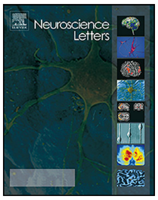
Neuroscience Letters 636 (2017) 77–82

Contents lists available at ScienceDirect

Neuroscience Letters

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



Research article

Sleep duration predicts behavioral and neural differences in adult speech sound learning

F. Sayako Earle[a,b,*], Nicole Landi[c,d], Emily B. Myers[a,c,d]

[a] *Department of Speech, Language, and Hearing Sciences, University of Connecticut, Storrs, CT, USA*
[b] *Department of Communication Sciences and Disorders, University of Delaware, Newark, DE, USA*
[c] *Department of Psychology, University of Connecticut, Storrs, CT, USA*
[d] *Haskins Laboratories, New Haven, CT, USA*

HIGHLIGHTS

- Sleep duration predicts improvement in measures of speech sound learning.
- Behavioral gains are associated with changes in ERP response magnitude overnight.
- Duration of sleep determines relative success of forming new perceptual categories.

ARTICLE INFO

Article history:
Received 24 June 2016
Received in revised form 12 October 2016
Accepted 22 October 2016
Available online 26 October 2016

Keywords:
Memory consolidation
Speech perception
Perceptual learning

ABSTRACT

Sleep is important for memory consolidation and contributes to the formation of new perceptual categories. This study examined sleep as a source of variability in typical learners' ability to form new speech sound categories. We trained monolingual English speakers to identify a set of non-native speech sounds at 8PM, and assessed their ability to identify and discriminate between these sounds immediately after training, and at 8AM on the following day. We tracked sleep duration overnight, and found that light sleep duration predicted gains in identification performance, while total sleep duration predicted gains in discrimination ability. Participants obtained an average of less than 6 h of sleep, pointing to the degree of sleep deprivation as a potential factor. Behavioral measures were associated with ERP indexes of neural sensitivity to the learned contrast. These results demonstrate that the relative success in forming new perceptual categories depends on the duration of post-training sleep.

1. Introduction

Perceptual categories are the foundation for mental organization of the external world and underlie complex cognitive operations such as object recognition and linguistic communication. The capacity to acquire *new* perceptual categories is fundamental to learning. A core question in cognitive science is what kind of experiences facilitates category acquisition. Recent work has pointed to the importance of memory consolidation during sleep in certain types of perceptual learning. The present study asks whether sleep behavior can account for the variability that exists in the ability to acquire new speech sound categories.

Memory processes that are involved in capturing and retaining acoustic information aggregate one's experience with spoken language into functional acoustic-phonetic units, speech sounds such as /d/ or /a/. Variations in the quality of these speech sound representations have been linked to differences in language and reading ability [1,2]. While differences in linguistic input are often cited as the source of such perceptual variability [3], we propose that variance in the representation of speech sounds amongst typical learners can in part be attributed to differences in sleep. Sleep's importance in memory consolidation has been documented across many domains [4], and sleep appears to play a role in the organization of speech information as well. Prior work suggests that a period of sleep following perceptual training is associated with generalization of speech information to unfamiliar lexical contexts in mapping synthetic tokens to native categories [5]. In the acquisition of nonnative speech sounds, the effect of sleep appears to be in protecting phonetic features from conflicting acoustic

---

* Corresponding author at: University of Delaware, STAR Health Sciences Complex, 540 S College Ave, Ste. 210BB, Newark, DE 19713, USA.
E-mail address: fsearle@udel.edu (F.S. Earle).

http://dx.doi.org/10.1016/j.neulet.2016.10.044
0304-3940/

information [6], and in the generalization of acoustic-phonetic features to an unfamiliar talker [7]. In these prior investigations, we have found the magnitude of the effect of offline consolidation on nonnative category acquisition to be highly variable between individuals. One possibility is that the magnitude of overnight gain is dependent on the amount of learning achieved during training. Another possibility, based on a study of word learning that showed that larger gains in learning were associated with longer sleep duration [8], is that the length of post-training sleep determines the degree of improvement. Non-rapid eye movement (NREM) sleep is considered critically important for the consolidation of hippocampal-encoded memory [9]. Specifically, it has been suggested that light sleep (NREM stages 1 and 2) is responsible for active potentiation of wake-state memory, in contrast with deep sleep's (NREM stage 3) role in maintaining homeostatic regulation [10].

The aim of the current study is to test the relationship between sleep behavior and non-native speech sound learning. At 8 P.M. on day 1, we trained monolingual English speakers with typical language and reading ability to identify voiced Hindi dental and retroflex consonants. The dental sound is produced with the tongue tip placed behind the teeth, and the retroflex with the tongue curled behind the alveolar ridge, and both sounds perceptually assimilate to the English (alveolar)/d/for most monolingual speakers of English. Participants were trained by the experimenter to use a sleep-monitoring headband [11] prior to leaving the lab on the first day (see [12] for reliability of the device), wore the device overnight, and returned at 8 A.M. the next morning for reassessment. We measured learning of the non-native contrast through two tasks: discrimination and identification. In discrimination, the listener compares two tokens to determine if the speech features are categorically similar or dissimilar, but this skill does not require awareness of category identity. In contrast, identification requires an explicit recall of a category label for tokens falling within a specific distribution of auditory stimuli. Therefore, the changes in performance on these perceptual tasks reflect different types of perceptual learning that are both commonly used to describe the representational quality of speech sounds. We tracked performance in each task on days 1 and 2 (see Fig. 1 for scheduling of tasks). We predicted that sleep duration would be positively associated with gains in behavioral performance.

It is possible that discrimination and identification tasks underestimate participants' ability to perceive differences among the trained sounds, or that motivational factors that vary as a function of fatigue might affect task performance. Thus, in order to obtain a task-independent measure of changes to perceptual sensitivity, we also collected electrophysiological recordings during an oddball paradigm to determine each participant's mismatch negativity (MMN) response to the trained contrast immediately before training on day 1, and after behavioral reassessment on day 2. The MMN component of the event-related potential (ERP) reflects pre-attentive responses to stimulus change, and has been used to measure phonological processing in both typical and atypical populations [13]. We predicted that overnight changes in the MMN response would be positively associated with changes in behavioral performance.

2. Materials and methods

2.1. Participants

All participants provided informed written consent in accordance with our human subjects research protocol approved by the University of Connecticut Institutional Review Board prior to participation. The data presented in this manuscript is from a subset from a larger dataset collected from monolingual, native speakers of American English, 18–24 years of age, who grew up in a household with only native speakers of American English. The participants included in the current report met additional criteria as follows. Participants reported no neurological, socio-emotional, or attention disorder, and a history of typical language, reading, and cognitive development. In addition, participants passed a pure tone hearing screening and obtained scores at or above the 25th percentile on the following standardized assessments of reading and nonverbal cognition at the time of the study: *Woodcock Reading Mastery Test-Revised* [14], *Test of Word Reading Efficiency* [15], *Wechsler Abbreviated Scale of Intelligence* [16]. These participants were further identified as 'typical' in language by the Modified Token Test and the 15-word spelling test, which are non-standardized measures devised to identify adults with a history of language impairment [17]. Twenty-seven participants (Mean: 20 years, 5 months; SD: 16 months; 17 female, 25 right-handed) met all criteria and completed the EEG recording and the perceptual training; of these, data from two participants were excluded from the current analyses for the reason that behavioral performance immediately after training was greater than 3 standard deviations above the mean; this brought their scores close to task ceiling (equal to or greater than 95% accurate) and as such, a meaningful measure of overnight improvement on perceptual tasks could not be obtained. Data on the remaining 25 participants contributed behavioral/EEG measures. For analyses involving sleep data, we excluded data points that revealed missing segments due to signal dropout during the recording period (n = 7). The remaining 18 (Mean: 20 years, 4 months; SD: 17 months; 11 female, 16 right-handed) completed sleep data collection.

Standardized assessments were administered and scored by the first author or one other graduate student, and rescored by one of two trained undergraduate students. Any discrepancies in scoring were flagged by the second scorer and resolved by the first author.

2.2. Stimuli

2.2.1. Nonnative phonetic training/assessment

A set of 10 naturally spoken tokens (5 each:/d̪ug/[dental],/ɖug/[retroflex]) were produced by a male, native speaker of Hindi in a sound attenuated audiology booth and recorded by an Edirol digital recorder [18]. Praat software [19] was used to cut tokens to the burst onset and to adjust mean amplitudes to 70 dB. For training, two novel visual objects [20] were used, each corresponding to one 'word' within the minimal pair:/d̪ug/-/ɖug/used during training.

2.2.2. Mismatch negativity (MMN) response

Two tokens from the above training set (one dental and one retroflex) were further modified in Praat to equate burst, vowel, and token duration to total 250 ms from onset to offset. The dental token was used as the frequently occurring Standard, and the retroflex token was used as the infrequently occurring Deviant.

2.3. Procedure

Participants were scheduled for the experiment on two consecutive days: from 7 to 9 P.M. on day 1, and 8 to 10 A.M. on day 2 (Fig. 1). At the end of the first session, participants were instructed on the use of the sleep-monitoring headband [11,12], and were sent home with the device. Participants returned at 8 A.M. the next day and were first reassessed behaviorally on their perception of the nonnative contrast. This was followed by a second EEG/ERP recording session, and then by administration of the remaining language and cognitive assessments.