

Learning repetition, but not syllable reversal

Elliott Moreton,¹ Brandon Prickett,² Katya Pertsova,¹ Josh Fennell,¹ Joe Pater,² and Lisa Sanders¹

¹ University of North Carolina, Chapel Hill, and ² University of Massachusetts, Amherst

1 Introduction

Reduplication is common, but analogous processes which reverse or otherwise permute the order of segments or syllables are rare, even in language games. This is true even though reversal involves nested dependencies and so is less complex on the Chomsky Hierarchy than reduplication, which involves crossed dependencies (i.e., context-free vs. context-sensitive).

We hypothesize that the explanation is as follows. Phonological learning and processing rely on phonological short-term memory, which stores items in temporal sequence. Repetitions can be recognized when they match and reactivate a stored trace in short-term memory. Recognizing a reversal, on the other hand, requires effortfully and time-consuming rearranging the input in working memory before attempting to match it to the stored trace. Repetitions can thus be recognized, and repetition patterns learned, implicitly, whereas reversals require explicit, conscious awareness on the part of the learner. The extra difficulty of reversals makes them harder to learn and so typologically rarer. In addition, the need for explicit effort (we hypothesize) makes reversal incompatible with natural language, but allows some reversal in language games.

This paper presents two experiments to test these hypotheses, within a larger research program into the relation between phonological and non-phonological learning, memory, and cognition.

1.1 Detecting symmetries Reduplication and reversal create symmetric outputs. A stimulus is symmetric with respect to a transformation if the transformation does not change the stimulus. The reduplicated string *abcabc* is symmetric with respect to the transformation that exchanges the two halves of the string, and the “reverplcated” string *abccba* is symmetric with respect to the transformation that mirror-reverses the string.

Humans do not treat all symmetries alike. A few can be detected rapidly and with little effort, while the rest require conscious manipulation using working memory. A well-studied example occurs with static visual images (reviewed in Wagemans 1995; Treder 2010; van der Helm 2015): Reflectional symmetry around a vertical axis (as in the letters A, T, Y) is detected faster and more accurately than reflectional symmetry around any other axis (e.g., B, E, K), rotational symmetry (e.g., S, Z), or repetition in space. Bilateral symmetry is detected early and automatically, before eye movements can be programmed and executed, and locally-symmetric regions attract fixations (Kootstra et al., 2011). Increasing figure complexity slows judgements of repetition symmetry (e.g., [I]) much more than those of mirror symmetry (e.g., [J]; Baylis & Driver 1994).

In auditory perception, with reflections and translations in time rather than space, the opposite occurs: Repetitions of musical phrases like *ADCB ADCB* are detected more accurately than reversals like *ADCB BCDA* (Krumhansl et al., 1987; Schulze et al., 2012; Schulze & Tillmann, 2013). Detecting musical palindromes is difficult even for trained musicians (Dowling, 1972; Balch, 1981; Dienes & Longuet-Higgins, 2004; Mongoven & Carbon, 2017; Petrović et al., 2017), whereas repetitions are ubiquitous (there is even standard musical notation for them, e.g., the repeat signs |: and :|, the verbal directions *da capo* and *dal segno*, and the measure-repeat sign %). The leading account of this difference is that repetition can be detected when the memory trace of the first presentation is matched and automatically reactivated by the second presentation, whereas detecting a reversal requires mentally rearranging one or the other (Dowling, 1972; Schulze et al., 2012; Li et al., 2013).

* The authors are indebted for comments and suggestions to Adam Aji, Kenneth J. Kurtz, Caitlin Smith, Hossep Dolatian, and the P-Side Caucus at UNC-Chapel Hill. This research was supported by U.S. National Science Foundation collaborative grant BCS 1651105/1650957 “Inside Phonological Learning”.

This view is supported by independent evidence for automatic, effortless matching of auditory stimuli to memory traces. Auditory word recognition (Marslen-Wilson, 1987; Norris & McQueen, 2008) and melody recognition (Dalla Bella et al., 2003) involve matching a stimulus to a trace in long-term memory. The electro- or magnetoencephalographic auditory mismatch negativity (MMN) is observed when repetitions of one stimulus, the “standard”, are followed by a different stimulus, the “deviant”, even when the stimuli are not attended. A mechanism based on short-term memory must automatically detect both the repetitions that establish the standard, and the non-repetition that distinguishes the deviant (Näätänen et al., 2007). The MMN has been observed for a range of auditory stimuli, including music and speech. For speech stimuli, the relevant level of representation is phonological rather than phonetic or auditory, as the MMN is more sensitive to contrastive than to subphonemic acoustic differences (Dehaene-Lambertz, 1997; Näätänen et al., 1997).

Evidence that reordering auditory stimuli is effortful comes from a variety of sources. A large body of work on list recall, reviewed in Hurlstone et al. 2014 and Donolato et al. 2017, shows that word lists presented by ear, unlike those presented by eye, are easier to recall forwards than backwards. Recall-memory spans are longer for sequences of digits, of tones, and of pseudowords in forward than in backward order (Benassi-Werke et al., 2012). Recalling a word list in the order in which it was presented is also easier than recalling it in a sorted order (Metcalfe & Sharpe, 1985). Mentally reversing a familiar melody recruits brain regions not involved in simply listening to it (Zatorre et al., 2010).

Assuming syllables within a wordlike stimulus are processed similarly to the notes within a melody or words within a list, we can expect a repetition within a pseudoword to be easier to detect than an internal reversal or any other rearrangement.

1.2 Explicit and implicit processes Participants in phonological-learning (“artificial-language”) experiments do not all use the same learning algorithm. Analysis of post-experiment debriefing questionnaires and individual learning curves shows that some rely on intuition; others search for an explicit rule, with varying results; others memorize individual stimuli; and still others use a combination of these approaches (Kimper 2016; Moreton & Pertsova 2016; Moreton & Pertsova, submitted). Implicit and explicit learning are not mutually exclusive, and may even facilitate each other by allowing either system to learn from the other (Ellis, 2005; Sun et al., 2007).

If repetition detection is automatic and effortless, whereas reversal detection requires conscious manipulation of the stimulus, it follows straightforwardly that a participants’ performance on a reversal task, but not necessarily on a repetition task, ought to depend on their use of explicit reasoning.

Since natural language appears to be learned and processed mainly implicitly in first-language and proficient second-language speakers (Ellis, 2005), the conscious effort required to process reversal may also explain why repetition is common, and reversal very rare, in natural language (Inkelas & Downing, 2015). It would also explain why reversal is more common, but still rare, in consciously-created cultural products such as music (Dowling, 1972; Kempf, 1996) and language games (Cowan et al., 1985; Bagemihl, 1989; Gil, 1996) than in natural language, and why cultural products which do rely on permutation of speech sounds, such as language games and anagrams, derive their amusement value precisely from their difficulty.

2 Experiment 1: Discovering reduplication and reversal

The first experiment compared inductive learning of a partial reduplication pattern with that of a syllable-order-reversing analogue to test two hypotheses: **H1**: Repetition will be easier to learn than reversal, regardless of whether learning is explicit or implicit (since both learning modes depend on phonological memory). **H2**: Reversal can *only* be learned explicitly (by using working memory to re-order the syllables).

2.1 Procedure Participants were recruited online, using the Prolific Academic platform (www.prolific.co), and randomly assigned to Red(uplication) or Rev(ersal) pattern groups. They were told that they would hear “words”, some of which fit a pattern and some of which did not, and that their task was to learn to distinguish between conforming and non-conforming “words”. On each of 50 trials, a seven-syllable pseudoword was presented, of the form *abcd_ _ _*. The syllables “_ _ _” were *abc* in conforming Red words, or *cba* in conforming Rev words (Figure 1). The conforming Red words thus simulated nonlocal reduplication in natural languages (Riggle, 2004). The non-participating syllable *d* was there to avoid a conspicuous *cc* sequence in Rev-

conforming stimuli. For the purposes of Hypotheses **H1** and **H2**, it did not matter whether a participant treated the first four or the last four syllables as the “base”, or even whether they distinguished a base at all.

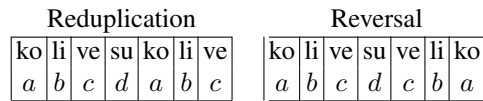


Figure 1: Stimulus structure for the Red(uplication) and Rev(ersal) conditions of Experiments 1 and 2.

On each trial, either a conforming stimulus or a non-conforming foil was presented. Participants would ideally make their response by deciding whether the stimulus conformed to the target pattern. However, they might, in reality, learn instead to recognize the schema used to generate the foils, in which case the experiment would no longer address the hypotheses. To minimize the chances that that would happen, the following steps were taken. The task was framed as learning to recognize conforming stimuli, an instruction which is known to promote a focus on the target pattern rather than on its complement (Gottwald, 1971; Peters & Denny, 1971). The foils were made by randomly transposing either Syllables 5 and 6, or else Syllables 6 and 7, from their conforming order, thus changing the smallest possible number of precedence relations and making the non-conformity as subtle as possible. Each participant thus experienced two different foil schemas. That made foils a less-homogeneous class than conforming stimuli, hence “ground” rather than “figure”, and forestalling any foil-detection strategy that focused on one specific position in each half of the stimulus (e.g., comparing Syllables 1 and 7), avoiding a problem noted by Li et al. (2013). Each syllable had the structure CV , where C was one of [p t k b d g f s j v z ʒ], and V one of [i e u o], making it phonotactically legal in English. The $abcd$ sequences were sampled from the same distribution in both Red and Rev conditions, independently for each participant. Twenty-five conforming and 25 non-conforming stimuli were randomly generated for each participant. Stimuli were synthesized using Tacotron 2 (Shen et al. 2018, fine-tuned by Adam Aji),

All participants were told, using written instructions, that the experiment was “about discovering patterns in the sounds of words”. They were instructed that “[a]t first you’ll just be guessing, but if you pay attention to the nonsense words (especially, how they sound), you’ll be able to figure out what the pattern is and get it right every time”.

After hearing a stimulus once, the participant judged whether the word fit the pattern (“Did that word sound like it followed the secret pattern? Yes/No”), received feedback (“Correct”/“Incorrect”, displayed for 1.5 s), and proceeded to the next stimulus. When all 50 trials were over, each participant completed a debriefing questionnaire, based on that used in Moreton and Pertsova (2016; submitted), which (inter alia) invited them to state an explicit rule.

Questionnaire responses were coded by two independent coders, using criteria based on those in Moreton & Pertsova 2020. Disagreements were resolved by a third independent coder. Cohen’s κ statistic for inter-rater reliability, shown in Table 1, was always 0.63 or greater, considered satisfactory by convention (Munoz & Bangdiwala, 1997; McHugh, 2012). Of 100 participants, 2 were excluded for taking notes, and 4 for interruption (at least one response time > 60 seconds), leaving 46 Red and 48 Rev.

Variable	Exp. 1			Exp. 2		
	N	disagreements	κ	N	disagreements	κ
Stated rule	100	16	0.679	104	16	0.696
Rule type (Red/Rev/other)	38	7	0.711	42	10	0.651
Stated correct rule	38	5	0.737	42	8	0.606
Focused on which word part	100	19	0.741	104	15	0.788
Compared first and last syllable	100	7	0.630	104	7	0.737
Took written notes	100	1	0.662	104	0	1

Table 1: Cohen’s κ for inter-rater reliability in scoring the post-experiment questionnaire.

2.2 Results

2.2.1 Self-reported strategy and insight into pattern structure Does the target pattern affect self-reported strategy? E.g., it could be that everyone tries rule-seeking first, and the Red group succeeds at that while the Rev group fails, so only the Rev group goes on to use memorization or intuition. It turned out, though, that in fact there were essentially no differences between the two groups in self-reported approach.

One part of the questionnaire offered check boxes (i.e., non-mutually-exclusive options) for self-reported strategy. Responses are shown in Table 2. Nearly everyone reported rule-seeking; about half reported using intuition; and about a third reported trying to memorize individual stimuli. The target pattern did not make any discernable difference in self-reported strategy.

	sought rule		intuition		memorized		Rule statement	Red	Rev
	Red	Rev	Red	Rev	Red	Rev			
FALSE	3	7	20	25	37	38	correct rule	14	7
TRUE	43	41	26	23	9	10	incorrect rule	12	8
							no rule	20	33

Table 2: Self-reported learning strategies and experimenter-scored rule accuracy, Experiment 1.

The free-response questions were coded for rule-stating and rule correctness as described in the previous section. The rate of rule-stating was significantly higher in the Red condition than in the Rev condition ($p = 0.02163$ by Fisher’s Exact Test, two-sided). Among those who did state a rule, there was no significant difference in rule correctness between the two conditions ($p = 0.7513$ by Fisher’s Exact Test, two-sided). Thus, although participants in both conditions sought rules at about the same rate, rule-seeking led to correct rule-stating more often in the Red condition than the Rev condition, supporting the hypothesis that the Red rule is easier to learn explicitly than the Rev rule (1). That could be because the explicit system is better at finding the Red rule than the Rev rule by itself. Alternatively, it could be that the implicit system succeeds better in the Red condition than the Rev condition, and then facilitates the explicit system’s rule search (Ellis, 2005; Sun et al., 2007).

2.2.2 Classification accuracy The explicit/implicit distinction was operationalized as stating vs. not stating the correct rule (i.e., stating an incorrect rule or, more often, no rule; Table 2). The left panel of Figure 2 compares those who stated the correct rule (dashed) vs. those who did not (solid) in the Red (red) and Rev (blue) conditions. The curves show an 11-trial moving-average window, with 95% bootstrap confidence intervals. Within each Pattern group, the Correct Staters outperformed the others, and the Red Correct Staters outperformed the rest starting very early.

A mixed-effects logistic-regression model was fit using the *glmer* function from R’s *lme4* package (Bates et al., 2015), with *Response* as the dependent variable, based on 4700 observations from 94 valid participants. The fixed predictors were *Pattern*, *Stated Correct Rule*, *Stimulus Conformity* (1 if the stimulus on that trial was pattern-conforming, 0 if not), and trial number, plus all possible two-, three-, and four-way interactions, plus by-participant random intercept and random slopes for *Stimulus Conformity* and trial number. The hypotheses were tested as planned comparisons with the *glht* function in the *multcomp* package in R (Hothorn et al., 2008), using the coefficient for *Stimulus Conformity* at the last trial in each condition as the measure of learning.

H1 predicted that within the Correct Staters, and again within the Not-Correct-Staters, Red would outperform Rev. Both predictions were supported ($\beta_{\text{Red Correct Staters}} - \beta_{\text{Rev Correct Staters}} = 1.65, p = 0.0047$; $\beta_{\text{Red Not-Correct-Staters}} - \beta_{\text{Rev Not-Correct-Staters}} = 0.99, p = 0.00524$).

H2 predicted that within the Rev group, the Not-Correct-Staters would be at chance. This (null) hypothesis was not rejected: $\beta_{\text{Rev Not-Correct-Staters}} = -0.047, 95\% z \text{ CI} = [-0.46, 0.37]$. The Rev Correct Staters’ estimate was $\beta_{\text{Rev Correct Staters}} = 1.59, 95\% z \text{ CI} = [0.53, 2.65]$. I.e., the Rev pattern was learned at best feebly by the Not-Correct-Staters, but it was learned better by the Correct Staters.

2.2.3 Reaction times According to the memory-based hypothesis underlying this experiment, distinguishing Red-positive from Red-negative stimuli can be done by simply checking whether the memory trace of the first part of the stimulus is or is not reactivated by the second part, whereas to distinguish Rev-positive

from Rev-negative, some part of the stimulus must be effortfully permuted in working memory by someone who is aware of what they are doing, i.e., by a Correct Stater. Permutation takes time; hence, for Rev Correct Staters, proportion correct should be positively correlated with reaction time.

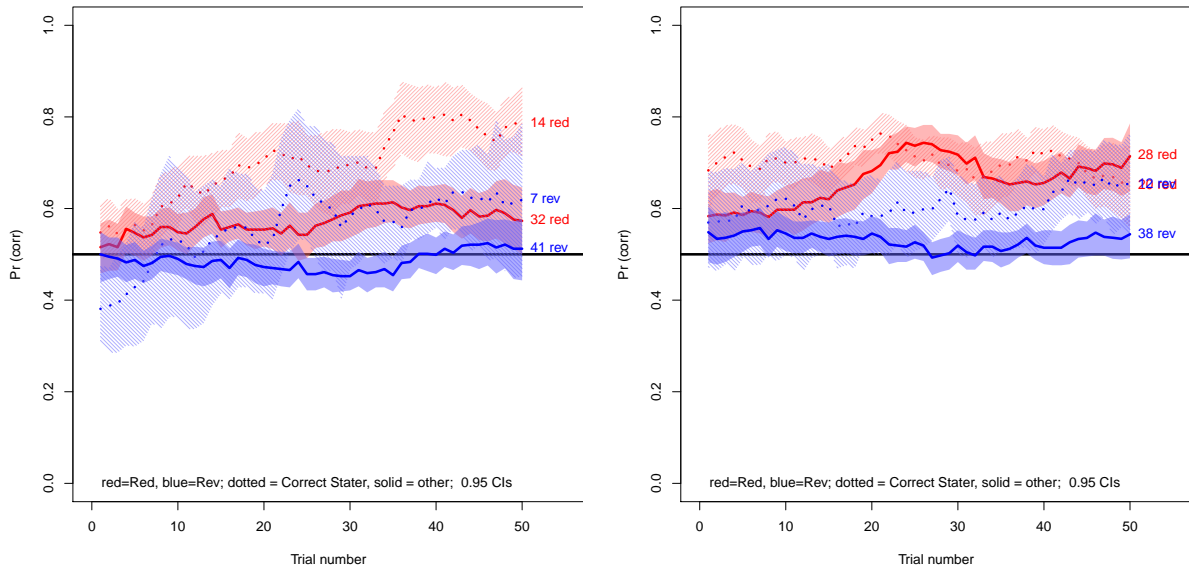


Figure 2: Proportion correct in Experiments 1 (left panel) and 2 (right panel) as a function of trial number, by Pattern (Reduplication vs. Reversal) and Correct Stater (true vs. false). The curve plots the moving average of points in a window of radius 5 trials. Confidence bands are 95% bootstrap intervals made by resampling participants. Labels at right give the number of participants in each of the four groups. The overlapping numbers in the right-hand panel are 20 Red and 12 Rev.

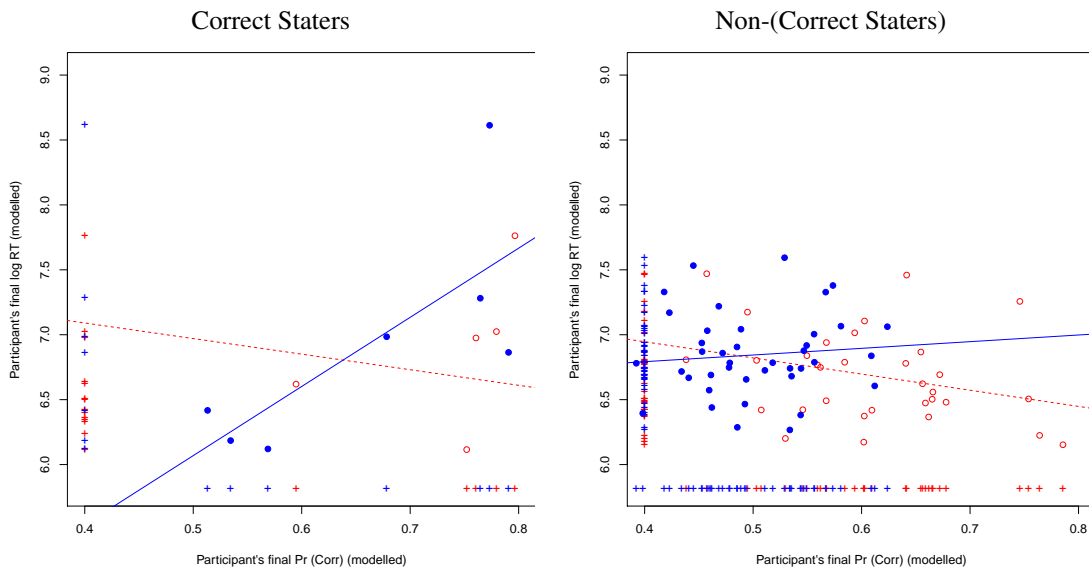


Figure 3: Relation between modelled final proportion correct and modelled final log RT, Experiment 1. Blue dots correspond to Rev participants, and red circles to Red participants. Marginal distributions are plotted with crosses.

For each participant, the response model described in the previous section was used to predict proportion correct on the last trial for a hypothetical positive and negative stimulus (i.e., regardless of whether that

participant’s actual last trial was positive or negative, the model was used to predict proportion correct both ways), and then the two were averaged together to yield a “modelled final proportion correct” for that participant. An analogous model with the same predictors was fit in the same way to predict log-transformed response time (RT), except that the (usually extra-slow) very first trial was omitted from the data, and trial number was log-transformed to improve linearity (the “power law of practice”, (Newell & Rosenbloom, 1981)). A “modelled final log RT” was computed by averaging together the predicted log RTs for a positive and a negative last stimulus. The results are scatter-plotted in Figure 3.

The strong correlation between final proportion correct and final log RT in the Rev Correct Staters, and *only* in the Rev Correct Staters, is apparent. A linear model was fit to the Correct Staters’ data using R’s `lm` function, with log RT as the response variable (Table 3). Pattern (Red = 0 vs. Rev = 1) and final modelled proportion correct were predictors. Proportion correct was centered by subtracting the overall mean across participants using R’s `scale` function.

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	6.6517	0.1600	41.563	$< 2e - 16$	***
PrCorr	-1.2001	1.7078	-0.703	0.4917	
Rev	0.8320	0.3033	2.743	0.0139	*
PrCorr:Rev	6.5263	2.3974	2.722	0.0145	*

Table 3: Linear model for participants’ modelled final log RT as a function of Pattern (Red vs. Rev) and final modelled proportion correct.

Log RT for the Red Correct Staters had no significant relation to PrCorr. Log RT for the Rev Correct Staters was significantly slower than for the Red Correct staters and was strongly correlated with PrCorr, as predicted. An analogous model, fit to the Non-(Correct Staters)’ data, found no significant effects of any sort beyond the intercept, although there was a marginally-significant *negative* PrCorr term.

2.3 Local discussion The picture that emerges from the results of Experiment 1 is that the Reversal pattern is more difficult than the Reduplication pattern and demands explicit computation. Naïve learners’ classification accuracy improved faster when trained on a reduplication pattern than when trained on one that reverses syllable order. This was true whether or not the learners explicitly stated the correct pattern at the end of the experiment when asked to do so. Final classification accuracy was significantly above chance in the Reduplication group whether they stated the correct pattern or not, but in the Reversal group, only the subgroup who correctly stated the pattern performed above chance on classification. In the Reversal-Correct Staters subgroup, and only in that subgroup, better classification performance at the end of the experiment was associated with longer reaction times, suggesting effortful, non-automatic processing of the Rev pattern.

3 Experiment 2: Recognizing repetition and reduplication

Participants in Experiment 1 had to both induce the pattern from, and apply it to, the training stimuli. It therefore does not distinguish between three possibilities: that the Rev pattern was harder to induce; that, once induced, it was harder to apply; or that it was harder to induce *because* it was harder to apply. This last possibility could arise, for instance, in an explicit hypothesis-testing learner, if difficulty in applying the reversal pattern caused difficulty in evaluating the evidence in favor of the reversal hypothesis. Experiment 2 eliminated the learning step by telling participants in advance what the pattern was, in order to test analogues of H1 and H2:

H1’: Repetition will be easier to *recognize* than reversal, regardless of whether *processing* is explicit or implicit (since both *processing* modes depend on phonological memory).

H2’: Reversal can *only* be *recognized* explicitly (by using working memory to re-order the syllables).

3.1 Procedure The procedure was identical to that of Experiment 1, except that participants in the Red condition were told that they had been assigned to the *repetition* pattern and that the first three syllables in each word that conformed to the pattern would be repeated at the word’s end. Those in the Rev condition were told that they had been assigned to the *reversal* pattern and that words that conformed to that pattern would have their first three syllables “repeating at the end [of the word] in the reverse order.” Participants in both

conditions were given color-coded examples of both a pattern-conforming and pattern-non-conforming word (e.g. **baritodatoriba** and **baritodatobari** in the Rev condition, respectively.) The questionnaire was adapted slightly from the one used in the first experiment, to take out any explicit references to “learning” or “finding a pattern”, since participants were given the pattern at the beginning of the experiment. Questionnaire responses were coded by the same coders, and using the same written scoring guide, as in Experiment 1. Cohen’s κ was at least 0.606, as shown in Table 1. 104 people completed the experiment. 2 were excluded for taking notes, and 4 for interruption (had at least one RT > 60 seconds), leaving 48 Red and 50 Rev.

3.2 Results

3.2.1 Self-reported strategy and insight into pattern structure In Experiment 1, nearly everyone reported rule-seeking; in Experiment 2, hardly anyone did (Table 4) — presumably because, having been told the rule in advance, they did not need to seek one. A third of Red and a half of Rev participants reported using intuition, but the difference between the two conditions was not significant by Fisher’s exact test ($p = 0.1058$). And a large majority of both groups reported memorizing stimuli (which is useless as a strategy, since each stimulus was only played once), again the opposite of Experiment 1. It could be that, with rule-seeking not an option, participants used the freed-up resources to memorize stimuli.

	sought rule		intuition		memorized		Rule statement	Red	Rev
	Red	Rev	Red	Rev	Red	Rev			
FALSE	48	48	32	25	8	9	correct rule	20	12
TRUE	0	2	16	25	40	41	incorrect rule	4	11
							no rule	24	27

Table 4: Self-reported learning strategies and experimenter-scored rule accuracy, Experiment 2.

Despite having been told the rule in advance, at the end of the experiment slightly more than half of the participants either stated it wrong or did not state a rule at all (Table 4). The Stating rate was nearly identical in the two conditions. However, the rate of Correct Stating among Staters was significantly lower in the Rev than the Red condition ($p = 0.03048$ by Fisher’s exact test). That could be either because people in the Rev condition forgot the rule they were supposed to be following (perhaps because it was a hard rule to follow), or because for some reason the Rev rule is verbally harder than the Red one to formulate to the satisfaction of the scorers. Thus, the overall rate of Correct Stating was lower in the Rev than the Red condition. (The same was true in Experiment 1. There, the Stating rate was lower in the Rev condition, but the Correct Stating rate among Staters did not differ between Red and Rev.)

3.2.2 Classification accuracy Despite being told the rule in advance and receiving right/wrong feedback after every response, no participant group reached 100% correct performance (right panel of Figure 2). Red Correct Staters were about 75% correct throughout, while Rev Correct Staters started at about 55–60% correct and ultimately converged with the Red Correct Staters. This convergence is consistent with the hypothesis that the Red pattern is inherently easier to recognize, that recognizing the Rev pattern requires effort and so improves with practice, and that processing difficulty contributed to the differences between Red and Rev Correct Staters in Experiment 1.

What the Non-Correct-Staters may be doing is unclear — they seem to have forgotten the instructions — but those in the Red condition start off in the 55–60% correct range and converge with the Red and Rev Correct Staters by the end, whereas those in the Rev condition, like the Rev Non-Correct-Staters in Experiment 1, never surpass the 55% mark. This, again, makes it appear that the Red pattern can be learned implicitly, while the Rev pattern cannot.

The same mixed-effects logistic-regression model was fit as in the analysis of Experiment 1, based on 4900 observations from 98 valid participants. **H1'** predicted that within the Correct Staters, and again within the Not-Correct-Staters, Red would outperform Rev by the end of the experiment. This prediction was only half confirmed. Among Correct Staters, the pattern type made no significant difference; among Non-Correct-Staters, however, those in the Red condition significantly outperformed those in the Rev condition ($\beta_{\text{Red Correct Staters}} - \beta_{\text{Rev Correct Staters}} = 0.56$, $p = 0.62$; $\beta_{\text{Red Not-Correct-Staters}} - \beta_{\text{Rev Not-Correct-Staters}} = 1.83$, $p < 0.001$). **H2'** predicted that within the Rev group, the Not-Correct-Staters would be at chance. This (null)

hypothesis was not rejected: $\beta_{\text{Rev Not-Correct-Staters}} = 0.23$, 95% z CI = $[-0.18, 0.64]$. The Rev Correct Staters' estimate was $\beta_{\text{Rev Not-Correct-Staters}} = 1.21$, 95% z CI = $[0.96, 1.95]$. I.e., the Rev pattern was learned at best feebly by the Not-Correct-Staters, but it was learned much better by the Correct Staters.

3.2.3 Reaction times Experiment 1 found that at the end of the experiment, proportion correct was positively correlated with reaction time for Rev Correct Staters, but for no other subgroup. When the data from Experiment 2 was plotted and analyzed the same way, the correlation persisted (Figure 4).

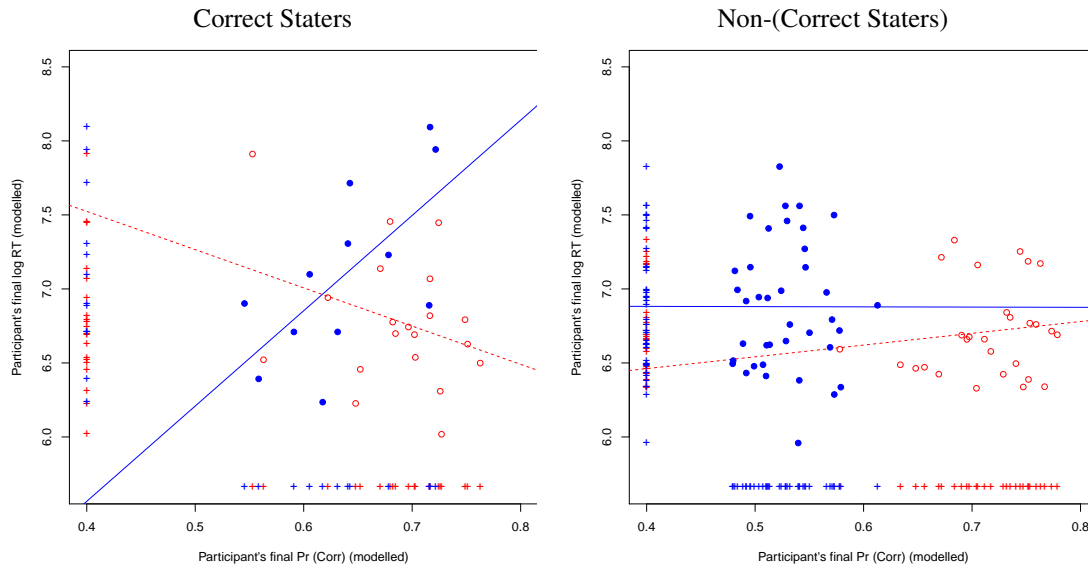


Figure 4: Relation between modelled final proportion correct and modelled final log RT, Experiment 2. Blue dots and a solid line correspond to Rev participants, red circles and a dotted line to Red participants.

The same effect was found as in Experiment 1: In the Rev Correct Staters group, and only in that group, at the end of the experiment, participants who were more accurate also had longer response times. A linear model like that in Section 2.2.3 replicated those results (Table 5).

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	6.8301	0.1051	64.966	$2e-16$	***
PrCorr	-2.5823	1.8089	-1.428	0.16448	
Rev	0.4638	0.1796	2.583	0.01531	*
PrCorr:Rev	9.0196	2.8884	3.123	0.00414	**

Table 5: Linear model for participants' modelled final log RT as a function of Pattern (Red vs. Rev) and final modelled proportion correct.

3.3 Local discussion The main difference between the outcomes of the two experiments was that among Correct Staters, final Red performance exceeded final Rev performance in Experiment 1, but not in Experiment 2. All of the other effects found with naïve learners in Experiment 1 were replicated with informed participants in Experiment 2, suggesting that the Rev pattern is harder to induce, whether implicitly or explicitly, in large part *because* it is harder to detect.

4 General discussion

Differences between within-word reduplication and reversal patterns in learning and detection performance, and in their interaction with implicit versus explicit cognitive processes, can be derived from general properties of auditory short-term memory established by studies of melodies and lists. This provides another instance in which results from “artificial-language” experiments parallel those of analogous concept-

formation experiments in other domains, including effects of abstract pattern structure (reviewed in Moreton & Pater 2012; see also Moreton et al. 2017; Gerken et al. 2019), inter- vs. intra-dimensional dependencies (Moreton, 2012), and implicit vs. explicit learning (Kemper 2016; Moreton & Pertsova 2016; submitted). These parallels suggest that the cognitive processes underlying pattern learning are homogeneous across domains, but heterogeneous within domains: Both implicit and explicit processes are available, within phonology and outside it. (They also highlight the methodological benefit of incorporating participants' self-reported approach into the analysis.)

4.1 The cognition-phonology interface Phonological theory in the rule-based and Optimality-Theory traditions provides formal means for implementing repetition while banning reversal and other rearrangements (Marantz, 1982; Downing & Inkelas, 2015). There is little formal motivation for this asymmetry beyond simply fitting natural-language typology *post hoc*.¹

The results of the present experiments provide a motivation: Suppose that the cognitive substrate for representations is phonological memory — long-term for the lexicon, short-term for surface representations, variables, and other ephemera — and that automatic storage and matching in phonological memory preserves temporal order, like a queue rather than a stack or other data structure (Section 1.1). Suppose further that natural-language constraints are restricted to using cognitive operations which can be performed automatically and unconsciously. Since constraints enjoining reversal and other reorderings do not meet this criterion, as shown by Experiments 1 and 2, those constraints, and those patterns, are found only in language games.

The memory-based proposal predicts a sharp difference between repetition on the one hand, and any kind of rearrangement on the other — not just whole-string reversal, but also reversal of a fixed-length substring, or jumbling syllables in a way that is not reversal. These differences are predicted to emerge in the lab, in natural-language typology, and in the typology of consciously-created cultural products like language games. The proposal additionally predicts analogous differences in analogous tasks in music and in other domains which rely on time-ordered memory.

4.2 Repetition and reversal in Formal Language Theory In Formal Language Theory (FLT) asymmetries that have to do with grammaticality or frequency of linguistic patterns can be explained in terms of the complexity of formal grammars that are used to generate them. The Chomsky hierarchy is an early hypothesis about a particular metric of grammar complexity (Chomsky, 1956). It predicts that unbounded crossing long-distance dependencies (leading to repetition patterns $A_1A_2A_3B_1B_2B_3$) are more complex than nested long-distance dependencies (leading to reversal-like patterns $A_1A_2A_3B_3B_2B_1$) because the latter can be generated by context-free grammars, while the former require grammars of higher complexity (context-sensitive) in the Chomsky hierarchy. This prediction contrasts with the natural language data, particularly in phonology — as we've discussed, phonological reversals (as well as center-embedding for that matter) are unattested, while reduplication is common. This paradox holds as long as reduplication really involves unbounded crossing-dependencies at the level of phonological segments. An alternative view is that unbounded reduplication is a morphological process happening at the word- or morpheme-level as proposed in the theory of morphological doubling (Inkelas & Zoll, 2005; Inkelas, 2008). In this case, reduplication does not present a problem for the Chomsky hierarchy since doubling at the morphological level does not involve crossing dependencies, while phonological cases of reduplication are always bounded and can be handled by regular grammars. However, any kind of evidence of whole-word reduplication that requires segmental correspondence (e.g., reduplicate-specific phonology) present a challenge to the theory of morphological doubling.

¹ One possible theoretical motivation is the following: A constraint requiring a “reverpicant” to be in reverse order from the base could not be a markedness constraint, because it must compare the base with the reverpicant, and could not be a faithfulness constraint, because a fully-faithful copy wouldn't satisfy it. If natural-language constraints must be one or the other (“classical OT”, Alderete 2001; Moreton 2004; De Lacy 2011), and natural-language learning is implicit, that would explain why reversal was only found explicitly in Experiments 1 and 2, and is found in language games but not in natural language.

However, entire families of non-markedness, non-faithfulness constraints have been proposed, such as Alignment (McCarthy & Prince, 1995), anti-faithfulness (Alderete, 2001), and *MAP (Zuraw, 2013). If they are admitted, there seems to be no theoretical obstacle to a constraint enjoining reversal.

If we accept that all reduplication, including whole-word reduplication, happens at the segmental level, then what exactly are the bounds on this process in terms of formal complexity? This question has been more extensively studied in syntax where instances of repetitions or copying are also ubiquitous. There are still debates about what subclass of context-sensitive grammars are appropriate for capturing repetitions in language (Stabler, 2004; Kobele, 2006; Kanazawa & Salvati, 2010; Clark & Yoshinaka, 2014). This debate depends on getting answers to such empirical questions as: how much structure (beyond the string-identity) can the copies have, do copies have to be exact, can copying be recursive, etc. The same questions have to be answered for reduplication. Recently, a series of papers by Dolatian and Heinz have investigated the typological space of reduplicative processes as well as the formal models of reduplication in the framework of FLT. Dolatian & Heinz (2020) develop their own complexity hierarchy of subclasses of 2-way Finite State Transducers (2-way FSTs) which they use to model the copying function. Finite State Transducers describe functions or relations between two sets (while grammars in the Chomsky hierarchy describe sets). Thus, claims about transducers are not directly comparable to claims about subclasses of context-sensitive grammars, but in any case the goal of this research program is to find the most restricted subclass of formal systems that can correctly account for all attested cases and rule out the unattested ones (with the caveat that some of the over- and under-generation that is due to extra-linguistic factors will still remain). Dolatian and Heinz's hierarchy identifies several subclasses of functions for different types of reduplication. In particular, they show that almost all cases of reduplication can be modeled with deterministic rotating 2-way FSTs (with most cases falling within more restricted subclasses). Interestingly, the rotating 2-way FSTs rule out mirror-images because of a restriction that the transducer cannot output anything while it is moving from right to left (in the opposite direction of the input tape). This is one example of how an asymmetry between two patterns can be accounted for by restricting the hypothesis space available to the learner in a very specific way. In other words, if human learners can only learn a subclass of functions generated by rotating 2-way FSTs we would predict that they should be able to learn repetitions, but not reversals.

An alternative explanation of the typological asymmetry between reversals and repetitions lies outside the realm of grammatical competence. One precedent for this approach can be found in Pulman (1986) who proposes a parser designed to explain why nested dependencies in syntax (like those in reversal) are hard to process.² Pullman's parser is an attempt to implement an idea that "restrictions on available parsing memory, and on our ability to operate properly when parsing recursive constructions" is what is responsible for the difficulties with center-embedding structures of degree more than two levels of embedding. However, he does not discuss whether his parser can handle crossed rather than nested dependencies [and I suspect that it cannot]. Our proposal can be viewed as a first step towards developing a similar processing explanation. This explanation rests on the assumption that auditory memory activation and retrieval proceed "forward" implementing something like identity matching of a memory trace to another representation, rather than "backward" by popping off stored representations off a stack. It is the same general feature of auditory memory that we think is responsible for the prevalence of repetitions over reversals not just in language, but in music. In non-auditory domains, however, it is possible that other mechanisms come into play, given what we have said earlier about findings in the literature of visual symmetry.

There is yet another reason for why at least some types of reversals are unattested. This reason has to do with the fact that reversing string order typically leads to violations of the grammatical constraints imposed on the linear order of linguistic units such as phonotactic constraints in phonology, word-order constraints in syntax, and possibly melodic constraints in music (Kempf, 1996). Nevertheless certain types of reversals (e.g., of syllables) should be possible.

References

- Alderete, John D. (2001). Dominance effects as transderivational anti-faithfulness. *Phonology* 18:2, 201–253.
- Bach, E., C. Brown & W. Marslen-Wilson (1986). Crossed and nested dependencies in German and Dutch: a psycholinguistic study. *Language and Cognitive Processes* 1:4, 249–262.
- Bagemihl, Bruce (1989). The Crossing Constraint and "backwards languages". *Natural Language and Linguistic Theory* 7:4, 481–549.
- Balch, W. R. (1981). The role of symmetry in the good continuation ratings of two-part tonal melodies. *Perception and Psychophysics* 29, 47–55.

² The empirical data suggesting processing difficulties of center-embedding come from studies by Bach et al. (1986).

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1, 1–48.
- Baylis, Gordon C. & Jon Driver (1994). Parallel computation of symmetry but not repetition within single visual shapes. *Visual Cognition* 1:4, 377–400.
- Benassi-Werke, Mariana E, Marcelo Queiroz, Rúben S Araújo, Orlando FA Bueno & Maria Gabriela M Oliveira (2012). Musicians' working memory for tones, words, and pseudowords. *Quarterly Journal of Experimental Psychology* 65:6, 1161–1171.
- Chomsky, Noam (1956). Three models for the description of language. *IRE Transactions on information theory* 2:3, 113–124.
- Clark, Alexander & Ryo Yoshinaka (2014). Distributional learning of parallel multiple context-free grammars. *Machine Learning* 96:1-2, 5–31.
- Cowan, Nelson, Martin DS Braine & Lewis A Leavitt (1985). The phonological and metaphonological representation of speech: Evidence from fluent backward talkers. *Journal of Memory and Language* 24:6, 679–698.
- Dalla Bella, Simone, Isabelle Peretz & Neil Aronoff (2003). Time course of melody recognition: a gating paradigm study. *Perception and Psychophysics* 65:7, 1019–1028.
- De Lacy, Paul (2011). Markedness and faithfulness constraints. van Oostendorp, Marc, Colin Ewen & Elizabeth Hume (eds.), *The Blackwell companion to phonology*, Wiley-Blackwell, 1–22.
- Dehaene-Lambertz, Ghislaine (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8, 919–924.
- Dienes, Zoltán & Christopher Longuet-Higgins (2004). Can musical transformations be implicitly learned? *Cognitive Science* 28, 531–558.
- Dolatian, Hossep & Jeffrey Heinz (2020). Computing and classifying reduplication with 2-way finite-state transducers. *Journal of Language Modelling* 8:1, 179–250.
- Donolato, Enrica, David Giofrè & Irene C Mammarella (2017). Differences in verbal and visuospatial forward and backward order recall: A review of the literature. *Frontiers in Psychology* 8, p. Article 663.
- Dowling, W. J. (1972). Recognition of melodic transformations: inversions, retrograde, and retrograde inversions. *Perception and Psychophysics* 12:5, 417–421.
- Downing, Laura J. & Sharon Inkelas (2015). What is reduplications? typology and analysis part 2/2: the analysis of reduplication. *Language and Linguistics Compass* 9:12, 516–528.
- Ellis, Nick C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition* 27:2, 305–352.
- Gerken, LouAnn, Carolyn Quam & Lisa Goffman (2019). Adults fail to learn a type of linguistic pattern that is readily learned by infants. *Language Learning and Development* .
- Gil, David (1996). How to speak backwards in Tagalog. *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Language and Linguistics, January 8-10, 1996, :Institute of Language and Culture for Rural Development, Mahidol University at Salaya*.
- Gottwald, Richard L. (1971). Effects of response labels in concept attainment. *Journal of Experimental Psychology* 91:1, 30–33.
- Hothorn, Torsten, Frank Bretz & Peter Westfall (2008). Simultaneous inference in general parametric models. *Biometrical Journal* 50:3, 346–363.
- Hurlstone, Mark J, Graham J Hitch & Alan D Baddeley (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological bulletin* 140:2, p. 339.
- Inkelas, Sharon (2008). The dual theory of reduplication. *Linguistics* 46:2, 351–401.
- Inkelas, Sharon & Laura J. Downing (2015). What is reduplications? typology and analysis part 1/2: the typology of reduplication. *Language and Linguistics Compass* 9:12, 502–515.
- Inkelas, Sharon & Cheryl Zoll (2005). *Reduplication: Doubling in morphology*, vol. 106. Cambridge University Press.
- Kanazawa, Makoto & Sylvain Salvati (2010). The copying power of well-nested multiple context-free grammars. *International Conference on Language and Automata Theory and Applications*, Springer, 344–355.
- Kempf, Davorin (1996). What is symmetry in music? *International Review of Aesthetics and Sociology of Music* 27, 155–165.
- Kimper, Wendell (2016). Asymmetric generalisation of harmony triggers. Hansson, Gunnar, Ashley Farris-Trimble, Kevin McMullin & Douglas Pulleyblank (eds.), *Proceedings of the 2015 Annual Meeting on Phonology*.
- Kobe, Gregory Michael (2006). *Generating Copies: An investigation into structural identity in language and grammar*. Ph.D. thesis, University of California, Los Angeles.
- Kootstra, Gert, Bart de Boer & Lambert R. B. Schomaker (2011). Predicting eye fixations on complex visual stimuli using local symmetry. *Cognitive Computation* 3, 223–240.
- Krumhansl, Carol L., Gregory J. Sandell & Desmond C. Sergeant (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception* 5, 67–113.
- Li, Feifei, Shan Jiang, Xiuyan Guo, Zhiliang Yang & Zoltan Dienes (2013). The nature of the memory buffer in implicit learning: learning Chinese tonal asymmetries. *Consciousness and Cognition* 22, 920–930.

- Marantz, Alec (1982). Re reduplication. *Linguistic Inquiry* 13:3, 435–482.
- Marslen-Wilson, William D. (1987). Functional parallelism in spoken word recognition. *Cognition* 25, 71–102.
- McCarthy, John J. & Alan M. Prince (1995). Faithfulness and reduplicative identity. Beckman, Jill N., Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, no. 18 in University of Massachusetts Occasional Papers, Graduate Linguistic Students' Association, Amherst, Mass., 249–384.
- McHugh, Mary L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica* 22:3, 276–282.
- Metcalfe, Janet & Donald Sharpe (1985). Ordering and reordering in the auditory and visual modalities. *Memory and Cognition* 13:5, 435–441.
- Mongoven, C. & C. C. Carbon (2017). Acoustic Gestalt: on the perceptibility of melodic symmetry. *Musicae Scientiae* 21:1, 41–49.
- Moreton, Elliott (2004). Non-computable functions in Optimality Theory. McCarthy, John J. (ed.), *Optimality Theory in phonology: a reader*, Blackwell, Malden, 141–164.
- Moreton, Elliott (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language* 67:1, 165–183.
- Moreton, Elliott & Joe Pater (2012). Structure and substance in artificial-phonology learning: Part I, structure. *Language and Linguistics Compass* 6:11, 686–701.
- Moreton, Elliott & Katya Pertsova (2016). Implicit and explicit processes in phonotactic learning. Scott, Jennifer & Deb Waughtal (eds.), *Proceedings of the 40th Boston University Conference on Language Development*, Cascadilla, Somerville, Mass., 277–290.
- Moreton, Elliott & Katya Pertsova (2020). Implicit and explicit processes in phonological pattern learning. MS, Department of Linguistics, University of North Carolina, Chapel Hill.
- Moreton, Elliott, Joe Pater & Katya Pertsova (2017). Phonological concept learning. *Cognitive Science* 41:1, 4–69.
- Munoz, Sergio R. & Shrikant I. Bangdiwala (1997). Interpretation of kappa and B statistics measures of agreement. *Journal of Applied Statistics* 24:1, 105–112.
- Näätänen, Risto, Anne Lehtokoski, Mietta Lennes, Marie Cheour, Minna Huottilainen, Antti Ilvonen, Martti Vainio, Paavo Alku, Risto J. Ilmoniemi, Aavo Luuk, Jüri Allik, Janne Sinkkonen & Kimmo Alho (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.
- Näätänen, R., P. Paavilainen, T. Rinne & K. Alho (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology* 118:12, 2544–2590.
- Newell, Allen & Paul S. Rosenbloom (1981). Mechanisms of skill acquisition and the law of practice. Anderson, J. R. (ed.), *Cognitive skills and their acquisition*, Erlbaum, Hillsdale, New Jersey, 1–55.
- Norris, Dennis & James M McQueen (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological review* 115:2, p. 357.
- Peters, K. G. & J. P. Denny (1971). Labelling and memory effects on categorizing and hypothesizing behavior for biconditional and conditional conceptual rules. *Journal of Experimental Psychology* 87:2, 229–233.
- Petrović, Milena, Gordana Ačić & Vera Milanković (2017). Sound of picture vs. picture of sound: musical palindrome. *New Sound* 50, 217–228.
- Pulman, Steven G (1986). Grammars, parsers, and memory limitations. *Language and Cognitive processes* 1:3, 197–225.
- Riggle, Jason (2004). Nonlocal reduplication. Moulton, K. & M. Wolf (eds.), *Proceedings of the 34th Annual Meeting of the North-East Linguistics Society (NELS)*, Graduate Linguistics Students' Association, Amherst, 485–496.
- Schulze, Katrin & Barbara Tillmann (2013). Working memory for pitch, timbre, and words. *Memory* 21:3, 377–395.
- Schulze, Katrin, W. Jay Dowling & Barbara Tillmann (2012). Working memory for tonal and atonal sequences during a forward and a backward recognition task. *Music Perception* 29:3, 255–267.
- Shen, Jonathan, Ruoming Pang, Ron J Weiss, Mike Schuster, Navdeep Jaitly, Zongheng Yang, Zhifeng Chen, Yu Zhang, Yuxuan Wang, Rj Skerrv-Ryan et al. (2018). Natural tts synthesis by conditioning wavenet on mel spectrogram predictions. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 4779–4783.
- Stabler, Edward P (2004). Varieties of crossing dependencies: structure dependence and mild context sensitivity. *Cognitive Science* 28:5, 699–720.
- Sun, Ron, Xi Zhang, Paul Slusarz & Robert Mathews (2007). The interaction of implicit learning, explicit hypothesis testing learning and implicit-to-explicit knowledge extraction. *Neural Networks* 20:1, 34–47.
- Treder, Matthias Sebastian (2010). Behind the looking-glass: a review on human symmetry perception. *Symmetry* 2, 1510–1543.
- van der Helm, Peter A. (2015). Symmetry perception. Wagemans, Johan (ed.), *The Oxford handbook of perceptual organization*, Oxford University Press, Oxford, England, chap. 6, 108–128.
- Wagemans, Johan (1995). Detection of visual symmetries. *Spatial Vision* 9:1, 9–32.
- Zatorre, Robert J., Andrea R. Halpern & Marc Bouffard (2010). Mental reversal of imagined melodies: a role for the posterior parietal cortex. *Journal of Cognitive Neuroscience* 22:4, 775–789.
- Zuraw, Kie (2013). *MAP constraints. MS, Department of Linguistics, UCLA, URL http://www.linguistics.ucla.edu/people/zuraw/dnldpprs/star_map.pdf.