



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The emergence of word-internal repetition through iterated learning

Citation for published version:

Ota, M, San José, A & Smith, K 2021, 'The emergence of word-internal repetition through iterated learning: Explaining the mismatch between learning biases and language design', *Cognition*, vol. 210, 104585. <https://doi.org/10.1016/j.cognition.2021.104585>

Digital Object Identifier (DOI):

[10.1016/j.cognition.2021.104585](https://doi.org/10.1016/j.cognition.2021.104585)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Cognition

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The emergence of word-internal repetition through iterated learning

Citation for published version:

Ota, M, San Jose, A & Smith, K 2020, 'The emergence of word-internal repetition through iterated learning: Explaining the mismatch between learning biases and language design', *Cognition*.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Cognition

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The emergence of word-internal repetition through iterated learning: Explaining the mismatch

between learning biases and language design


Mitsuhiko Ota^a, Aitor San Jose^{b,c}, and Kenny Smith^a


^a University of Edinburgh


^b Max Planck Institute for Psycholinguistics

^c International Max Planck Research School for Language Sciences

Author Note

Mitsuhiko Ota  <https://orcid.org/0000-0003-3693-8490>

Aitor San Jose  <https://orcid.org/0000-0002-3404-1420>

Kenny Smith  <https://orcid.org/0000-0002-4530-6914>

We have no conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Mitsuhiko Ota, 3 Charles

Street, Dugald Stewart Building, Edinburgh EH8 9AD, United Kingdom. Email:

mits@ling.ed.ac.uk

Abstract

The idea that natural language is shaped by biases in learning plays a key role in our understanding of how human language is structured, but its corollary that there should be a correspondence between typological generalisations and ease of acquisition is not always supported. For example, natural languages tend to avoid close repetitions of consonants within a word, but developmental evidence suggests that, if anything, words containing sound repetitions are more, not less, likely to be acquired than those without. In this study, we use word-internal repetition as a test case to provide a cultural evolutionary explanation of when and how learning biases impact on language design. Two artificial language experiments showed that adult speakers possess a bias for both consonant and vowel repetitions when learning novel words, but the effects of this bias were observable in language transmission only when there was a relatively high learning pressure on the lexicon. Based on these results, we argue that whether the design of a language reflects biases in learning depends on the relative strength of pressures from learnability and communication efficiency exerted on the linguistic system during cultural transmission.

Keywords: learnability; language typology; cultural transmission; iterated learning; sound repetition

The emergence of word-internal repetition through iterated learning: explaining the mismatch between learning biases and language design

1. Introduction

1.1 Learning biases and linguistic generalisations

One of the main theses underpinning our understanding of fundamental properties of human language is that languages are shaped by their learners: that is, linguistic features which are easy to learn are common crosslinguistically, and crosslinguistically rare features are uncommon because they are hard to learn. This idea has been invoked by many authors of a wide range of theoretical persuasions to account for crosslinguistic generalisations about multiple linguistic domains, leading to a variety of proposed learning biases, including constraints on human perception, attention and memory (Creel, Newport, & Aslin, 2004; Pacton & Perruchet, 2008), constraints on the biomechanics of speech production (MacNeilage & Davis, 2000), general tendencies toward attending to or remembering certain regularities in the input (Endress, Nespors, & Mehler, 2009; Moreton, 2012), and biases towards structural simplicity or regularity (e.g., Chomsky & Halle, 1968; Culbertson & Kirby, 2016; Culbertson, Smolensky, & Legendre, 2012; Feldman, 2003; Smith & Wonnacott, 2010). The key idea motivating all these accounts is that such biases in learning, which make some systems easier or harder to learn, may act to restrict the space of possible languages. Languages that contain a pattern that is dispreferred in learning are either unlearnable or more likely to be learned inaccurately and are therefore more likely to change as language is passed from generation to generation via learning. As a result, all these accounts predict a close match between biases in learning and the types of languages we see in the world, with harder-to-learn structures being relatively rare crosslinguistically.

If language is shaped by learning in this way, then we might also expect that linguistic forms and structures that are more common in human languages should be more readily learned than those that are rare. This prediction is indeed consistent with many observations from research on language learning and development. For instance, linguistic patterns that are found widely across languages (e.g., ‘CV’ syllables consisting of a consonant and a vowel) tend to emerge in children’s early production before patterns that are crosslinguistically more restricted (e.g., CCV or CVC syllables) (Jakobson, 1968; Levelt, Schiller, Levelt, 2000). Typologically common linguistic patterns are also preferentially learned in language experiments (Moreton & Pater, 2012a; Wilson, 2006; see Culbertson, 2012 for a review). For example, natural languages tend to place numerals and adjectives on the same side of the modified noun (Greenberg, 1963), and artificial language learning experiments show that both 6- to 7-year-old children and adults exhibit a bias toward systems that consistently order modifiers before or after nouns in this way (Culbertson & Newport, 2015; Culbertson, Smolensky, & Legendre 2012). Experiments with infants and adults also demonstrate learning biases that align with the observation that segmental inventories in natural languages tend to be defined by as few phonological features as necessary (Clements, 2003), which, in turn, may be a manifestation of a more general bias in favour of simpler category systems (e.g., Feldman, 2003; Moreton & Pater, 2012a). For instance, 9-month-olds learn a class of segments defined by one feature (e.g., voice) more readily than a class defined by two features (e.g., voice and place of articulation) (Saffran & Thiessen, 2003). Similarly, adult speakers are better at learning a novel class of segments defined by two features than a class defined by three features (Kuo, 2009; Skoruppa & Peperkamp, 2011).

It is important to note here that evidence of learning biases that align with crosslinguistic generalisations is observed not only in prelinguistic infants and young children, but also in older children and adults. This suggests that some of these learning patterns are not merely a product

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

of immature cognitive systems, but rather a reflection of a maturation-independent predisposition in human perception or cognition. Such learning biases not only constrain early language development but may also continue to exert their effects on the use and transmission of language into adulthood.

However, learning biases and crosslinguistic typological asymmetries are not always congruent. A well-known example of a mismatch is the case of grammatical gender. Classification of nouns into different grammatical genders (or noun classes) can be correlated with other properties of the noun, such as its phonological shape (e.g., the final segment) and semantics (e.g., male/female, animate/inanimate) (Corbett, 1991; Comrie, 1999). Of these, the primary correlate of gender from a typological point of view is semantics, since all known gender assignment systems show some reference to semantic cues, but not necessarily to phonological ones (Aikhenvald, 2000). The prediction is, therefore, that learners should rely more on semantic cues to classify nouns. Yet studies have repeatedly shown that children are more responsive to phonological than to semantic cues in learning a gender system (Culbertson, Gagliardi & Smith, 2017; Culbertson, Jarvinen, Haggarty & Smith, 2019; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Mills, 1986; Müller, 2000; Pérez-Pereira, 1991). Most explanations for this puzzling misalignment between a learning bias and crosslinguistic asymmetry appeal to the differences between phonological and semantic cues in the context of learning. For instance, phonological cues may be privileged over semantic cues because phonological patterns are inherently more salient than semantic information (Gagliardi, Feldman, & Lidz, 2016) or because the phonological properties of words are learned before their meanings and are therefore likely to form earlier associations with other information such as grammatical gender (Culbertson et al., 2017, 2019; Gagliardi & Lidz, 2014). Although the precise mechanism behind this learning bias is still a matter of debate, the case of grammatical gender shows that crosslinguistic generalisations may not map directly on to biases in learning

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

when different linguistic domains (e.g., phonological and semantic) are involved in the characterisation of the linguistic pattern.

Another dimension of mismatch that has been discussed in the literature relates to the acquisition of phonological rules (Glewwe, 2019; Hayes & White, 2013; Moreton, 2008; Moreton & Pater, 2012a, 2012b). Phonological rules in natural languages exhibit two types of typological asymmetry. First, as mentioned above, languages favour phonological alternations and relations that can be defined by a simple rule (e.g., involving fewer features) over ones that require a complex rule (e.g., involving more features). Second, phonetically-natural rules (e.g., palatalization before front vowels) are more common than phonetically-unnatural ones (e.g., palatalization before back vowels). Laboratory experiments investigating learning biases for the former type of typological asymmetry have found substantial evidence that simple phonological rules are learned more easily than complex ones by infants (Saffran & Thiessen, 2003; Cristia & Seidl, 2008; Onishi, Chamber & Fisher, 2002) and adults (Pycha, Nowak, Shin, & Shosted, 2003; Skoruppa & Peperkamp, 2011). However, similar attempts at demonstrating a learning advantage for phonetically-natural rules (a notion known as substantive bias) have produced mixed or null results (e.g., Glewwe, 2019; Pycha, Nowak, Shin, & Asted, 2003; Skoruppa & Peperkamp, 2011; Wilson, 2006, although see Martin & White, 2019). For example, the prevalence of vowel harmony in human languages, as opposed to vowel disharmony, may be attributable to the phonetic naturalness of vowel harmony, which can develop from vowel-to-vowel coarticulation. Yet, adult speakers can learn vowel disharmony in an artificial language as easily as vowel harmony (Pycha et al., 2013; Skoruppa & Peperkamp, 2011). These observations suggest that there may be only certain types of substantive biases (e.g., perceptual naturalness but not articulatory naturalness) that give rise to typological asymmetries (Glewwe, 2019). Alternatively, typological asymmetries that are related to phonetic factors may not originate in biases in learning, but rather in systematic

misapprehension that reflects constraints on speech production and perception (i.e., channel biases, e.g., Moreton, 2008; Ohala 1993).

1.2 Word-internal sound repetition

In this study, we turn our attention to a case of potential mismatch between a learning bias and a typological asymmetry that is purely phonological in nature, thus abstracting away from the complexity of learning-typology relationships that involve more than one linguistic domain. It is also a pattern for which there is good evidence for a learning bias – only one that is inconsistent with the typological observations. The case in question relates to proximate repetition of similar or identical phonological units within a lexical item. Most languages impose restrictions on word-internal co-occurrences of similar elements (Berkley, 1994; McCarthy, 1986; Pozdniakov & Segerer, 2007; Suzuki, 1998).¹ For example, the first two consonants in Arabic and Hebrew verb roots cannot be identical or homorganic in place of articulation (e.g., *m_m..., *b_m..., *g_k...) (Greenberg, 1950). In English, place-sharing consonants can co-occur in proximity (e.g., *bib*, *mop*, *pop*), but such sequences are statistically less frequent than would be expected (Berkley, 1994; see also Monaghan & Zuidema, 2015 for confirmation of the same pattern in Dutch, French and German, and Pozdniakov & Segerer, 2007 in languages across 15 families and isolates). This avoidance of word-internal repetition is subject to effects of similarity and proximity; stronger co-occurrence avoidance between two sounds is observed when they are more similar in type and closer in distance. The effect is therefore most evident between identical units that are adjacent to each other, where adjacency is defined in terms of ‘tiers’, or separate sequences of sounds of the same type (e.g., a consonant tends not to be identical to the next consonant in a word, regardless of any intervening vowels) (McCarthy, 1986). In theoretical phonology, this pattern is captured under the Obligatory

¹ This generalisation relates to phonological elements that make up the root of a word and does not necessarily apply to morphologically-introduced elements, where reduplication is a common process (Inkelas & Zoll, 2005). However, avoidance of certain phonological repetition is seen even in the context of morphological reduplication (Wedel, 1999; Yip, 1998).

Contour Principle (OCP), which formally prohibits consecutive identical phonological features in the underlying representations (Leben, 1973; Goldsmith, 1976; McCarthy, 1986). Avoidance of word-internal repetition also targets certain elements more often than others. Thus, we find widespread tendencies for languages to avoid repetition of consonants and tones, but fewer cases of repetition avoidance of vowels (Suzuki, 1998; Gordon, 2016). In fact, many languages exhibit the opposite effect for vowels, preferring proximate vowels to be similar in major features such as height, backness and roundedness (i.e., ‘vowel harmony’). In contrast, a preference for consonants in a word root to share major place features, such as labial and dorsal, is hardly attested (Rose & Walker, 2011), and the opposite pattern – *avoidance* of shared major place features between consonants in a word root – is typologically widespread and hypothesised to be a universal property (Pozdniakov & Segerer, 2007).

Avoidance of proximate consonant repetitions within a word is also reflected in adult speakers’ intuitions. When asked to choose a label for a novel object, adult English speakers prefer one that has no consonant repetitions (e.g., *slafmak*) over one with repetitions (e.g., *slaflaf*) (Berent, Bat-El, Brentari, Dupuis, & Vaknin-Nusbaum, 2016). Dutch adult speakers prefer to segment sequences in which place-agreeing consonants are separated by another consonant (e.g., /**bo**depo/) over sequences in which those consonants are in tier-adjacent positions (e.g., /de**bo**pe/) (Boll-Avetisyan & Kager, 2014).

In stark contrast to these linguistic and psycholinguistic observations, there is evidence that proximate sound repetitions are preferred, not avoided, in the context of word segmentation and word learning by young learners. For example, when 9-month-olds are exposed to passages containing novel words with or without repetitions of consonant-vowel (CV) sequences (e.g., *neenee, foofoo* vs. *neefoo, foonee*), they later recognise the former better than the latter. This suggests that infants are predisposed to process phonological strings containing sound repetitions as a unit (Ota & Skarabela, 2018). Furthermore, 18-month-olds are better at learning

associations between a novel word and an unfamiliar object when the novel word contains CV repetitions than when it does not (Ota & Skarabela, 2016). These results are consistent with other developmental findings demonstrating the role of sound repetition in early speech processing. For instance, neonates exhibit higher neural activation in response to input with adjacent CV repetitions (e.g., *mubaba*) in comparison to input without any repetitions (e.g., *mubage*) or with non-adjacent repetitions (e.g., *mubamu*) (Gervain, Macagno, Cogoi, Peña, & Mehler, 2008, Gervain, Berent, & Werker, 2012). Six-month-olds are better at discriminating trisyllabic stimuli when they share repeated syllables (cf. *bakoko/dukoko* vs. *batiko/dutiko*; Goodsitt, Morse, ver Hoeve, & Cowan, 1984). There is therefore a range of evidence suggesting that strings that contain repeated elements are more likely to be noticed and retained by human learners (Endress, Nespors, & Mehler, 2009).²

Another observation that demonstrates a preference for sound repetition during early development is the tendency for children to repeat sound structures in word production (e.g., *water* [wɔwɔ], *scissors* [dɪdɪ], *blanket* [baba]; Schwartz, Leonard, Wilcox, & Folger, 1980; Fee & Ingram, 1982; Menn, 1971) Strikingly, by far the most common pattern of such sound repetition involves the assimilation of consonants in major place of articulation (e.g., *doggy* [gɔgi], *nipple* [mɪpəl], *kitchen* [gɪgən]; Berg, 1992; Fikkert & Levelt, 2008; Vihman, 1978; Pater & Werle, 2003; Smith, 1973). As discussed earlier, this is descriptively the very same pattern that is avoided in adult language. Indeed, there is a long-standing question in the literature on whether this ‘consonant harmony’ in children’s word production should be interpreted within the same phonological frameworks proposed for adult language or treated as a distinct phenomenon (Hansson, 2010; Vihman, 1980; see also references immediately above). We return to this point below.

² Here we are discussing evidence that sound repetitions improve detection/retention of words that contain them. This should not be confused with evidence that it is easier to learn a phonological *rule* that requires sounds in a word to be similar or identical compared to one that requires sounds to be different, for which there is some disagreement in the current literature (see Moreton & Pater, 2012b, Martin & White, 2019, and Pycha et al., 2013).

1.3 Possible explanations

How can we explain this incongruence between learning and crosslinguistic generalisations, where word-internal repetition proffers advantages in the former at least in the context of early learning, but is avoided in the latter, specifically for consonants? One possibility is that the mismatch simply reflects different phonological units that have been investigated in these contexts. Most of the developmental findings showing an advantage for word-internal repetition involve consonant-vowel (CV) units (e.g., *neenee*). Because harmony between vowels is a common phenomenon in natural languages, it is possible that the learning bias for forms such as *neenee* is driven by infants' preference for vowel harmony. Indeed, a recent study by Mintz, Walker, Welday and Kidd (2018) showed that infants tend to treat strings sharing the same vowel (e.g., *tokobo*) as being contiguous. Under this account, there may be a bias against consonant repetition in early learning, consistent with the avoidance of consonant repetition in human languages, but the anti-repetition bias may be masked by a stronger preference for vowel repetition in experiments using stimuli that contain both consonant and vowel repetitions.

A second way to reconcile these contrasting findings is to consider the potential impact of developmental changes in the way linguistic input is processed. Recent developmental work on attention shows that there is a shift during infancy and childhood in the type of attentional mechanisms that are recruited in spatial and temporal cognitive processing (Colombo, 2002; Colombo & Cheatham, 2006; Richards, Reynolds, & Courage, 2010). Initially, infants' attention is primarily exogenous in nature, automatically oriented towards inherently salient properties of the stimuli (e.g., flashing light). However, towards the middle of the second year in life, another type of attention becomes available to children, which is endogenous and guided by internal expectations for the stimuli (e.g., regularity in the signals or symbolic relationships). This delayed development of endogenous attention can have important implications for

language development (de Diego-Balaguer, Martinez-Alvarez, & Pons, 2016; Martinez-Alvarez, Pons, & de Diego-Balaguer, 2017). Specifically, adjacent repetitions in the stimuli, which may be inherently salient, can attract the primarily exogenous attention of very young infants (de Diego-Balaguer et al., 2016; see also Endress et al., 2009). But this exogenously-driven attention is later offset by an increase in endogenous attention, which draws the learner's attention instead to the distributional and transitional properties of the input. Such properties in the input, if anything, should result in an expectation against sound repetitions, as the probability of sound units repeating themselves in the input to a learner is usually below chance-level (Ota & Skarabela, 2016; see also Gerken, Dawson, Chatila, & Tenenbaum, 2015). In short, the learning bias for adjacent sound repetition attested in infants and the avoidance of sound repetitions in lexicons do not constitute a contradiction if the learning bias is present only during early childhood.

More generally, there is some debate in the literature on whether learning biases in early childhood are relevant to explaining language change and crosslinguistic generalisations. One of the arguments against early language acquisition as the main source of language change (and ultimately, of crosslinguistic generalisations) is that the types of changes that young learners make to the adult system are not always the same as those found in language change (Bybee & Slobin, 1982, Croft, 2000, and Slobin, 2002). In fact, the tendency for children to produce place-assimilated forms (e.g., [gɔgi] for *doggy*) and the rarity of equivalent consonant harmony processes in adult language are often cited as a prime example of such a mismatch that is problematic for the view that language change is induced by early language acquisition (Vihman, 1980; Croft, 2000). According to this view, the main locus of language change is rather in proficient speakers who are also subject to biases/constraints on perception, learning, and use of language. For example, there is growing evidence that phonetic biases in perception and articulation can shift adults' long-term lexical representations, leading to systematic

historical changes in phonological patterns (e.g., Blevins, 2004; Bybee, 2001; Ohala, 1981; Sósokuthy, 2015; Wedel, 2006; see also references in Hall, Hume, Jaeger, & Wedel, 2018). From this perspective, the observed bias for sound repetitions in early acquisition has little relevance to the processes that mould the shape of natural languages.

These arguments that explain away the learning-typology mismatch by appealing to a bias in infancy which is short-lived or just not relevant to language change would of course be weakened if adults have similar repetition-favouring biases as young children. Linzen and Gallagher (2017) report a series of experiments examining the extent to which adult learners generalise phonotactic patterns based on a short exposure to a set of artificial language words, some of which speak to this issue. In one experiment they exposed their participants to eight novel CVCV words, half of which contained identical consonants (e.g., *pipa*) and half of which did not (e.g., *kesa*). The participants later judged unheard items with consonant repetitions (e.g., *kuka*) more likely to be part of that language than novel items without a repetition (e.g., *pina*). In another experiment, participants were exposed to a set of eight CVCV words, all of which had nonidentical consonants (e.g., *kupe*). This time, the participants did not show higher endorsements for words with consonant repetitions, indicating, as argued by Linzen and Gallagher (2017), that the result of the first experiment is not due to a pre-existing preference for words with identical consonants; that is, adults and young children do perhaps have different biases with respect to consonant repetition). Note, however, that while these outcomes show that adults' readiness to learn a phonotactic pattern with consonant repetition cannot be attributed to their inherent preference for consonant repetitions, they do not rule out the possibility that there is an underlying bias to learn words with consonant repetitions if the exposure items contained some instances of identical consonants.

A third possible explanation for the learning-typology mismatch — the account that we aim to test in this study — is that the inductive bias for repetition of all types of sounds remains

active throughout our lifespan but its effects are counteracted by a competing demand from language *use* that favours nonrepetition of linguistic forms. There are at least two reasons why language use may give rise to such a countervailing pressure against sound repetition. One is the impact that sound repetition can have on the phonological space in the lexicon, as constraints of phonological structure can determine the set of possible distinct word forms in the language (see Winter & Wedel, 2016). A bias for certain shapes in linguistic forms limits the signal space, and in the extreme, a strong preference for word-internal repetition means wherever possible, words must consist of repeated phonological structures, a radical reduction in the phonological space available for the lexicon. While a bias *against* repetition would also reduce the available phonological space, this reduction is far less dramatic. For instance, a strict ban on consonant repetition still permits each consonant to be followed by any of the other consonants in the language. In contrast, a strict requirement for consonant repetition means that only one type of consonant can follow the first one. A second potential source for anti-repetition effects in language may be found in perceptual and memory biases in serial processing (Boersma, 1998; Frisch, Pierrehumbert, & Broe, 2004; Leivada, 2017; Walter, 2007). Perceptually, immediately repeated items are subject to a deficit in detecting repetitions in visual stimuli (Kanwisher, 1987; Kanwisher, Driver, & Machado, 1995; Buffat, Plantier, Roumes, & Lorenceau, 2013) and auditory stimuli (Miller & MacKay, 1994, 1996; Soto-Faraco & Sebastián-Gallés, 2001), a phenomenon known as repetition blindness/deafness. A similar deficit for repetition is found in the memory domain (i.e., the Ranschburg effect), where short-term recall is inhibited for items repeated in a string (Jahnke, 1969; Henson, 1998). Both of these effects should have a negative impact not only on learning but also on language use because words containing sound repetition will always run a greater risk of being misperceived. As a result, repeated sounds in words may be subjected to a channel bias, or systematic errors during inter-speaker communication.

This last explanation is related to the idea that language systems are shaped not only by pressures for learnability, favouring simpler systems, but also by counteracting demands for an informative and efficient system for communication (e.g., Hall et al., 2018; Rosch, 1978; Kemp, Xu, & Regier, 2018, Zipf, 1935). For example, while languages differ in the number of categories for kinship, colours, spatial relations and numerals, all systems in these semantic domains still optimally balance the need to constrain the number and complexity of linguistic distinctions (i.e., simplicity) and the need to provide the maximum amount of information (i.e., informativeness) (e.g., Kemp & Regier, 2012; Regier, Kemp, & Kay, 2015). In relation to the phonological lexicon, we see similar interactions between pressures to reduce articulatory or perceptual complexity and pressures to accurately communicate lexical contrasts. Thus, words that are used the most frequently in a language tend to be phonologically simpler (Zipf, 1935; Mahowald, Dautriche, Gibson, & Piantadosi, 2018; Piantadosi, Tily, & Gibson, 2011) and sounds that occur in the most lexically predictable or the least informative contexts tend to undergo reduction or neutralisation (Cohen Priva, 2015, 2017; Seyfarth, 2014). The case of sound repetition can be understood under the same light, with simplicity favouring word forms with sound repetitions, and informativeness holding sound repetition at bay.

One way to explore the interaction of competing demands for simplicity and informativeness in the context of language transmission is to examine how an artificial language changes over the course of iterated transmission between language users. This type of empirical paradigm allows independent manipulation of pressures for learning, which should favour simplicity, and pressures for efficient communication, which should favour informativeness (e.g., Kirby, Tamariz, Cornish, and Smith, 2015; Carr, Smith, Culbertson, & Kirby, 2020). Kirby et al. (2015) demonstrate the trade-off between these two sources through a set of computational models and experiments simulating three types of cultural transmission. In their transmission-only condition, a linguistic system is learned and then reproduced, with the

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

reproduced system being passed on to the next group of learners who learn and reproduce it in turn, modelling a scenario in which the language is under substantial pressure to be learned. In their communication-only condition, pairs of individuals repeatedly communicate with each other using the linguistic system, modelling a scenario where the language is primarily under pressure for effective communication, with the learnability pressure associated with repeated transmission being removed. In the learning-and-transmission model, individuals use the language to communicate with each other (as in the communication-only condition), but the language produced during communication forms the input to subsequent groups of learners who reproduce it in the same way (i.e., the language is repeatedly learned, as in the transmission-only condition). Only the last scenario results in languages which are adapted to the constraints of both learning and use. In particular, they find that (as shown in for example Experiment 1 of Kirby, Cornish & Smith, 2008, or Cornish, 2010) in the transmission-only conditions, the artificial languages rapidly degenerate as they are passed from learner to learner. While the initial languages passed to the first generation of learners provide a distinct label for each event to be communicated, the languages lose distinctions at a high rate, collapsing down to a state where a small number of underspecified or ambiguous forms are used (minimally, one form). Such degenerate languages are extremely easy to learn but (irrelevantly, in the transmission-only condition) not useful for communication. In the communication-only condition, they find that the initial languages are more or less maintained: the languages tend not to degenerate, maintaining as many distinct forms as required for efficient communication, but also not developing any structural regularities that would facilitate learning (and which are not required in the communication-only condition). Finally, in the learning-and-communication condition they find that compositional languages emerge, where each label consists of component parts that convey the separate dimensions of the objects they refer to. Compositional languages,

unlike degenerate languages, allow unambiguous communication but contain regularities which facilitate their learning.

1.4 Purpose of the study

Here we investigate the role of competing pressures from learning and language use as an explanation of the misalignment between learning biases and language design with regards to word-internal repetition. In Experiment 1, we examine how mappings between novel words and unfamiliar objects are transmitted among adult participants in order to test two models of cultural transmission, one that features both learning and communicative interaction and one that features mostly communicative interaction with a limited role for learning.³ Following the terms used in Kirby et al. (2015), we call the experiment condition that iterates learning and communicative interaction ‘chains’ (because pairs of participants form a chain of transmission, where a language is learned and then passed onto the next pair) and the experimental condition with mostly communicative interaction ‘closed groups’ (because the same pair of participants learn and communicate with each other throughout the experiment). The key idea in this paradigm is that we can test whether cultural transmission of language results in the emergence of different patterns of within-word repetition when the pressure to learn is relatively high (chains) or low (closed groups) relative to the pressure to communicate.

The main goal of Experiment 1 was to test our hypothesis that a learning bias for repetition exists in all human learners including adults and that the mismatch between development and the lexicons of natural language is due to competing demands from learnability and effectiveness in communication. We thus expected the occurrence of within-

³ Note that we do not run a learning-only condition as in Kirby et al. (2008). Although all other things being equal we would expect the effects of sound repetition on learnability to be most salient in the absence of competing pressures from communication, any pressures favouring sound repetition are likely to be swamped by the drive to degeneracy which seems to be the overwhelming pressure in such experiments. In other words, we were not confident that biases in learning favouring sound repetition would be strong enough to still be apparent in lexicons that are reduced down to a single word or a handful of words, with the choice of the eventual ‘winning’ words being highly dependent on chance events occurring early in transmission.

word repetitions to be modulated by the relative importance of pressures from learning and communication. Specifically, we predicted that repetitions should increase more in the chains (both learning and communicative interactions) than in the closed groups (mostly communicative interactions). Our predictions can be contrasted with those that follow from the two alternative explanations of the learning-typology mismatch discussed in Section 1.3. If the explanation of the mismatch is that there is in fact a learning bias against consonant repetition that has been overlooked in previous research, then we expect to find the effects of this bias to emerge in response to an increased pressure to learn. Under this scenario, we predict word-internal repetitions to be more suppressed in the chains than in the closed groups. If the explanation of the mismatch is that learning biases for sound repetitions are relevant only to young children, our adult participants would be subject to pressures from communication but not to the learning biases. Therefore, we should predict a decline in sound repetitions in the closed groups, as sound repetitions are hypothesised to be detrimental to communication.

In testing the key predictions described above, we also considered the role of the lexicon size under pressure from learning and communication. If within-word repetitions are dispreferred in communication, we expect them to be suppressed when the number of lexical items in the lexicon is high and the communicative challenge for our participants is greatest. Alternatively, because a larger lexicon adds to the pressure for learning, we might expect sound repetition to *increase* as the size of the lexicon increases, if learnability is the dominant pressure shaping the lexicon. We tested the role of the lexicon by crossing the transmission conditions (chains versus closed groups) with two lexicon sizes (a 18-word ‘large’ lexicon versus a 12-word ‘small’ lexicon). However, as we could not predict which of the two possible effects described above would manifest itself during iterated learning, we left this as an open-ended research question without any directional predictions.

As we will show below, the results of Experiment 1 support the hypothesis that word-internal repetitions of both consonants and vowels are preferred, not avoided, in the context of learning by adult speakers, but the effect of this preference on the transmitted linguistic system is only visible when the system is under learning pressure (i.e., we see more repetition developing in chains than in closed groups). However, the results also indicate that consonant repetitions and vowel repetitions increase during language transmission via different mechanisms. Specifically, examination of the novel words reproduced by participants immediately after training reveals that accuracy of learning is improved by the presence of consonant repetitions in the word but not by the presence of vowel repetitions, whereas errors in learning lead to more vowel repetitions but not consonant repetitions. In Experiment 2 we verify those results using a non-iterated version of Experiment 1 in which the participants' task was simply to learn the novel words as labels for unfamiliar objects rather than use those learned word-object associations to communicate with other participants.

2. Experiment 1

2.1 Method

2.1.1 Overview. The method employed in the current experiment was modelled after the paradigm developed by Kirby et al. (2015) with some modifications. The task given to participants was to learn novel words assigned to unfamiliar objects and to engage in a communication task that uses those word-object mappings. The experiment had four conditions, which were generated by crossing two factors: lexicon size (large vs. small) and group type (chain vs. closed group). The lexicon size refers to the number of object-word associations that participants had to learn. In the large lexicon conditions, there were 18 object-word associations, and in the small lexicon conditions, there were 12 associations. These numbers were set based on the results of studies using similar paradigms (e.g., Kirby et al., 2015). The

group type refers to the way in which the participants were involved in learning and communicating the association pairs.

Both group types had five *generations* or *rounds* (see Figure 1, overleaf). In the *chain* condition, each generation consisted of a pair of participants who learned the same set of novel word-object mappings during the training phase. In the subsequent interaction phase, they engaged in a communication game in which they took turns in playing the *director*, who tried to name each object using the associated novel word, and *matcher*, who tried to select the object the director was naming. The words produced by each pair in the final block of interaction formed the training set for the next generation in the chain (i.e., a new pair of participants who underwent the same training and interaction phases). In this way, the word-object mappings of a given generation were passed on to the following generation until they reached the fifth generation. In the *closed group* condition, the same pair of participants had one training phase during which they learned the novel word-object associations and then five rounds of interaction, where they played the communication game repeatedly. Note that unlike in the method used by Kirby et al. (2015), participants in this condition were not retrained on their own past productions after each round of interaction.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

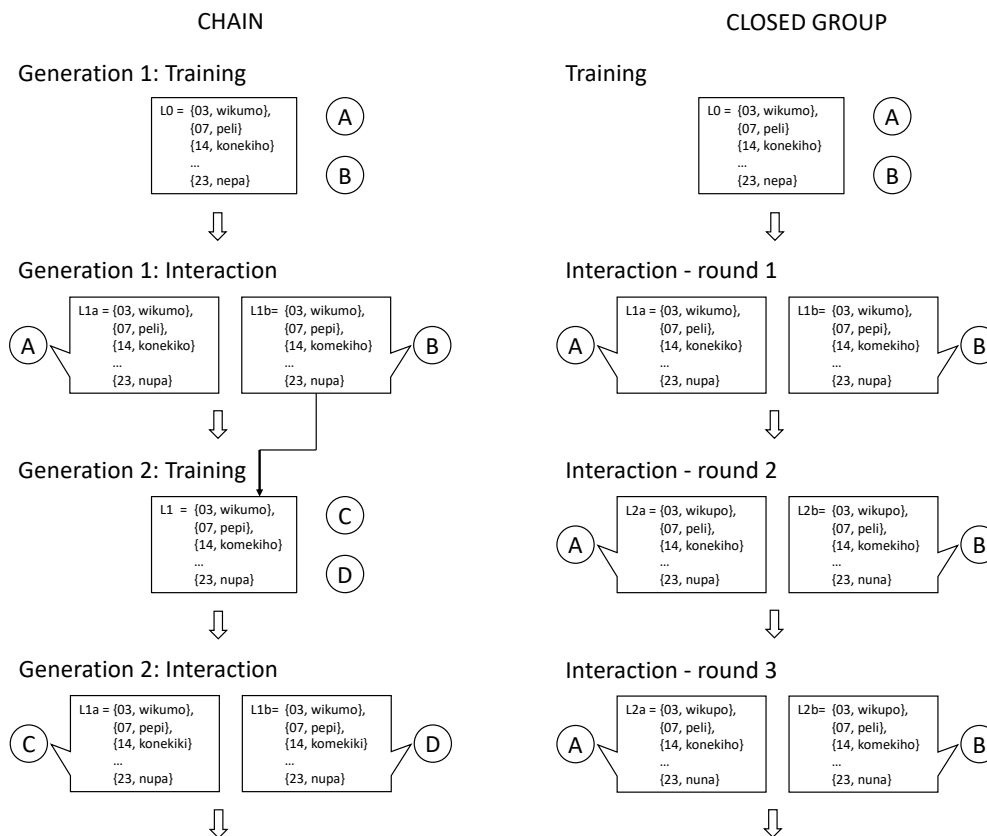


Figure 1. Diagram illustrating the initial flow of the experiment in the chain condition (left) and closed group condition (right). Circled letters (A, B, C, D) stand for unique participants. L_x indicates a language, which consists of pairs of associations between an unfamiliar object (here, indicated by an arbitrarily assigned item number) and a name of the object (e.g., {03, wikumo} indicates an association between object 03 and the label *wikumo*). In a chain, each pair of participants underwent training and interaction, and the outputs of the final interaction were learned by the next pair, with the language being passed down five generations of interacting pairs of participants. In a closed group, a single pair of participants underwent training once and then repeated the interaction phase for five rounds. In the large vocabulary condition, there were 18 word-object associations, and in the small vocabulary condition, there were 12 associations. Note that object names can differ between members of a pair and can change between generations/rounds due to mislearning and/or miscommunication.

2.1.2 Participants. The participants were 144 students from a British university. They were native speakers of English with normal or corrected-to-normal vision. An additional 18 participants took part in the experiment but were not included in the analysis because they failed to comply with the instructions (e.g., by using existing English words as labels of novel

objects).⁴ These individuals were replaced with new participants until the predetermined number of participants was reached. All participants received a financial compensation of £7/hour. As an incentive for participants to focus on the task, an additional £20 was given to the pair with the highest communicative accuracy score (£10 each). Informed consent was obtained from all participants.

The participants were randomly assigned to one of the four conditions: chain/small, chain/large, closed/small or closed/large. We ran six independent groups in each condition. In chains, each group required 10 participants (5 generations, 2 participants per generation) and therefore 60 participants per chain condition (i.e., chain/small and chain/large). In the closed group condition, each group consisted of a single participant pair, for a total of 12 participants per closed group condition (i.e., close/small and closed/large).

2.1.3 Materials. The materials consisted of photographs of unfamiliar objects and novel orthographic words as described below.

2.1.3.1 Unfamiliar objects. We used photographs of unfamiliar objects from the Novel Object and Unusual Name (NOUN) database (Horst and Hout, 2016). To avoid iconicity effects, we excluded objects that visually invoked the notion of repetition (e.g., an object with stripes in alternating colours). We also chose objects with low familiarity scores (i.e., low percentages of adults who indicated they had seen the object before) and low nameability scores (i.e., low percentages of adults who spontaneously came up with the same name for the object). For the large lexicon condition, 18 pictures were selected with a mean familiarity score of 19.8 (SD = 9.8) and a mean nameability score of 32.9 (SD = 10.6). The corresponding small-lexicon set was created by excluding the objects whose removal from the large lexicon would yield the

⁴ The instructions stated: “We are interested in how well people can learn a novel language, so please do not use words from any existing languages, such as English or Spanish. If we see you using words that at least RESEMBLE English or any existing language we can identify, we will be forced to ask you to leave, so please make sure you understand the rules before continuing.” When at least one member of a pair did not follow this rule, both members were excluded from the analysis.

smallest difference in both familiarity and nameability between the small and large set. The mean difference between the resulting picture sets was 1.0 for familiarity and 0.0 for nameability. Pictures of the 18 objects are shown in Appendix A.

2.1.3.2 Novel words. We used orthographic words instead of spoken words as labels for the unfamiliar objects. This allowed us to avoid technical issues associated with the use of auditory stimuli in this experimental paradigm. In order to use spoken words, productions by participants would have to be audio-recorded and edited or synthesised each time before being passed on to the next generation/round so that phonetic and acoustic variation (e.g., differences in speech rate and loudness) and potential extraneous cues for word-object associations (e.g., disfluency, noise) were removed. This would have made the administration of the experiment substantially more difficult. Our use of orthographic words was also justified by robust evidence that visual word recognition is mediated by access to phonological representations (Coltheart et al., 2001; Frost, 1998; Van Orden, Johnston, & Hale, 1988), even in a foreign language (Ota, Hartsuiker, & Haywood, 2009). Our assumption therefore was that fluent readers of English should process ‘consonant’ and ‘vowel’ letters as if they were consonant and vowel sounds in learning orthographic novel words.

The novel orthographic word forms that were initially assigned to the unfamiliar objects as labels consisted of two to four syllables, and were generated by combining a predetermined set of consonants (C) and vowels (V) into CV structures.⁵ To do this, we first created a set of ‘syllables’ that included all possible CV combinations between eight consonants <g, h, k, l, m, n, p, w> and five vowels <a, e, i, o, u> (e.g., *ge, ho, li*). From this, nine CV syllables were

⁵ The motivation for using 2 to 4-syllable words was two-fold. First, the variation in phonological structure prevented the novel words from being overly confusable. Memory of orthographic word lists is known to be negatively affected by phonological similarities between the words (Baddeley, 1968; Conrad & Hull, 1964), and with the restricted segmental inventories used in the stimuli, we were concerned that a list of novel words with a uniform syllable count may have been too challenging to learn. Second, it struck a balance between the competing demands for ease of learning (which favours shorter words) and opportunities for sound repetitions to occur (which favours longer words).

sampled randomly without replacement to generate a seed syllable set. The novel words were generated by concatenating syllables sampled randomly from this syllable set. The resulting words could be two-, three- or four-syllables long, with each word length having equal probability within a language. Since we sampled repeatedly with replacement from a set of nine syllables, these randomly-generated labels could contain repetitions.

Following this procedure, we first generated six sets of 18 word forms (i.e., labels for the objects), which were used for the large lexicon conditions. For each word set, a 12-word small lexicon counterpart was created by systematically excluding the six words whose removal caused the smallest differences between the large and small version of the word set in terms of adjacent repetition of CV syllables (i.e., whether the word string contained at least one CV unit that was adjacently repeated, e.g., *wawagu*) and non-adjacent repetition (i.e., whether the word string contained at least one CV syllable that was non-adjacently repeated, e.g., *waguwa*). A complete list of items is given in Appendix B. Mean adjacent CV repetition was 0.081 (SD = 0.034) for small-lexicon sets and 0.096 (SD= 0.030) for large-lexicon sets. Mean non-adjacent CV repetition was 0.083 (SD= 0.037) for small-lexicon sets and 0.060 (SD= 0.027) for large-lexicon sets. The training word-object pairs for each experimental group were generated by randomly combining the object set and word set of the same size, with the constraint that one word set was shared between one chain and one closed group. For example, one of the chains learning a large lexicon and one of the closed group learning a large lexicon had the same set of 18 labels and the same set of 18 objects, although the word-object associations were different. Hereafter, we refer to an initial mapping between novel words and objects as a seed language.

2.1.4 Procedure. The two participants making up a pair sat in separate sound-attenuated booths equipped with networked computers which they used to learn the object-word associations in the training phase and communicate with each other in the interaction phase.

The session lasted for around 40 minutes for a pair in the chain (who underwent one round of interaction), and between 1 and 2.5 hours for a pair in the closed group conditions (who underwent 5 rounds of interaction).⁶ The experiment was written in Python using the PsychoPy library (Peirce, 2009).

2.1.4.1 Generations/Rounds. In the closed group conditions, training occurred only once at the beginning of the experiment – participants were trained on one of the seed languages as described above, and after training, participants repeated the five rounds of interaction without retraining. In chain conditions, the first generation was trained on a seed language generated as described above. For subsequent generations in chains, following Kirby et al. (2015), each generation was trained on the language created by one of the two participants in the previous generation (the ‘transmitter’), who was chosen at random. To generate the transmitter’s language, we took the last label produced for each object by the transmitter.

2.1.4.2 Training phase. Both members of the pair were trained simultaneously but separately. Training consisted of six blocks where each word-object pair in the training language was presented once. Participants were instructed to learn the label for each novel object. On each trial, the object and its corresponding orthographic word (in lowercase letters) were presented in the centre of the screen for 2 seconds.

2.1.4.3 Interaction phase. The interaction phase consisted of a series of alternating director and matcher trials. The two members of the pair were randomly assigned to either a director or a matcher for the initial trial, and switched roles after every trial. On director trials, participants saw an object on the screen and typed in the object’s name (see Figure 2a overleaf). Participants were told not to use words in any language known to both participants (See footnote 3). Emphasis was made to avoid English-like words that would enable their partner to identify

⁶ The duration varied more among the pairs in the closed groups than in the chains because the difference in pace was higher during the interaction phase than the training phase. The proportion of time spent on the interaction, as opposed to the training, was higher in a closed-group pair than a chain pair.

the object for its visual characteristics (e.g. *longu* for a long object). In the event of forgetting the word, they were told to make the closest guess possible. On matcher trials, participants were presented with the word that the director had typed alongside an array of six objects, one of which was the target object (see Figure 2b). The matcher's task was to click on the object which they believed the word corresponded to. Feedback ('Correct' or 'Incorrect') was given after each trial to both director and matcher. A selection was 'correct' if the matcher selected the object that the director was labelling. Trial order was semi-randomized with the condition that the same object could not be allocated to two consecutive pairs of director-matcher trials. The selection of distractor objects and the position of the target object in the matcher array was randomized. Each round of interaction consisted of 64 trials in the large lexicon conditions and 48 trials in the small lexicon conditions, with each participant acting as director twice for each object. In the chain conditions, each participant pair completed a single round of interaction, with the labels produced during interaction forming the basis for the training language passed to the next generation (see below). In the closed group conditions each participant pair completed five rounds of interaction. In the analyses that follow, we compare generations in the chain conditions to rounds of interaction in the closed group conditions and refer to this factor as 'generation' or 'round' as appropriate.

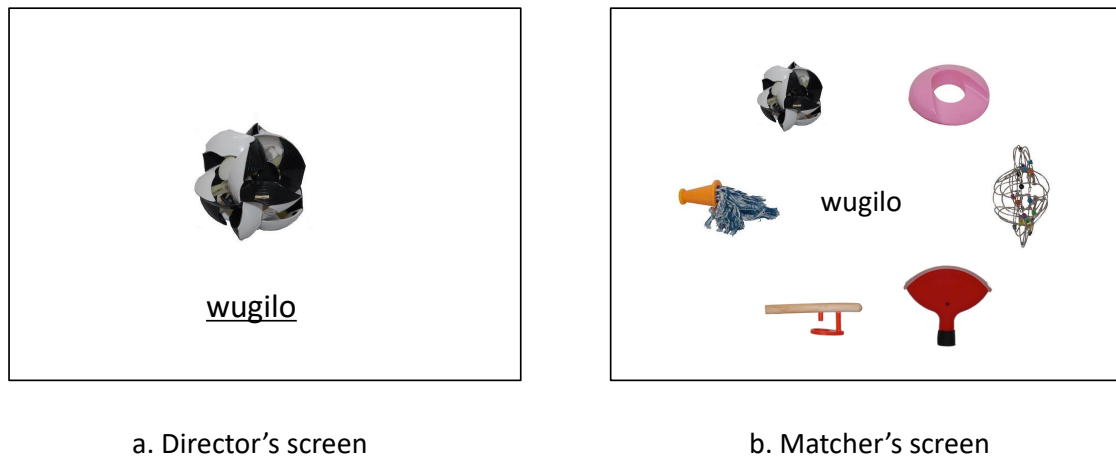


Figure 2. Director's and Matcher's screen during the interaction phase. In this example, the Director has typed *wugilo* as the name of the object shown on their screen. The Matcher sees the Director's input (*wugilo*) and chooses the matching object from an array of six options.

2.2 Main Results

In this section, we report the results of planned analyses on the amount of word-internal repetition in the words observed during the task. Our primary measures of interest related to word-internal adjacent repetition, which we measured in terms of consonants, vowels, and syllables. The precise definitions of these measures are explained below in the relevant subsection. We included syllables in our analyses as well as consonants and vowels because the seed words were manipulated by repeating CV combinations and there was a possibility that the unit of repetitions were perceived to be at the syllable level. The analyses of repetition were carried out on data from the five generations/rounds plus the seed language (which was treated as generation/round 0).

The key predictors of the analyses were Generation/Round, Group Type (chains versus closed group) and Lexicon Size (small versus large). We also included two other predictor

variables in order to control for extraneous factors.⁷ One of them was Word Length. Repetitions may naturally increase as a function of word length simply due to the fact that longer words offer more opportunities for identical sounds to co-occur by chance. We therefore added Word Length as a factor in case there is a tendency for novel words to become longer when they are repeatedly used in interaction or transmitted to new participants. Word Length was operationalised as the number of characters in a word.

The second control variable was Homonymy, or the use of the same word to label more than one object. Inspection of the data indicated that some participants began collapsing distinctions between words, effectively introducing homonymy in the lexicon, as is commonly seen in iterated learning experiments (e.g. Kirby et al., 2008). If words with repeated sounds are more likely to be used in reference to multiple objects, then this process, rather than changes within words, could result in an increase in within-word repetitions. To guard against this possibility, we included in our model a measure of homonymy, which was calculated by subtracting the number of unique word forms from the overall number of word-object pairs and dividing that by the overall number of word-object pairs. For example, if there is a small (i.e., 12 word) lexicon that contains 11 word forms that uniquely refer to different objects but one word form that is used to label two different objects, Homonymy was $(12 - 11)/12 = 0.083$.

In case the effects of Generation/Round were best predicted by a nonlinear model, the data were initially submitted to growth curve analysis (Mirman, 2016). For this procedure, the Generation/Round variable was first transformed to orthogonal linear (ot1), quadratic (ot2) and cubic (ot3) terms, and centred to the mean. The two categorical factors, Group Type and Lexicon Size, were sum-coded. Word Length and Homonymy were centred and scaled. The dependent variables – consonant/vowel/syllable repetition – were binary coded (i.e., each label

⁷ An analysis that does not include these two additional predictors produces the same pattern of results, including, most importantly, the same significant interaction between Group Type and Generation/Round for our measures of consonant and vowel repetition.

either included or did not include repetition of the relevant type), and subjected to binomial analyses. Analyses were carried out using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2018). All analyses began with three maximal models with different sets of polynomial Time terms: first-order (ot1 only), second-order (ot1+ot2) and third-order (ot1+ot2+ot3). The fixed effects of these initial models included the Time terms (ot1, ot1+ot2, or ot1+ot2+ot3), Group Type, Lexicon Size, Word Length, and Homonymy and also the interactions among the Time terms, Group Type and Lexicon Size. The random factors consisted of random intercepts and random slopes for Time by Participant and Object. As these models resulted in singular fit, their structure was reduced by removing the random slopes. For example, the model syntax for the third-order model for consonant repetition was: $CRep \sim (ot1+ot2+ot3)*GroupType*LexiconSize + Word\ Length + Homonymy + (1|Participant) + (1|Object)$. Model fit comparison between the three polynomial time models was based on ΔAIC (calculated as the difference between the model Akaike Information Criterion and the lowest AIC of the three models) and AIC weights. In all analyses, the first-order (linear) model had the best fit. We therefore report only the first-order model for each analysis below, and simply use ‘Generation/Round’ to refer to the linear time term (ot1).

In total, we carried out three first-order analyses, one for each dependent variable (consonant/vowel/syllable repetition) and each based on 4,295 observations. Table 1 summarises the predictions for these analyses. Note that two opposing predictions are listed for the lexicon size effects on repetition. This is because we did not have pre-experimental arguments to determine which of the two counteracting forces would exhibit its effects: the added pressure for learning may cause more repetitions in the large lexicon in comparison to the small lexicon, or the added pressure to maintain lexical distinction during communication may attenuate the amount of repetitions in the large lexicon in comparison to the small lexicon.

Table 1

Outline of main analyses, relevant section numbers, and predicted effects. Predictions are numbered (e.g., P1, P2 ...) for later reference

	Predicted effects	
	Group type	Lexicon size
Repetition in transmission		
§2.2.1 (consonants)	P1: More consonant repetitions emerge in chains than closed groups	P4: More repetitions emerge in the large than small lexicon (as learnability drives preference for repetition); or
§2.2.2 (vowels)	P2: More vowel repetitions emerge in chains than closed groups	P5: Fewer repetitions emerge in the large than small lexicon (as communicative efficiency drives dispreference for repetition)
§2.2.3 (syllables)	P3: More syllable repetitions emerge in chains than closed groups	

Prior to these analyses on word-internal repetition, we also examined how accurately participants were able to communicate with each other during the interaction phase. As these results are not the main focus of our experiment, they are reported in Appendix D. Accuracy level was higher in the closed groups than in the chains, and higher in the small lexicon than in the large lexicon, but participants in both group types and lexicons became more accurate over rounds or generations. Thus overall, the languages in the experiment permitted successful communication about objects, i.e. they remained informative.

2.2.1 Adjacent consonant repetition: C(V)C. We defined adjacent consonant repetition as the presence of identical consonants that are directly adjacent to each other or separated only by vowels. This variable had the value 1 if the word string contained any $C_1V^*C_2$ sequence where C_1 and C_2 were identical and adjacent with any number of intervening vowels, including none (e.g., *papule*, *kokkiwo*, *woww*).⁸ Otherwise, it was 0. The results are summarised in Figure 3.

Table 2 gives the first-order model of this data. There were significant main effects of Generation/Round and Word Length and an interaction between Generation/Round and Group Type. The model indicates that adjacent consonant repetition increased with generation/round (Generation/Round: $\beta = 0.410$, $SE = 0.173$, $p = 0.002$) and that this effect differed between the group types (Generation/round \times Group type: $\beta = 0.415$, $SE = 0.172$, $p = 0.016$), reflecting a higher rate of consonant repetitions in the chains compared to the closed groups in later generations/rounds (Fig. 3). Neither the main effect of Lexicon Size or its interaction with Generation/Round was statistically significant. These results support our prediction (see P1 in Table 1) that the amount of consonant repetitions should become higher in chains, but do not provide evidence for the impact of lexicon size on repetitions (P4 and P5 in Table 1).

⁸ This meant that strings with double consonants were included (e.g., *tkitkitt*, *woww*, *gomello*). While it is possible that some of these forms were influenced by the English orthographic convention of using letter doubling to signal a vowel contrast (e.g., *diner* vs. *dinner*) rather than sound repetition, we included them in our analyses as there was no principled way to determine the phonological import of such doubling. In any case, there were only 10 unique labels with double consonants, which accounted for 0.051% of the 1953 unique forms produced by our participants.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

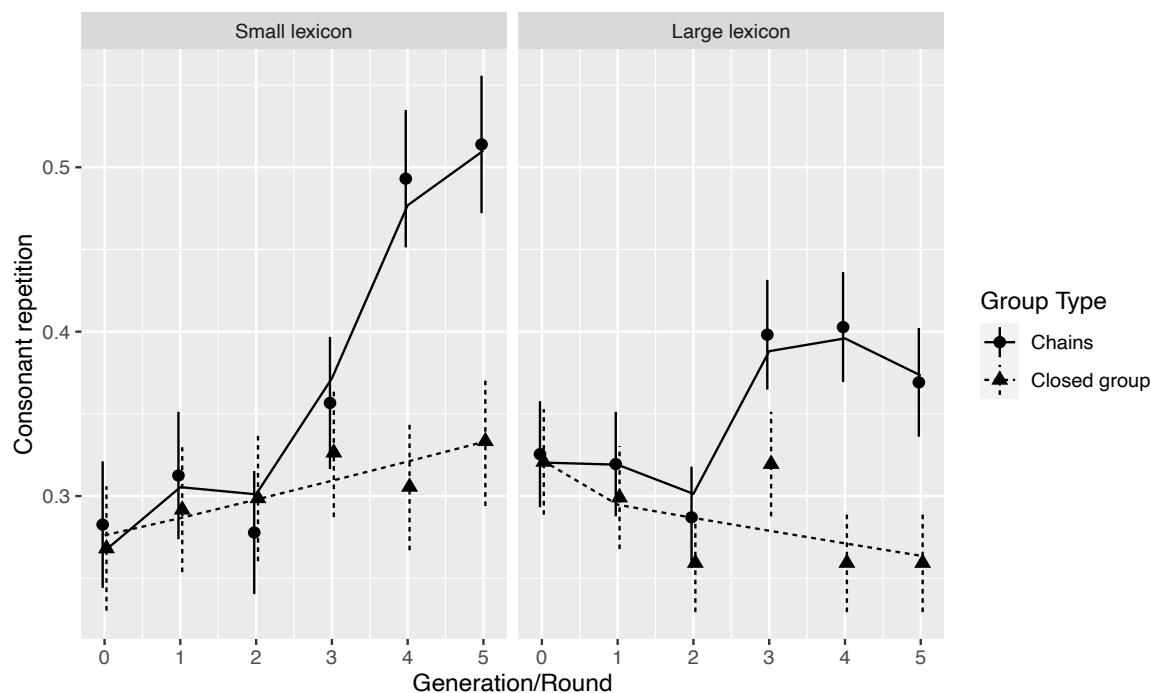


Figure 3. Adjacent consonant repetitions across generations/rounds in Experiment 1. Points represent descriptive means. Error bars are standard errors. Lines represent model predictions.

Table 2

Fixed factor coefficients for adjacent consonant repetition in Experiment 1.

	Estimate	Std. Error	<i>z</i> value	Pr(> <i>z</i>)
Intercept	-0.905	0.168	-5.392	< .001***
Generation/round	0.410	0.173	2.374	< .002**
Group type	0.132	0.112	1.179	.238
Lexicon size	-0.069	0.114	-0.606	.544
Word Length	0.536	0.043	12.357	< .001***
Homonymy	-0.044	0.122	-0.361	.718
Generation/round × Group type	0.415	0.173	2.402	.016*
Generation/round × Lexicon size	-0.324	0.172	-1.878	.060
Group type × Lexicon size	0.034	0.112	0.307	.759
Generation/round × Group type × Lexicon size	-0.111	0.173	-0.643	.520

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

2.2.2 Adjacent vowel repetition: V(C)V. We defined adjacent vowel repetition as the presence of identical vowels that are directly adjacent to each other or separated only by consonants. This variable had the value 1 if the word string contained any $V_1C^*V_2$ sequence where V_1 and V_2 were identical and adjacent with any number of intervening consonants, including none (e.g., *patale*, *kaapiwa*).⁹ Otherwise, it was 0. The results are summarised in Figure 4.

Table 3 gives the first-order model of this data. There were significant main effects of Generation/Round and Word Length, and interactions between Generation/Round and Group Type and between Group Type and Lexicon Size. Adjacent vowel repetition increased with Generation/Round ($\beta = 0.296$, $SE = 0.102$, $p = 0.009$) and, as with the case with consonant repetitions, this effect was larger in the chains than in the closed groups (Generation/Round \times Group type: $\beta = 0.310$, $SE = 0.124$, $p = 0.013$). The interaction between Group Type and Lexicon Size ($\beta = 0.115$, $SE = 0.057$, $p = 0.044$) reflects the difference in the way the chains and the closed group diverge. As can be seen in Figure 4, the difference between the chains and the closed group emerges earlier in the large lexicon (around generation/round 2) than in the small lexicon (around the last generation/round). Again, these results support our prediction (see P2 in Table 1) that there should be more repetitions (in this case, vowels) emerging in the chains in comparison to the closed groups. It also shows that this effect is larger in the large lexicon than the small lexicon, consistent with the possibility that repetitions increase in response to the learning pressure imposed by a larger size of vocabulary (P4 in Table 1).

⁹ As with double consonants, words with double vowels (*leeewu*, *hupoowu*, *pagaglogoo*, *wopoomoki*) were included even though in English orthography, double ‘e’ and double ‘o’ are used to represent a single vowel (e.g., *red* vs. *reed*, *rot* vs. *root*). Again, this was because we were not able to determine the phonological intention of the doubling in these cases. However, there were only 4 unique labels with double ‘e’ or ‘o’, which accounted for 0.02% of the 1953 unique forms produced by our participants.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

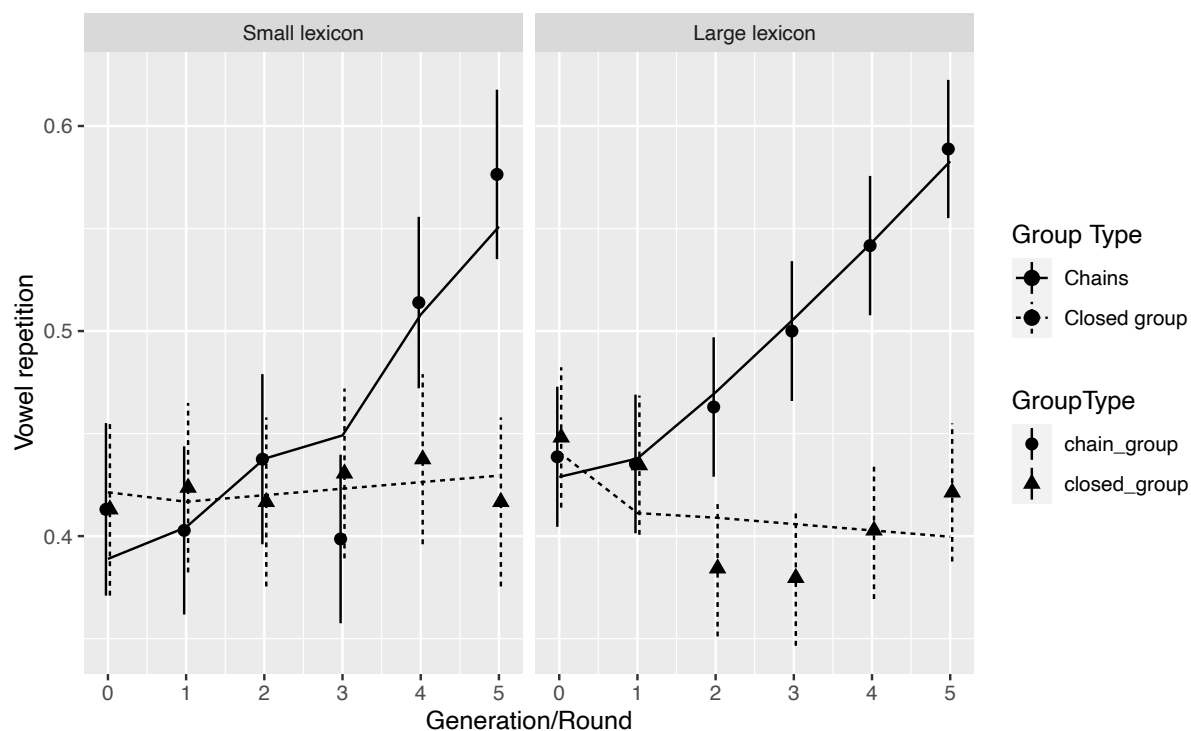


Figure 4. Adjacent vowel repetitions across generations/rounds in Experiment 1. Points represent descriptive means. Error bars are standard errors. Lines represent model predictions.

Table 3

Fixed factor coefficients for adjacent vowel repetition in Experiment 1.

	Estimate	Std. Error	<i>z</i> value	<i>p</i>
Intercept	-0.279	0.083	-3.345	< .001***
Generation/round	0.304	0.124	2.451	.014*
Group type	0.111	0.057	1.946	.052
Lexicon size	0.023	0.059	0.405	.686
Word Length	0.944	0.042	22.482	< .001***
Homonymy	-0.021	0.108	-0.197	.844
Generation/round × Group type	0.310	0.124	2.492	< .013*
Generation/round × Lexicon size	-0.049	0.124	-0.394	.694
Group type × Lexicon size	0.115	0.057	2.017	.044*
Generation/round × Group type × Lexicon size	-0.021	0.125	-0.165	.869

Note. * $p < .05$, *** $p < .001$

2.2.3 Adjacent syllable repetition. Adjacent syllable reduplication was 1 if the word string contained at least one syllable that was adjacently repeated (e.g., *wawagu*), and otherwise 0. A syllable was defined theory-neutrally and was allowed to take one of the following structures: V, CV, VC, and CVC, where each symbol type can repeat its type up to three times (e.g., ‘V’ = {*a, ai, aio ...*}, ‘CV’ = {*ma, pwa, plwa, pai, plai ...*}). The results are summarised in Figure 5.

Table 4 gives the first-order model of this data. Apart from Word Length, the only significant fixed effect was Generation/Round, showing an overall increase in the amount of syllable reduplication ($\beta = 0.440$, $SE = 0.186$, $p = 0.018$). Thus, there was an overall tendency for a consonant-vowel combination to be repeated in later rounds/generations, but there was no evidence that the rate of increase differed between the two group types or lexicon sizes (see P3-5 in Table 1). This indicates that the group type effects found in consonant repetitions and vowel repetitions are not due to participants’ repetition of syllable-size units even though the repetition of CV combinations in the seed language could have been perceived as syllable repetitions.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

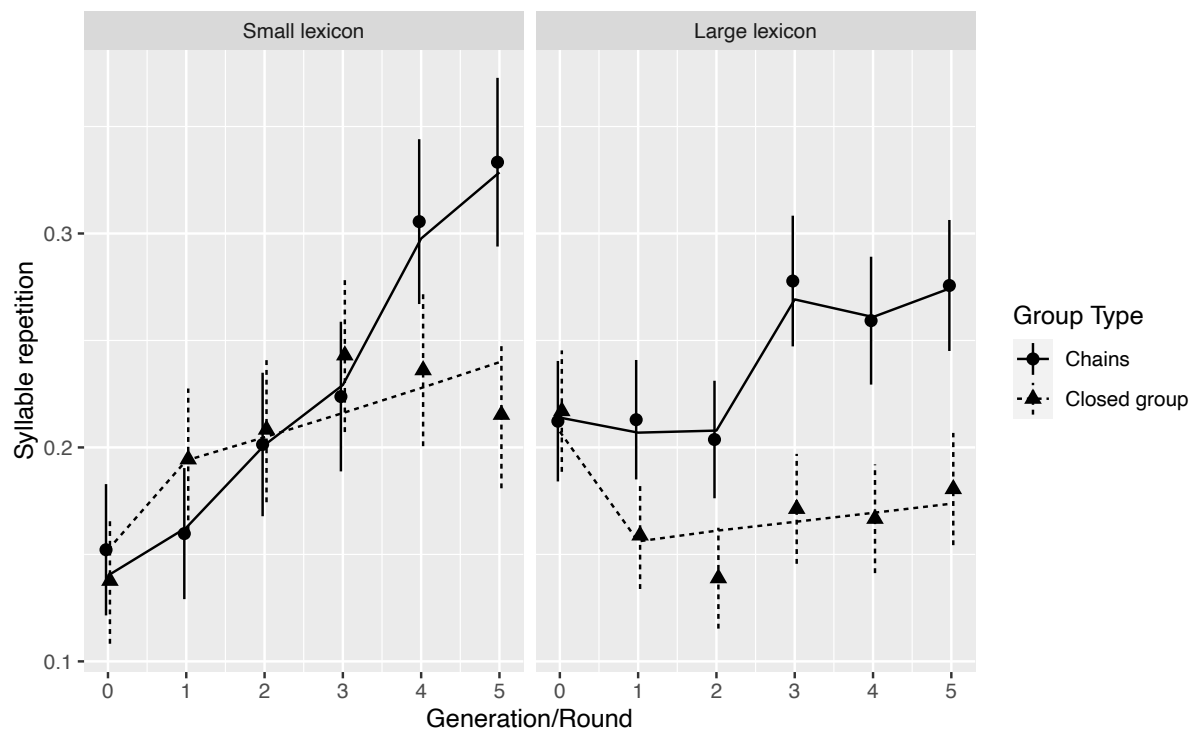


Figure 5. Adjacent syllable repetitions across generations/rounds in Experiment 1. Points represent descriptive means. Error bars are standard errors. Lines represent model predictions.

Table 4

Fixed factor coefficients for adjacent syllable repetition in Experiment 1

	Estimate	Std. Error	<i>z</i> value	Pr(> <i>z</i>)
Intercept	-1.674	0.147	-11.424	< .001***
Generation/round	0.440	0.186	2.362	.018*
Group type	0.106	0.118	0.899	.369
Lexicon size	-0.039	0.119	-0.328	.743
Word Length	0.632	0.049	12.839	< .001***
Homonymy	0.076	0.135	0.561	.575
Generation/round × Group type	0.226	0.187	1.213	.225
Generation/round × Lexicon size	-0.188	0.186	-1.008	.185
Group type × Lexicon size	0.156	0.118	1.324	.186
Generation/round × Group type × Lexicon size	-0.211	0.187	-1.130	.258

Note. * $p < .05$, *** $p < .001$

2.3 Posthoc Analyses

The analyses reported in the previous section show that both consonant and vowel repetition increased more rapidly in the chains than in the closed groups. As the word-object associations had to be learned afresh by a new pair of participants every generation in the chains but only once by pairs in the closed groups, the results support the hypothesis that word-internal repetition increases in response to pressure to learn. This interpretation could be further augmented by evidence showing that string-internal repetitions facilitated the learning of novel words. Although our experiment was not designed to directly examine this effect, we expect to see the impact of repetition on learning in the labels produced by the participants during the interaction phase of the first generation/round, when they were attempting to reproduce the labels from the training phase. To this end, we ran two posthoc analyses of the changes in the labels between the training and the first interaction phase.

2.3.1 Repetition and accuracy of learning. Our hypothesis was that when the pressure to learn is high, word-internal repetition increases due to a learning bias for repeated elements in words. If words that contain repetitions are easier to learn than words that do not contain repetitions, then participants' learning of novel labels should be more accurate when they feature repetitions. To explore this effect, we examined the accuracy of label reproduction in the first generation/round by measuring the Levenshtein distance between each seed word and the participant's production for that word (i.e., the minimum number of character edits required to change one into the other). This is an error measure, and therefore will be lower when the reproduction is more accurate. Figures 6 and 7 display the average Levenshtein distance for each participant depending on the size of the lexicon and whether the seed word contained consonant repetitions (Fig. 6) or vowel repetitions (Fig. 7).

The linear mixed-effects model we used for the analysis included Lexicon Size, Consonant Repetition, Vowel Repetition and their interactions as fixed effects, and random

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

intercepts for Participant and Item. Lexicon Size was sum-coded, and both repetition measures were binary (presence/absence of repetition). The dependent variable was the Levenshtein distance, normalised for word length (number of characters). Group Type was not included as a factor because there is no difference between the chains and closed groups during this phase when participants in both groups had just completed training with the same set of seed languages. The results are presented in Table 5. There was a significant negative main effect of Consonant Repetition ($\beta = -0.121$, $SE = 0.058$, $p = 0.037$), indicating that the labels reproduced were less distant from the training labels (i.e., more accurate) when the training labels contained a consonant repetition (i.e., words containing consonant repetition were more accurately learned). There was no main effect of Vowel Repetition ($\beta = 0.056$, $SE = 0.039$, $p = 0.154$), but a significant interaction was found between Lexicon Size (small), Consonant Repetition and Vowel Repetition ($\beta = -0.299$, $SE = 0.114$, $p = 0.009$), showing that accuracy was higher in the small lexicon than in the large lexicon when both consonant and vowel repetitions were present.

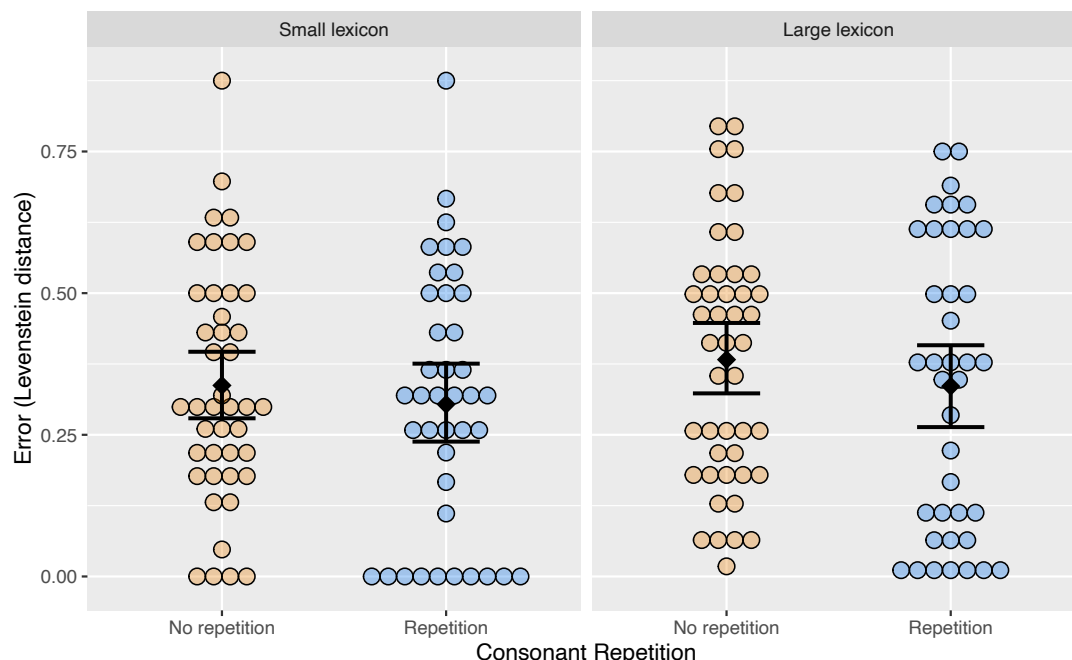


Figure 6. Mean error of the first production in Experiment 1 depending on the absence/presence of consonant repetitions in the word. The left panel shows the results for the small lexicon and the right panel for the large lexicon. Coloured circles indicate means for individual participants, black points indicate grand means and error bars indicate bootstrapped 95% confidence intervals.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

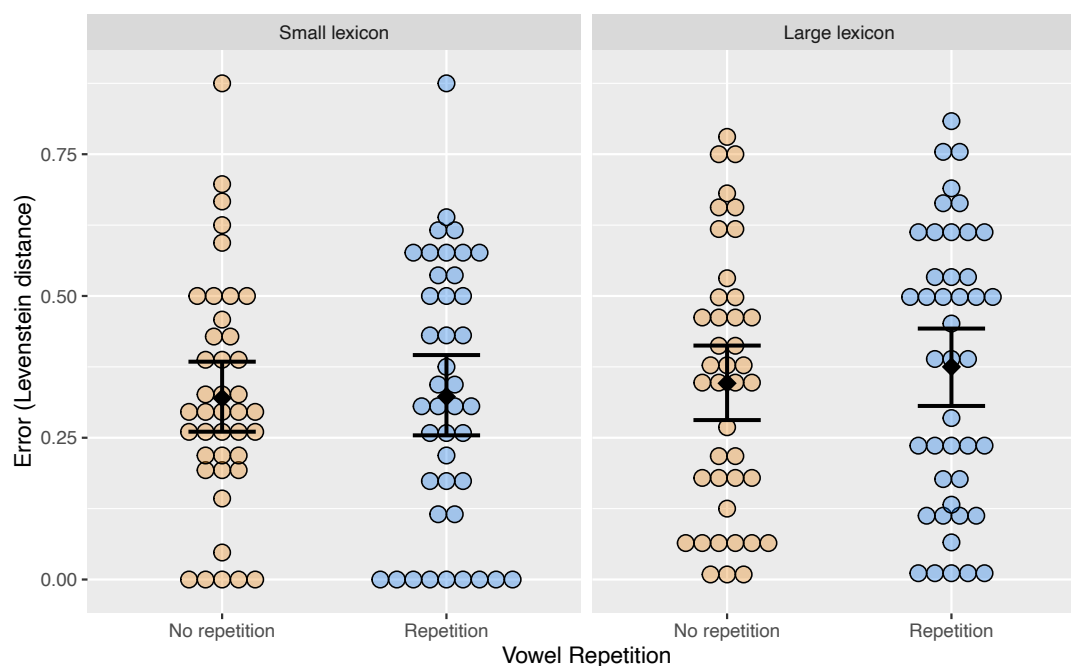


Figure 7. Mean error of the first production in Experiment 1 depending on the absence/presence of vowel repetitions in the word. The left panel shows the results for the small lexicon and the right panel for the large lexicon.

Table 5

Fixed factor coefficients for production accuracy in the first communication task in Experiment 1.

	Estimate	Std. Error	<i>df</i>	<i>t</i> value	<i>p</i>
Intercept	0.339	0.038	66.7	8.83	< .001***
Lexicon size	-0.035	0.055	73.0	-0.65	.520
Consonant repetition	-0.120	0.058	692.0	-2.08	.038*
Vowel repetition	0.056	0.039	686.1	1.42	.154
Lexicon size × Cons rep	0.147	0.092	688.1	1.60	.109
Lexicon size × Vowel rep	0.101	0.060	689.0	1.67	.096
Cons rep × Vowel rep	0.102	0.071	677.6	1.44	.151
Lexicon size × C rep × V rep	-0.299	0.114	681.4	-2.62	.009**

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

2.3.2 Introduction of sound repetitions. Our main results show that the forms of the novel words changed gradually as they were repeatedly transmitted across generations of learners, and that process brought in more repetitions of both consonants and vowels. The posthoc analysis in the previous section (2.3.1) showed that words containing consonant repetitions were more accurately learned than those without consonant repetitions, consistent with the idea that features that facilitate learning tend to accumulate during language transmission. However, the same analysis showed that vowel repetition improves accuracy in learning only when there is also consonant repetition and the lexicon is small. It is therefore unclear why there is a general increase in vowel repetitions during language transmission independent of consonant repetitions and lexicon size, as demonstrated in Section 2.2.2. One possibility is that there is a tendency for words to acquire vowel repetitions in them when they are (mis)learned; that is, mistakes during recall tend to increase vowel repetition even though vowel repetition in itself does not facilitate learning. To address this question, we examined the labels produced by the participants during the interaction phase of the first generation/round again, but this time analysed how often the participants *introduced* consonant or vowel repetitions in their reproduction when the seed words did not contain any consonant/vowel repetitions. The presence of adjacent consonant or vowel repetitions was established in the same way as in the main analyses (see Sections 2.2.2 and 2.2.3 for the precise operationalisations) Figure 8 (left panel) shows the proportion of words to which the participants introduced consonant or vowel repetitions. As a comparison, the right panel of the same figure shows the proportion of words in which consonant or vowel repetitions in the seed words were retained in the reproduction.

We submitted the results to a linear mixed-effects model with the presence of consonant or vowel repetition in the reproduced words as the dependent variable. Fixed effects included as factors were Lexicon Size, Repetition Type Produced (i.e., whether the repetition produced

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

was consonants or vowels), and Repetition in Seeds (i.e., whether the seed word contained repetitions). The model also included interactions between the fixed effects, random intercepts for Participant and random slopes for Participants by Repetition Type Produced and Repetition in Seeds. Lexicon Size and Repetition Type were sum-coded, but Repetition in Seeds was coded as a treatment contrast with no repetitions set as 0. There was a main effect of Repetition in Seeds ($\beta = 2.969$, $SE = 0.286$, $p < 0.001$), simply confirming that seed words that contained repetitions were more likely to be reproduced with repetitions. There was also a significant main effect of Repetition Type in the direction of vowels ($\beta = 1.168$, $SE = 0.268$, $p < 0.001$) and a significant negative interaction between Repetition in Seeds and Repetition Type ($\beta = -1.502$, $SE = 0.348$, $p < 0.001$) that counteracted the main effect of Repetition type. Thus, participants were more likely to introduce vowel repetitions than consonant repetitions when the seed words contained no repetitions (see Fig. 8, left panel) but when the seed words contained repetitions, the asymmetry between consonants and vowels was attenuated or reversed (see Fig. 8, right panel).

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

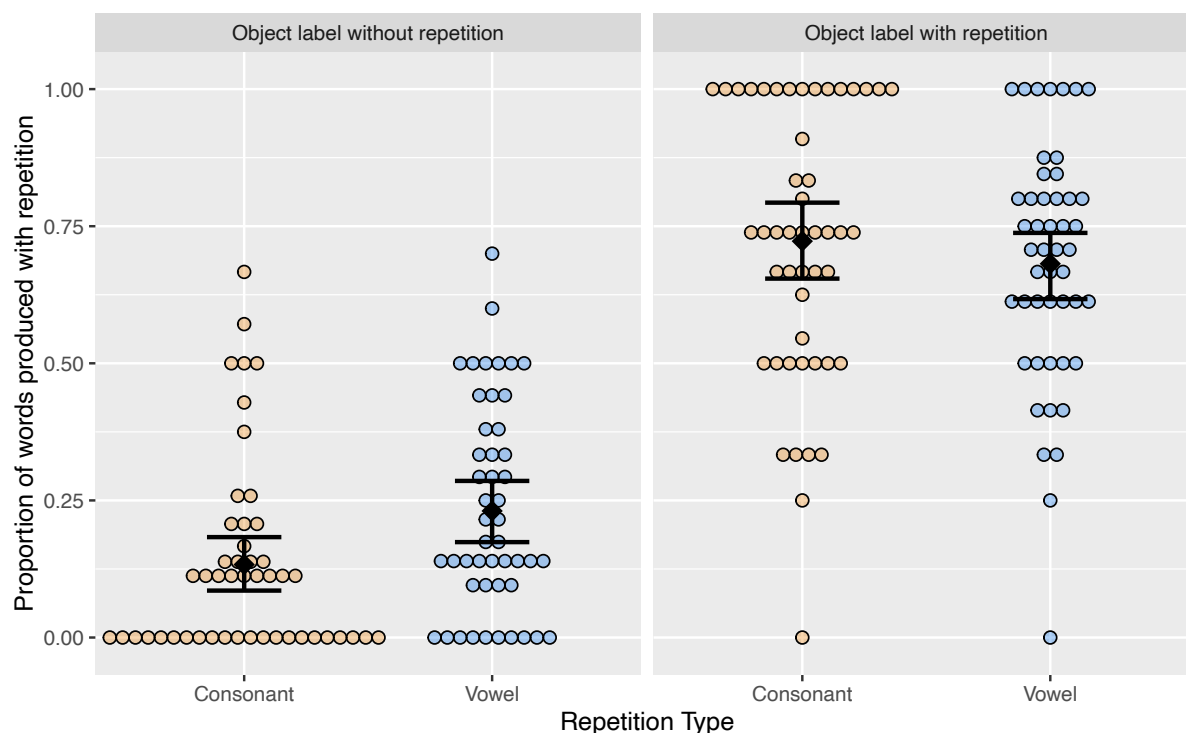


Figure 8. Proportion of words with repetitions produced by participants immediately after the training phase in Experiment 1. The left panel shows the proportion of consonant or vowel repetitions introduced to seed words that did not contain repetitions. The right panel shows the proportion of consonant or vowel repetitions maintained in seed words that did contain repetitions.

Table 6

Fixed factor coefficients for repetitions produced for words during the first communication phase in Experiment 1.

	Estimate	Std. Error	Z	p
Intercept	-1.970	0.202	-9.728	< .001***
Repetition in seeds	2.969	0.286	10.373	< .001***
Repetition type	1.168	0.268	4.358	< .001***
Lexicon size	0.246	0.381	0.645	.519
Repetition in seeds × Rep type	-1.502	0.348	-4.312	< .001***
Repetition in seeds × Lex size	-0.267	0.553	-0.484	.629
Repetition type × Lex size	-0.346	0.437	-0.791	.429
Rep in seeds × Rep type × Lex size	0.008	0.608	0.013	.989

Note. *** $p < .001$

2.4 Discussion

In Experiment 1, we tested the effects of learning on word-internal repetition in two models of cultural transmission with different levels of learning pressure – languages were either learned once and then used repeated for communication (closed groups) or learned anew at each generation before being used for communication (chains). There were two key results in the main analyses. First, and most importantly, both consonant repetitions and vowel repetitions in the transmitted words increased faster over time in the chains compared to the closed groups (Sections 2.2.1 and 2.2.2), and these effects targeted segments (i.e., consonants and vowels) rather than syllables, as indicated by the lack of group difference in the syllable repetition analysis (Section 2.2.3). Second, at least in the vowel repetition analysis (Section 2.2.2), the difference between chains and closed groups was larger in the large lexicon than the small lexicon. These outcomes support our hypothesis that repetitions increase in response to learning pressures acting on the language.

However, our posthoc analyses of the word forms produced immediately after the initial training phase suggests that, in the context of word learning, consonant repetitions and vowel repetitions behave in a different way. On the one hand, the presence of consonant repetitions made the reproduction more accurate across the board, while the presence of vowel repetitions aided accuracy of reproduction only under specific conditions (when the lexicon size was small and the seed word contained a consonant repetition) (Section 2.3.1). On the other hand, seed words had a stronger tendency to acquire vowel repetitions than consonant repetitions (Section 2.3.2). These differences between consonants and vowels imply that they play different roles in word learning: consonant repetitions can aid word learning by facilitating retention of the word form that includes the repetition, while vowel repetitions tend to be added (in error) to word forms that are imperfectly learned. If so, sound repetitions are involved in two separate

types of bias, one that makes a particular pattern more likely to be learned and one that makes a particular pattern more likely to be introduced during learning or production.

A caveat of this conclusion is that it is based on posthoc analyses of the outcomes of an experiment that was designed to look at the effects of transmission on word-internal repetitions, rather than the effects of a single iteration of learning. Although our analysis of learning effects looked only at words that were reproduced immediately after the training phase, those words were produced in the communication phase of the experiment and therefore might have been affected by the communicative goal of the task, which was to lead the matcher to the correct object rather than to reproduce the seed words as accurately as possible. We therefore conducted a follow-up experiment that allowed us to re-examine the effects of consonant and vowel repetition in a learning-only task, in order to test the robustness of the findings of these posthoc analyses.

3. Experiment 2

3.1 Overview

In order to explore the apparent differential effects of consonant versus vowel repetition on learning, we ran a second experiment testing a second group of participants on their ability to learn the same set of languages that was used to train Generation 1 participants in Experiment 1. We ran this experiment online. The method was as close as possible to that used to train participants in Experiment 1, but with a noncommunicative recall test and with modifications to make the method more suitable to online testing, in particular, more active involvement of participants during training.

3.2 Method

3.2.1 Participants. We recruited 45 participants via Amazon Mechanical Turk (mTurk). We used participants who were based in the US, who had a 97% approval rate or higher on 1000+ HITs (as indicated by location codes and qualifications provided by mTurk), and who had not completed this or similar experiments run by us (managed using custom qualifications).¹⁰ Participants were self-reported native speakers of English. The participants were randomly assigned to one of two conditions: small (23 participants) or large (22 participants) lexicon. All participants who completed the experiment were paid US\$4. One additional participant began the experiment but failed on the catch trials in block 1 (see below), was paid \$1, and was excluded from the analysis. Informed consent was obtained from all participants.

3.2.2 Materials. As in Experiment 1, participants were trained on associations between pictures of unfamiliar objects and novel orthographic words, then labelled those unfamiliar objects in a final test. We used the same objects and words as in Experiment 1, and used the same set of object-word associations that were used to train Generation 1 participants in Experiment 1. This provided a set of 12 small and 12 large input lexicons, each of which was used during training for at least one and at most two participants in Experiment 2. As in Experiment 1, small lexicons consisted of 12 object-label associations, large lexicons consisted of 18 associations. Because repetitions in the words from Experiment 1 were controlled for CV combinations but not for consonants and vowels independently, there were differences between consonant repetitions (mean: 0.30, SD: 0.14) and vowel repetitions (mean: 0.43, SD: 0.09) in the label words.

3.2.3 Procedure. After accepting the HIT on mTurk, participants completed the experiment in a web-browser, running custom javascript code. Prior to commencing the

¹⁰ A HIT (Human Intelligence Task) is a unit of task used in Amazon Mechanical Turk.

experiment, we asked participants not to take written notes, and then at the end of the experiment we asked participants to self-declare if they had done so.¹¹ No participants indicated they had taken written notes. The experiment took on average around 12 minutes to complete in the small-lexicon condition, and 16 minutes in the large-lexicon condition.

The experiment consisted of 3 blocks, each block consisting of a passive training phase, comprehension training, and a test.

3.2.3.1 Passive training phase. Each object-word pair in the language was presented once in each block of passive training. Participants were instructed simply to watch and learn. On each trial, the object and its corresponding orthographic word were presented in lowercase in the middle of the screen for 6 seconds.

3.2.3.2 Comprehension training phase. Each object-word pair in the language was presented once in each block of comprehension training. On each trial, one label and 4 objects were presented on-screen and participants were asked to click on the object corresponding to the label. The 4 objects included the target object and 3 other objects selected randomly from the set of objects in the target lexicon. Participants were given a running score during each comprehension block. Correct responses (i.e., clicking on the correct object) were rewarded with 10 points, and incorrect responses received 0 points. Regardless of whether the participant responded correctly or incorrectly, after they made their selection all incorrect objects were removed and the label plus its corresponding object were presented on-screen for 3 seconds.

In addition to training on the novel object-label associations, in Block 1 the comprehension training phase included 3 catch trials, randomly interspersed with the other training trials. Each catch trial consisted of an English word as the label (*moustache, guitar,*

¹¹ Prior to commencing: “Please do not take written notes! In this experiment we are interested in what your brain can do, not what your brain plus a notebook can do, so please don’t write anything down. Just do your best - we are interested in what you can’t learn as well as what you can.”. On completing the experiment: “Did you write stuff down or take notes during the task? Please be honest - it won’t affect your payment, we promise, and if you tell us now we can correct for this in our analysis without affecting the validity of our experiment”. Participants responded using a yes/no radio button.

umbrella) and an array containing 4 icons of familiar objects, the target plus 3 foils; all icons were obtained from thenounproject.com. Catch trials were scored as per normal trials. Participants were warned that these catch trials would be included, and participants answering any catch trial incorrectly were redirected to the exit page after the end of the comprehension training phase and paid a reduced amount due to the shortened duration of the experiment (the hourly rate was equivalent to that obtained by participants completing the full experiment).

3.2.3.3 Test phase. At the end of each block, participants were tested on their ability to provide (i.e., type) labels for objects. At the end of Blocks 1 and 2, this test phase consisted of 4 trials on 4 randomly-selected objects. This brief test was simply intended to make it clear to participants that they would need to learn the language sufficiently well to be able to type labels for objects from memory. At the end of Block 3, we tested participants on all the objects in their target lexicon (i.e., 12 or 18 trials depending on lexicon size). On each trial, participants were simply prompted with an object and asked to type the corresponding label into a text box. Participants received no feedback on the accuracy of their responses.

The total exposure (combined across 3 blocks of passive training and comprehension training) was equal to that used in Experiment 1, with each object-label association being presented 6 times. Unlike in Experiment 1, the final test was a pure recall test, rather than being embedded in a communicative task, and participants were tested on each object only once.

3.3 Results

Our main interest in the data was the accuracy of the labels participants produced during testing, which we assessed using normalised Levenstein distance between the target label and the label produced, as for Experiment 1. Since the test trials on blocks 1-2 contained only a small number of trials, we only analysed the production data for the third and final block, when participants were required to label all objects in the set. We also examined participants'

accuracy in selecting the correct referent during comprehension training; as these results are not central to the investigation here, they are reported in Appendix E.

3.3.1 Sound repetition and accuracy on production trials. Figures 9 and 10 show how accurately participants reproduced the word forms, again, depending on the size of the lexicon and the presence in the assigned label of consonant repetitions (Fig. 9) or vowel repetitions (Fig. 10). As accuracy here is measured by means of error (or Levenstein distance), a lower score indicates higher accuracy. These figures are the equivalent of Figures 6 and 7, which reported the same measures for our exploratory analysis of our Experiment 1 data.

We carried out a mixed effects analysis on the error using the lmer function on R with Lexicon size, Consonant repetition and Vowel repetition as fixed effects. The data consisted of 672 observations. To achieve convergence and avoid singularity, the random effect structure was reduced to the intercept of participants only. The results (Table 7) showed a significant negative (i.e. lower error) main effect of Consonant repetition. This effect is visible in Figure 9, which shows lower error means for words with consonant repetitions both in the small and the large lexicons. Thus, as our exploratory analysis for Experiment 1 showed, training labels containing consonant repetition were more faithfully reproduced than those without consonant repetitions. No other effects were significant. In particular, labels containing vowel repetition were *not* more accurately produced, which echoes the finding of the exploratory analysis from Experiment 1. The three-way interaction between lexicon size, consonant repetition and vowel repetition that was detected in Experiment 1 was not replicated in the current analysis. The reason for this cannot be established from the available data, but it is possible that the tendency to reproduce combined repetitions of consonants and vowels when the lexicon is small may be a product of the more communicative nature of the task in Experiment 1 (i.e., participants were reproducing the labels for their partners rather than for themselves) and not something that is induced by pressures to learn.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

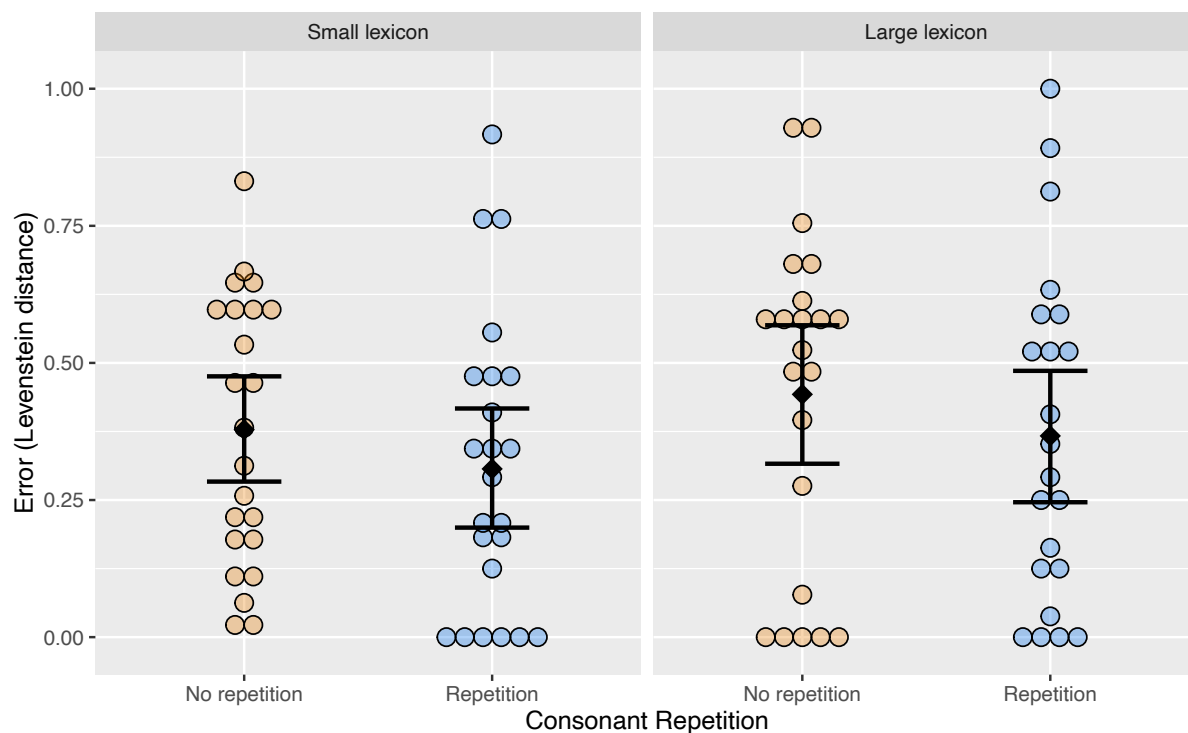


Figure 9. Mean error of the last production trial in Experiment 2 depending on the absence/presence of consonant repetitions in the word. The left panel shows the results for the small lexicon and the right panel for the large lexicon.

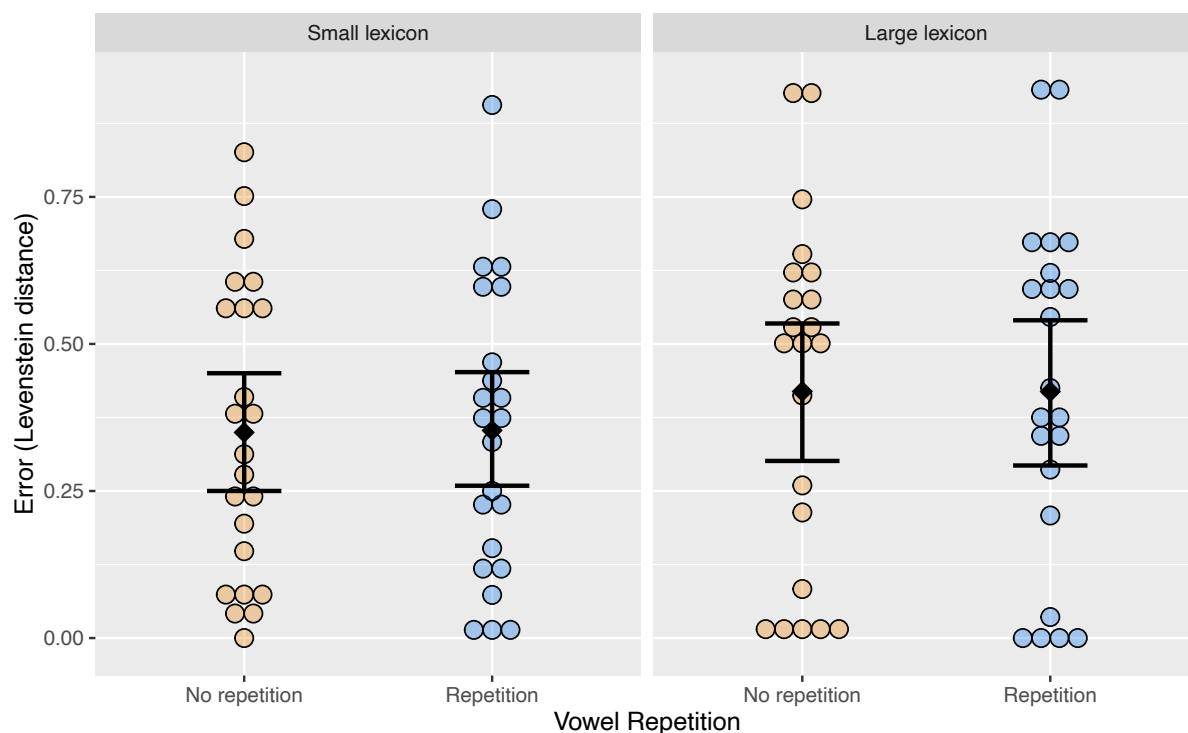


Figure 10. Mean error of the last production trial in Experiment 2 depending on the absence/presence of vowel repetitions in the word. The left panel shows the results for the small lexicon and the right panel for the large lexicon.

Table 7

Fixed factor coefficients for production accuracy in Experiment 2.

	Estimate	Std. Error	<i>df</i>	<i>t</i> value	<i>p</i>
Intercept	0.373	0.039	46.2	9.47	< .001***
Lexicon size	0.073	0.079	46.2	0.92	.363
Consonant repetition	-0.097	0.026	631.5	-3.75	< .001***
Vowel repetition	0.041	0.025	625.7	1.62	.106
Lexicon size × Cons rep	0.002	0.052	631.5	0.04	.969
Lexicon size × Vowel rep	0.002	0.050	625.7	0.04	.967
Cons rep × Vowel rep	-0.021	0.050	624.1	-0.42	.674
Lexicon size × C rep × V rep	0.003	0.099	624.1	0.03	.975

Note. *** $p < .001$

3.3.2 Introduction of sound repetitions. As our posthoc analysis of data from Experiment 1 indicated that vowel repetitions were more likely than consonant repetitions to be introduced to the novel words, we examined the results from Experiment 2 to see if the same effect can be observed in this experiment. Figure 11 displays the proportion of words with consonant or vowel repetitions that participants produced during the final block. The data are organised by the presence or absence of repetitions in the words used in training. This figure is the equivalent of Figure 8 for our exploratory analysis of Experiment 1.

As with the corresponding analysis from Experiment 1, we submitted the results to a linear mixed-effects model with consonant or vowel repetition in the words produced by the participants as the dependent variable (see Table 8). Fixed effects included as factors were Lexicon size, Repetition type produced, and Repetition in seeds. The model also included interactions between the fixed effects, random intercepts for Participant and random slopes for Participants by Repetition type produced and Repetition in seeds. Lexicon size and Repetition Type were sum-coded, but Repetition in seeds was coded as a treatment contrast with no

repetitions set as 0. There were significant effects of Repetition in seeds ($\beta = 3.46$, $SE = 0.278$, $p < 0.001$) and Repetition type ($\beta = 0.89$, $SE = 0.427$, $p = 0.036$). These main effects show that participants are more likely to produce repetitions when the trained form included repetitions, and were more likely to introduce vowel repetitions than consonant repetitions when the seed words did not contain any repetitions. This difference between vowel and consonant repetition diminished when the seed words contained repetition, as indicated by a significant negative interaction between Repetition in seeds and Repetition type ($\beta = -1.13$, $SE = 0.572$, $p < 0.049$). These results are consistent with the equivalent posthoc analysis of Experiment 1 data.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

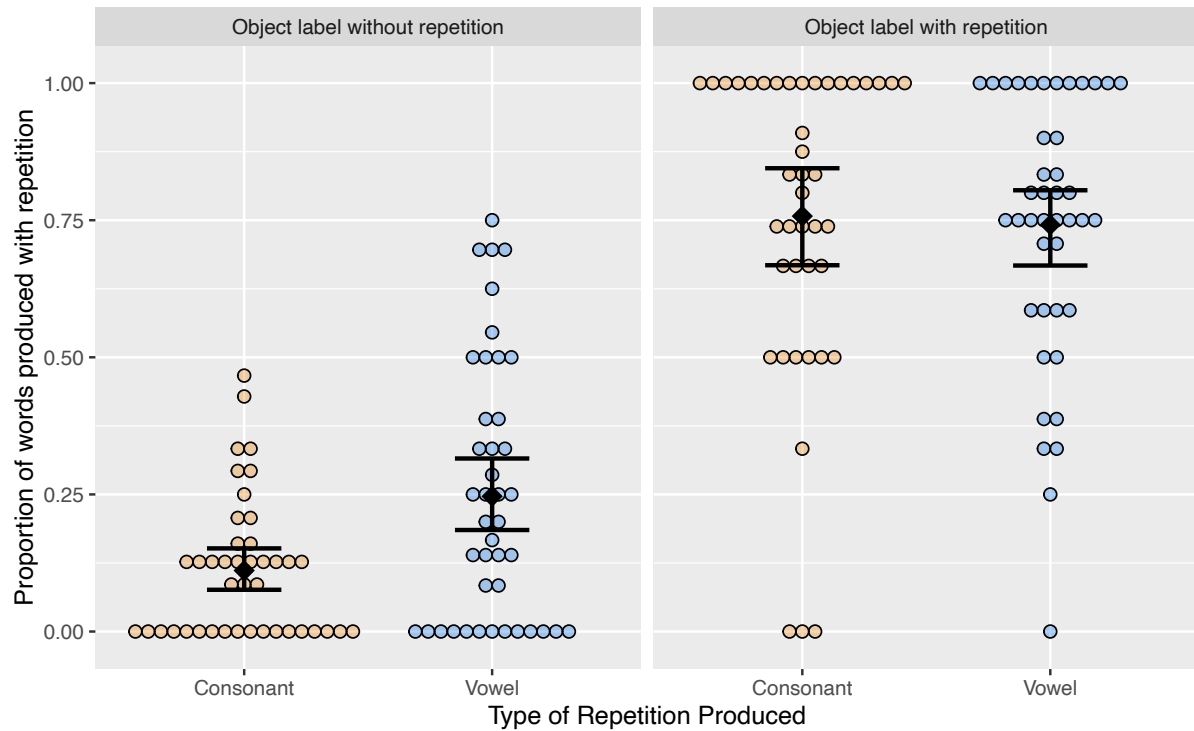


Figure 11. Proportion of words with repetitions produced by participants during the last block in Experiment 2. The left panel shows the proportion of consonant or vowel repetitions introduced to object labels that did not contain repetitions. The right panel shows the proportion of consonant or vowel repetitions maintained in object labels that did contain repetitions.

Table 8

Fixed factor coefficients for repetitions produced for words in Experiment 2.

	Estimate	Std. Error	Z	p
Intercept	-2.179	0.278	-7.823	< .001***
Repetition in seeds	3.463	0.478	7.248	< .001***
Repetition type	0.894	0.427	2.095	.036*
Lexicon size	0.486	0.374	1.301	.193
Repetition in seeds × Rep type	-1.127	0.572	-1.972	.049*
Repetition in seeds × Lex size	-0.150	0.659	-0.227	.820
Repetition type × Lex size	0.003	0.536	0.005	.996
Rep in seeds × Rep type × Lex size	0.087	0.697	0.125	.901

Note. * $p < .05$, *** $p < .001$

3.4 Discussion

The findings from the exploratory analyses for Experiment 1 were largely replicated in Experiment 2. Participants were more accurate in reproducing words assigned as object labels when the words contained consonant repetitions, but reproduction accuracy was not aided by vowel repetitions (Section 3.3.1). Participants also introduced more vowel repetitions than consonant repetitions in the words they reproduced. Unlike in Experiment 1 where the participants were engaged in an ostensibly communicative task, the participants in this experiment were only asked to learn the word-object mappings and the words themselves (Section 3.3.2). As such, Experiment 2 confirms that these differences between consonant and vowel repetitions emerge in the process of learning and recall, rather than being driven by communication.

4. General Discussion

There is a tendency across languages to avoid word-internal proximate repetition of sounds (in particular, of consonants), but previous developmental research suggests that words containing repetitions are learned more easily. This constitutes a puzzling deviation from the widely attested alignment between learning and typological generalisations. Here we examined a cultural evolutionary explanation of this discrepancy by testing the hypothesis that biases in learning favouring word-internal repetition are present in adult speakers but can be offset by pressures in language use, which disprefers repetition because it reduces the distinctiveness of words and potentially impedes communication. The results of the two experiments are summarised in Table 9.

Table 9

Summary of analyses and results

Experiment and analysis	Results
Experiment 1: Iterated learning and communication	
Repetition in transmission (§2.2.1-2.2.3)	More rapid increase of repetitions in chains (vs closed groups) for consonants and vowels (but not for syllables). Larger chain-closed group difference for large (vs small) lexicons.
Repetition and accuracy of learning (§2.3.1)	Words with consonant repetitions are reproduced more accurately than those without.
Introduction of repetition during learning (§2.3.2)	Vowel repetitions are more likely than consonant repetitions to be introduced in reproduction of learned words.
Experiment 2: Non-iterated learning	
Repetition and production accuracy (§3.3.1)	Words with consonant repetitions are reproduced more accurately than those without, confirming results from Exp 1.
Introduction of repetition (§3.3.2)	Vowel repetitions are more likely than consonant repetitions to be introduced in reproduction of learned words, confirming results from Exp 1.

The central findings of Experiment 1 come from the analyses of consonant and vowel repetitions in Sections 2.2.1 and 2.2.2. The amount of both adjacent consonant and vowel repetitions increased over time in chains; that is, when learning pressure was high, because word-object mappings were transmitted to a fresh pair of individuals who had to learn the mappings from the previous pair. In comparison, this effect was suppressed in closed groups, whose transmission featured only one phase of learning and multiple rounds of interaction between the same pair of individuals. Therefore, when learning pressure in the transmission system is below a certain threshold, or downweighted relative to communication, patterns favoured by a learning bias do not proliferate in the language.

Interestingly though, the posthoc analyses of Experiment 1 data (Sections 2.3.1-2.3.2) provided evidence that consonants and vowels behave differently in the context of learning. Adjacent consonant repetitions, but not vowel repetitions, had a general facilitation effect on

learning in that words containing consonant repetitions were more faithfully reproduced than those without consonant repetitions. In contrast, vowel repetitions had a stronger tendency than consonant repetitions to be introduced to words not featuring those repetitions. These findings indicate that both consonant and vowel repetitions increase across generations in response to pressures to learn, but through different mechanisms. These findings were confirmed in Experiment 2, where participants engaged in a pure learning task (Section 3.3.1-3.3.2).

Why would consonant and vowels behave differently with respect to adjacent repetitions in word learning? The more straightforward case of the two is adjacent vowel repetitions, which, in our experiments, tended to infiltrate into word forms during learning; errors in learning lead to more vowel repetitions than consonant repetitions in the acquired form, which are then passed on to other speakers. This could be because the inventory of vowels (or vowel letters) is smaller than that of consonants – a pattern that is true not only in our experimental setup but also in virtually all known phonological systems in the world's languages (Maddieson, 2013) – making accidental introduction of identical sounds more likely for vowels than for consonants. Relatedly, because of the inventory size asymmetry between consonants and vowels, all other things being equal, consonants contribute more information than vowel towards the identify of a word (Nespor, Peña, & Mehler, 2003). As studies on sound change indicate that listeners encode less phonetic information of a sound that is more contextually predictable (e.g., Cohen Priva, 2015; Seyfarth, 2014; Wedel, Jackson, & Kaplan, 2013), they may be more prone to make errors in reproducing vowels than consonants, raising the likelihood of repetition for vowels. It is also possible that vowel repetitions are more likely to be introduced during learning due to an inherent preference towards identity relationships in adjacent vowels. There is evidence that infants exposed only to English, a language without a vowel harmony system, nevertheless tend to segment strings containing identical vowels as discrete units (Mintz et al., 2018), suggesting the possibility that we may be predisposed to

expect word-internal vowels to be identical or similar. Any of the possibilities mentioned above can cause learning to add more vowel repetitions than consonant repetitions.

The mechanism driving the increase of consonant repetitions during language transmission is different, however. The evidence emerging from our experiments is that words containing consonant repetitions are more accurately remembered compared to words without them. Therefore, while consonant repetitions are less likely than vowel repetitions to be added to a word by error, once they are introduced, they tend to remain as part of the representation of that word. In other words, while vowel repetitions increase due to a bias in the direction of changes that occur during the reproduction phase of learning, consonant repetitions increase due to a bias in the retention of forms. We are not aware of any other studies which have been able to disentangle the contribution of spontaneous innovation and learnability advantages in this way, or which have shown such a clean separation between these two mechanisms (although see Tamariz & Kirby, 2015 for a related phenomenon in a nonlinguistic task). The technique we lay out here may be of use in building a more detailed picture of how – for example – compositional structure (e.g. Kirby et al., 2008) or iconicity (Tamariz, Roberts, Martínez, & Santiago, 2018) develops through language transmission.

At this juncture, it is worth revisiting the other two potential explanations for the learning-typology mismatch for word-internal repetition that was discussed in Section 1.3. One possibility was that the advantages in detecting, remembering and recognising words with sound repetitions reported in the developmental literature can be attributable to the stimuli used in those studies, which typically featured repetitions of both consonants and vowels (e.g., *neenee*). There is no contradiction between early learning and typological generalisations if the preference for vowel repetitions (demonstrated by Mintz et al., 2018) is robust enough to give rise to the preference for CV repetitions on its own, overriding a potential learning bias against consonant repetitions. However, the current study presents evidence that consonant repetitions

make words more learnable and increase during language transmission, which contradicts the possibility that there is an underlying bias against consonant repetitions during learning. Another potential explanation was that the preference for sound repetitions exhibited by young learners is a developmental phase, possibly driven by a primarily exogenous attention system that infants and young children have. However, we find a similar preference for word-internal repetitions in the participants of the current study, who were adults. The learning advantages afforded by repetitions do not disappear after early childhood.

We now turn to results of our experiment that raise potential issues for the cultural evolutionary account pursued here. Recall that the ultimate question we were trying to address is why learning biases are not always reflected in typological generalisations to the extent that some linguistic patterns privileged in learning can be typologically avoided, as is the case with consonant repetitions. Our evolutionary explanation is that during the course of language transmission, patterns favoured in learning may be counteracted by constraints on communication. In the case of word-internal repetitions, we identified two potential sources of communicative pressures that may suppress them: the reduction in lexical contrast space that could result from word-internal segmental identity and the perceptual confusion that could arise from repetition blindness/deafness. One might therefore expect the amount of consonant repetitions to *decrease* in the closed groups, where these communicative pressures were higher, simulating the avoidance of proximate consonant repetitions that is pervasive in natural languages. But this was not the case. There are two possible explanations for why this effect was not observed. Firstly, the lexicons used in the experiment might have been too small to show the putative effect of contrast space reduction. The phonological space in the seed language was defined by words that varied in length from 2 to 4 syllables, which in turn were drawn from a pool of 9 CV syllables. Although a prohibition of repeating any of the 9 syllables would result in approximately 51% reduction in the phonological space, it would still leave

3,600 different disyllabic to quadrisyllabic combinations as potential words. With only 12 or 18 lexical distinctions required, the lexicon in Experiment 1 space might have been impervious to such an effect. Secondly, the nonserial nature of our stimulus presentation could have blocked any potential effects of repetition blindness/deafness. Our participants viewed the novel words presented in intact orthographic representation (e.g., *wawagu*) without necessarily having to serially process subcomponents of the words, which is the condition under which repetition blindness/deafness obtains.

Another manipulation in Experiment 1 that did not yield clear effects on the amount of proximate sound repetitions was the size of the lexicon. We considered two possibilities. One was that an increase in the size of the lexicon adds pressure for learning, and that, in turn, would lead to an increase in sound repetitions for the same reason that more increase in repetitions were observed in the chains than the closed groups. Another was that an increase in the size of the lexicon adds pressure to maintain contrasts between the lexical items and therefore should lead to a *decrease* in repetitions, because repetitions can incur communicative cost by constraining the phonological space. The only effect of lexicon size that was statistically verified in Experiment 1 was its interaction with group type on vowel repetitions (See Fig. 4). The pattern there showed that there were more vowel repetitions in chains compared to closed groups when the lexicon was larger, a result that is more consistent with the learnability-driven alternative. However, it is possible that our lexicon size manipulation was too subtle to induce demonstrable effects of contrast maintenance. Evidence of contrast enhancement in speech production shows that speakers try to maintain distinctness of phonologically similar words, but those effects are typically found between close phonological neighbours such as minimal pairs (Baese-Berk & Goldrick, 2009; Nelson & Wedel, 2017; Wedel, Nelson, & Sharp, 2018). Given the vast number of less similar words that can be held in the hypothetical phonological

space of our lexicons, increasing their size from 12 words to 18 words might not have been sufficient to capture the communicative costs arising from sound repetitions.

A caveat against drawing broad conclusions about the role of phonological repetition based on these findings is that our stimuli were orthographic, not spoken, words. As already discussed above, the synchronous nature of orthographic stimulus presentation limits the interpretation of the results with respect to effects of repetition blindness/deafness that are typical of serial processing. Caution must also be applied to the assumption that consonant/vowel letters are comparable to consonant/vowel sounds in the context of orthographic word processing. Despite robust evidence that written word recognition is mediated by phonological representation, not every aspect of grapheme processing is translatable to segmental processing. Nonetheless, what is crucial to our study is that many of the constraints that apply to repetition of sounds in spoken words, such as the reduction of signal space in the lexicon, also apply to graphemes in written words. Furthermore, our results bear out the predictions that follow from observations in spoken word learning and processing. Future work using auditory stimuli will be able to confirm the generalisability of our findings to spoken words.

A few other questions remain unanswered at present. While our results indicate that different type of biases are involved in consonant and vowel repetitions, further work is required to understand the precise nature and source of these biases. In particular, it remains unclear whether the contrasting behaviours of consonants and vowels arise from substantive differences in the phonetic characteristics of the two sound types, or from probabilistic differences in retaining or selecting sounds from phonological sets that vary in inventory size. A study in which the ratio of consonants and vowels is reversed can shed light on this issue. An aspect of real-life language transmission that is not captured in the setup of the current experiment is the asymmetry between adult and child speakers whereby learning occurs almost exclusively

through transmission from adult speakers (caregivers) to children (learners), while the main locus of communication lies between adult speakers. This type of asymmetry can be implemented, for example, by first training ‘adult’ participants on a language, who will then interact with other ‘adults’ or with ‘child’ participants during which they provide them with model input to learn that language. Including such a ‘life cycle’ of language interaction may change some fundamental dynamics of modelled cultural transmission of language.

This last point is particularly important given the observation that the structure of the vocabulary is not uniform within a language. Across speech communities, there is typically a set of lexical items (often termed ‘baby-talk words’) that are unique to the register addressed to infants and young children, with phonological characteristics that are different from the adult lexicon. Intriguingly, one such feature of baby-talk words is the preponderance of sound repetitions (Ferguson, 1964, 1977; Gervain & Werker, 2008; Endress, et al., 2009). For example, a number of languages have a baby-talk word for ‘sleep/nap’ that contains repetitions of sounds despite the lack of sound repetition in the corresponding adult word (e.g., Basque: *lolo* (cf. adult word: *lo egin*), Cree: *miimii* (cf. *nipaah*), Czech: *nini* (cf. *spat*), French: *dodo* (cf. *dormir*), Hungarian: *csicsi* (cf. *alvas*), Japanese: *nenne* (cf. *neru*), Swedish: *nanna* (cf. *sova*)). The contrast between child-directed vocabulary and adult vocabulary is consistent with the notion that language responds to learning biases in a different way depending on the degree of learning pressure the language (or a given subsystem of the language) is under. It is exactly where the pressure to learn is high (i.e., the linguistic environment of infants and children) that we find sound repetitions which are otherwise absent in the adult words. This is also the context where the pressure for efficient communication is low; the small size of children’s lexicons means that the reduced contrasts due to the sound repetitions in these words have a relatively low cost, as caregivers can easily infer the intended meanings of the produced words.¹² The

¹² We thank Andrew Wedel for making this observation, which further highlights the distinct nature of the learner’s linguistic environment where word-internal repetition emerges as a common property.

case of baby-talk words demonstrates the dynamic nature of learning biases. Their effects may be visible only under certain contexts within a linguistic system, depending on the balance between learnability and communication efficiency.

This study contributes to a growing body of work providing a new way to understand and explore the relationship between language learning and language design. One obvious reason why we do not always find natural language typology to be in perfect correspondence with learning biases is that language is also shaped by factors outside learning (e.g., channel bias: Ohala, 1993; Moreton, 2008). Additionally, a potential learning bias that favours the use of a certain type of linguistic information over another may not be observed in children because those aspects of the language follow different developmental timetables (e.g., phonology versus semantics: Culbertson, Gagliardi, & Smith, 2017; Gagliardi, Feldman, & Lidz, 2017). We have demonstrated here that there is another way in which the correspondence between learning biases and language design can be disrupted: Some learning biases can be suppressed during the process of cultural transmission of language because their effects are overridden by countervailing biases favoured by language use or communication.

The possible existence of underlying learning biases whose effects are not readily detectable in language design because of other pressures at work raises some intriguing broader questions for the study of language and cognition. What other types of biases in learning may exist which have no apparent signature in linguistic systems? What strategies can be used to uncover such inductive biases in the face of concealed evidence? Do these inductive biases have intrinsic properties that separate them from those that are more conspicuous? These questions define a fruitful area for further work on the role of learning in human language and cognition.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 681942), and from the School of Philosophy, Psychology and Language Sciences at the University of Edinburgh. Many thanks to Andy Wedel and an anonymous reviewer for their detailed and very helpful comments.

References

- Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun categorization devices*. Oxford University Press.
- Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology*, 20(3), 249–264.
<https://doi.org/10.1080/14640746808400159>
- Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, 24(4), 527–554.
<https://doi.org/10.1080/01690960802299378>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Berent, I., Bat-El, O., Brentari, D., Dupuis, A., & Vaknin-Nusbaum, V. (2016). The double identity of linguistic doubling. *Proceedings of the National Academy of Sciences*, 113(48), 13702–13707. <https://doi.org/10.1073/pnas.1613749113>
- Berg, T. (1992). Phonological harmony as a processing problem. *Journal of Child Language*, 19(2), 225–257. <https://doi.org/10.1017/S0305000900011405>
- Berkley, D. M. (1994). The OCP and gradient data. *Studies in the Linguistic Sciences*, 24(1/2), 59–72.
- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511486357>
- Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Holland Academic Graphics.
- Boll-Avetisyan, N., & Kager, R. (2014). OCP-PLACE in speech segmentation. *Language and Speech*, 57(3), 394–421. <https://doi.org/10.1177/0023830913508074>

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

Bybee, J. (2001). *Phonology and language use*. Cambridge University Press.

<https://doi.org/10.1017/CBO9780511612886>

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 58(2), 265–289. <https://doi.org/10.1353/lan.1982.0021>

Carr, J. W., Smith, K., Culbertson, J., & Kirby, S. (2020). Simplicity and informativeness in semantic category systems. *Cognition*, 202, 104289.

<https://doi.org/10.1016/j.cognition.2020.104289>

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. MIT Press.

Clements, G. N. (2003). Feature economy in sound systems. *Phonology*, 20(3), 287–333.

<https://doi.org/10.1017/S095267570400003X>

Cohen Priva, U. (2015). Informativity affects consonant duration and deletion rates.

Laboratory Phonology, 6(2). <https://doi.org/10.1515/lp-2015-0008>

Cohen Priva, U. (2017). Informativity and the actuation of lenition. *Language*, 93(3), 569–597. <https://doi.org/10.1353/lan.2017.0037>

Colombo, J. (2002). Infant attention grows up: The emergence of a developmental cognitive neuroscience perspective. *Current Directions in Psychological Science*, 11(6), 196–200.

<https://doi.org/10.1111/1467-8721.00199>

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior*, 34, 283–

322. [https://doi.org/10.1016/S0065-2407\(06\)80010-8](https://doi.org/10.1016/S0065-2407(06)80010-8)

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204–256. <https://doi.org/10.1037/0033-295X.108.1.204>

Comrie, B. (1999). Grammatical gender systems: A linguist's assessment. *Journal of*

Psycholinguistic Research, 28(5), 457–466. <https://doi.org/10.1023/A:1023212225540>

- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, 55(4), 429–432. <https://doi.org/10.1111/j.2044-8295.1964.tb00928.x>
- Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Cornish, H. (2010). Investigating how cultural transmission leads to the appearance of design without a designer in human communication systems. *Interaction Studies*, 11(1), 112–137. <https://doi.org/10.1075/is.11.1.02cor>
- Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5), 1119–1130. <https://doi.org/10.1037/0278-7393.30.5.1119>
- Cristia, A., & Seidl, A. (2008). Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development*, 4(3), 203–227. <https://doi.org/10.1080/15475440802143109>
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. Longman.
- 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., Gagliardi, A., & Smith, K. (2017). Competition between phonological and semantic cues in noun class learning. *Journal of Memory and Language*, 92, 343–358. <https://doi.org/10.1016/j.jml.2016.08.001>
- Culbertson, J., & Kirby, S. (2016). Simplicity and specificity in language: Domain-general biases have domain-specific effects. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01964>

- Culbertson, J., Jarvinen, H., Haggarty, F., & Smith, K. (2019). Children's sensitivity to phonological and semantic cues during noun class learning: Evidence for a phonological bias. *Language*, 95(2), 268–293. <https://doi.org/10.1353/lan.2019.0031>
- Culbertson, J., & Newport, E. L. (2015). Harmonic biases in child learners: In support of language universals. *Cognition*, 139, 71–82. <https://doi.org/10.1016/j.cognition.2015.02.007>
- Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, 122(3), 306–329. <https://doi.org/10.1016/j.cognition.2011.10.017>
- de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00044>
- Endress, A. D., Nespors, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, 13(8), 348–353. <https://doi.org/10.1016/j.tics.2009.05.005>
- Fee, J., & Ingram, D. (1982). Reduplication as a strategy of phonological development. *Journal of Child Language*, 9(1), 41–54. <https://doi.org/10.1017/S0305000900003603>
- Feldman, J. (2003). The simplicity principle in human concept learning. *Current Directions in Psychological Science*, 12(6), 227–232. <https://doi.org/10.1046/j.0963-7214.2003.01267.x>
- Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist*, 66(6_PART2), 103–114.
- Ferguson, C. A. (1977). Baby talk as a simplified register. In Snow, Catherine E. & C. A. Ferguson (Eds.), *Talking to children* (pp. 209–235). Cambridge University Press. <http://philpapers.org/rec/FERBTA-2>

- Fikkert, P., & Levelt, C. C. (2008). How does Place fall into place?: The lexicon and emergent constraints in children's developing phonological grammar. In P. Avery, B. E. Dresher, & K. Rice (Eds.), *Contrast in phonology: Theory, perception, acquisition* (pp. 231-270). Mouton de Gruyter.
- Frisch, S. A., Pierrehumbert, J. B., & Broe, M. B. (2004). Similarity avoidance and the OCP. *Natural Language & Linguistic Theory*, 22(1), 179–228.
<https://doi.org/10.1023/B:NALA.0000005557.78535.3c>
- Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, 123(1), 71–99. <https://doi.org/10.1037/0033-2909.123.1.71>
- Gagliardi, A., Feldman, N. H., & Lidz, J. (2017). Modeling statistical insensitivity: Sources of suboptimal behavior. *Cognitive Science*, 41(1), 188–217.
<https://doi.org/10.1111/cogs.12373>
- Gagliardi, A., & Lidz, J. (2014). Statistical insensitivity in the acquisition of Tsez noun classes. *Language*, 90(1), 58–89.
- Gerken, L., Dawson, C., Chatila, R., & Tenenbaum, J. (2015). Surprise! Infants consider possible bases of generalization for a single input example. *Developmental Science*, 18(1), 80–89. <https://doi.org/10.1111/desc.12183>
- Gervain, Judit, Berent, I., & Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24(3), 564–574. https://doi.org/10.1162/jocn_a_00157
- Gervain, J., Macagno, F., Cogoi, S., Pena, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, 105(37), 14222–14227. <https://doi.org/10.1073/pnas.0806530105>

Gervain, J., & Werker, J. F. (2008). How infant speech perception contributes to language acquisition. *Language and Linguistics Compass*, 2(6), 1149–1170.

<https://doi.org/10.1111/j.1749-818X.2008.00089.x>

Glewwe, E. (2019). Bias in phonotactic learning: Experimental studies of phonotactic implicationals [Doctoral dissertation]. UCLA. Retrieved from

<https://escholarship.org/uc/item/4456s1j0>

Goldsmith, J. (1976). *Autosegmental phonology* [Doctoral dissertation]. MIT.

Goodsitt, J. V., Morse, P. A., Hovee, J. N. V., & Cowan, N. (1984). Infant speech recognition in multisyllabic contexts. *Child Development*, 55(3), 903.

<https://doi.org/10.2307/1130141>

Gordon, M. K. (2016). *Phonological typology*. Oxford University Press.

<https://doi.org/10.1093/acprof:oso/9780199669004.001.0001>

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (Ed.), *Universals of language* (pp. 73–113). MIT Press.

Greenberg, J. H. (1950). The patterning of root morphemes in Semitic. *Word*, 6(2), 162–181.

<https://doi.org/10.1080/00437956.1950.11659378>

Hall, K. C., Hume, E., Jaeger, T. F., & Wedel, A. (2018). The role of predictability in shaping phonological patterns. *Linguistics Vanguard*, 4(s2). [https://doi.org/10.1515/lingvan-](https://doi.org/10.1515/lingvan-2017-0027)

[2017-0027](https://doi.org/10.1515/lingvan-2017-0027)

Hansson, G. Ó. (2010). *Consonant harmony: Long-distance interaction in phonology*.

University of California Press.

Hayes, B., & White, J. (2013). Phonological naturalness and phonotactic learning. *Linguistic Inquiry*, 44(1), 45–75. https://doi.org/10.1162/LING_a_00119

https://doi.org/10.1162/LING_a_00119

Henson, R. N. A. (1998). Item repetition in short-term memory: Ranschburg repeated.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 24(5), 1162–1181. <https://doi.org/10.1037/0278-7393.24.5.1162>

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database:

A collection of novel images for use in experimental research. *Behavior Research Methods*, 48(4), 1393–1409. <https://doi.org/10.3758/s13428-015-0647-3>

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1(2), 151–195. <https://doi.org/10.1080/15475441.2005.9684215>

Inkelas, S., & Zoll, C. (2005). *Reduplication: Doubling in morphology*. Cambridge University Press.

Jahnke, J. C. (1969). The Ranschburg effect. *Psychological Review*, 76(6), 592–605. <https://doi.org/10.1037/h0028148>

Jakobson, R. (1968). *Child language aphasia and phonological universals*. Mouton.

Kanwisher, N., Driver, J., & Machado, L. (1995). Spatial repetition blindness is modulated by selective attention to color or shape. *Cognitive Psychology*, 29(3), 303–337. <https://doi.org/10.1006/cogp.1995.1017>

Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, 27(2), 117–143. [https://doi.org/10.1016/0010-0277\(87\)90016-3](https://doi.org/10.1016/0010-0277(87)90016-3)

Karmiloff-Smith, A. (1979). *A functional approach to child language*. Cambridge University Press.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, 336(6084), 1049–1054. <https://doi.org/10.1126/science.1218811>

- Kemp, C., Xu, Y., & Regier, T. (2018). Semantic typology and efficient communication. *Annual Review of Linguistics*, 4(1), 109–128. <https://doi.org/10.1146/annurev-linguistics-011817-045406>
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31), 10681–10686. <https://doi.org/10.1073/pnas.0707835105>
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102. <https://doi.org/10.1016/j.cognition.2015.03.016>
- Kuo, L.-J. (2009). The role of natural class features in the acquisition of phonotactic regularities. *Journal of Psycholinguistic Research*, 38(2), 129–50. <https://doi.org/10.1007/s10936-008-9090-2>.
- Leben, W. R. (1973). *Suprasegmental phonology* [Doctoral dissertation]. MIT.
- Leivada, E. (2017). What's in (a) label? Neural origins and behavioral manifestations of identity avoidance in language and cognition. *Biolinguistics*, 11, 221–250.
- Levelt, C. C., Schiller, N. O., & Levelt, W. J. (2000). The acquisition of syllable types. *Language Acquisition*, 8(3), 237–264. https://doi.org/10.1207/S15327817LA0803_2
- Linzen, T., & Gallagher, G. (2017). Rapid generalization in phonotactic learning. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 8(1), 24. <https://doi.org/10.5334/labphon.44>
- MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science*, 288(5465), 527–531.
- Maddieson, I. (2013). Consonant-vowel ratio. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structure*. <https://wals.info/chapter/3>

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

- Mahowald, K., Dautriche, I., Gibson, E., & Piantadosi, S. T. (2018). Word forms are structured for efficient use. *Cognitive Science*, 42(8), 3116–3134.
<https://doi.org/10.1111/cogs.12689>
- Martin, A., & White, J. (2019). Vowel harmony and disharmony are not equivalent in learning. *Linguistic Inquiry*, 1–20. https://doi.org/10.1162/ling_a_00375
- Martinez-Alvarez, A., Pons, F., & de Diego-Balaguer, R. (2017). Endogenous temporal attention in the absence of stimulus-driven cues emerges in the second year of life. *PLOS ONE*, 12(9), e0184698. <https://doi.org/10.1371/journal.pone.0184698>
- McCarthy, J. J. (1986). OCP Effects: Gemination and antigemination. *Linguistic Inquiry*, 17(2), 207–263.
- Menn, L. (1971). Phonotactic rules in beginning speech. *Lingua*, 26(3), 225–251.
[https://doi.org/10.1016/0024-3841\(71\)90011-8](https://doi.org/10.1016/0024-3841(71)90011-8)
- Miller, M. D., & MacKay, D. G. (1994). Repetition deafness: Repeated words in computer-compressed speech are difficult to encode and recall. *Psychological Science*, 5(1), 47–51. <https://doi.org/10.1111/j.1467-9280.1994.tb00613.x>
- Miller, M. D., & MacKay, D. G. (1996). Relations between language and memory: The case of repetition deafness. *Psychological Science*, 7(6), 347–351.
<https://doi.org/10.1111/j.1467-9280.1996.tb00387.x>
- Mills, A. E. (1986). Acquisition of the natural-gender rule in English and German. *Linguistics*, 24(1). <https://doi.org/10.1515/ling.1986.24.1.31>
- Mintz, T. H., Walker, R. L., Welday, A., & Kidd, C. (2018). Infants' sensitivity to vowel harmony and its role in segmenting speech. *Cognition*, 171, 95–107.
<https://doi.org/10.1016/j.cognition.2017.10.020>
- Mirman, D. (2016). *Growth curve analysis and visualization using R*. CRC Press.

- Monaghan, P., & Zuidema, W. H. (2015). General purpose cognitive processing constraints and phonotactic properties of the vocabulary. In H. Little (Ed.), *The evolution of phonetic capabilities: Causes, constraints, consequences – Proceedings of the 18th International Congress of Phonetic Sciences* (pp. 20–24).
https://pure.mpg.de/rest/items/item_2351030/component/file_2351029/content#page=24
- Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, 25(1), 83–127.
<https://doi.org/10.1017/S0952675708001413>
- Moreton, E. (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language*, 67(1), 165–183.
<https://doi.org/10.1016/j.jml.2011.12.003>
- 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>
- Müller, N. (2000). Gender and number in acquisition. In B. Unterbeck, M. Rissanen, T. Nevalainen, & M. Saari (Eds.), *Gender in grammar and cognition: Part I: Approaches to gender* (pp. 351–400). Walter de Gruyter.
- Nelson, N. R., & Wedel, A. (2017). The phonetic specificity of competition: Contrastive hyperarticulation of voice onset time in conversational English. *Journal of Phonetics*, 64, 51–70. <https://doi.org/10.1016/j.wocn.2017.01.008>

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, 2, 203–230.

<https://doi.org/10.1418/10879>

Ohala, J. J. (1981). Articulatory constraints on the cognitive representation of speech. In *Advances in Psychology*, 7, 111–122. [https://doi.org/10.1016/S0166-4115\(08\)60183-1](https://doi.org/10.1016/S0166-4115(08)60183-1)

Ohala, J. J. (1993). Sound change as nature's speech perception experiment. *Speech*

Communication, 13(1), 155–161. [https://doi.org/10.1016/0167-6393\(93\)90067-U](https://doi.org/10.1016/0167-6393(93)90067-U)

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, 83(1), B13–B23. [https://doi.org/10.1016/S0010-0277\(01\)00165-2](https://doi.org/10.1016/S0010-0277(01)00165-2)

Ota, M., Hartsuiker, R. J., & Haywood, S. L. (2009). The KEY to the ROCK: Near-homophony in nonnative visual word recognition. *Cognition*, 111(2), 263–269.

<https://doi.org/10.1016/j.cognition.2008.12.007>

Ota, M., & Skarabela, B. (2016). Reduplicated words are easier to learn. *Language Learning and Development*, 12(4), 380–397. <https://doi.org/10.1080/15475441.2016.1165100>

Ota, M., & Skarabela, B. (2018). Reduplication facilitates early word segmentation. *Journal of Child Language*, 45(1), 204–218. <https://doi.org/10.1017/S0305000916000660>

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 80–96. <https://doi.org/10.1037/0278-7393.34.1.80>

Pater, J., & Werle, A. (2003). Direction of assimilation in child consonant harmony. *The Canadian Journal of Linguistics / La Revue Canadienne de Linguistique*, 48(3), 385–408. <https://doi.org/10.1353/cjl.2004.0033>

Peirce J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2(10), 1-8. <https://doi.org/10.3389/neuro.11.010.2008>

- Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, 18(3), 571–590. <https://doi.org/10.1017/S0305000900011259>
- Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9), 3526–3529. <https://doi.org/10.1073/pnas.1012551108>
- Pozdniakov, K., & Segerer, G. (2007). Similar place avoidance: A statistical universal. *Linguistic Typology*, 11(2). <https://doi.org/10.1515/LINGTY.2007.025>
- Pycha, A., Nowak, P., Shin, E., & Shosted, R. (2003). Phonological rule-learning and its implications for a theory of vowel harmony. In M. Tsujimura & G. Garding (Eds.), *Proceedings of WCCFL 22* (pp. 101-114). Cascadilla Press.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org/>
- Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In B. MacWhinney & W. O'Grady (Eds.), *The Handbook of Language Emergence* (pp. 237–263). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118346136.ch11>
- Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science*, 19(1), 41–46. <https://doi.org/10.1177/0963721409360003>
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Lawrence Erlbaum.
- Rose, S., & Walker, R. (2011). Harmony systems. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (pp. 240–290). Wiley-Blackwell. <https://doi.org/10.1002/9781444343069.ch8>

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners.

Developmental Psychology, 39(3), 484–494. <https://doi.org/10.1037/0012-1649.39.3.484>

Schwartz, R. G., Leonard, L. B., Wilcox, M. J., & Folger, M. K. (1980). Again and again:

Reduplication in child phonology. *Journal of Child Language*, 7(1), 75–87.
<https://doi.org/10.1017/S0305000900007030>

Seyfarth, S. (2014). Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition*, 133(1), 140–155.

<https://doi.org/10.1016/j.cognition.2014.06.013>

Skoruppa, K., & Peperkamp, S. (2011). Adaptation to novel accents: Feature-based learning of context-sensitive phonological regularities. *Cognitive Science*, 35(2), 348–366.

<https://doi.org/10.1111/j.1551-6709.2010.01152.x>

Slobin, D. I. (2002). Language evolution, acquisition, diachrony: Probing the parallels. In T. Givón & B. F. Malle (Eds.), *Typological Studies in Language* (Vol. 53, pp. 375–392).

John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.53.20slo>

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge University Press.

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated

learning. *Cognition*, 116(3), 444–449. <https://doi.org/10.1016/j.cognition.2010.06.004>

Sóskuthy, M. (2015). Understanding change through stability: A computational study of sound change actuation. *Lingua*, 163, 40–60.

<https://doi.org/10.1016/j.lingua.2015.05.010>

Soto-Faraco, S., & Sebastián-Gallés, N. (2001). The effects of acoustic mismatch and

selective listening on repetition deafness. *Journal of Experimental Psychology: Human*

- Perception and Performance*, 27(2), 356–369. <https://doi.org/10.1037/0096-1523.27.2.356>
- Suzuki, K. (1998). *A typological investigation of dissimilation* [Doctoral dissertation, University of Arizona]. <https://rucore.libraries.rutgers.edu/rutgers-lib/38312/>
- Tamariz, M., & Kirby, S. (2015). Culture: Copying, compression, and conventionality. *Cognitive Science*, 39(1), 171–183. <https://doi.org/10.1111/cogs.12144>
- Tamariz, M., Roberts, S. G., Martínez, J. I., & Santiago, J. (2018). The interactive origin of iconicity. *Cognitive Science*, 42(1), 334–349. <https://doi.org/10.1111/cogs.12497>
- Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 371–386. <https://doi.org/10.1037/0278-7393.14.3.371>
- Vihman, M. (1978). Consonant harmony: Its scope and function in child language. In J. H. Greenberg (Ed.), *Universals of human language, vol 2: Phonology* (pp. 281–334). Stanford University Press.
- Vihman, M. (1980). Sound change and child language. In E. C. Traugott, R. Labrum, & S. C. Shepherd (Eds.), *Current Issues in Linguistic Theory* (Vol. 14, p. 303). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.14.31vih>
- Walter, M. A. (2007). *Repetition avoidance in human language* [Doctoral dissertation, MIT]. <http://hdl.handle.net/1721.1/41694>
- Wedel, A. (1999). Turkish emphatic reduplication. *UC Santa Cruz: Working Papers*. Retrieved from <https://escholarship.org/uc/item/6sm3953w>
- Wedel, A. B. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, 23(3). <https://doi.org/10.1515/TLR.2006.010>

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, 128(2), 179–186.

<https://doi.org/10.1016/j.cognition.2013.03.002>

Wedel, A., Nelson, N., & Sharp, R. (2018). The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language*, 100, 61–88.

<https://doi.org/10.1016/j.jml.2018.01.001>

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, 30(5), 945–982.

https://doi.org/10.1207/s15516709cog0000_89

Winter, B., & Wedel, A. (2016). The co-evolution of speech and the lexicon: The interaction of functional pressures, redundancy, and category variation. *Topics in Cognitive Science*, 8(2), 503–513.

<https://doi.org/10.1111/tops.12202>

Yip, M. (1998). Identity avoidance in phonology and morphology. In S. G. Lapointe, D. K. Brentari, & P. M. Farrell (Eds.), *Morphology and its relation to phonology and syntax*, (pp. 216–246). Stanford, CA.: CSLI Publications.

Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Houghton Mifflin.

Appendix A

Unfamiliar objects used in the experiments. Source: NOUN database (Horst and Hout, 2016).



Appendix B

Lists of nonwords used as labels of objects

Set	Items (small lexicon)
1	wikawawo pewa wakiwa wakoko peku nupewiko wikowa kakiki wiwa kikawi kukuko kopekowi
2	woguhepi makomo kagu makoma mopiwa mawa mokohe wowo koka kawohe pikamawa kawa
3	wega welulaga moga gomowemo goga melugomo lumo gomela gamo wila luhimola wewe
4	wunene nehune wunele pupomoki kilepo leki lenepe mohuwu hupewupo nepu powu wumopele
5	wuko kikika kolukolu kalu kokika wakaki lumo kolu wunumo wululowu wunuka mowa
6	hihiko lahoka pahiwinu mila palaka mikomimi hika panu komimi kahi kawi kowi
Set	Items (large lexicon)
1	pewa kopekowi kukuko peku wakoko kukuwipe wokopepe wikawawo kika wakiwa kikawi kakawa wikowa kako kiwo wiwa nupewiko kakiki
2	makoma kowo kagu kowa kopimako wowomahe kawohe mawa moguko mokohe wowo kawa woguhepi pikamawa koka makomo pimokahe mopiwa
3	gomowemo gomela luhimola goga lawimo gowe welulaga mohi wewe wela lumegola moga melugomo lumo gamo wega wila wewiwi
4	wunele wuwunepo pelepo mopuwu mohuwu hupewupo wumopele powu lenepe leki wunene kilepo nelemomo lehu pupomoki nehune nepu mohulepo
5	wunumo kokika lumo kakowa kawa lolokimo lomo wululowu wakaki mowa lolomo kakokolo wuko kolukolu kikika kalu wunuka kolu
6	wikanumi panu mikomimi palaka hika pahiwini lahoka hihiko mihi kowi kopahika komimi pahiwinu mila kawi hoka hokakami kahi

Appendix C

Supplementary information

The data of this study are publicly available via the Open Science Framework

(<https://osf.io/g5p8q/>).

Appendix D

Communication accuracy of interaction in Experiment 1

Accuracy during the interaction phase in Experiment 1 was coded as a binary variable with the value 1 if the matcher successfully selected the object the director was describing and 0 if they did not. The results are summarised in Figure D1.

Table D1 presents the first-order model of this data. There were significant main effects of Generation/Round, Group Type, Lexicon Size and Word Length. Lexicon Size also interacted with Generation/Round. The model shows that accuracy increased as a function of generation/round (Generation/Round: $\beta = 0.944$, $SE = 0.129$, $p < .001$). Communication accuracy was lower in the chains than in the closed group (Group Type: $\beta = -0.306$, $SE = 0.101$, $p = 0.002$). Participants communicated less accurately in the large lexicon conditions than in the small lexicon conditions (Lexicon Size: $\beta = -0.318$, $SE = 0.102$, $p = 0.002$), and this difference also increased over time (Generation/round \times Lexicon size: $\beta = -0.275$, $SE = 0.128$, $p = 0.031$). In sum, accuracy increased in all groups but more so in the small lexicon than in the large lexicon, and accuracy was also higher in the closed group compared to the chains and.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

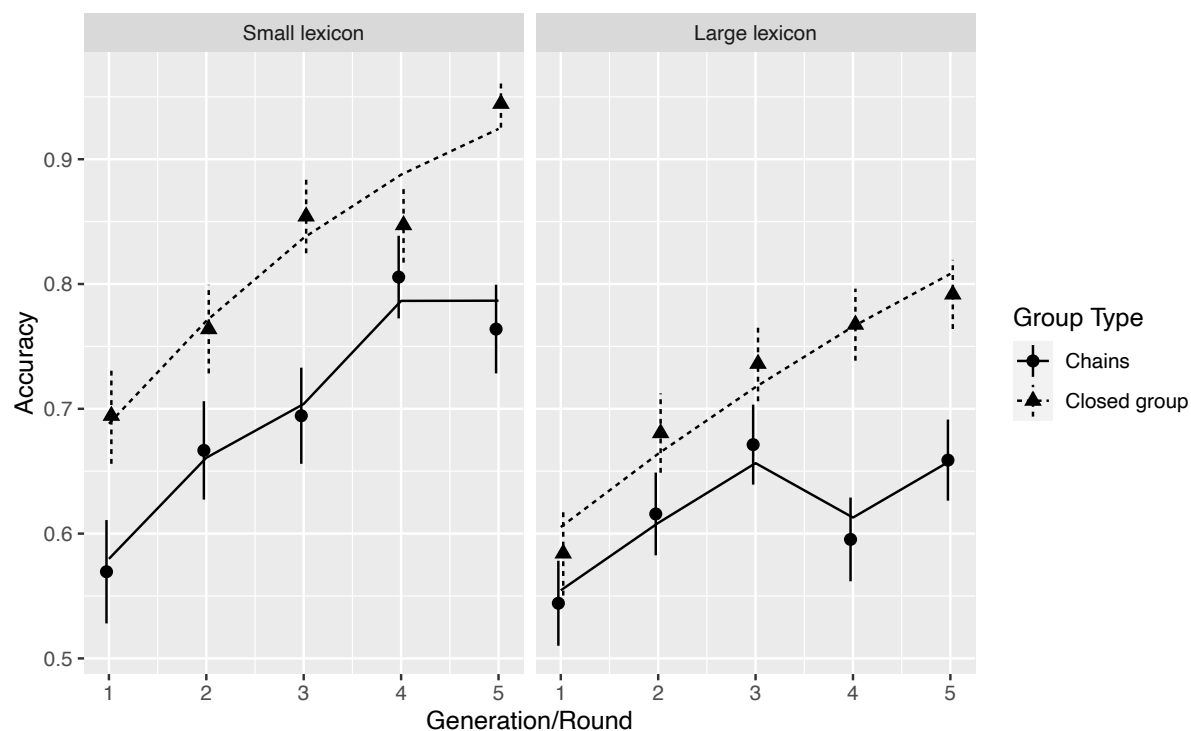


Figure D1. Communication accuracy across generations/rounds in Experiment 1. Points represent descriptive means. Error bars are standard errors. Lines represent model predictions.

Table D1

Fixed factor coefficients for accuracy scores in Experiment 1.

	Estimate	Std. Error	<i>z</i> value	Pr(> <i>z</i>)
Intercept	1.147	0.109	10.553	< .001***
Generation/Round	0.944	0.129	7.346	< .001***
Group type	-0.306	0.101	-3.027	< .001***
Lexicon size	-0.318	0.102	-3.128	.002**
Word length	-0.441	0.043	-10.142	< .001***
Homonymy	-0.122	0.161	-0.758	.448
Generation/round Group type	× -0.238	0.128	-1.863	.062
Generation/round Lexicon size	× -0.276	0.128	-2.155	.031*
Group type × Lexicon size	0.070	0.101	0.694	.488
Generation/Round Group type × Lexicon size	× -0.001	0.128	-0.006	.996

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Appendix E

Accuracy of object selection in Experiment 2

We examined how accurately participants selected the matching object for each label in the comprehension trials in Experiment 2. Accuracy was coded as a binary variable with the value 1 if the participant successfully selected the correct object and 0 if they did not. The results are summarised below depending on the size of the lexicon and the presence in the assigned label of consonant repetitions (Figure E1) or vowel repetitions (Figure E2).

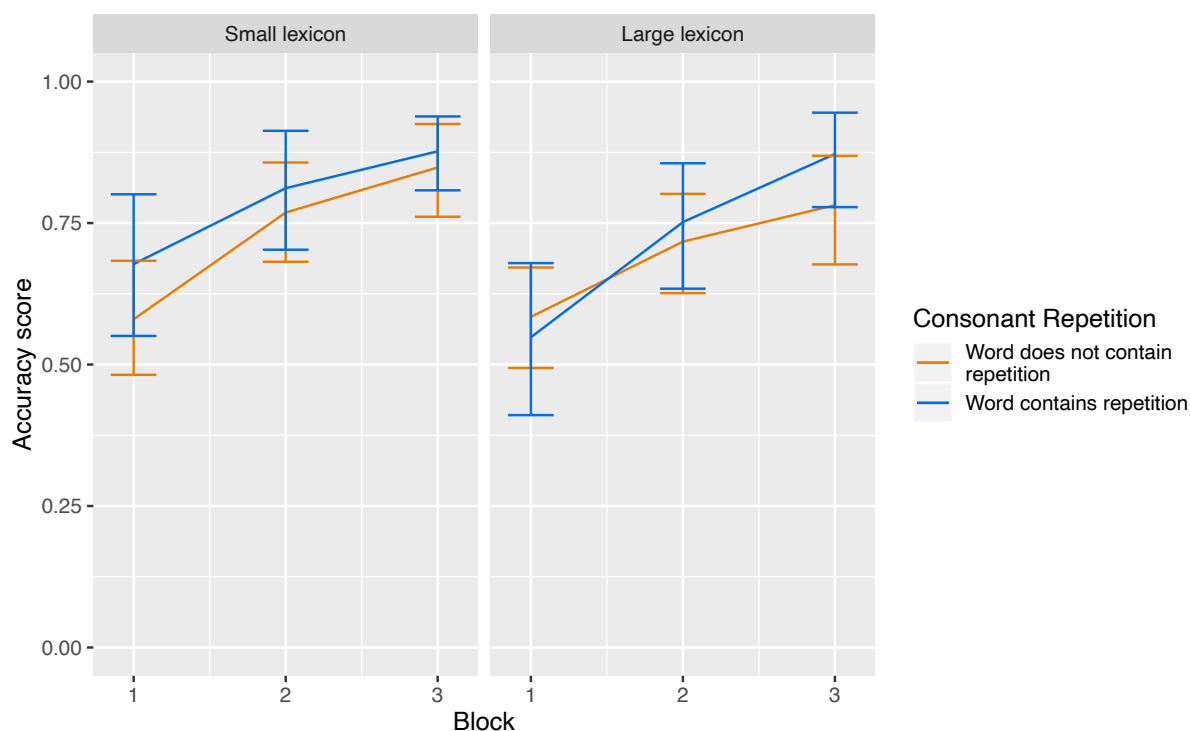


Figure E1. Mean accuracy of comprehension trials in Experiment 2 depending on the absence/presence of consonant repetitions in the word. The left panel shows the results for the small lexicon and the right panel for the large lexicon.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

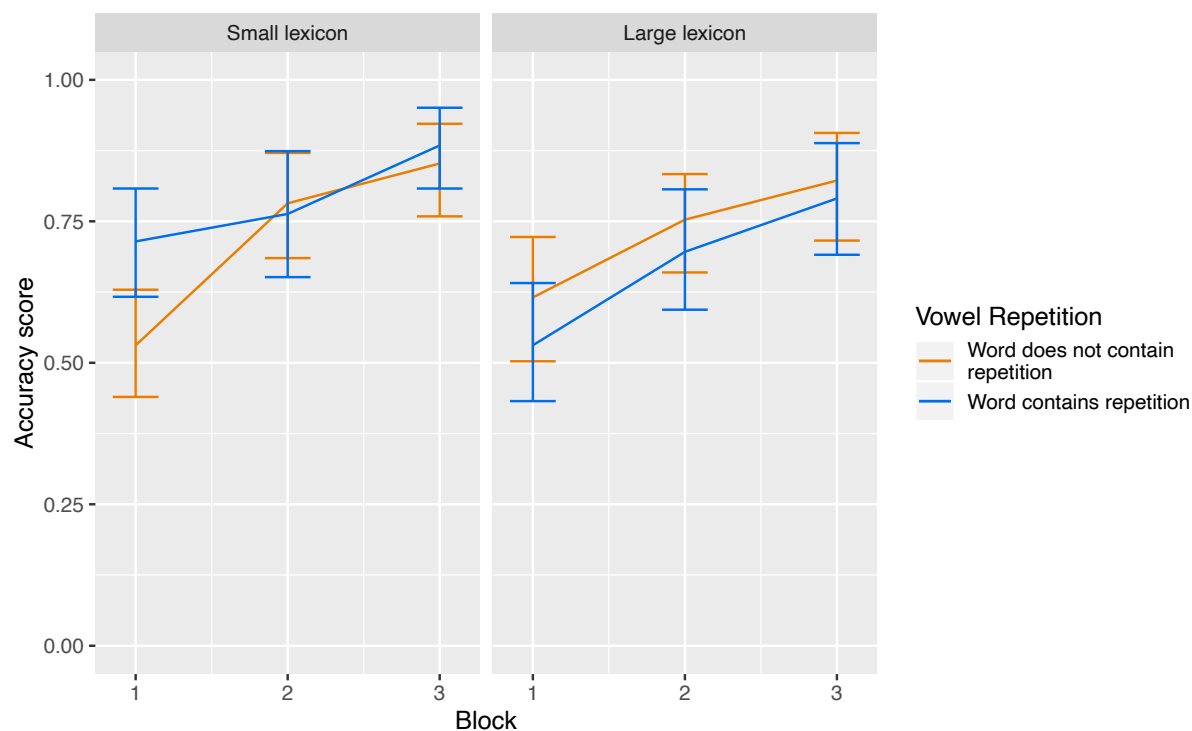


Figure E2. Mean accuracy of comprehension trials in Experiment 2 depending on the absence/presence of vowel repetitions in the word. The left panel shows the results for the small lexicon and the right panel for the large lexicon.

To examine the effects of sound repetitions and lexicon size on accuracy, we conducted a growth-curve analysis similar to that employed in Experiment 1 using the `glmer` function in R. Model comparison indicated that the linear model had a better fit than the quadratic or cubic model, and some of the random slopes had to be removed to achieve convergence. The results of the linear model we report here in Table E1 had the following structure: Accuracy \sim Lexicon size * Consonant repetition * Vowel repetition * Block + (1 + Block | Participant). There was a significant positive main effect of Block ($\beta = 2.164$, SE = 0.323, $p < 0.001$), indicating that participants' overall accuracy improved with more training. There was also a significant interaction between Lexicon size and Vowel repetition ($\beta = -0.880$, SE = 0.295, $p = 0.003$). As can be seen in Figure D2, this effect reflects the higher accuracy for words with vowel repetition in the small lexicon compared to those in the large lexicon. However, there were no main effects showing advantage of consonant or vowel repetitions for comprehension accuracy.

EMERGENCE OF REPETITION THROUGH ITERATED LEARNING

Table E1

Fixed factor coefficients for comprehension accuracy in Experiment 2.

	Estimate	Std. Error	Z	p
Intercept	1.157	0.171	6.757	< .001***
Block	2.164	0.323	6.696	< .001***
Lexicon size	-0.309	0.338	-0.915	.360
Consonant repetition	0.228	0.152	1.501	.133
Vowel repetition	-0.001	0.147	-0.006	.995
Block × Lex size	-0.192	0.620	-0.310	.756
Block × C rep	0.715	0.404	1.770	.077
Lex size × C rep	0.194	0.304	0.637	.524
Block × V rep	-0.731	0.396	-1.847	.065
Lex size × V rep	-0.880	0.295	-2.985	.003**
C rep × V rep	0.368	0.293	-1.256	.209
Block × Lex size × C rep	0.868	0.810	1.072	.284
Block × Lex size × V rep	0.724	0.792	0.915	.360
Block × C rep × V rep	-0.852	0.791	-1.077	.281
Lex size × C rep × V rep	0.162	0.585	0.277	.782
Block × L size × C rep × V rep	0.677	1.577	0.430	.667

Note. ** $p < .01$, * $p < .05$, *** $p < .001$