuli and Simpson (2012) have reported a significant correlation (albeit weak) between the ability of both children and adults to detect dependencies in a sequence of visual stimuli, and their reading abilities in L1. Recently, Frost and his colleagues tracked the acquisition of literacy in Hebrew as L2 by native English speakers, reporting that native speakers of English who more accurately picked up the implicit statistical structure embedded in the continuous stream of nonsense visual shapes better assimilated the Semitic structure of Hebrew words, as reflected in several reading tasks (Frost, Siegelman, Narkiss, & Afek, 2013). In addition to these studies, that show a link between IL or SL and linguistic abilities in the normal population, recent studies have also provided evidence for the existence of poor IL/SL abilities for individuals with language disorders compared to matched controls (both for children with SLI, Evans et al., 2009; Hsu, Tomblin, & Christiansen, 2014, and for agrammatic aphasics, Christiansen, Louise Kelly, Shillcock, & Greenfield, 2010).

Taken together, all of the above seems to provide support for the basic theoretical link between SL as an individual ability, and abilities related to the learning of linguistic regularities. From this perspective, language is taken as a rich environment that is characterized by a variety of statistical correlations (see Frost, 2012, for discussion), and SL is taken as a general ability that can be tapped by a variety of possible experimental tasks that could potentially predict ease or difficulty of language acquisition (see Conway & Pisoni, 2008; for a discussion). Thus, the choice of a specific SL task (i.e., the selection of units that compose the input sequence, the decision regarding the embedded transitional probabilities, the extent of similarity or dissimilarity between test trials and foils, etc.) to predict a given linguistic ability in a study, is often incidental. In fact, few discussions have focused on whether different possible dimensions could underlie the theoretical construct of SL (but see Thiessen et al., 2013,