

Sleep Underpins the Plasticity of Language Production



M. Gareth Gaskell¹, Jill Warker², Shane Lindsay¹,
Rebecca Frost¹, James Guest¹, Reza Snowdon¹, and
Abigail Stackhouse¹

¹University of York and ²University of Scranton

Psychological Science
2014, Vol. 25(7) 1457–1465
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797614535937
pss.sagepub.com
 SAGE

Abstract

The constraints that govern acceptable phoneme combinations in speech perception and production have considerable plasticity. We addressed whether sleep influences the acquisition of new constraints and their integration into the speech-production system. Participants repeated sequences of syllables in which two phonemes were artificially restricted to syllable onset or syllable coda, depending on the vowel in that sequence. After 48 sequences, participants either had a 90-min nap or remained awake. Participants then repeated 96 sequences so implicit constraint learning could be examined, and then were tested for constraint generalization in a forced-choice task. The sleep group, but not the wake group, produced speech errors at test that were consistent with restrictions on the placement of phonemes in training. Furthermore, only the sleep group generalized their learning to new materials. Polysomnography data showed that implicit constraint learning was associated with slow-wave sleep. These results show that sleep facilitates the integration of new linguistic knowledge with existing production constraints. These data have relevance for systems-consolidation models of sleep.

Keywords

sleep, plasticity, phonotactic constraints, speech errors, production, slow-wave sleep, learning, open data, open materials

Received 8/28/13; Revision accepted 4/4/14

Recent studies have provided numerous demonstrations of the enduring plasticity underlying language processing. This plasticity allows people to learn languages (Mårtensson et al., 2012), as well as new words in their language (Gaskell & Dumay, 2003), and to adapt to new accents or environments (Bradlow & Bent, 2008; Norris, McQueen, & Cutler, 2003). Some studies have revealed that plasticity in language may involve memory consolidation during sleep (e.g., Dumay & Gaskell, 2007; Gomez, Bootzin, & Nadel, 2006). These studies fit with a broader literature highlighting a critical role for sleep in the consolidation of many forms of memory, including paired-associate learning (Fenn & Hambrick, 2012; Marshall, Helgadóttir, Mölle, & Born, 2006), procedural knowledge (Plihal & Born, 1997), and statistical learning (Durrant, Taylor, Cairney, & Lewis, 2011).

Studies of sleep and plasticity in language have focused on two particular areas: word learning (Dumay

& Gaskell, 2007; Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010) and acquisition of grammar (Gomez et al., 2006; Hupbach, Gomez, Bootzin, & Nadel, 2009). The grammar data suggest that sleep can assist in the abstraction of structure from new instances of sequences. The word-learning data show that sleep can facilitate the embedding of novel words in the mental lexicon. Both areas are noteworthy because they provide evidence that sleep does more than merely strengthen new memories (Walker & Stickgold, 2010). Rather, they show that sleep facilitates an integrative process that situates the newly learned information in its appropriate context in long-term memory. This is a key prediction of systems models

Corresponding Author:

M. Gareth Gaskell, Department of Psychology, University of York,
York YO10 5DD, United Kingdom
E-mail: gareth.gaskell@york.ac.uk

of sleep and memory consolidation (Diekelmann & Born, 2010; Walker & Stickgold, 2006), which propose that slow-wave sleep (SWS) in particular provides a means of integrating episodic hippocampal memories with long-term neocortical memory (Alvarez & Squire, 1994).

However, whether these findings can be applied more generally remains in question, both within the language domain and in the memory-consolidation literature. Within language, it is important to determine what other aspects of learning and plasticity are related to sleep and memory consolidation. More broadly, there is a particular need for further paradigms that involve relating newly learned information to existing stable memories, given the value of such paradigms for the predictions of systems-consolidation models. To further elucidate the relationship between language plasticity and memory consolidation, the experiment presented here employed an implicit-learning paradigm in which recent experiences are incorporated into the language-production system and constrain the nature of subsequent speech errors.

All languages have phonotactic constraints, which set restrictions on the acceptable placement of phonemes or combinations of phonemes. First-order constraints concern the syllabic positions of isolated phonemes. For example, in English, /ŋ/ is acceptable at syllable coda (e.g., *sing*) but not at syllable onset (e.g., *ngis*). Second-order constraints involve phoneme combinations—for example, /r/ is acceptable in English at syllable onset in combination with /d/, as in *drink*, but not with /s/, as in *slink*; for /l/, the acceptable combinations are reversed (e.g., *dlink* vs. *slink*). These constraints affect performance in both perception (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999) and production, with speech errors that violate the phonotactic constraints of a speaker's language (e.g., *slink*) being extremely rare (Wells, 1951).

The susceptibility of phonotactic constraints to recent experience was exposed in a seminal study by Dell, Reed, Adams, and Meyer (2000). They tested whether speech errors might be influenced by laboratory-induced first-order constraints, in which certain consonants were artificially restricted to particular syllable positions over several sessions. Speech errors began to mirror these constraints, which suggests that participants had begun to incorporate new phonotactic constraints into their productions as a consequence of their recent experience. These first-order effects, based on simple occurrence frequency of phonemes, emerged quickly on the first day of testing. This learning appears to be implicit, in that informing participants of the nature of the constraints made no difference to their speech errors, and uninformed participants were unable to describe the constraints when prompted.

Warker and Dell (2006) found that second-order constraints could also influence speech errors, but the effects of these more subtle constraints tended to emerge later in the experiment, on the second day of testing. Warker (2013) looked at whether this delayed emergence was a consequence of extra training or the opportunity to consolidate the initial learning, and she found that exposure level had a modest effect, but the opportunity for consolidation was particularly important. Although Warker could not implicate sleep over time in any consolidation process that resulted in constraint acquisition, there are clear precedents in the literature favoring such an account, such as studies showing that sleep can enhance statistical learning (Durrant et al., 2011).

In the current work, we used the phonotactic-learning paradigm to address both the mechanisms of language plasticity and the nature of memory consolidation. The use of this paradigm allows examination of the extent to which plasticity in speech production makes use of sleep to consolidate new phonotactic information. Furthermore, it addresses the extent to which newly learned materials (phonotactic constraints) affect the operation of a well-established neocortical mechanism (i.e., the generation of errors in the speech-production system). The study of speech errors has had a long history of informing the understanding of the normal mechanisms of word production (Levelt, 1999).

We examined the association between sleep and acquisition of second-order phonotactic constraints using a design based on that of Warker (2013). We asked two groups of participants to repeat a short block of syllable sequences containing second-order constraints in which the syllable position of two phonemes (/f/ and /s/) was restricted based on the identity of the syllable's vowel. One group then slept in our lab while polysomnographic data were recorded. The other group remained awake for a similar period, and then both groups were tested for the effect of constraints on their speech errors by undergoing two further blocks of syllable repetition.

Given Warker's (2013) results, we expected that the wake group would show no evidence of the within-experiment, second-order constraints in their speech errors, but our prediction for the sleep group was that sleep would facilitate consolidation of the newly learned phonotactic knowledge, leading to a change in speech-error patterns. We further predicted, on the basis of existing models of declarative-memory consolidation (Diekelmann & Born, 2010), that SWS duration would be correlated with the strength of this change. In a posttest, we looked at whether sleep influenced the extent to which participants could generalize their newly acquired phonotactic knowledge in a perceptual task, given prior evidence for the role of sleep in abstraction of statistical knowledge (Durrant et al., 2011). In this generalization

task, participants selected between untrained syllables that either corresponded to or violated the within-experiment constraints.

Method

Participants

Data were collected for 38 adults (20 male, 18 female) with a mean age of 22 years (range 18–27); one further participant was excluded because of experimenter error (see Additional Methodology and Results in the Supplemental Material available online for further details about participants).

Design and materials

Sleep was manipulated between participants (sleep interval vs. wake interval), and all participants received two sessions of syllable repetition (pre- and postinterval). The preinterval session (one block of 48 sequences) was intended to provide sufficient information about the within-experiment constraints but be short enough not to affect speech errors within that session. The postinterval session was longer (two 48-sequence blocks), but previous research suggested that the wake group would still show no effect of constraints on speech errors (Warker, 2013).

The sequences comprised four consonant-vowel-consonant (CVC) syllables presented visually to be read aloud in time with a metronome (e.g., *baf kan sang gam*). The vowel (/æ/ or /ɪ/) remained the same in each sequence but alternated between sequences. Of the eight consonants used, /h/ was always at syllable onset, /ŋ/ was always at syllable coda, the positions of /f/ and /s/ depended on the vowel (restricted), and the remaining consonants (/n/, /m/, /g/, and /k/) could appear at either syllable onset or coda (unrestricted). All consonants were used once per sequence. Two versions of the experiment language counterbalanced the constraints between participants. In one version, /s/ was restricted to onset and /f/ to coda when the vowel was /æ/, whereas /f/ was restricted to onset and /s/ to coda when the vowel was /ɪ/. Version 2 reversed these constraints. A sequence generator applied these rules pseudorandomly to create the syllable sequences (Warker, 2013).

A two-alternative forced-choice (2AFC) recognition task was used to assess explicit memory of the syllables. Twenty syllables containing restricted consonants were taken from one language version and were randomly paired with 20 equivalent syllables from the second version (e.g., *fim*, *bif*). The two vowels were equally represented. Participants were presented with both syllables and asked to select the syllable from the experiment.

A 2AFC generalization task was used to assess participants' ability to make judgments about the correspondence between novel, visually presented CVC sequences and the repetition materials. Forty CVC syllables were used in which the vowel was either *a* or *i* (20 of each). In each syllable, one consonant came from the restricted set (*s* or *f*; 20 of each), and the other was from a larger set containing only new consonants that were not used in the repetition task. Twenty target CVC syllables adhered to the second-order constraints used in the repetition task, with each constraint equally represented. The foils followed the opposite constraints. Targets were pseudorandomly paired with foils, and participants had to select which syllable best fitted the language they had seen.

Because of experimenter error, the set of “new” consonants contained one consonant that had been used in the repetition task (*m*) alongside four correctly chosen new consonants (*t*, *d*, *l*, *z*). This actually added an interesting complexity to the task. In the majority of trials (12 or 13, depending on the list), both items in the generalization pairs contained one restricted consonant and one novel consonant (categorized as *both unfamiliar*; e.g., *das*, *fid*). For the remainder, one of the syllables used only consonants from the repetition task, whereas the other had a novel consonant (e.g., *mas*, *fid*). Half these items (three or four) used the unfamiliar consonant in the target syllable (*target unfamiliar*), and half used it in the foil (*target familiar*). All these items were retained for analysis, but the data were categorized by the type of trial.¹ (Information relating to assessment of sleepiness, explicit knowledge of the sequences, and reading ability are provided in Additional Methodology and Results.)

Procedure

On arrival to the lab, participants completed a reading-ability test and were then pseudorandomly allocated to the sleep or wake condition. Sleep participants were wired up using a nine-electrode electroencephalogram (EEG) montage (see Additional Methodology and Results). During the first session of syllable repetition (48 sequences), participants sat in individual booths, with testing administered using PsychoPy software (Peirce, 2007). Participants were instructed to read each sequence aloud in time to a metronome, once slowly (1 Hz) and three times at a faster pace (2.5 Hz). The metronome clicks began 1 s after the sequence appeared on screen and were delivered via headphones. The interstimulus interval was 1 s. Spoken responses were recorded using a Sony MP3 recorder.

Participants were told whether they would be napping only after training. Sleep participants were given 90 min to nap in a bedroom, followed by an interval of at least

Table 1. Mean Number of Errors and Percentages of Errors That Preserved Syllable Position in Training and Testing

Session and measure	Sleep group		Wake group	
	Unrestricted consonants	Restricted consonants	Unrestricted consonants	Restricted consonants
Training				
Same-position errors only	85% (11%)	80% (8%)	79% (10%)	90% (14%)
Total number of errors	45.3 (14.1)	9.5 (4.7)	50.2 (28.5)	8.8 (5.4)
Test Block 1				
Same-position errors only	84% (11%)	92% (9%)	85% (11%)	86% (15%)
Total number of errors	32.3 (15.4)	4.9 (3.8)	39.8 (17.3)	6.1 (4.7)
Test Block 2				
Same-position errors only	84% (9%)	94% (17%)	84% (8%)	85% (16%)
Total number of errors	28.3 (25.4)	6.0 (5.4)	29.5 (25.6)	6.8 (4.9)

Note: Standard deviations are given in parentheses.

20 min to reduce sleep inertia. Wake participants stayed in the testing booth and watched a 105-min video followed by a 5-min break. The video had little dialogue and was chosen to occupy the participants while minimizing interference with the language tasks. In the test session (25 min), participants completed two 48-sequence blocks of syllable repetition followed by the recognition and generalization tests.

As in Warker (2013), speech errors were coded based on whether a slipping consonant slipped to the same syllable position in a given sequence, creating a *same-position error*, or to a different syllable position, creating a *different-position error* (see Additional Methodology and Results).

Results

Speech errors

Speech errors for 1 participant who did not complete the task were excluded from analysis (see Table 1 for the remaining data). As in previous studies, the primary measure was the percentage of errors (per participant) involving a particular consonant type that were classed as same-position errors (i.e., that did not change syllable position). To provide the cleanest estimate of production-error performance given small numbers of errors, we pooled data from the postinterval blocks for analysis. For unrestricted consonants, the percentage of same-position errors provides a baseline measure of the extent to which a participant's speech errors preserve syllable position. For restricted consonants, the key question is whether the measure rises significantly above the unrestricted baseline rate, which would suggest that newly acquired phonotactic constraints are influencing production errors. We termed this difference between restricted and unrestricted same-position percentages the *phonotactic-learning score*, for which positive values suggest that phonotactic learning has taken place.

Because participants and conditions varied substantially in the numbers of speech errors observed and because the dependent measure was binary, we ran a mixed-effects logistic regression using lme4 in R on the error data, with the syllabic position of the incorrect consonant (same or different) as the binary dependent variable. Session (preinterval vs. postinterval), restriction (restricted consonant vs. unrestricted consonant) and group (sleep vs. wake) were all effect coded (-.5, .5). Random effects for participants were incorporated with maximal inclusion of random slopes for within-participants variables and their interactions (Barr, Levy, Scheepers, & Tily, 2013), and *p* values were calculated using Wald *z*. The analysis revealed a significant three-way interaction between session, restriction, and group ($\beta = 1.97$, $SE = 0.50$, $z = 3.31$, $p < .001$), as well as less important effects of session ($\beta = 0.30$, $SE = 0.14$, $z = 2.14$, $p = .03$), restriction ($\beta = 0.65$, $SE = 0.20$, $z = 3.34$, $p < .001$), and a group-by-session interaction ($\beta = 0.57$, $SE = 0.28$, $z = 2.01$, $p = .044$).

Planned comparisons examining the phonotactic-learning score showed no effect prior to the interval for the sleep group ($\beta = -0.09$, $SE = 0.35$, $z = -0.28$, $p = .78$) but a clear effect after sleep ($\beta = 1.37$, $SE = 0.41$, $z = 3.38$, $p < .001$). When broken down by block, the effect was as strong in the first block after sleep (phonotactic-learning score = 14%; $\beta = 1.40$, $SE = 0.40$, $z = 2.48$, $p = .013$) as it was in the second block (phonotactic-learning score = 13%; $\beta = 1.02$, $SE = 0.43$, $z = 2.35$, $p = .019$). Surprisingly, the wake group did show an effect prior to the interval ($\beta = 1.11$, $SE = 0.33$, $z = 3.35$, $p < .001$), but this effect did not remain significant after the interval ($\beta = 0.21$, $SE = 0.28$, $z = 0.74$, $p = .46$). The effect for the wake group prior to the interval was unexpected for several reasons. Prior experiments (Warker, 2013; Warker & Dell, 2006) suggest that such effects should not be present after exposure to so few sequences. Furthermore, the sleep group received identical training and showed no such effect at that point. Finally, the effect in the wake group

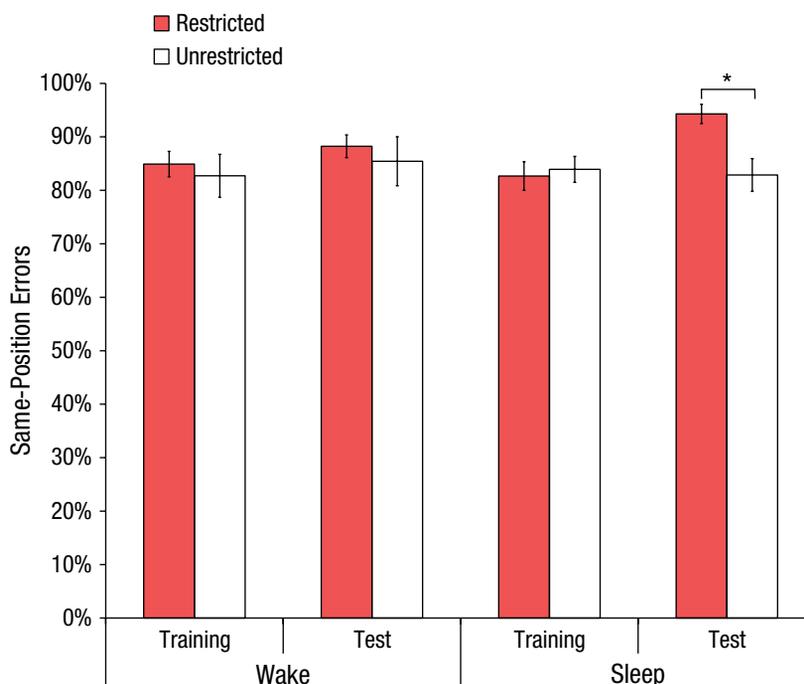


Fig. 1. Mean percentage of same-position errors as a function of group, session, and consonant restriction. Error rates are shown for the subset of participants matched on training performance. The asterisk indicates a significant difference between sequences with restricted and unrestricted consonants ($p < .05$). Error bars indicate standard errors.

was not observed after the interval, when the increase in the number of sequences provided greater power to observe such an effect. It seems likely, therefore, that it was a Type I error. It is worth noting that the number of restricted different-position errors that wake participants made was near constant across the three blocks (training: 0.9, Test Block 1: 0.8, Test Block 2: 1.1), whereas there was a clear drop in this error type between training and test for the sleep participants (training: 1.9, Test Block 1: 0.4, Test Block 2: 0.4).

To examine whether the groups showed a different effect of the interval when matched on performance prior to the interval, we selected a subset of 28 of the 38 participants by progressively excluding outliers from both groups on the basis of their preinterval phonotactic-learning score (see Fig. 1). Unsurprisingly, for this subset, there was no difference in the nature of their errors in training (i.e., no effect of group, restriction, or an interaction between group and restriction; all $ps > .17$). However, the Group \times Restriction \times Session interaction remained significant for this subset ($\beta = 1.35$, $SE = 0.61$, $z = 2.20$, $p = .028$), with the Restriction \times Session interaction found for participants who slept ($\beta = 1.15$, $SE = 0.50$, $z = 2.31$, $p = .021$) but not those who remained awake ($\beta = -0.66$, $SE = 0.44$, $z = -1.50$, $p = .13$). After the interval, there was still no effect for the wake group ($\beta = -0.19$, $SE = 0.24$, $z = -0.81$, $p = .41$), whereas the sleep group showed a clear effect ($\beta = 1.04$, $SE = 0.44$, $z = 2.38$, $p = .017$).

2AFC tests

Recognition performance was uniformly good, with no differences between groups (sleep group = 88%, wake group = 87%) in correct response rates. For the generalization test, mixed-effects logistic regression was run on the correct response rates, with group (sleep vs. wake) and trial type (target familiar vs. target unfamiliar vs. both unfamiliar) as effect coded independent variables (see Table 2). This showed that the familiarity of the two response options affected selection in that performance in both the target-familiar condition ($\beta = 2.95$, $SE = 0.52$, $z = 5.70$, $p < .001$) and the target-unfamiliar condition ($\beta = -1.94$, $SE = 0.31$, $z = -6.29$, $p < .001$) differed from the overall mean performance level. More important, scores were higher in the sleep group (69% correct) than in the wake group (56% correct; $\beta = 1.11$, $SE = 0.53$, $z = 2.08$, $p = .038$). In fact, only the sleep group showed familiarity with the

Table 2. Mean Percentage of Correct Responses in the Generalization Test for Each Trial Type

Trial type	Sleep group	Wake group
Target familiar	99% (3%)	92% (3%)
Target unfamiliar	47% (7%)	24% (7%)
Both unfamiliar	60% (4%)	52% (4%)

Note: Standard deviations are given in parentheses.

Table 3. Sleep-Stage Data and Their Association With Changes in Phonotactic-Learning Score

Sleep stage	Mean duration (min)	β^a	r^b
Stage 1	8.6 (6.9)	-0.06 (0.04)	-.21
Stage 2	43.2 (18.9)	-0.03 (0.02)	-.06
Slow-wave sleep	14.6 (13.5)	0.06** (0.02)	.47*
REM	4.3 (6.8)	-0.07* (0.03)	-.49*
Sleep period	82.5 (8.2)	-0.06* (0.03)	-.22

Note: For sleep-stage durations, standard deviations are given in parentheses; for betas, standard errors are given in parentheses. Tests of significance were conducted without Bonferroni correction.

^aThis column presents logistic regression coefficients for the interaction of sleep-stage duration with session and restriction in predicting phonotactic-learning score. ^bThis column presents by-participant correlations between sleep-stage duration and change in phonotactic-learning score across the interval between training and test.

* $p < .05$. ** $p < .01$.

constraints learned during the experiment and an ability to generalize those constraints, as assessed by comparisons with chance for the both-unfamiliar items (sleep group: $\beta = 0.43$, $SE = 0.17$, $z = 2.59$, $p = .0096$; wake group: $\beta = 0.09$, $SE = 0.15$, $z = 0.61$, $p = .54$). Note, however, that no participant was able to describe explicitly any of the constraints that he or she had learned (see Additional Methodology and Results), which shows that even the participants who could generalize their knowledge did so in a somewhat implicit way. Interactions between group and trial type were not significant (target familiar: $\beta = 0.79$, $SE = 1.04$, $z = 0.77$, $p = .44$; target unfamiliar: $\beta = -0.03$, $SE = 0.62$, $z = -0.05$, $p = .96$).

Correlation analyses

Our key finding was that sleep participants showed a shift in their phonotactic-learning score across the interval, which suggests that their use of phonotactic constraints in speech production had been altered by recent experience. Given the a priori evidence for an involvement of SWS in memory consolidation, we predicted that sleep participants would show an association between SWS and this behavioral shift. A logistic regression analysis looking at the interaction of Session \times Restriction \times SWS Duration did indeed reveal a significant relationship ($\beta = 0.06$, $SE = 0.02$, $z = 3.10$, $p = .002$; see Table 3).

When logistic regression interactions with other stages were examined in a post hoc analysis, unexpected effects in the opposite direction were found for REM sleep duration and overall sleep period, perhaps partly because SWS and REM sleep durations were themselves negatively correlated ($r = -.51$, $p = .026$). However, a regression equation that included all the three-way interactions with sleep parameters left only the beta value for SWS

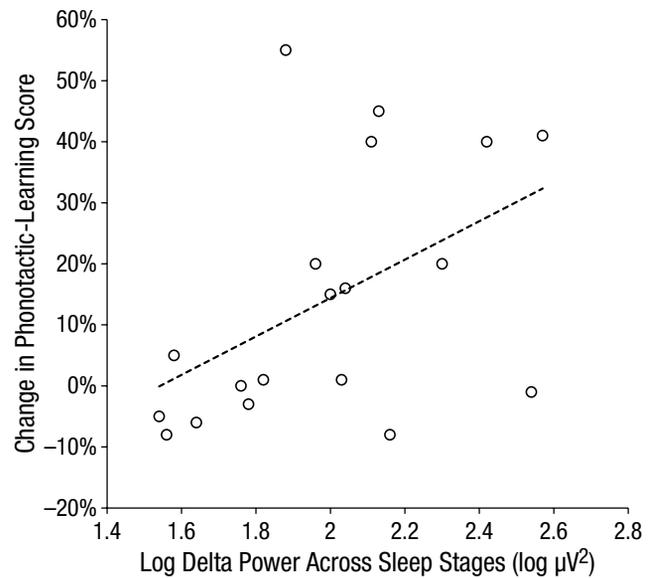


Fig. 2. Scatter plot (with best-fitting regression line) showing the correlation between log delta power (averaged across central and frontal electrodes in all sleep stages) and change in phonotactic-learning score across the interval between training and test for the sleep participants.

duration as a significant independent predictor ($\beta = 0.05$, $SE = 0.03$, $z = 2.15$, $p = .03$).

To provide additional evidence, we took another measure of slow-wave activity: the average log power in the delta band (0.3–4.0 Hz) across all sleep stages. This measure includes oscillatory activity for epochs that do not qualify as SWS. Delta activity, which is independent of sleep duration, was significantly correlated with time in SWS ($r = .79$, $p < .001$) but not time in REM sleep ($r = -.35$, $p = .14$). This measure also showed a significant interaction with Session \times Restriction ($\beta = 0.007$, $SE = 0.003$, $z = 2.64$, $p = .008$), which supports the proposition that the incorporation of new phonotactic constraints in the production system is associated with slow waves (see Fig. 2 for equivalent correlation).

Given the group difference in generalization ability, correlation analyses were run for this measure. There was no association between SWS duration and generalization ability ($r = .011$, $p = .965$), and post hoc analyses suggested that no other stage was associated with this effect.

Discussion

Our experiment revealed a clear association between sleep and the engagement of new phonotactic knowledge in the production system. Participants who stayed awake between a short block of syllable repetition and a longer test session showed no influence of the recent experience on their pattern of speech errors during test. On the other hand, participants who slept in the

intervening period demonstrated implicit effects of this recent experience in their speech errors, with this effect being present in both blocks of the test session. Furthermore, when confronted with two novel syllables, sleep participants were better able to select items that fit with the constraints from the experiment than wake participants, who remained at chance. Whereas this latter effect was not associated with any particular component of sleep, the implicit effect on speech errors showed a positive correlation with time in SWS and a negative correlation with time in REM and sleep duration. On balance, the correlation with SWS is more likely to have been causally implicated in the consolidation effect, given further correlations with EEG delta power.

The speech-error result is compatible with a growing body of evidence that sleep has a particular role to play in the transformation rather than stabilization of recent memories, such that after sleep, the new knowledge is in a more useful state for future actions (Stickgold & Walker, 2013). This argument has been made on the basis of several studies in which participants were required to generate inferences based on the integration of a set of new facts (Djonlagic et al., 2009; Ellenbogen, Hu, Payne, Titone, & Walker, 2007). Other studies have found that sleep benefits the integration of new memories with existing long-term memories (Dumay & Gaskell, 2007, 2012; Gaskell & Dumay, 2003).

The current paradigm has some facets of both the above forms of integration. The phonotactic constraints can be formed only by integrating statistical knowledge across many production sequences (cf. Djonlagic et al., 2009; cf. Ellenbogen et al., 2007). Furthermore, the implicit phonotactic effect was observed in normal speech errors, which are clearly also influenced by long-term knowledge of phonotactic constraints in the speaker's native language (Dell et al., 2000). This aspect can therefore be explained as integration of new experiences with long-term memory (cf. Dumay & Gaskell, 2007). More specifically, a participant enters the experiment with existing phonotactic knowledge from years of experience speaking English. During the experiment, the participant experiences a different phonological distribution that is not predicted by the existing phonotactic knowledge (e.g., the participant implicitly learns that /f/ tends to occur at syllable onset only if followed by /i/). This distribution becomes encapsulated with the experimental context, allowing the recently experienced constraints to become integrated with existing phonotactic knowledge in long-term memory without overwriting that knowledge. During sleep, the association between syllable position and vowel identity is consolidated and becomes embedded in the language-production system, which means that when the participant returns to the experimental context, the newly acquired constraints influence

production errors in combination with more established language-wide constraints. This ensures that the participant's language model is as up-to-date as possible, optimizing the nature of future linguistic interactions.

It is worth noting that integration as just described may not be the end of the integration process. Potentially, after further weeks or months of consolidation, the new knowledge may become decontextualized so that it has a more general (and yet less dramatic) effect on speech production. A further point is that first-order constraints (e.g., /f/ is always at syllable onset) show integration effects without the need for sleep or extended experience. In this case, it may be that the simplicity of the constraint (observable in simple positional frequency measures) allows it to become influential prior to sleep. Performance on a transitive-inference task has shown a similar dissociation. Ellenbogen et al. (2007) found that simpler inferences (single mediating item) did not benefit from sleep over wake, but there was a sleep-specific boost in performance on more complex inferences over two mediating items. The relative complexity or context-dependence of the second-order constraints here may be the reason that sleep facilitates their emergence in production errors.

Participants presumably made use of their updated language model during the forced-choice generalization task. Their ability to apply implicitly acquired phonotactic knowledge to novel syllables in a separate, explicit decision suggests that participants did not merely memorize the syllables they experienced during the production task. Rather, it suggests that participants learned a more abstract form of the second-order constraints (but one still largely unavailable to verbal description). Given that those who slept performed better at this task than those who remained awake, sleep may play a particular role in the reorganization and formation of abstract knowledge of information about the experimental constraint, allowing novel sequences to be accepted as matching those constraints (cf. Durrant et al., 2011).

The current data establish that the acquisition of these second-order constraints is associated with sleep rather than simply the passing of time (cf. Warker, 2013). This effect is not predicted by Warker and Dell's (2006) connectionist model of plasticity in speech production. However, it can be accommodated by a complementary systems variant (Davis & Gaskell, 2009; McClelland, McNaughton, & O'Reilly, 1995), in which the existing network is supplemented by a complementary network (modeled on the hippocampus) that specializes in the temporary acquisition of new mappings that cannot be easily acquired in the existing network (within-experiment, second-order phonotactic constraints). As in other studies of this nature, the evidence that SWS is related to restructuring is correlatory and does not imply a causal explanation. Nonetheless, converging evidence from a

wide range of methods (e.g., Ngo, Martinetz, Born, & Mölle, 2013) fits parsimoniously with an account in which sleep, and more particularly SWS, provides a mechanism for transfer of knowledge to the main network just as in the systems-consolidation models referred to above (e.g., Diekelmann & Born, 2010). McClelland (2013) demonstrated how a complementary systems model explains the incorporation of new memories into existing schema (Tse et al., 2007; van Kesteren et al., 2013). The current paradigm can be described in the same terms. Participants had pre-established long-term knowledge of the likelihood of encountering various phoneme combinations in different syllable positions. The syllable-repetition task altered these expectations in specific ways, and sleep provided the means for off-line consolidation of this knowledge, updating the phonotactic schema.

In conclusion, we have shown that sleep-associated consolidation is beneficial to the integration of complex phonotactic constraints in speech production. After a small amount of exposure to sequences that subtly conflicted with participants' phonotactic knowledge of English syllables, sleep provided a means of updating their language models. Participants who slept exhibited changes in their pattern of speech errors, particularly those with high levels of SWS, whereas those who stayed awake showed no change. These results fit with a general model of language learning and processing, in which complementary systems underpin the substantial plasticity that remains in the system even during adulthood.

Author Contributions

M. G. Gaskell developed the study concept. All authors were involved in the design of the experiment. R. Frost, J. Guest, R. Snowdon, and A. Stackhouse ran the experiment and conducted the preliminary analyses under the supervision of M. G. Gaskell and S. Lindsay. M. G. Gaskell ran the final analyses and wrote the manuscript, based on an initial draft by R. Frost. J. Warker and S. Lindsay provided critical feedback. All authors approved the final version of the manuscript for submission.

Acknowledgments

S. Lindsay is now at the University of Dundee. R. Frost is now at the University of Lancaster.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

The research was supported by Economic and Social Research Council grants awarded to M. G. Gaskell.

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>

Open Practices



All data and materials have been made publicly available via Open Science Framework and can be accessed at osf.io/zqg9y. The complete Open Practices Disclosure for this article can be found at <http://pss.sagepub.com/content/by/supplemental-data>. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at <https://osf.io/tvyxz/wiki/view/> and <http://pss.sagepub.com/content/25/1/3.full>.

Note

1. One item in one list used *m* in both alternatives, which made it equivalent to a recognition trial. This item was excluded from analyses.

References

- Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Sciences, USA*, *91*, 7041–7045.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml.2012.11.001
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707–729. doi:10.1016/j.cognition.2007.04.005
- Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 3773–3800. doi:10.1098/rstb.2009.0111
- Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1355–1367.
- Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*, 114–126. doi:10.1038/nrn2762
- Djonlagic, I., Rosenfeld, A., Shohamy, D., Myers, C., Gluck, M., & Stickgold, R. (2009). Sleep enhances category learning. *Learning & Memory*, *16*, 751–755.
- Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*, 35–39.
- Dumay, N., & Gaskell, M. G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition*, *123*, 119–132. doi:10.1016/j.cognition.2011.12.009
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1568–1578.
- Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*, 1322–1331.

- Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences, USA*, *104*, 7723–7728.
- Fenn, K. M., & Hambrick, D. Z. (2012). Individual differences in working memory capacity predict sleep-dependent memory consolidation. *Journal of Experimental Psychology: General*, *141*, 404–410.
- Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*, 105–132.
- Gomez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science*, *17*, 670–674.
- Hupbach, A., Gomez, R. L., Bootzin, R. R., & Nadel, L. (2009). Nap-dependent learning in infants. *Developmental Science*, *12*, 1007–1012. doi:10.1111/j.1467-7687.2009.00837.x
- Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*, 223–232. doi:10.1016/S1364-6613(99)01319-4
- Marshall, L., Helgadóttir, H., Mölle, M., & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature*, *444*, 610–613.
- Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, *63*, 240–244. doi:10.1016/j.neuroimage.2012.06.043
- McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General*, *142*, 1190–1210.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.
- Ngo, H.-V. V., Martinetz, T., Born, J., & Mölle, M. (2013). Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. *Neuron*, *78*, 545–553. doi:10.1016/j.neuron.2013.03.006
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*, 204–238.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.
- Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience*, *9*, 534–547.
- Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience*, *16*, 139–145.
- Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience*, *30*, 14356–14360. doi:10.1523/jneurosci.3028-10.2010
- Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., . . . Morris, R. G. M. (2007). Schemas and memory consolidation. *Science*, *316*, 76–82.
- van Kesteren, M. T., Beul, S. F., Takashima, A., Henson, R. N., Ruiters, D. J., & Fernández, G. (2013). Differential roles for medial temporal and medial prefrontal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, *51*, 2352–2359.
- Walker, M. P., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annual Review of Psychology*, *57*, 139–166.
- Walker, M. P., & Stickgold, R. (2010). Overnight alchemy: Sleep-dependent memory evolution. *Nature Reviews Neuroscience*, *11*, 218. doi:10.1038/nrn2762-c1
- Warker, J. A. (2013). Investigating the retention and time course of phonotactic constraint learning from production experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 96–109.
- Warker, J. A., & Dell, G. S. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 387–398. doi:10.1037/0278-7393.32.2.387
- Wells, R. (1951). Predicting slips of the tongue. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 82–87). The Hague, The Netherlands: Mouton. (Reprinted from *The Yale Scientific Magazine*, XXVI, 9–30)