Schwitzgebel, D., & Peperkamp, S. (2026). Linguistic typology and learning biases: A study of syllable-copy reduplication across modalities. *Glossa Psycholinguistics*, 5(1): 17, pp. 1–40. DOI: https://doi.org/10.5070/G6011.53912

Linguistic typology and learning biases: A study of syllable-copy reduplication across modalities

David Schwitzgebel, Laboratoire de Sciences Cognitives et Psycholinguistique/Institut Jean Nicod, École normale supérieure – Paris Sciences et Lettres (ENS-PSL), david.schwitzgebel@ens.psl.eu

Sharon Peperkamp, Laboratoire de Sciences Cognitives et Psycholinguistique, École normale supérieure – Paris Sciences et Lettres (ENS-PSL, EHESS, CNRS), sharon.peperkamp@ens.psl.eu

Do the typological patterns of reduplication reflect underlying learning biases in *perception* and *production*? In morphological reduplication, reduplicants are rarely defined in terms of the syllable structure of the base, instead doubling a segmental melody. Using an artificial grammar learning paradigm with written stimuli, Haugen et al. (2022) reported a learning bias in line with the typological facts: a bias against syllable-copy reduplication. Here, we extend the methods of Haugen et al. (2022) to six new experiments, investigating whether their original findings replicate with novel sets of carefully controlled written stimuli, recorded auditory stimuli, and both *perception* and *production* tasks. Across all experiments, we included 545 native speakers of Italian, a language that has no morphological reduplication. In each experiment, participants were randomly assigned to a typologically rare syllable-copy condition (CVX) or a typologically common string-defined condition (CV or CVC). Although we replicated the bias against the syllable-copy condition for our first set of written stimuli, a subsequent experiment found that this does not generalize to stimuli with differing written lengths. In experiments with recorded audio stimuli, we observed a modality-specific bias: there were no significant differences between the syllable-copy and string-copy conditions when participants performed a *perception* task, but participants performed worse in the syllable-copy condition than in the string-copy conditions, given a *production* task – despite identical exposure phases. Subsequent experiments reproduced this pattern of results and ruled out an alternative explanation. Broadly, our results corroborate the original claim that the typological rarity of syllable-copy reduplication is driven by an underlying learning bias; however, this bias is selectively driven by *production*, rather than *perception*.

1. Introduction

Why are some grammatical phenomena more common than others? Cross-linguistic typological patterns often share properties that cannot merely be attributed to historical accident – rather, they are driven by cognitive biases. These biases can reflect non-arbitrary factors, such as formal simplicity (Culbertson & Kirby, 2016; Lambert et al., 2021) and naturalness (phonetic: Blevins, 2008; semantic: Jackendoff, 2000; conceptual: Strickland, 2017), innate grammatical constraints ("Universal Grammar"/UG: Chomsky, 1965), or some combination of cognition-general and language-specific influences. However, before disambiguating among specific mechanisms, the first step in linking typology to cognition is verifying that the link exists (Culbertson, 2012; Moreton, 2008). In order to do so, a growing body of research uses artificial grammar learning (AGL) paradigms to investigate whether typological patterns are manifested as parallel learning biases. By testing participants on experimentally-controlled novel grammars, it is possible to identify systematic differences in the learnability of linguistic patterns. In turn, observed differences in learnability can be leveraged as explanatory factors for linguistic typology: ceteris paribus, more easily learnable patterns are likely to be transmitted and acquired with less error than less easily learnable patterns. Much of this research has found evidence for learning biases that are consistent with linguistic typology, independently of cognition-external factors (for reviews, see Culbertson, 2012, 2023; Moreton & Pater, 2012a, 2012b).

One topic in linguistics that has received relatively little experimental attention is reduplication – a phenomenon in which the meaning of a word is altered through the repetition of all or part of the word. Despite the immense quantity of theoretical literature in morphology and phonology discussing reduplication and its typology (Marantz, 1982; McCarthy & Prince, 1995; Moravcsik, 1978; Raimy, 2000; Zukoff, 2017; inter alia), to our knowledge, only two attempts have been made to apply experimental methods to investigate the cognition-typology link in reduplication: Haugen, Ussishkin, and Dawson (2022) and Wang (2024). The lack of more experimental research on this topic is particularly noteworthy, given that reduplication exhibits an extensively-discussed and highly persistent typological pattern: there is only one known natural language in which the syllable structure of a reduplicated affix is directly dependent on the syllable structure of its base word (Haugen, 2003, 2014; Haugen & Kennard, 2011). Although the two aforementioned experimental studies provide some insight into the cognitive mechanisms driving the rarity of the latter type of reduplication, they also leave open many methodological and theoretical questions. The present article aims to replicate and extend existing findings, with the goal of rigorously investigating the relationship between the observed typological pattern, phonological perception and production, and the presence of underlying learning biases (or the lack thereof).

1.1 Background

Partial reduplication often depends on a segmental melody: a sequence of segment variables present in the reduplicative form are filled by a corresponding segmental sequence in the base form (Marantz, 1982), regardless of syllable structure. That is, the segmental melody doubled in the reduplicative form may cross syllabic boundaries in the base form. For example, Agta has a reduplication process by which a CVC string is copied, regardless of syllable structure (Healey, 1960):

(1)	a.	karúba – karkarúba	'neighbor – neighbors'
	b.	balatáng – balbalatáng	'girl – little girl'

In contrast, Tohono 0'odham has a reduplication process for pluralization by which a CV string is copied, again regardless of syllable structure (Raimy, 2000):

(2)	a.	gimai – gi-gimai	'a braggart'
	b.	hon – ho-hon	'the body'

Although many formal theories of reduplication make reference to syllable structure, they generally do not permit direct copying of the syllable from the base to the reduplicant (e.g., McCarthy & Prince, 1995). This aligns with the observed typological pattern: "syllable-copy reduplication" is largely unattested in natural languages. However, studies by Haugen (2003, 2014) and Haugen and Kennard (2011) provide documentation and analysis of the only known case of syllable-copy reduplication: Hiaki. In this language, the habitual verb form involves the copying of the complete initial syllable from the base. Thus, if the base-initial syllable is CV, the reduplicant will be CV; if it is CVC, the reduplicant will be CVC:

(3)	a.	hewite – hehewite	'agree'
	b.	vamse – vamvamse	'hurry'

Therefore, the theoretical constraint described above – that the syllable structure of the base is not directly copied – is violated by Hiaki. A critique of existing theories is beyond the scope of this article; the crucial point is that syllable-copy reduplication, while attested, is exceedingly rare.

Intuitively, it is remarkable that there is a cross-linguistic tendency for partial reduplicants to discount – and often override – the syllable structure of the base form. After all, the syllable is a well-defined constituent in phonology, and many phonological rules and constraints make reference to it (for an overview, see Blevins, 1995). Furthermore, there is evidence from language games, or ludlings, that speakers spontaneously manipulate syllables. For example, In the Colombian version of the Spanish language game Jerigonza, /p/ and a copy of the preceding vowel are added after every syllable; for instance, /ma.es.tro/ 'teacher' becomes /ma.pa.es.pe.

tro.po/ (Piñeros, 1998). Thus, under the reasonable assumption that the syllable is a cognitively accessible unit, syllable-copy reduplication is not more complex than string-copy reduplication. Similarly, syllable-copy reduplication does not appear to be phonetically unnatural; on the contrary, repeating a syllable is economical in terms of both planning and articulatory effort. This raises the following question: despite the apparent mismatch between the rarity of syllable-copy reduplication and the naturalness of syllable repetition in general, might the typology still be driven by an underlying learning bias?

1.2 Previous work

To our knowledge, only one study has applied AGL to directly investigate cognitive mechanisms that may drive the rarity of *syllable-copy* reduplication: Haugen et al. (2022). This study investigated the learnability of *syllable-copy* reduplication using a training + feedback paradigm. In this paradigm, participants perform a forced-choice task and get feedback (correct/incorrect) on every trial. Thus, they start by making guesses; the question is to what extent and how fast they learn the pattern, based on the feedback. In their study, monolingual native English speakers were presented with written trisyllabic nonce words, the first syllable of which was either of the form CV or CVC. For each item, participants had to choose among three reduplicative options, one of which was full reduplication of the base word, one a partial reduplication of the first two segments, and one a partial reduplication of the first three segments. In half of the trials, participants had to identify the plural form; in the other half, the augmentative form. For instance, for the item 'dog' they were asked either 'What is the word for dogs?' or 'What is the word for big dog?'. The correct plural form was always the one with full reduplication; this condition served as a baseline. The correct augmentative form depended on the participant's group. In two *string-copy* groups – CV and CVC – it was determined by copying a segmental melody of the base form, regardless of syllable boundaries; in the *syllable-copy* group – CVX – it was determined by directly copying the first syllable of the base, similar to Hiaki.[1] Thus, in the CV and CVC groups, the reduplicant always consisted of either a light or a heavy syllable, respectively, whereas in the CVX group, some reduplicants consisted of a light syllable and others, of a heavy syllable, depending on the syllable structure of the base. Examples are shown in **Table 1**. Note that, in case of *syllable-copy*ing, the augmentative form copies CV if the first syllable of the base is CV (*nekola*) and CVC if the first syllable of the base is CVC (*vamseta*).

[1] Haugen et al. (2022) use a different terminology: *base-dependent reduplication* and *consistent [CV/CVC] reduplication*, respectively. We have renamed the conditions in the present article for clearer framing.

Table 1: Sample stimuli from Haugen et al. (2022).

base form	plural	augmentative		
	all groups	CVC string-copy	CV string-copy	syllable-copy
nekola	nekola-nekola	nek-nekola	ne-nekola	ne-nekola
vamseta	vamseta-vamseta	vam-vamseta	va-vamseta	vam-vamseta

The results showed no difference between the full reduplication and the two string-copy patterns, which were all learned faster and better than the syllable-copy pattern. Given the crucial difference between string-copy and syllable-copy reduplication, the authors concluded that an underlying learning bias disfavors syllable-copy reduplication. In turn, they argued that this cognitive constraint against syllable-copy reduplication may drive its observed typological rarity.

A second set of AGL experiments concerning reduplication was conducted for the dissertation of Wang (2024); while these experiments were not designed to directly investigate the rarity of syllable-copy reduplication, they have bearing on the topic. The experiments focused on how native English speakers generalize an artificial reduplicative grammar to novel base structures. Stimuli were presented auditorily, and participants performed a production task. This study did not directly compare the learnability of syllable-copy reduplication to string-copy reduplication, but some of the experiments incidentally converged on results that are largely consistent with those of Haugen et al. (2022). Specifically, in Experiments 1a-c, participants were exposed to CVC reduplication (e.g., [dɔvgə] → [dɔv-dɔvgə]). During the test phase, a strong majority of participants generalized CVC reduplication to novel stimuli, including both those with an initial CVC-syllable and those with an initial CV-syllable (e.g., [dəbeɪ] → [dəb-dəbeɪ]). Only a small minority produced syllable-copy reduplication (i.e., reduplicate CVC when the initial syllable of the base form is CVC, reduplicate CV when the initial syllable of the base form is CV). Vice versa, in Experiment 2c, participants were exposed to CV reduplication (e.g., [pif] → [pipif]), and during the test phase, the majority of them produced CV-copy reduplication for novel test items, including those with an initial CVC-syllable ([dəbgɪv] → [də-dəbgɪv]); only a small minority consistently copied the first syllable of the base forms. Thus, participants across these experiments seemed to generalize string-copy reduplicative forms to the test phase, in line with both the observed typology and the results of Haugen et al. (2022): responses were suggestive of a cognitive constraint against syllable-copy reduplication.

1.3 The present study

Although the previous findings indeed suggest that the typological rarity of syllable-copy reduplication may be driven by an underlying learning bias, Wang's (2024) experiments do not directly address this issue, and Haugen et al.'s (2022) findings are limited by a number of methodological and theoretical considerations. For example, Haugen et al. (2022) relied exclusively on written stimuli, whereas the phenomenon under consideration is a morphophonological one, not an orthographic one. Biases in the written modality do not necessarily reflect the auditory modality, particularly with regard to phonology: languages are generally acquired and transmitted via auditory (or, in the case of sign languages, visual) input. Therefore, findings observed for orthographic stimuli might not generalize to linguistic cognition more broadly – and, if so, may be insufficient to account for the typological pattern. Haugen et al. (2022) acknowledge this shortcoming. Here, we include experiments that test for the presence of a learning bias in the auditory modality.

A second important issue concerns potential differences between biases in speech perception and biases in speech production. The cognitive pressures that operate on speech perception are distinct from those that operate on speech production (e.g., minimizing articulatory effort is relevant for speakers, but not for listeners), and, indeed, some prior AGL studies in the realm of phonology have found differing patterns of results in perception versus in production (Do & Havenhill, 2021; Peperkamp et al., 2006). This raises the possibility that typological trends are selectively driven by one or the other. Haugen et al. (2022) relied on perception, while Wang (2024) relied on production; neither study compared the two. Here, we directly compare performance on a perception and a production task, using the exact same exposure.

In addition to these substantial extensions, we implement several smaller methodological improvements. Since Haugen et al. (2022) tested English speakers, their stimuli were written in English orthography, which is notoriously opaque. Consequently, the syllable boundaries in many of their stimuli were ambiguous – e.g., the CVC reduplicative form of *hitmolu* was *hithitmolu*, which could be interpreted as *hi.thit.mo.lu* or *hit.hit.mo.lu*. To address this, we test speakers of a language with a transparent orthography, i.e., Italian, on items that are phonotactically legal in their language. In addition, whereas Haugen et al.'s (2022) participants were undergraduate linguistics students, we draw from a more diverse sample of participants recruited online via Prolific. Furthermore, they provided participants with explicit feedback after each trial to prompt learning, which has the advantage of enabling the researchers to model learning curves. However, exposure to large quantities of explicit, negative feedback is rare in naturalistic learning settings (Chomsky, 1965, 1980); therefore, in our design, we implement separate exposure and test phases, both without feedback.

A final consideration concerns response strategies. In Haugen et al. (2022), word length was perfectly predictable: base forms were either of the form CVC.CV.CV or CV.CV.CV. Therefore, base forms with a CV initial syllable were always six letters; base forms with a CVC initial syllable were always seven letters. Correspondingly, the CVC-reduplicated form was always either ten or nine letters, and the CV-reduplicated form was always nine or eight letters, contingent on the initial syllable of the base form. Therefore, it may have been relatively easy for participants to attend to spurious cues as an effective response strategy. For instance, in the CV condition, the correct augmentative form of a six-letter base word was always the eight-letter option. We examine whether the findings of Haugen et al. (2022) would generalize to stimuli with higher variability – and, hence, whether their original results may have been influenced by a spurious response strategy.

Below, we report six online experiments. Experiments 1 and 2 use written stimuli, aiming to determine whether Haugen et al.'s (2022) results generalize when accounting for the smaller methodological issues mentioned above. Experiments 3-6 use auditory stimuli and directly contrast perception and production tasks. In all the experiments, our primary research question is whether syllable-copy reduplication is more difficult to learn than string-copy reduplication, where – like Haugen et al. (2022) – we implement the latter as CV- and CVC-reduplication for two different participant groups. Given this research question, we aim at equal sample sizes in the syllable-copy condition, on one hand, and the string-copy condition, on the other hand, even though the string-copy condition includes two subconditions. However, one may also wonder whether there is a learnability difference between these subconditions, i.e. CV-or CVC-reduplication. Even though the sample sizes for this comparison are necessarily smaller, we consider this secondary research question as well, and, hence, additionally compare the two string-copy subconditions to one another. Note, though, that we have no hypothesis regarding a potential difference between these two subconditions, and, hence, formulate no predictions – these additional analyses are exploratory.

2. Experiment 1

Experiment 1 aims to reproduce Haugen et al.'s (2022) findings with several methodological modifications: we test Italian participants drawn from a more diverse sample, construct items with unambiguous syllable structure, and use separate exposure and test phases. Furthermore, as we are only interested in the potential differences among the conditions with partial reduplication, we did not include their baseline condition with full reduplication. We expect that these modifications will not affect the pattern of results, and, hence, we predict that we will observe, like Haugen et al. (2022), better performance in the string-copy condition compared to the syllable-copy condition.

2.1 Methods

2.1.1 Open access

All experiment data, code, and stimuli are available at the OSF repository identified in the data accessibility statement. This experiment was not preregistered.

2.1.2 Participants

Fifty native speakers of Italian were recruited online via Prolific. Participants were asked to report which languages they spoke or had studied, but no additional data on language background (AoA, LoR, proficiency, etc.) were collected. In this and subsequent experiments, all participants reported at least one language other than Italian; one participant in this experiment mentioned languages whose morphology contains productive affixal reduplication (Ancient Greek and Latin). For this and subsequent experiments, we do not exclude participants who mentioned languages with productive affixal reduplication (see Appendix A for a complete list); however, we verified that the results are qualitatively the same without their data.

2.1.3 Stimuli

We generated 60 base words, all of which were phonotactically legal nonce words written in Italian orthography. For each participant, forty of the base words were randomly selected for the exposure phase, and the remaining 20 were used for the test phase. The base words were modeled after the prosodic structure of the stimuli from Haugen et al. (2022): half were of the form CVC.CV.CV, and the other half were of the form CV.CV.CV. Also in line with their study, reduplicative forms were generated by prefixing either the initial CVC or CV onto the base form, creating a total of 180 stimuli. However, the words were written via Italian orthography, which is extremely transparent, and all stimuli (both base and reduplicative forms) had completely unambiguous syllable boundaries. There was only one systematic source of ambiguity: as word stress in Italian falls on one of the last three syllables and is marked orthographically only when word-final, either the penultimate or the antepenultimate syllable could be interpreted as carrying stress in both the base and reduplicated forms.[2] Examples are shown in **Table 2**.

Table 2: Sample stimuli from Experiment 1.

base form	**reduplicated form**	
	CV-prefixed	**CVC-prefixed**
pu.mo.ca	**pu.**pu.mo.ca	**pum.**pu.mo.ca
re.li.co	**re.**re.li.co	**rel.**re.li.co
dan.tu.pe	**da.**dan.tu.pe	**dan.**dan.tu.pe
til.ma.vo	**ti.**til.ma.vo	**til.**til.ma.vo

One hundred and twenty images were included to illustrate the singular/plural distinction in the exposure phase. These images were black-and-white images of simple objects (e.g., candles, buttons). 60 images contained a single object (to represent the singular), and 60 images contained

[2] Penultimate heavy syllables attract stress, but, in our items, the penultimate syllable is always light.

two identical objects, side-by-side (to represent the plural; generated by replicating the object in the corresponding singular object image). All images were presented at a fixed height of 150 pixels. For the exposure phase, a subset of 40 of these objects were used (20 randomly selected singular objects, and 20 distinct randomly selected plural objects).

2.1.4 Procedure

Upon agreeing to participate in the study, participants were directed to the online experiment (implemented in JavaScript/HTML/CSS). Participants were randomly assigned, in approximately equal numbers, to three reduplication conditions: CVC (string-copy), CV (string-copy), and CVX (syllable-copy).

Upon providing consent, participants were instructed that they would be learning a new language: Martian.[3] In the exposure phase, participants were told that they would see a sequence of words, each alongside a corresponding image. The instructions made clear that participants did not need to learn the meanings of each word; rather the images simply highlighted the singular/plural distinction. They were asked to learn how to form the plural in Martian, based on these examples.

The exposure phase consisted of 40 trials. During these trials, participants were presented with a written word and a corresponding image (**Figure 1**). Upon viewing the word and image, participants were asked to press a key to continue to the next trial. Half of the trials included *singular* items, including a base word and a *singular* image; the other half included *plural* items, including the reduplicated form of a base word and a *plural* image. Singular and *plural* trials were randomly interleaved, and in order to further increase difficulty, each item appeared only once, either in the *singular* or in the *plural* form. Crucially, the reduplicated form of the *plural* trials varied depending on condition. In the string-copy conditions, the *plural* always had either the CV reduplicated form (in the CV condition) or else the CVC reduplicated form (in the CVC condition). In the syllable-copy condition, the *plural* always had the CVX reduplicated form: the first syllable in the base form was copied, whether CV or CVC.

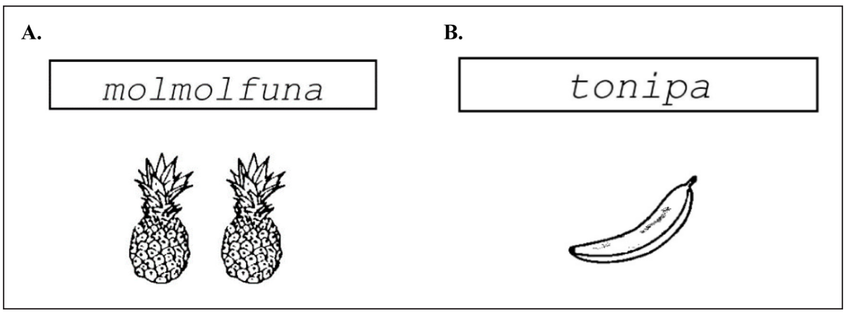

Figure 1: Sample plural **(A)** and singular **(B)** exposure phase trials.

[3] All methodological descriptions are translated from Italian.

Following the exposure phase, participants were instructed to complete a test phase. This phase consisted of 20 trials, composed of the base form of a word and two corresponding reduplicated forms, one with a CVC reduplicative prefix and the other with a CV reduplicative prefix (**Figure 2**). Participants were asked to select the correct plural form of the base form by pressing a key. After making a selection, they were immediately shown the next test trial. Accuracy was determined by condition, in line with the exposure phase: in the string-copy conditions, the correct reduplicated form was based on copying a consistent segmental melody; in the syllable-copy condition, it was based on copying the initial syllable. Unlike Haugen et al. (2022), participants were not provided with feedback during this phase; they were asked to do their best based on what they had learned during the exposure phase. After completing the test phase, participants responded to a brief posttest questionnaire about response strategies.

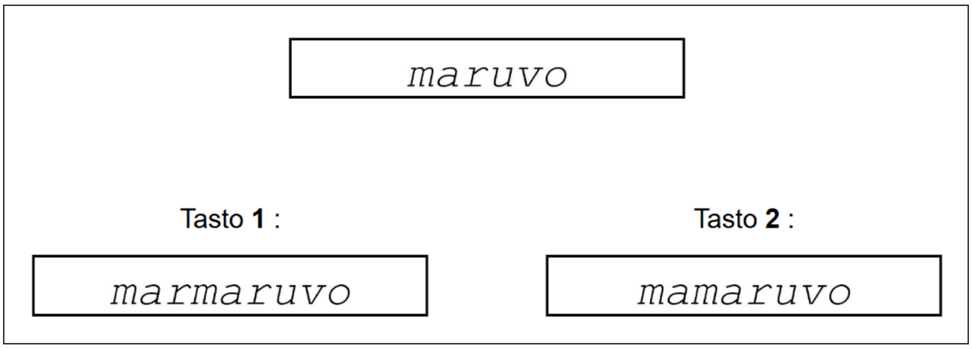

Figure 2: Sample test phase trial.

2.1.5 Analytic method

To test for the presence of a learning bias, we used the *lme4* package (Bates et al., 2015) in R to construct a primary generalized linear mixed-effects model (GLMM) on all data: a logistic regression (binary logit model) with fixed factor Main Type (string-copy vs. syllable-copy) and random intercepts for subject and item. In order to test for differences between string-copy subconditions, we constructed a secondary GLMM on the data from the string-copy conditions, with fixed factor Type (CV vs. CVC) and random intercepts for subject and item. In these and all subsequent analyses, categorical variables were dummy (treatment-) coded, the BOBYQA optimizer (Powell, 2009) was used when it was necessary to obtain model convergence, and odds ratios (OR) were obtained as a measure of effect size (Harrison & Pius, 2021). To assess statistical significance, we followed the standard procedure described in Levy (2014): in addition to the full models (described above), we fit base models without the fixed effect for comparison. The comparisons were conducted using the built-in *anova* function in R, which – when called with two mixed-effects models as parameters – conducts a likelihood ratio test over the fits of the two models (yielding a Chi-square statistic derived from their log-likelihoods).

Descriptive statistics per condition for this and all subsequent experiments can be found in Appendix B.

2.2 Results and discussion

The primary GLMM revealed that accuracy was lower in the syllable-copy condition than in the string-copy condition ($\beta = 1.66$, OR = 5.24, SE = 0.65, $\chi^2 = 6.54$, $p < .05$). The secondary GLMM identified no significant difference between the string-copy subconditions ($\chi^2 < 1$). These results (**Figure 3A**) are consistent with Haugen et al. (2022) and Wang (2024): participants exhibit a learning bias against syllable-copy reduplication.

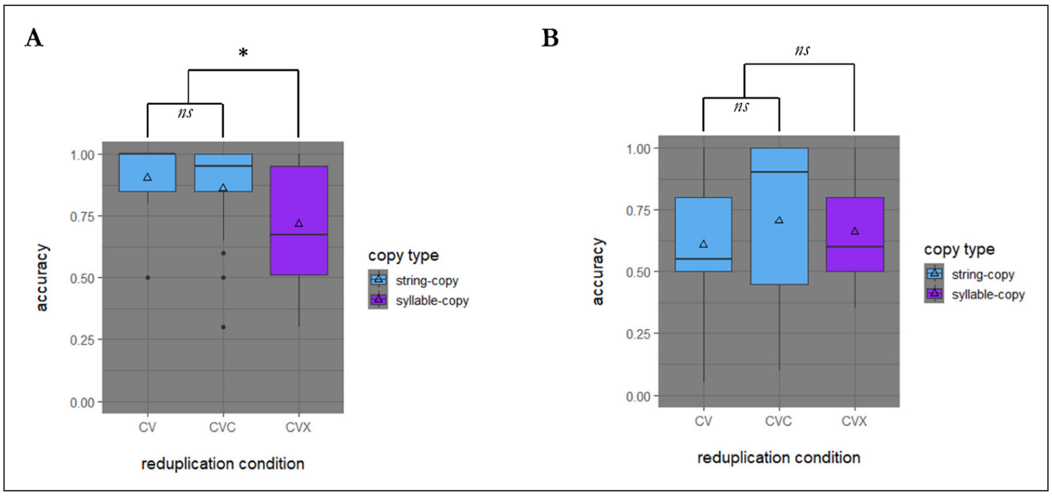

Figure 3: Boxplots illustrating the results of Experiment 1 **(A)** and Experiment 2 **(B)**. The y-axis represents accuracy in the test condition, the x-axis represents reduplication condition (including string-copy subconditions), and color represents main type. The triangles represent condition mean.

In their posttest responses, participants reported relying on a wide variety of strategies in the test phase. For example, consider the following responses to (all translated from Italian) "What strategy did you use to choose between the options? Did you find a specific rule?"

(4) a. "Which of the two words sounds better and is shorter to pronounce"
b. "No definite rule, I was going by sound because I also read the first words aloud."

These participants report pronouncing the words (either overtly or covertly); hence, they were attending to how the words sound. However, other participants reported relying on general syllabification rules:

(5) a. "repeat the first syllable."
b. "singular words = trisyllabic; plural words = first syllable (x2) + 2 final syllables"

For these participants, it is ambiguous whether they actually attended to how the words would be pronounced, as they may have referred to the syllabification of written words. Moreover, consider the following three responses:

(6) a. "For the plural we repeated the first 2 or 3 letters of the word, depending on whether the amount of letters was even or odd, if even we repeated the first 2, if odd we repeated the first 3"
b. "Repeat the first 3 letters of the word"
c. "doubling the first pair of letters."

These participants paid attention exclusively to orthography rather than inferred phonology. Presenting the stimuli auditorily may yield different results for these participants in particular. Furthermore, the relative simplicity of some of the strategies (e.g., simply repeating the first letters) highlights the concern that responses to the stimuli may be influenced by their lack of variability. Experiment 2 attempts to directly address this latter concern.

3. Experiment 2

Experiment 2 aims to determine whether the results of Experiment 1 – and Haugen et al. (2022) – generalize to more variable stimuli. It is possible that the result in the previous experiment was due to participants paying attention to the number of letters; as in the original Haugen et al. (2022) article, for participants in the CV condition, the reduplicated form of a six-letter base word was always eight letters (see example (6a) for a participant who explicitly identified this contingency). By adding variability in the length of the items, we remove this confound and provide a better test of the learnability of the reduplicative conditions per se.

3.1 Methods

3.1.1 Open access

This experiment was preregistered on AsPredicted: https://aspredicted.org/vscv-3gdc.pdf. All experiment data, code, and stimuli are available at the OSF repository identified in the data accessibility statement.

3.1.2 Participants

One hundred participants were recruited online via Prolific, according to the following criteria: they had to be L1 Italian speakers who had not participated in Experiment 1. Preregistered exclusion criteria included: responding positively to posttest questions asking whether they took notes during the exposure phase and whether they had studied (psycho-)linguistics. No participants were excluded under these criteria.

3.1.3 Stimuli

Unlike Experiment 1, each participant was presented the same set of base words in the exposure and test phases. For the test phase, twenty base words were chosen from among the items used in Experiment 1. We kept the variability of the items in the test phase the same, as they are simply meant to assess (rather than teach) the reduplication patterns. To determine whether the findings of Experiment 1 depend on the complexity of the items used for learning, we modified the exposure phase. We increased the variability of the stimuli in this phase by generating a new set of 40 base words (all phonotactically legal nonce words written in Italian orthography) containing more complex phonological and/or orthographic features in one or more syllables: complex onsets, consonant clusters, and digraphs. Crucially, the words still followed the prosodic structure of Experiment 1 (i.e., three syllables, with the first syllable including a coda consonant for half of the words). Furthermore, the syllable boundaries of all of the words – including reduplicated forms – were completely unambiguous. As a consequence, reduplicated forms ranged from 9 to 14 letters.

In line with Experiment 1, reduplicated forms copied a consistent segmental melody. Due to the presence of more complex onsets, this melody often included onset clusters of more than one consonant/letter; however, to keep naming conventions consistent, reduplicative forms are still classified as "CV" or "CVC". Note that copying onset clusters in toto is typologically common in reduplicative languages (Kim, 2020; McCarthy & Prince, 1995). Therefore, our more variable stimuli remain naturalistically valid for the purposes of comparing string copy (e.g., copying the first onset and vowel, including both C and CC onsets) and syllable copy. Examples are shown in **Table 3**.

Table 3: Sample stimuli from Experiment 2.

base form		**reduplicated form**	
		CV-prefixed	**CVC-prefixed**
chia.ni.ta	/kjanita/	**chia.**chia.ni.ta	**chian.**chia.ni.ta
dri.no.vla	/drinovla/	**dri.**dri.no.vla	**drin.**dri.no.vla
dan.tru.po	/dantrupo/	**da.**dan.tru.po	**dan.**dan.tru.po
quir.cio.pa	/kwirtʃopa/	**qui.**quir.cio.pa	**quir.**quir.cio.pa

The images presented alongside the words were identical to those from Experiment 1. As in Experiment 1, a random subset of the images was selected for each participant in the exposure phase.

3.1.4 Procedure

The procedure of Experiment 2 was largely identical to Experiment 1. However, upon being directed to the experiment, participants were assigned to one of the two main conditions via block randomization: string-copy (50) or syllable-copy (50). Within the string-copy condition, participants were randomly assigned to one of two subconditions (25 CV, 25 CVC).

3.1.5 Analytic method

The dependent variable, primary GLMM, and secondary GLMM were identical to those from Experiment 1. In the preregistered analyses for this and subsequent experiments, we did not include random slopes for condition. Following reviewer feedback, we repeated our tests with by-subject slopes for condition and found the same qualitative pattern of results (i.e., significance) across all comparisons.

3.2 Results and discussion

The primary GLMM identified no significant difference between the string-copy and syllable-copy conditions ($\chi^2 < 1$), and the secondary GLMM identified no significant difference between the string-copy subconditions ($\beta = 1.00$, OR $= 2.72$, SE $= 0.60$, $\chi^2 = 2.71$, $p = .099$). These results (**Figure 3B**) contradict both Experiment 1 and Haugen et al (2022): participants did not exhibit a learning bias against syllable-copy reduplication. In particular, the numeric difference between mean accuracy in the string-copy and syllable-copy conditions was negligibly small in this experiment ($<1\%$), despite a larger sample size and higher power than Experiment 1. This suggests that the failure to replicate the finding in this experiment is due to the difference between designs rather than, e.g., noise in significance-based comparisons (Vasishth et al., 2018). Moreover, whereas the difference between the two string-copy subconditions was not significant, the magnitude of the effect (as quantified by the odds ratio) was moderate (Rosenthal, 1995); for subsequent exploratory analyses, see 8.2.

At first glance, these results seem to suggest that there is a complex interdependence between typology, learning biases, and syllable-copy reduplication: the learning bias against syllable-copy reduplication applied to the highly simple items in Experiment 1, but it did not generalize to the more variable ones in this experiment. This suggests that previous findings have been the spurious result of participants adopting metalinguistic strategies that are easier to apply to simple stimuli, casting doubt on prior findings about the relationship between reduplicative typology and cognitive learning biases. However, there is a crucial factor that neither Experiment 1 nor Experiment 2 accounted for: modality. What results do we observe when we use audio recordings, rather than written words? A compelling prediction is that, given more variable stimuli, the underlying learning bias is masked by the limitations of written stimuli – and will reemerge with auditory input.

4. Experiment 3

Experiment 3 aims to determine whether there is a learning bias when stimuli are presented in auditory form rather than in written form. If the extreme cross-linguistic rarity of syllable-copy reduplication is due to a perceptual learning bias, performance should be better in the string-copy conditions than in the syllable-copy condition.

4.1 Methods

4.1.1 Open access

This experiment was preregistered on AsPredicted: https://aspredicted.org/k3jh-qnnk.pdf. All experiment data, code, and stimuli are available at the OSF repository identified in the data accessibility statement.

4.1.2 Participants

One hundred participants were recruited online via Prolific, according to the following criteria: they had to be L1 Italian speakers who had not participated in Experiments 1 or 2. No participants were excluded for taking notes, but 5 participants were excluded without replacement, due to having studied (psycho-)linguistics, leaving a final sample of 95 participants.

4.1.3 Stimuli

The exposure and test items for Experiment 3 were the same as those in Experiment 2. They were recorded by the second author, a fluent Italian speaker. She produced all 180 items twice, once with penultimate stress and once with antepenultimate stress.

4.1.4 Procedure

The procedure was largely identical to Experiment 2. However, a number of changes were implemented, in order to reflect the new stimulus modality. Firstly, participants were randomly assigned (in approximately equal numbers) to one of the two stress conditions. Secondly, during the exposure phase, the recording of the word for each trial was played once, immediately after the appearance of the corresponding image. Participants still proceeded from trial to trial via a keypress, with trials starting 300 ms after the keypress. Finally, during the test phase, the recordings of the words were played in the following sequence (with an ISI of 1000 ms): base, reduplicated form 1 (randomly CV or CVC), reduplicated form 2. After listening to each word, participants were asked to press a key to indicate which of the reduplicated forms was correct. After making a selection, there was a 500 ms delay before the next test trial began.

4.1.5 Analytic method

The dependent variable, primary GLMM, and secondary GLMM were preregistered and identical to those in Experiments 1 and 2, except that the random intercept for item was removed from the secondary model, in order to avoid a convergence failure (Bates et al., 2015) that could not be resolved by adding the BOBYQA optimizer. In addition, we analyzed the effect of stress condition in a GLMM on all data (specifically, a logistic regression), with the fixed factor Stress Condition (penultimate vs. antepenultimate) and random intercepts for subject and item.

4.2 Results and discussion

The results (**Figure 4A**) are consistent with Experiment 2 rather than Experiment 1: accuracy was not significantly lower in the syllable-copy condition than in the string-copy condition ($\chi^2 < 1$), and there was no significant difference between string-copy subconditions ($\beta = 0.76$, OR $= 2.15$, SE $= 0.46$, $\chi^2 = 2.65$, $p = .104$). Notably, the effect sizes of both comparisons were comparable to those in Experiment 2 (negligible for the former, and moderate for the latter). There were also no significant differences between the two stress conditions ($\chi^2 < 1$). This suggests that the results of Experiment 2 generalize to the auditory modality. More broadly, they corroborate the conclusion that the perceptual learning bias against syllable-copy reduplication fails to replicate, given more variable stimuli.

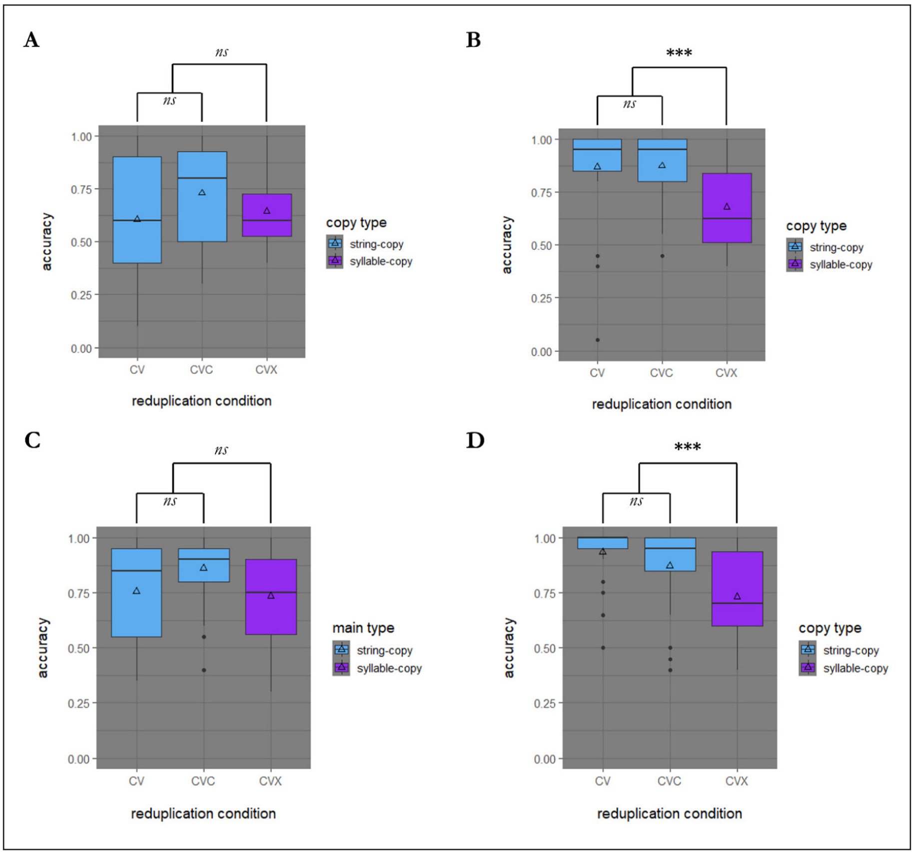

Figure 4: Boxplots illustrating the results of Experiments 3 **(A)**, 4 **(B)**, 5 **(C)**, and 6 **(D)**. The y-axis represents accuracy in the test condition, the x-axis represents reduplication condition (including string-copy subconditions), and color represents main type. The triangles represent condition mean.

However, this seems to introduce an empirical contradiction: Wang (2024) found results consistent with Haugen et al. (2022), despite using relatively variable stimuli; why are our results inconsistent with their convergent findings? One possibility is that Wang's (2024) alignment with Haugen et al. (2022) was spurious. In line with this, recall that she did not directly test for a learning advantage of string-copy reduplication over syllable-copy reduplication; this finding was secondary to her primary empirical targets. Another possibility lies in the fact that she used a speech production task rather than a perception task. As discussed in 1.3, the cognitive mechanisms underpinning speech perception and production are distinct, and AGL studies have revealed disparities between phonological learning biases as a function of task modality (Do & Havenhill, 2021; Peperkamp et al., 2006). Thus, before concluding that there is no learning bias that can account for the extreme cross-linguistic rarity of syllable-copy reduplication, we examine in the next experiment if such a bias can be observed in a speech production task.

5. Experiment 4

Experiment 4 aims to determine whether the null pattern of results observed in Experiment 3 generalizes to a production task or not. If the extreme cross-linguistic rarity of syllable-copy reduplication is due to a learning bias (even if that bias is not detectible in a perception task), we predict that performance will be better in the string-copy condition than in the syllable-copy condition for this production task.

5.1 Methods

5.1.1 Open access

This experiment was preregistered on AsPredicted: https://aspredicted.org/kkrn-5cz9.pdf. Coded experiment data, code, and stimuli are available at the OSF repository identified in the data accessibility statement. In order to preserve participant anonymity, raw audio recordings will not be made publicly accessible.

5.1.2 Participants

Participants were recruited online via Prolific, and all exclusions were filled with replacements, until a final sample of 100 participants was reached. Participants were prescreened according to the following criteria: they had to be L1 Italian speakers who had not participated in Experiments 1, 2, or 3. Furthermore, they had to complete an automated microphone check before beginning the experiment. Forty-nine participants were excluded and replaced for taking notes (N = 2), inaudible recordings in the test trials (N = 3), or failing to produce any word-initial reduplication in at least 80% of trials (N = 44).[4] (No participant reported having studied (psycho-)linguistics.)

[4] Of these 44 participants, all but two failed to produce any reduplication. The vast majority either used standard Italian pluralization or repeated the words in their base form.

5.1.3 Stimuli

The stimuli were identical to those in Experiment 3.

5.1.4 Procedure

The exposure phase proceeded identically to Experiment 3. The test phase again consisted of 20 randomly ordered trials, with each trial corresponding to one of the novel base forms. Participants were instructed to record themselves producing the plural form; they had an unlimited amount of time to produce this form, and the experiment would not proceed until after it had detected microphone input above an audible threshold.

5.1.5 Response coding and analytic method

Each recording was manually coded for content, blinded to experiment condition. First, the responses for each trial were transcribed in Italian orthography. Based on these transcriptions, each trial was then coded as one of three possible reduplicant types: CV, CVC, or other (encompassing no reduplication and reduplications other than CV/CVC). During this step, trials in which participants replaced one or more phonemes from the base word and/or reduplicated with a phoneme of the corresponding consonant or vowel category (e.g., /v/ → /b/, /i/ → /o/) were coded as CV or CVC if they still followed that reduplication pattern. For example, given the base word *mulfona*, a trial in which the participant produced "munmunfona" would be coded as CVC. Given multiple produced forms in the recording from a single trial, the coding was based on the final form. For example, if a participant in this hypothetical trial produced "mu*mulfona*… [pause]…mul*mulfona*," that trial would be coded as CVC. Trials in which participants did not produce any reduplicated form or produced a reduplicated form that did not correspond to CV/CVC were coded as *other*, which was automatically counted as incorrect, regardless of condition.

All three preregistered GLMMs were identical to Experiment 3.

5.2 Results and discussion

Among the data from participants included in the experiment, 2.65% of transcriptions were coded as "other" (i.e., failing to produce any reduplicative form or producing a reduplicative form other than CV/CVC). These trials (automatically counted as incorrect) were included in the analysis; excluding them does not yield any differences in model outcome direction or significance.

Accuracy was significantly lower in the syllable-copy condition than in the string-copy condition ($\beta = 1.64$, OR = 5.17, SE = 0.28, $\chi^2 = 31.13$, $p < .001$), but there was no significant difference between string-copy subconditions ($\chi^2 < 1$) (**Figure 4B**). There was also no significant difference between stress conditions ($\beta = 0.36$, OR = 1.43, SE = 0.32, $\chi^2 = 1.25$, $p = .263$). This provides a striking contrast to Experiment 3: the learnability of the syllable-copy condition was strongly reduced, relative to both of the string-copy subconditions (direct comparisons between these two experiments are described in 8.1). Crucially, these two experiments shared identical exposure phases, so the disparity is most likely due to the methodological difference in the test phases: production versus perception. This suggests that the typological rarity of syllable-copy reduplication is driven by an underlying cognitive learning bias, but that bias selectively operates in speech production. In turn, this provides a clear resolution to the seemingly conflicting findings between our previous three experiments and the previous two studies: when accounting for stimulus variability, the finding originally attested by Haugen et al. (2022) does not replicate in perception tasks, but does replicate in production tasks, in line with Wang (2024) – a surprising finding, that is, however, reminiscent of two other studies that compared perception and production in phonological rule learning. First, Peperkamp et al. (2006) observed better learning of a phonetically natural rule, compared to a matched unnatural one, when using a production task, but not when using a perception task. Second, using only a perception task, Do and Havenhill (2021) found better learning of a phonetically natural rule only when participants had to repeat items out loud during the exposure phase.

Beyond a simple analysis of accuracy, phonological production data allow for a richer analysis of potential biases. Specifically, it is possible that errors in the syllable-copy condition are directionally biased: individual participants may tend to rely on either a CV-copy or a CVC-copy strategy, resulting in a mix of incidentally accurate responses (e.g., when a participant using a CV-copy strategy encounters a CV-initial base form) and inaccurate responses (e.g., when a participant using a CV-copy strategy encounters a CVC-initial base form). If there is indeed a systematic bias in favor of string-copy reduplication over syllable-copy reduplication, we would predict that, for a given participant in the syllable-copy condition, errors are disproportionately biased towards either CV or CVC forms (rather than a mix of the two). This leads to a predicted relationship between accuracy and response bias in the syllable-copy condition: for participants with high accuracy, there will be an even mix of CV and CVC responses (by necessity, as the correct CVX reduplicated forms require a mix of the two); as accuracy decreases, responses diverge towards mostly CV responses and mostly CVC responses.

In order to test this, we ran a non-preregistered exploratory analysis: We computed for each participant in the syllable-copy condition a *copy strategy bias index* (CSBI), defined as the number of trials in which participants erroneously produced a CVC reduplicated form minus the number of trials in which participants erroneously produced a CV reduplicated form. For example, a maximum score of 10 indicates that a participant erroneously produced CVC forms for all 10 trials where CV forms would have been correct, but correctly produced CVC forms for all 10 trials where CVC forms were correct – in other words, always produced CVC. The minimum score of –10 indicates the opposite (i.e., that participants always produced CV). A score of 0 indicates that a participant either made no errors or made an even mix of CVC and CV errors. The expected pattern is evident in a visualization of the CSBI data: as accuracy decreases, CSBI diverges towards 10 – reflecting a CVC-copy strategy – and –10 – reflecting a CV-copy strategy (**Figure 5**). This also leads to a specific quantitative prediction: the absolute value of CSBI should be inversely correlated with accuracy. This prediction was borne out in a Pearson correlation analysis ($r = –0.90$, $SE = 0.06$, $t(48) = 14.08$, $p < .001$). Therefore, beyond our main results, we even see a bias towards string-copy reduplication strategies within the syllable-copy condition.[5]

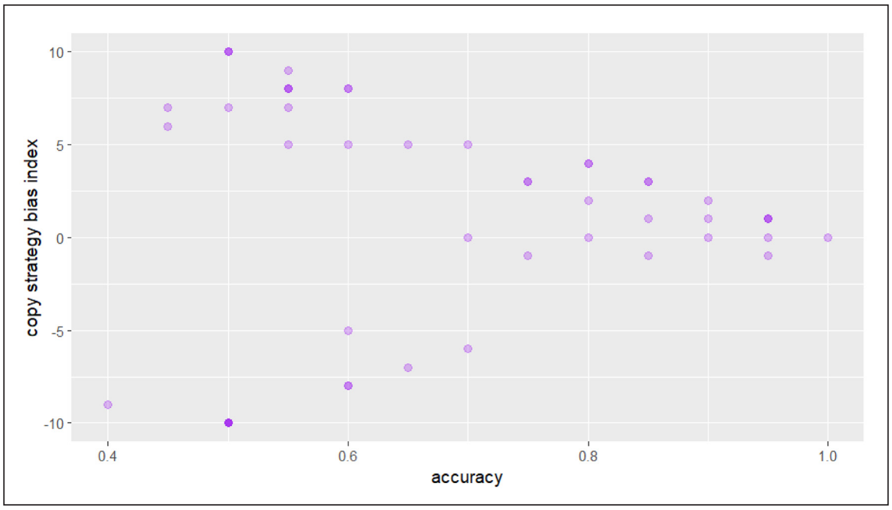

Figure 5: A scatterplot illustrating the exploratory analysis of errors in the syllable-copy condition of Experiment 4. Opacity represents participant density at a given point. As accuracy decreases, participants systematically diverge towards errors indicative of either CVC-copy (positive index) or CV-copy (negative index) reduplication strategies.

The learning bias against syllable-copy reduplication revealed by this experiment calls for a reconsideration of data from our earlier orthography-based experiments. Because the stimuli from those experiments were written words, some participants may have either "mouthed" the words silently or pronounced the words out loud – thereby producing the plural forms. If this is the case, we would predict that these participants may have selectively been subject to the learning bias observed here, even if it was absent in the experiment as a whole. One of the posttest strategy questions posed to participants in Experiment 1[6] has bearing on this possibility (translated from Italian): "During the first stage, did you say the words (silently or aloud)?" The

[5] A standard way to analyze this sort of bias is via Signal Detection Theory, which enables the computation of a "criterion" index representing participants' relative tendency to produce false alarms, as opposed to misses (Green & Swets, 1966, inter alia). We do not present such an analysis here, since the CSBI index is substantially more intuitive in the context of our experiment. However, we verified that the results of the two types of analysis are consistent.

[6] No similar question was included in the questionnaire of Experiment 2.

response options were: "yes, always" (selected by 36% of participants); "yes, only for singular" (0%); "yes, only for plural" (12%); "yes, sometimes" (32%); and "no" (20%).

Based on their responses, participants can be divided into two groups: one group which consistently produced the plural form (encompassing "yes, always," and "yes, only for plural"; N = 24) and one group which did not, or only inconsistently (encompassing "no" and "yes, sometimes"; N = 26). We carried out two separate exploratory analyses, one for each group: a logistic regression (binary logit model) with fixed factor Main Type (string-copy vs. syllable-copy). The regression for the "no/yes, sometimes" group included random intercepts for subject and item; the random intercept for item was removed from the "yes, always/yes, only for plural" group, in order to avoid a convergence failure that could not be resolved by adding the BOBYQA optimizer. These models (**Figure 6**) revealed a significant bias against syllable-copy reduplication in the "yes, always/yes, only for plural" group (β = 3.70, OR = 40.40, SE = 1.19, χ^2 = 9.59, $p < .01$), but not in the "no/yes, sometimes" group ($\chi^2 < 1$). This suggests that the production-dependent learning bias revealed by the present experiment is observable even in our first experiment, but was originally concealed by the inconsistency of strategies used by participants in its orthography-based paradigm.

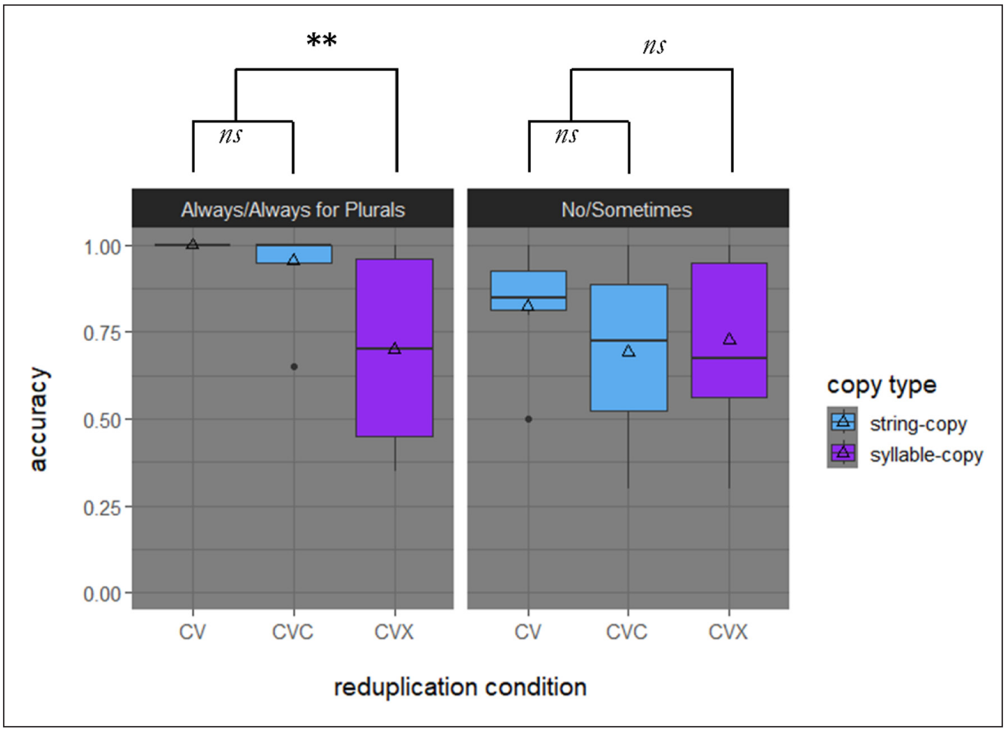

Figure 6: Boxplots illustrating the exploratory reanalysis of Experiment 1, split into "consistent plural production" (Always/Always for Plurals) and "inconsistent plural production" (No/Sometimes) groups.

Our experiments so far are indicative of the presence of a selectively production-driven learning bias in favor of string-copy over syllable-copy reduplication. However, beyond the production-perception distinction, there is one methodological concern that remains unaddressed: sampling bias. Despite the inclusion of a total of 20 base forms and 20 reduplicative forms in the exposure phase, a remarkable number of participants (44) were excluded and replaced, due to failing to produce reduplication in the majority of trials. In the previous experiments, the two-alternative forced-choice task did not allow us to identify participants who may not have learned any sort of reduplication in the exposure phase. Therefore, it is possible that in the present experiment, we ended up with a more "conscientious" sample of participants than in the previous ones, leading to an asymmetric comparison. In order to address this issue, it is necessary to create an easier production experiment, such that fewer participants will fail to produce a substantial amount of reduplicated forms, as well as a matched perception experiment to allow for an appropriate comparison between perception and production. To this end, Experiments 5 and 6 are matched perception and production experiments that differ from Experiments 3 and 4 in the exposure phase.

6. Experiment 5

Experiment 5 modifies the methodology of Experiment 3 (perception) by presenting stimuli in singular-plural pairs during the exposure phase, thus making reduplication particularly salient. We expect to observe the same result as in Experiment 3, i.e., no difference between the syllable-copy and string-copy conditions.

6.1 Methods

6.1.1 Open access

This experiment was preregistered on AsPredicted: https://aspredicted.org/4jmx-58js.pdf. All experiment data, code, and stimuli are available at the OSF repository identified in the data accessibility statement.

6.1.2 Participants

Participants were recruited online via Prolific, and all exclusions were filled with replacements, until a final sample of 100 participants was reached. Participants were prescreened according to the following criteria: they had to be L1 Italian speakers who had not participated in Experiments 1, 2, 3, or 4. One participant was excluded and replaced for taking notes, and three participants were excluded and replaced, due to having studied (psycho-)linguistics.

6.1.3 Stimuli and procedure

All stimuli drew from the stimulus set from Experiments 3 and 4. As neither Experiment 3 nor Experiment 4 found any significant differences between stress conditions, only the stimuli with penultimate stress[7] were used, in order to simplify design and reduce noise in the analyses. For the exposure phase, participants were told that, on each trial, they would hear first the singular form and then the plural form of a word in Martian. Similarly to the previous experiments, they were also told that they did not need to memorize the words, but that their task was to learn how to make plurals in Martian. Contrary to the previous experiments, the items were presented without accompanying pictures, since, due to the pairwise presentation, it was not necessary to visually illustrate the singular/plural distinction.

The presentation of pairs of base and reduplicated forms may – in addition to highlighting reduplication – make it easier to identify the correct reduplication strategy; therefore, the exposure phase for each participant was reduced to 10 trials, to avoid ceiling performance. For each participant, we randomly selected a subset of 10 out of the 40 base words from the exposure set (5 CV-initial and 5 CVC-initial) and the corresponding reduplicated form. On each trial, participants first heard the base form, and, 500 ms later, the reduplicated form. The test phase proceeded identically to that in Experiment 3.

6.1.4 Analytic method

The dependent variable, primary GLMM, and secondary GLMM were identical to those for Experiment 3, and all were preregistered. (Because all participants were assigned to the same stress condition, we did not conduct the prosody analysis from previous experiments.)

6.2 Results

The results (**Figure 4C**) are consistent with Experiment 3: accuracy was not significantly lower in the syllable-copy condition than in the string-copy condition ($\beta = 0.57$, OR = 1.77, SE = 0.30, $\chi^2 = 3.48$, $p = .062$), and there was no significant difference between string-copy subconditions ($\beta = 0.74$, OR = 2.09, SE = 0.44, $\chi^2 = 2.63$, $p = .105$). This suggests that the results of Experiment 3 generalize when the relevant grammatical feature – reduplication – is made more explicit in the exposure phase.

7. Experiment 6

Experiment 6 combines the exposure of Experiment 5 (pairwise presentation of base and reduplicated forms) with the production task of Experiment 4. We expect that, with this simplified design, fewer participants will have to be excluded for lack of applying word-initial reduplication. Then, if the better performance in the string-copy condition observed in Experiment 4 was not due to a sampling bias, we should observe the same result.

[7] In Italian, penultimate stress is the default. This holds, in particular, for words that match the structure of our items, i.e., trisyllabic nouns in which the last two syllables have no coda consonant: the electronic corpus phonItalia (Goslin et al., 2013) contains 5550 such nouns, 74% of which have penultimate stress.

7.1 Methods

7.1.1 Open access

This experiment was preregistered on AsPredicted: https://aspredicted.org/rcn4-gv7q.pdf. Coded experiment data, code, and stimuli are available at the OSF repository identified in the data accessibility statement. In order to preserve participant anonymity, raw audio recordings will not be made publicly accessible.

7.1.2 Participants

Participants were recruited online via Prolific, and all exclusions were filled with replacements, until a final sample of 100 participants was reached. Participants were prescreened according to the following criteria: they had to be L1 Italian speakers who had not participated in Experiments 1, 2, 3, 4, or 5. Furthermore, all participants had to complete an automated microphone check before beginning the experiment. Participants were excluded and replaced for the following reasons: taking notes (N = 2), having studied (psycho-)linguistics (N = 4), inaudible recordings in the test trials (N = 2), failing to produce any word-initial reduplication in at least 80% of trials (N = 9).

7.1.3 Stimuli, procedure, coding, and analytic method

The exposure phase proceeded identically to Experiment 5. In line with Experiment 5, participants were uniquely assigned to the penultimate stress condition. The test phase proceeded identically to Experiment 4. Responses were coded according to the same criteria and procedure as in Experiment 4, and preregistered analyses were identical, except for the absence of the stress condition analysis.

7.2 Results

The results (**Figure 4D**) are consistent with Experiment 4: accuracy was significantly lower in the syllable-copy condition than in the string-copy condition ($\beta = 1.86$, OR = 6.40, SE = 0.35, $\chi^2 = 29.41$, $p < .001$), and there was no significant difference between string-copy subconditions ($\beta = 1.39$, OR = 4.00, SE = 0.75, $\chi^2 = 3.65$, $p = .056$). This suggests that the findings of Experiment 4 generalize when highlighting the salience of reduplication in the exposure phase. Crucially, the present experiment – in tandem with Experiment 5 – was not subject to the same asymmetric exclusion volumes as Experiments 3 and 4. Only nine participants in the present experiment were excluded due to a failure to produce reduplication in the majority of trials; this suggests that our manipulation of the exposure phase was successful in making reduplication more explicit as the relevant grammatical feature. Therefore, this experiment has a comparable sample to Experiment 5, in turn corroborating the pattern observed in Experiments 3 and 4: we reliably identify a selectively production-driven learning bias against syllable-copy reduplication. In fact, this experiment yielded a numerically stronger effect size for the string-copy vs. syllable-copy comparison than Experiment 4 (6.40 vs. 5.17), confirming that the prior finding was not inflated by the higher exclusion rate. The present experiment also exhibited the same pattern of error bias as Experiment 4: in the syllable-copy condition, participants systematically diverge towards CVC-copy and CV-copy reduplicative strategies as accuracy decreases (**Figure 7**). This is supported by a Pearson correlation analysis, identifying an inverse relationship between accuracy and absolute CSBI ($r = -0.90$, $SE = 0.06$, $t(48) = 14.68$, $p < .001$). As in Experiment 4, this further illustrates the robustness of the underlying learning bias: even in the syllable-copy condition, participants still gravitate towards string-copy reduplicative strategies.

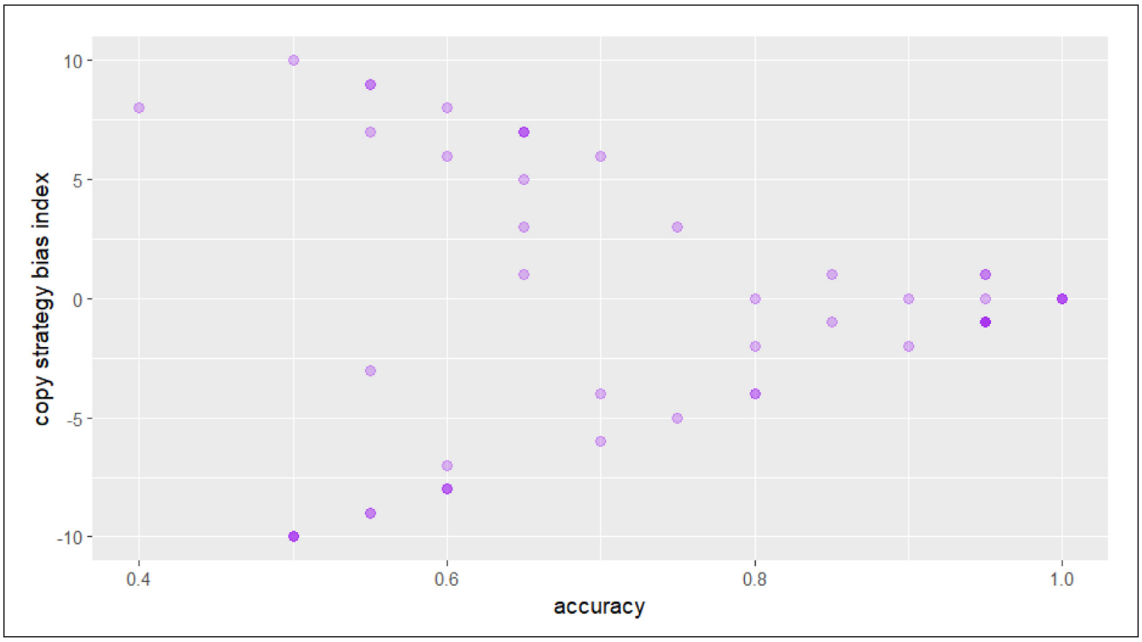

Figure 7: A scatterplot illustrating the exploratory analysis of errors in the syllable-copy condition of Experiment 6. Opacity represents participant density at a given point. As accuracy decreases, participants systematically diverge towards errors indicative of either CVC-copy (positive index) or CV-copy (negative index) reduplication strategies.

8. Exploratory composite analyses

8.1 Perception vs. production

In order to compare performance in perception vs. in production, we constructed two GLMMs (logistic regressions), one over the combined results of Experiments 3 and 4, and one over those of Experiments 5 and 6. The models had Task (perception vs. production), Main Type (string-copy vs. syllable-copy), and their interaction as fixed factors, and included random intercepts for

Item and Participant. For Experiments 3 and 4, the model revealed a main effect of Task (β = 1.58, OR = 4.85, SE = 0.28, χ^2 = 21.13, $p < .001$), with higher accuracy in production, and a significant interaction between Task and Main Type (β = 1.39, OR = 4.00, SE = 0.38, χ^2 = 13.14, $p < .001$). Post-hoc analyses carried out with the *emmeans* package (Lenth et al., 2025) showed that, for string-copy reduplication, accuracy was higher in production than in perception (EMM = 1.58, SE = 0.28, z = 5.71, $p < .001$), while for syllable-copy reduplication, there was no difference between tasks (EMM = 0.19, SE = 0.26, z = 0.76, p = .871). The analysis of Experiments 5 and 6 likewise yielded a main effect of Task (β = 1.16, OR = 3.18, SE = 0.34, χ^2 = 5.55, $p < .05$), again with better performance in production, and a significant interaction between Task and Main Type (β = 1.25, OR = 3.50, SE = 0.45, χ^2 = 7.78, $p < .01$). While effect sizes were marginally smaller for these comparisons than the corresponding comparisons for Experiments 3 and 4, this may be due to higher accuracy across conditions in Experiments 5 and 6 (i.e., a ceiling effect – although both the main effect and interaction are still significant). Once more, for string-copy reduplication, accuracy was higher in production (EMM = 1.16, SE = 0.34, z = 3.44, $p < .01$), while there was no difference for syllable-copy reduplication (EMM = 0.10, SE = 0.30, z = 0.32, p = .988).

The fact that performance was better in production than in perception is remarkable, as generating a correct form is intuitively more difficult than merely recognizing one. And, indeed, in a study of phonological rule learning mentioned earlier, higher performance was found in a perception task, compared to a production task, despite identical exposure (Peperkamp et al., 2006). Why we found the opposite pattern is unclear. The large number of participant exclusions in Experiment 4 may have yielded a more conscientious group than in Experiment 3, which could account for the higher overall accuracies in the former. However, this explanation cannot account for the same difference between Experiments 5 and 6, which were designed to ensure comparable samples. Therefore, it is possible that some aspect of reduplication itself facilitates production over perception – e.g., perhaps being forced to produce the plural forms (or anticipating, during the exposure phase, that the test phase will involve production) draws attention to the phonological make-up of the words that is more likely to be ignored in a purely perceptual task. However, this explanation is purely speculative, and it is unclear how it may account for the additional finding that the difference between production and perception is selectively driven by the string-copy condition.

8.2 CV vs. CVC

None of our experiments identified a significant difference between the string-copy subconditions (CV and CVC). Yet the two perception experiments (3 and 5) followed a consistent trend: participants in the CVC condition had numerically higher accuracy than participants in the CV condition (**Figure 8**, left panel). Moreover, the effect size for this comparison was moderate in both of these experiments – suggesting that the difference between the two subconditions was not detected due to insufficient power. Indeed, a logistic regression over the data of the string-copy subconditions from these two experiments taken together does show significantly higher accuracy in the CVC condition, compared to the CV condition ($\beta = 0.77$, OR = 2.16, SE = 0.34, $\chi^2 = 5.00$, $p < .05$). In the two production experiments, no similar trend is apparent, and the corresponding analysis for Experiments 4 and 6 (**Figure 8**, right panel) shows no significant difference between the two string copy conditions ($\beta = 0.75$, OR = 2.12, SE = 0.47, $\chi^2 = 2.68$, $p = .102$).

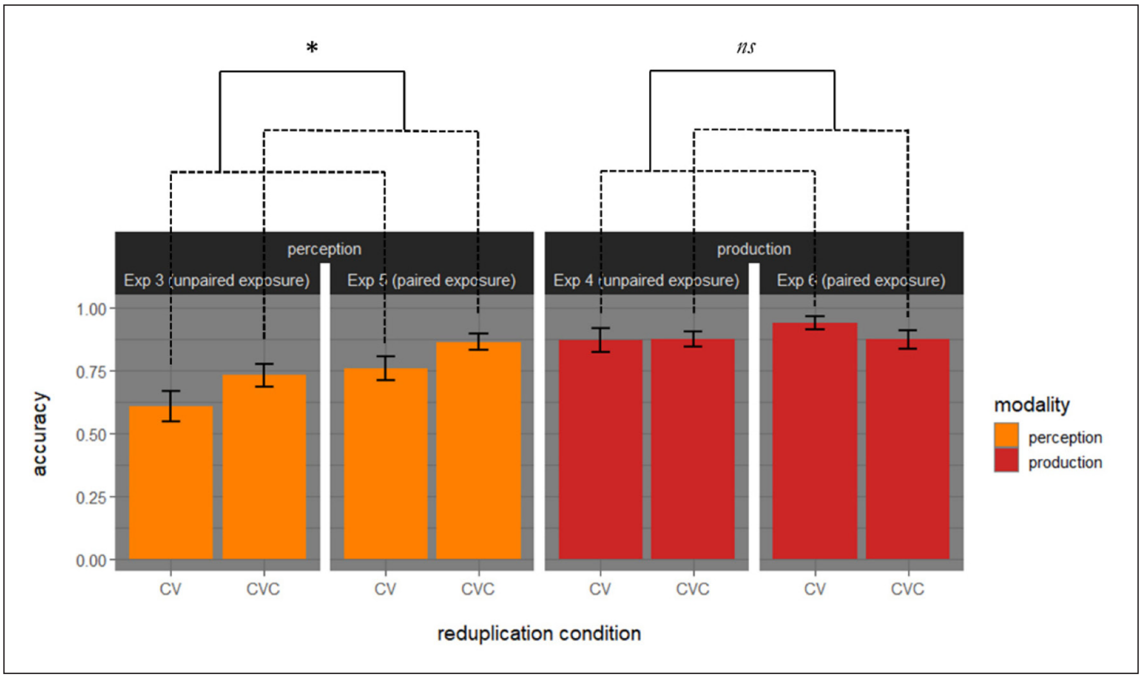

Figure 8: Barplots illustrating accuracy in the string-copy subconditions (CV and CVC) in Experiments 3–6.

Furthermore, recall the CSBI analyses in these experiments: participants tended to either systematically produce a CV or a CVC reduplicated form in the syllable-copy condition, rather than an even mix of the two. However, there is no evidence that these participants were more biased to produce CVC than CV: when combining the data from the syllable-copy conditions of Experiments 4 and 6, the median CSBI is exactly 0 (i.e., neither in the direction of a CV copy bias nor a CVC copy bias). The reasons for this perception-selective pattern are unclear: speculative explanations will be considered in Section 9.

9. Conclusion

Our results paint a complex picture of the relationship between reduplicative typology, cognitive learning biases, stimulus variability, and modality. Experiments 1-2 suggest that the original findings of Haugen et al. (2022) may have been limited by their relatively low-variability, written stimuli. With increased stimulus variability, their original finding disappeared, seemingly suggesting that the extreme rarity of syllable-copy reduplication may not have been due to the proposed learning bias. However, these experiments concerned the written modality, contrasting with methodologies that target phonological perception and production per se. By extending the AGL design to auditory stimuli, we were able to identify a persistent perception/production disparity. Across both of our perception experiments, there were no significant differences between the string-copy and syllable-copy conditions, whereas across both of our production experiments, there was a clear learning bias against syllable-copy reduplication, relative to string-copy reduplication. This leads us to a nuanced conclusion: when accounting for stimulus variability, the typological rarity of syllable-copy reduplication is selectively driven by a modality-specific learning bias in speech production.

In a very broad sense, our results do corroborate the central claim of Haugen et al. (2022): reduplicative typology is, in fact, linked to an underlying learning bias. The present article provides a crucial clarification: the learning bias is much less general than was originally proposed, and it is not reliably detectable with a perception task. As mentioned previously, there is a precedent for learning biases identified only when production is involved (Do & Havenhill, 2021; Peperkamp et al., 2006). However, our experiments do not directly address why this disparity might exist for the learning bias under scrutiny. One of the most straightforward reasons for a perception/production disparity in AGL experiments is that perception tasks are less challenging than production tasks; that is, in perception, participants may reach ceiling performance faster (which obscures potential learning biases), or they may be able to allocate sufficient cognitive resources to learning even a complex or unnatural rule (Peperkamp et al., 2006). However, this type of explanation can be ruled out for our findings, since overall accuracy was lower in our perception experiments than our production experiments, despite strictly identical exposure phases. Therefore, it is possible that there is a deeper psycholinguistic mechanism underlying the perception/production bias that we observed.

One argument that deserves scrutiny concerns the idea that minimizing production effort is relevant for speakers but not for listeners. We mentioned in Section 1 that syllable-copy reduplication is arguably less effortful than string-copy reduplication, because repeating a syllable is economical in terms of both planning and articulatory effort. In the context of our production experiments, however, where multiple reduplicated forms had to be produced one after the other, perhaps participants were motivated to keep the syllable type of the reduplicants across trials consistent – i.e., the necessary switching between CV and CVC reduplicants in our syllable-copy condition may require more effort (whereas the same pressure would not exist in perception). This argument, while purely speculative, can additionally account for the finding that in the syllable-copy condition, participants with low accuracy tended to apply either CV- or

CVC-copy reduplication throughout, rather than mixing the two types randomly. Yet, it cannot easily accommodate the finding that even in Experiment 1 (written input), we observed a learning bias when we restricted the analysis to those participants who produced the words during the exposure phase. That is, despite the fact that participants in the syllable-copy condition of this experiment thus gained experience with producing a mix of CV- and CVC-reduplication, their performance was worse, compared to those in the string-copy condition. However, there is no a priori reason to believe that switching between CV and CVC is more effortful than switching between full and partial syllable reduplication, and the theoretical argument, thus, hinges on the status of syllables versus segmental melodies in speech production.

Therefore, a more fruitful approach to understanding the observed learning bias may be to consider the relative complexity of syllable-based vs. string-based reduplication. Formal simplicity has been argued by Chater and Vitányi (2003) to be a driving factor in many areas of cognition, including language learning. Yet, as noted by the same authors, it can be hard to pin down, because there is no consensus about the mental representation of many perceptual and linguistic stimuli. The syllable is a case at hand: While it has traditionally played a central role in phonological theory (see, e.g., Blevins, 1995) and is referred to in paralinguistic domains such as poetic meter (Fabb & Halle, 2008) and language games (Vaux, 2011), its internal structure may not be universal, its phonetic correlates are elusive, and its necessity in accounting for segmental and suprasegmental phenomena has been called into question (for an overview of these empirical problems, see Côté, in press). At the neural level, speech is analyzed in different temporal integration windows, one of which coincides with syllable-sized units (Poeppel, 2003). Yet this is not evidence of the neural reality of the syllable, as segmental strings made up of a single vowel and any number of consonants are also “syllable-sized”. In the same vein, the widely shared intuition that speech consists of a sequence of syllable-sized units is not evidence of the psychological reality of the syllable. Perhaps, then, syllables are nothing more than an epiphenomenon (see, e.g., Ohala, 2008).

In psycholinguistic research, the question of whether the syllable, as such, is an online coding unit has been studied for decades, both in speech perception (reviewed in Dumay & Content, 2012) and in speech production (reviewed in Cholin, 2011). In speech perception, an early landmark study presented French-speaking participants with a sound sequence and had them indicate whether it was contained in a following word. This sequence detection task revealed a syllable congruency effect; e.g. /ba/ was easier to detect in *ba.lance* ‘balance’ than in *bal.con* ‘balcony’, while the reverse held for /bal/ (Mehler et al., 1981). This effect was later reinterpreted in terms of phonetic similarity (e.g., consonant realizations differ between onset and coda position, and vowel realizations, likewise, differ between open and closed syllables; see Kolinsky, 1998). The same holds for seemingly syllabic effects observed with different paradigms, such as phoneme monitoring and cross-modal semantic priming. In the latter paradigm, for instance, Italian listeners' recognition of a visually presented word (e.g., *rumore* 'noise') was facilitated by an auditorily presented CVC fragment when it was cut out from a semantically related word (*si.lenzio* 'silence'), but not when it was cut out from an unrelated word that starts with the same sequence, yet has a different syllable structure (*sil.vestre* 'sylvan'; Tabossi et al., 2000). Studies using dichotic listening, however, did find genuine evidence for the syllable as a coding unit. In one such study, Mattys and Melhorn (2005) showed that syllables can migrate (and do so more than individual vowels), such that, for instance, when presented with 'con.rool.tric' in one ear and 'bing.sen.crep' in the other, English listeners report hearing the word *concentric*.

In speech production, there is, likewise, evidence both for and against the hypothesis that the syllable is a genuine coding unit. Studies using fragment priming have failed to provide convincing evidence that syllables are activated in lexical retrieval. In this paradigm, participants are presented with a letter string followed by a picture that they have to name; the letter string either coincides or does not coincide with the initial syllable of the word to be named (e.g., prime BA or BAL, target balance or balcony). After years of conflicting findings, the current consensus is that of a segmental length effect: CVC primes yield faster response latencies than CV primes, regardless of the syllabic affiliation of the prime's post-vocalic consonant in the target word (e.g., Schiller & Costa, 2006). In the context of our production experiments, this suggests that the syllable-copy strategy may, indeed, be more complex: participants would have to attend to a unit that is not automatically activated. Yet, various other paradigms with a production task have shown that a shared syllable structure is beneficial (Costa & Sebastián-Gallés, 1998; Ferrand & Segui, 1998; Sevald et al., 1995). For instance, participants are faster to produce repetitions of non-word sequences (always a monosyllable followed by a disyllable) in which the first sequence has the same structure as the initial syllable of the second sequence, e.g., KEM-TIL.FER and KEMP-TILF.NER versus KEM-TILF.NER and KEMP-TIL.FER (Sevald et al., 1995). Moreover, the benefit of shared syllable structure is larger than that of shared segmental structure alone. For instance, in a study with an implicit form preparation paradigm, where Dutch participants learned sets of prompt-response pairs, a gradient preparation effect was observed: compared to completely heterogeneous sets, sets that shared both segmental and syllabic structure (e.g., *spui.en*, *spui.de*, *spui.er*, *spui.end* – all related to the stem *spui* 'drain') yielded faster response latencies than sets that only shared segmental structure (*spoe.len*, *spoel*.de, *spoe.ler*, *spoe.lend* – all related to the stem *spoel* 'rinse'); note the boldface consonant in a coda rather than an onset position. This suggests that the syllable is an abstract phonological encoding unit, making it hard to interpret our finding of worse performance in the production of syllable-copy reduplication compared to string-copy reduplication.

Finally, in both speech perception and speech production, syllabic effects – when they arise – can be subject to language-specific differences: they are sometimes observed in languages with clear syllable boundaries and not observed (or to a lesser extent) in languages with relatively unclear syllable boundaries. All in all, though, the diverging results have led to the conclusion that syllables are coding units at some, but not all, processing stages in both modalities (Cholin, 2011; Dumay & Content, 2012). Unfortunately, this makes it difficult to draw conclusions about the mechanisms that underlie our results. Rather, it calls for future research to investigate the extent to which our results persist across a wider variety of tasks, stimuli, and languages.

Another question concerns the directionality of the observed learning bias relative to baseline. Throughout our article, we have generally referred to a bias “against syllable-copy reduplication”. In the context of our experiments, this is functionally equivalent to a bias “in favor of string-copy reduplication”, since we only compared the learnability of the two types of reduplication to each other. However, in principle, the typological pattern could be driven by an exclusive learning bias against syllable-copy reduplication, an exclusive bias in favor of string-copy reduplication, or some combination of the two. Although all of these possibilities could explain a tendency for string-copy reduplication to be more frequent and relatively more easily learnable than syllable-copy reduplication, they have subtly different implications, particularly with regard to learnability relative to the baseline. An exclusively positive bias would predict that string-copy patterns are more learnable than a baseline comparison, whereas an exclusively negative bias would predict that syllable-copy patterns are less learnable than a baseline comparison; if both biases are present, we would predict a complex, possibly gradient, relationship between learnability and reduplication pattern (Herce et al., 2023). The analyses comparing performance in perception and production hint at the possibility that in our experiments, the bias is driven by a facilitation of string-copy reduplication: the better performance in production was due to enhanced performance in the string-copy condition (see 8.1 for details). However, our experiment did not include any baseline comparison. Future experiments may directly investigate the extent to which performance in production tasks is degraded by syllable-copy reduplication versus enhanced by string-copy reduplication.

Finally, there remains the issue of the apparent tendency for participants to favor CVC reduplication over CV reduplication in the perception experiments. Although the effect size for this tendency is small, and it could be explained away as a statistical artifact (it was identified via an exploratory, composite analysis), it is worthy of further scrutiny. Assuming that this effect is real and reproducible, there is, to our knowledge, no existing reduplicative or phonological theory that could satisfactorily explain this pattern. This is particularly noteworthy, given that it complements the production results: the bias against syllable-copy reduplication was exclusively observed in the production experiments, whereas the bias in favor of CVC reduplication was exclusively observed when taking a composite of the perception experiments. In principle, one could argue that CVC melodies may be more salient as a relatively “maximal” syllable, in line with the finding that it is more easily learnable; however, a comparably plausible counterargument is that CV melodies are simpler and should, therefore, be easier to learn, at odds with our finding. Speculatively, the explanation might not be due to any general reduplicative or phonological bias, but rather to our sample of native Italian speakers. Although Italian has no productive reduplication process, it does have some instances of low-productivity "discourse-level" reduplication. For example, some adjectives may be emphasized or intensified by full reduplication (Thornton, 2023), and these modified forms may be truncated to a CVC-string. The complete, or perhaps near-complete, list is shown in (7).

(7)	a.	pian piano < piano piano	'slow – little-by-little'
	b.	man mano < mano mano	'hand – gradually'
	c.	bel bello < bello bello	'nice – nice and pretty'
	d.	ben bene < bene bene	'well – very well'

The existence of the above pattern may explain the slight tendency for L1 Italian speakers to perform better in the CVC than in the CV string-copy subcondition – although this alone could not explain why the pattern is selective to the perception experiments. Future studies may investigate instances of language-specific sensitivity to reduplication in AGL tasks performed by users of languages with limited, non-productive reduplication.[8]

To conclude, our findings corroborate the existence of a cognitive learning bias as the underlying source of the typological rarity of syllable-copy reduplication. However, in contrast to previous AGL studies of reduplication (Haugen et al., 2022; Wang, 2024), our experiments identified a modality-specific effect: the learning bias is selectively driven by production, rather than perception. More broadly, our results suggest that future experimental research on reduplication – and, most generally, any research that seeks to implement AGL paradigms to study cognitive biases in language – should seek to disentangle the modality-specific effects that may underpin learning.

[8] French, for instance, has semi-productive reduplication in the realm of affectionate or childish language, as well as slang, with the reduplicant being a CV-string (e.g., *Bébelle* < *Isabelle*, *Bébert* < *Robert*, *faire dodo* < *dormir* 'to sleep', *fifille* < *fille* 'girl' *cracra* < *crasseux* 'filthy', *coco* < *cocaine* or *communiste*. In an AGL task like the present one, we would, thus, expect French participants to perform better on CV string-copy reduplication than on CVC string-copy reduplication.

Appendix A Participants who mentioned, among the languages they spoke or had studied, a language whose morphologies contain productive affixal reduplication

Experiment 1

N = 1: Ancient Greek and Latin

Experiment 2

N = 1: Ancient Greek and Latin

N = 1: Hebrew

N = 2: Korean

N = 2: Latin

Experiment 3

N = 1: Hebrew

N = 1: Korean

N = 1: Latin

N = 1: Turkish

Experiment 4

N = 1: Hebrew

N = 1: Korean

Experiment 5

N = 1: Latin

N = 3: Korean

N = 1: Persian

Experiment 6

N = 1: Hebrew

N = 1: Turkish

Appendix B Descriptive statistics

Experiment 1

Condition	Mean accuracy[9]	Standard error[10]
string-copy	87.86%	3.61
CV	*90.45%*	*4.64*
CVC	*86.18%*	*5.21*
syllable-copy (CVX)	71.82%	5.28

Experiment 2

Condition	Mean accuracy	Standard error
string-copy	65.80%	4.26
CV	*61.00%*	*5.31*
CVC	*70.60%*	*6.62*
syllable-copy (CVX)	66.10%	2.87

Experiment 3

Condition	Mean accuracy	Standard error
string-copy	66.56%	3.83
CV	*60.60%*	*5.87*
CVC	*73.04%*	*4.57*
syllable-copy (CVX)	64.47%	2.28
Condition	**Mean accuracy**	**Standard error**
stress condition A	65.21%	3.11
stress condition B	65.85%	3.24

[9] computed as 100*proportion accuracy.
[10] computed as 100*se_proportion_accuracy.

Experiment 4

Condition	**Mean accuracy**	**Standard error**
string-copy	87.20%	2.74
CV	*87.00%*	*4.65*
CVC	*87.40%*	*3.00*
syllable-copy (CVX)	67.80%	2.43
Condition	**Mean accuracy**	**Standard error**
stress condition A	74.40%	3.04
stress condition B	80.60%	2.76

Experiment 1 reanalysis

	“yes, always”/ “yes, only for plural”	
Condition	**Mean accuracy**	**Standard error**
string-copy	96.88%	2.18
CV	*100.00%*	*0*
CVC	*95.45%*	*3.12*
syllable-copy (CVX)	70.00%	10.09
	“no”/ “yes, sometimes”	
Condition	**Mean accuracy**	**Standard error**
string-copy	75.83%	6.57
CV	*82.50%*	*7.16*
CVC	*69.17%*	*10.99*
syllable-copy (CVX)	72.86%	6.26

Experiment 5

Condition	**Mean accuracy**	**Standard error**
string-copy	80.90%	2.85
CV	*75.60%*	*4.62*
CVC	*86.20%*	*3.07*
syllable-copy (CVX)	73.60%	2.93

Experiment 6

Condition	Mean accuracy	Standard error
string-copy	90.40%	2.26
CV	*93.60%*	*2.57*
CVC	*87.20%*	*3.66*
syllable-copy (CVX)	73.10%	2.51

Data accessibility statement

Experimental data, code, analyses, and stimuli are available at the OSF repository: https://osf.io/ny8zj.

Ethics and consent

Study procedures were approved by the Institutional Review Board of the Inserm 20-733 bis. All participants included in this study provided informed consent prior to the experiment.

Acknowledgements

We would like to thank Adam Ussishkin for stimulating discussion during the early stages of this project and Manon Dhuicque for research assistance. We also thank Alexander Martin, as well as the audiences at the Laboratoire de Phonétique et Phonologie in Paris and the 23rd Old-World Conference in Phonology held in Cambridge.

Competing interests

The authors have no competing interests to declare.

Authors' contributions

D.S.: Conceptualization, Resources, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing (original draft), and Writing (review & editing). S.P.: Conceptualization, Methodology, Funding acquisition, Project administration, Supervision, Writing (original draft), Writing (review & editing).

ORCiD IDs

David Schwitzgebel: https://orcid.org/0000-0001-9323-767X

Sharon Peperkamp: https://orcid.org/0000-0001-5985-8878

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blevins, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Blackwell. https://doi.org/10.1111/b.9780631201267.1996.00008.x

Blevins, J. (2008). Natural and unnatural sound patterns: A pocket field guide. In K. Willems & L. De Cuypere (Eds.), *Naturalness and iconicity in language* (pp. 121–148). John Benjamins. https://doi.org/10.1075/ill.7.08ble

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences, 7*(1), 19–22. https://doi.org/10.1016/S1364-6613(02)00005-0

Cholin, J. (2011). Do syllables exist? Psycholinguistic evidence for the retrieval of syllabic units. In C. E. Cairns & E. Raimy (Eds.), *Handbook of the syllable*. Brill. https://doi.org/10.1163/ej.9789004187405.i-464.69

Chomsky, N. (1965). *Aspects of the theory of syntax* (50th ed.). The MIT Press. https://doi.org/10.21236/AD0616323

Chomsky, N. (1980). *Rules and representations*. Columbia University Press. https://doi.org/10.1017/S0140525X00001515

Costa, A., & Sebastián-Gallés, N. (1998). Abstract phonological structure in language production: Evidence from Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 886–903. https://doi.org/10.1037//0278-7393.24.4.886

Côté, M.-H. (in press). The syllable: Epicenter or epiphenomenon? In A. Jardine & P. de Lacy (Eds.), *The Cambridge handbook of phonology* (2nd ed.). Cambridge University Press.

Culbertson, J. (2012). Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass, 6*(5), 310–329. https://doi.org/10.1002/lnc3.338

Culbertson, J. (2023). Artificial language learning. In J. Sprouse (Ed.), *The Oxford handbook of experimental syntax* (pp. 271–300). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198797722.013.9

Culbertson, J., & Kirby, S. (2016). Simplicity and specificity in language: Domain-general biases have domain-specific effects. *Frontiers in Psychology, 6*, 1964. https://doi.org/10.3389/fpsyg.2015.01964

Do, Y., & Havenhill, J. (2021). Production and substantive bias in phonological learning. In R. Bennett, R. Bibbs, M. L. Brinkerhoff, M. J. Kaplan, S. Rich, N. Van Handel, & M. W. Cavallaro (Eds.), *Proceedings of the 2020 Annual Meeting on Phonology*. Linguistic Society of America. https://doi.org/10.3765/amp.v9i0.4925

Dumay, N., & Content, A. (2012). Searching for syllabic coding units in speech perception. *Journal of Memory and Language, 66*, 680–694. https://doi.org/10.1016/j.jml.2012.03.001

Fabb, N., & Halle, M. (2008). *Meter in poetry: A new theory*. Cambridge University Press. https://doi.org/10.1017/CBO9780511755040

Ferrand, L., & Segui, J. (1998). The syllable's role in speech production: Are syllables chunks, schemas, or both? *Psychonomic Bulletin & Review, 5*, 253–258. https://doi.org/10.3758/BF03212947

Goslin, J., Galluzzi, C., & Romani, C. (2013). PhonItalia: A phonological lexicon for Italian. *Behavior Research Methods, 46*. https://doi.org/10.3758/s13428-013-0400-8

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Harrison, E., & Pius, R. (2021). *R for health data science* (1st ed.). Chapman & Hall/CRC.

Haugen, J. D. (2003). Allomorphy in Yaqui reduplication. In L. M. Barragan & J. D. Haugen (Eds.), *Studies in Uto-Aztecan (MIT Working Papers on Endangered and Less Familiar Languages* No. 5, pp. 75–103). MIT Working Papers in Linguistics.

Haugen, J. D. (2014). Hiaki (Yaqui) reduplication redux: Morphological and prosodic targets in syllabic reduplication. *International Journal of American Linguistics, 80*(4), 507–532. https://doi.org/10.1086/677913

Haugen, J. D., & Kennard, C. (2011). Base-dependence in reduplication. *Morphology, 21*, 1–29. https://doi.org/10.1007/s11525-010-9154-5

Haugen, J. D., Ussishkin, A., & Dawson, C. R. (2022). Learning a typologically unusual reduplication pattern: An artificial language learning study of *base-dependent reduplication*. *Morphology, 32*, 299–315. https://doi.org/10.1007/s11525-022-09396-y

Healey, P. (1960). An Agta grammar. *The Journal of Asian Studies, 21*(4), 569–570.

Herce, B., Saldana, C., Mansfield, J., & Bickel, B. (2023). Naturalness is gradient in morphological paradigms: Evidence from positional splits. *Glossa, 8*(1). https://doi.org/10.16995/glossa.9280

Jackendoff, R. (2000). Foundations of language. Oxford University Press.

Kim, H. S. (2020). A new look at onset transfer in Indo-European reduplication: Dissimilation of consonant clusters. *Language Research, 56*(1), 1–27. https://doi.org/10.30961/lr.2020.56.1.1

Kolinsky, R. (1998). Spoken word recognition: A stage-processing approach to language differences. *European Journal of Cognitive Psychology, 10*, 1–40. https://doi.org/10.1080/713752266

Lambert, D. J., Rawski, J., & Heinz, J. (2021). Typology emerges from simplicity in representations and learning. *Journal of Language Modelling, 9*(1), 151–194. https://doi.org/10.15398/jlm.v9i1.262

Lenth, R. V., Banfai, B., Bolker, B., Buerkner, P., Giné-Vázquez, I., Hervé, M., Jung, M., Love, J., Miguez, F., Piaskowski, J., Riebl, H., & Singmann, H. (2025). *emmeans*: Estimated marginal means, aka least-squares means (Version 1.11.1) [R package]. https://CRAN.R-project.org/package=emmeans

Levy, R. (2014). *Using R formulae to test for main effects in the presence of higher-order interactions*. arXiv: Methodology.

Marantz, A. (1982). Re-reduplication. *Linguistic Inquiry, 13*, 435–482.

Mattys, S., & Melhorn, J. (2005). How do syllables contribute to the perception of spoken English? Evidence from the migration paradigm. *Language and Speech, 48*, 223–253. https://doi.org/10.1177/00238309050480020501

McCarthy, J. J., & Prince, A. S. (1995). Faithfulness and reduplicative identity. In J. Beckman & S. Urbanczyk (Eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers 18, pp. 249–384). Graduate Linguistic Student Association.

Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20*, 298–305. https://doi.org/10.1016/S0022-5371(81)90450-3

Moravcsik, E. A. (1978). Reduplicative constructions. In J. H. Greenberg (Ed.), *Universals of human language*: Vol. 3. Word structure (pp. 297–334). Stanford University Press.

Moreton, E. (2008). Analytic bias and phonological typology. *Phonology, 25*(1), 83–127. https://doi.org/10.1017/S0952675708001413

Moreton, E., & Pater, J. (2012a). Structure and substance in artificial phonology learning, Part I: Structure. *Language and Linguistics Compass, 6*(11), 686–701. https://doi.org/10.1002/lnc3.363

Moreton, E., & Pater, J. (2012b). Structure and substance in artificial phonology learning, Part II: Substance. *Language and Linguistics Compass, 6*(11), 702–718. https://doi.org/10.1002/lnc3.366

Ohala, J. (2008). The emergent syllable. In B. Davis & K. Zajdo (Eds.), *The syllable in speech production* (pp. 179–186). Lawrence Erlbaum.

Peperkamp, S., Skoruppa, K., & Dupoux, E. (2006). The role of phonetic naturalness in phonological rule acquisition. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of BUCLD 30* (pp. 464–475). Cascadilla Press.

Piñeros, C. (1998). Prosodic morphology in Spanish: Constraint interaction in word formation [Doctoral dissertation, Ohio State University].

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as "asymmetric sampling in time." *Speech Communication, 41*, 245–255. https://doi.org/10.1016/S0167-6393(02)00107-3

Powell, M. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives* (Report No. DAMTP 2009/NA06). Centre for Mathematical Sciences, University of Cambridge.

Raimy, E. (2000). *The phonology and morphology of reduplication* (Studies in Generative Grammar 52). Mouton de Gruyter. https://doi.org/10.1515/9783110825831

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192. https://doi.org/10.1037/0033-2909.118.2.183

Schiller, N., & Costa, A. (2006). Activation of segments, not syllables, during phonological encoding in speech production. *The Mental Lexicon, 1*(2), 231–250. https://doi.org/10.1075/ml.1.2.04sch

Sevald, C., Dell, G., & Cole, J. (1995). Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language, 34*, 807–820. https://doi.org/10.1006/jmla.1995.1035

Strickland, B. (2017). Language reflects "core" cognition: A new theory about the origin of cross-linguistic regularities. *Cognitive Science, 41*, 70–101. https://doi.org/10.1111/cogs.12332

Tabossi, P., Collina, S., Mazzetti, M., & Zoppello, M. (2000). Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 758–775. https://doi.org/10.1037//0096-1523.26.2.758

Thornton, A. M. (2023). Repetition and reduplication in Italian. In J. P. Williams (Ed.), *Expressivity in European languages* (pp. 231–268). Cambridge University Press. https://doi.org/10.1017/9781108989084.011

Vasishth, S., Mertzen, D., Jäger, L. A., & Gelman, A. (2018). The statistical significance filter leads to overoptimistic expectations of replicability. *Journal of Memory and Language, 103*, 151–175. https://doi.org/10.1016/j.jml.2018.07.004

Vaux, B. (2011). Language games. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), *The handbook of phonological theory* (2nd ed., pp. 722–750). Wiley-Blackwell. https://doi.org/10.1002/9781444343069.ch22

Wang, Y. (2024). *Studies in morphophonological copying* [Doctoral dissertation, UCLA].

Zukoff, S. (2017). The reduplicative system of ancient Greek and a new analysis of Attic reduplication. *Linguistic Inquiry, 48*(3), 459–497. https://doi.org/10.1162/ling_a_00250