

# Distinguishing cognitive from historical influences in phonology

Gašper Beguš

University of California, Berkeley  
To appear in *Language*

## Abstract

Distinguishing cognitive influences from historical influences on human behavior has long been a disputed topic in behavioral sciences, including linguistics. The discussion is often complicated by empirical evidence being consistent with both the cognitive and the historical approach. This paper argues that phonology offers a unique test case for distinguishing historical and cognitive influences on grammar and proposes an experimental technique for testing the cognitive factor that controls for the historical factor. The paper outlines a model called catalysis for explaining how learnability influences phonological typology and designs experiments that simulate this process. Central to this discussion are unnatural phonological processes, i.e. those that operate against universal phonetic tendencies and that require complex historical trajectories to arise. Using statistical methods for estimating historical influences, mismatches in predictions between the cognitive and historical approaches to typology can be identified. By conducting artificial grammar learning experiments on processes for which the historical approach makes predictions that differ from the cognitive approach, the experimental technique proposed in this paper controls for historical influences while testing cognitive factors. Results of online and fieldwork experiments on two languages, English and Slovenian, show that subjects prefer postnasal devoicing over postnasal fricative occlusion and devoicing in at least a subset of places of articulation which aligns with the observed typology. The advantage of the proposed approach over existing experimental work is that it experimentally confirms the link between synchronic preferences and typology that is most likely not influenced by historical biases. Results suggest that complexity avoidance is the primary influence of the cognitive bias on phonological systems in human languages. Applying this technique to further alternations should yield new information about those cognitive properties of phonological grammar that are not conflated with historical influences.

Keywords: cognitive influences, historical bias, phonology, artificial grammar learning experiments, experimental fieldwork, sound change

## 1. Introduction

Distinguishing historical (also called cultural or emergent) from cognitive (also called innate) influences on human behavior has been a topic of discussion in any discipline dealing with hu-

man cognition, ranging from psychology to musicology (Altarriba 1993, Gauvain 1995, Nisbett & Norenzayan, Cross 2012). The equivalent dichotomy in linguistics (Kirby et al. 2007, 2014, Griffiths et al. 2008, Reali & Griffiths 2009, Haynie & Bowern 2016, Ferdinand et al. 2019), and specifically, phonology (de Lacy 2006b, de Lacy & Kingston 2013a), is the discussion between approaches that explain recurrent patterns in the sound systems and phonological alternations of world's languages primarily from synchronic grammatical constraints or learnability (COGNITIVE BIAS henceforth; Kiparsky 1995, 2006, 2008, Wilson 2006, Moreton 2008, Hayes et al. 2009, Finley & Badecker 2009, Becker et al. 2011, Moreton & Pater 2012a,b, Baer-Henney & van de Vijver 2012, Finley 2012) and approaches that explain these recurrent patterns as emergent from historical transmission of language in speech communities in time and space (HISTORICAL BIAS henceforth; Hyman 1975, Greenberg 1978, Ohala 1981, 1983, 1993, Blevins 2004, Hansson 2008, Morley 2012, Garrett & Johnson 2013).<sup>1</sup>

One phonological process that exemplifies the discussion around the historical versus cognitive pressures in phonology is postnasal voice alternation. Postnasal voicing is a phonological alternation where voiceless stops /p, t, k/ turn to the corresponding voiced stop [b, d, g] when they appear after nasal sounds [m, n, ɳ] (represented as T → D / N\_).<sup>2</sup> Greek /ton topo/ meaning ‘the place’ surfaces as [tondopo], because /t/ appears after a nasal and turns into a voiced [d] (Pater 2004). Its opposite process is postnasal devoicing (PND; represented as D → T / N\_), where voiced stops /b, d, g/ turn to voiceless stops [p, t, k] after nasal sounds [m, n, ɳ], e.g. in Shekgalagari /χɔmbóná/ meaning ‘to see me’ is realized as [χɔmpóná] Solé et al. 2010). Distribution of phonological alternations across the world’s languages is not uniform – some alternations are substantially more frequent than others. Postnasal voicing (PNV), for example, is a comparatively widespread alternation, occurring in approximately 28 of 629 languages surveyed (Mielke 2018). Postnasal devoicing, on the other hand, is comparatively rare, occurring in only 2–3 languages as a productive alternation among approximately 600 languages surveyed (Hyman 2001, Beguš 2019). Distributional asymmetries like these highlight the debate on whether these asymmetries are driven by cognitive or historical factors.

Discussing cognitive and historical influences in linguistics is complicated both by terminology as well as by matters of substance (Section 1.1). Here, we attempt to clarify several concepts that enter the discussion. We describe with the term COGNITIVE BIAS any influence of both domain-specific grammatical and domain-general cognitive mechanisms that result in typological asymmetries and operate as synchronic tendencies in individual speakers. While the line between domain-specific grammatical and domain-general cognitive influences is sometimes blurred, it is possible to divide between the two in some cases. Cognitive bias influences have long been divided into the SUBSTANTIVE BIAS and the COMPLEXITY BIAS. The first states that phonetically unmotivated processes are dispreferred by the grammar compared to phonetically motivated processes and are, as such, predicted to be less frequent. Complexity bias, on the other hand, states that complex alternations (e.g. those involving more features or those being perceptually more distant) are

---

<sup>1</sup>Various alternative names exist for the two approaches in the literature: The cognitive bias is also known as analytic or learnability bias; the historical bias as the channel bias (Moreton 2008).

<sup>2</sup>The following capital letters represent different groups of segments: N – nasals, such as [m, n, ɳ]; T – voiceless stops, such as [p, t, k]; D – voiced stops, such as [b, d, g]; S – voiceless fricatives, such as [f, θ, x]; Z – voiced fricatives, such as [v, ð, y]; C – consonant; V – vowel. All symbols representing sounds of language follow International Phonetic Alphabet conventions.

dispreferred or more difficult to learn. Complexity bias has been confirmed in numerous studies (see Moreton & Pater 2012a); complex alternations are therefore predicted to be typologically less frequent (for ambiguous experimental outcomes of the substantive bias, see Section 1.1).

Domain-specific grammatical and domain-general cognitive influences occasionally align with substantive and complexity biases. For example, there is no clear cognitive domain-general reason for why the unnatural  $D \rightarrow T / N_{-}$ , e.g. in  $/χumbóná/ \rightarrow [χompóná]$ ) would be dispreferred compared to its exact opposite process  $T \rightarrow D / N_{-}$ , e.g.  $/ton topo/ \rightarrow [tondopo]$ ). Under the influential proposal of Phonetically Based Phonology (Hayes 1999, Hayes & Steriade 2004), grammatical constraints are phonetically grounded, so ultimately, it is the phonological GRAMMAR that makes the unnatural alternations dispreferred under this proposal. For example, postnasal devoicing is dispreferred by the substantive bias because it operates against phonetic naturalness, but it is not more complex than the natural postnasal voicing. In the absence of domain-general cognitive explanation, this particular dispreference can be understood as a domain-specific bias. On the other hand, dispreference against complex alternations that target more than a single feature and that also involve perceptually distant allophones, such as postnasal fricative occlusion and devoicing (PFOD;  $Z \rightarrow T / N_{-}$ , e.g.  $/m̪sona/ \rightarrow [mpona]$  in Pedi; Dickens 1984) can be explained with domain-general cognitive mechanisms. Moreton et al. (2017) adopt concept learning and argue that complexity dispreference in phonology might have the same underlying mechanisms as other cognitive processes, where concepts that require more features to be described are more difficult to learn.

These grammatical dispreferences against phonetically unnatural (substantive bias) or complex processes (complexity bias) can result from several mechanisms. For example, learning asymmetries (Wilson 2006) can underly both complexity and substantive bias. Featurally complex concepts are difficult to learn not only in phonology, but in other domains as well (Moreton et al. 2017), which would explain the relative rarity of complex alternations. On the other hand, it has been assumed that phonetically unnatural processes are difficult to learn; thus, learning difficulties are also the basis of the substantive bias. Another mechanism that can underlie the substantive bias is the inability of the grammar to accommodate an unnatural process (e.g. as a tendency to reanalyze an unnatural process as a process that conforms to naturalness; Kiparsky 2006, 2008).

Finally, perceptual forces can influence typology. Whether perceptual influences should be analyzed as part of complexity or substantive bias is an open question. While this distinction is primarily terminological in nature and does not crucially affect the results of this paper, we analyze perceptual forces as part of the complexity bias. First, perceptual distance can often be directly analyzed in terms of formal featural complexity. Even if two alternations are featurally equally complex, the tendency towards minimizing some distance can be understood as part of complexity (somewhat diverging from the literature in Wilson 2006 and White 2014): similarity is less complex than dissimilarity. Finally, while the substantive bias has traditionally been assumed to be limited to domain-specific processes, the preference towards perceptual similarity is not domain specific, but likely a domain-general cognitive mechanism (e.g. in the visual domain, see Schloss & Palmer 2011). A tendency to keep phonology perceptually minimal (P-map; Steriade 2001) can thus be analyzed as part of the complexity bias, while substantive bias is reserved for a domain-specific dispreference against phonetically unmotivated or unnatural processes.

With HISTORICAL BIAS, on the other hand, we describe those properties that emerge when articulatory and perceptual tendencies in speech production and perception operate in language use when language is transmitted in space and time across generations of speakers and accumulate

in phonological or phonetic processes. Production of speech is a highly variable process. Variation in speech production is biased. For example, stops [p, t, k] tend to be produced with a higher degree of voicing (more like [b, d, g]) if they appear after nasal stops [n, m, ɳ] (Hayes & Stivers 2000, Davidson 2016, 2018). Such minor biased phonetic variation, motivated by articulatory or perceptual factors, gives rise to phonological alternations via sound change and the process of phonologization (Hyman 2013).

Articulatory and perceptual forces differ in terms of their relatedness to cognitive influences. There is little cognitive influence on articulatory forces: automatic articulatory tendencies are often motivated by motor planning mechanisms dissociated from higher level cognitive processes. For example, the reason for why postnasal voicing (/ton topo/ → [tondopo]) is motivated is purely mechanic: when the velum rises to close the nasal cavity (from [n] to [t]), volume of the oral cavity increases and as the velum is rising, airflow can continue for some period of time. Because increased volume and airflow promote voicing, stops in postnasal position feature more voicing into closure compared to other positions (Hayes & Stivers 2000, Coetzee & Pretorius 2010, Davidson 2016, 2018). These articulatory mechanisms are purely mechanical influences and have little connection to cognition. On the other hand, the historical bias includes perceptual influences as well. Perception is not dissociated from cognition. However, perceptual influences from the historical bias perspective are distinct from synchronic perceptual influences. First, the perceptual mechanisms in Ohala's (1981) terms result in sound change by operating gradually in a speech community in space and time (i.e. historical bias influences). Second, perceptual mechanisms often (but not always) require some minor phonetic variation that originates in non-cognitive articulatory forces. For example, longer vowel duration before voiced (vs. voiceless) stops is likely caused by perceptual enhancement (Kluender et al. 1988), but the initial distribution on which perceptual forces operate to enhance it likely stems from articulatory factors (Beguš 2017).

In sum, perceptual influences can be part of both the cognitive and historical biases. It is possible that perceptual forces result in typological patterns because of cognitive synchronic preferences for phonological alternations to be perceptually minimally distant (P-map; Steriade 2001) or because of hypo- and hypercorrection (Ohala 1981). The advantage of our proposal is that we can disambiguate between the two by comparing experimental results against independent historical samples and against those samples in which potential perceptual forces are related to a synchronic phonological alternation. For example, Ohala's (1981) perceptual forces should operate at equal rates in systems in which target segments are not part of a synchronic alternation as well as on those in which it is. Because our results suggest the opposite, we can argue that, even if the observed results are part of perception, they are driven by cognitive factors, not perceptual mechanisms operating in space and time (see Section 2.3).

The impact of these factors on phonological typology is central to phonology and linguistics in general and has far-reaching consequences. It is likely that both influences affect phonological typology (Hyman 2001, Myers 2002, Moreton 2008, Moreton & Pater 2012a,b, de Lacy & Kingston 2013b), but to distinguish the two major influences has been a challenging task, primarily because empirical evidence tends to support both approaches equally well (for non-experimental attempts, see de Lacy 2006b; de Lacy & Kingston 2013a). Distinguishing cognitive factors from those aspects of phonology that are emergent from the historical transmission of language in space and time is a desirable task: it would yield a better understanding of which properties of phonology and consequently of human language capacity are influenced by cognitive processes and should therefore be captured by models of grammar.

## 1.1 The duplication problem

Empirical research in the discussion of cognitive vs. historical forces in phonology is complicated by several confounding factors: (i) ambiguous empirical evidence, (ii) difficulty in disassociating cognitive influence from phonetic variation, and (iii) lack of elaborate models that link learning and typology.

First, empirical evidence often supports both approaches equally well. Numerous experimental studies have established a link between learnability and typological distribution: Phonological alternations that are more difficult to learn in laboratory settings are typologically less frequent (for an overview, see Moreton & Pater 2012a,b). On the other hand, phonetic variation, motivated by articulatory or perceptual mechanisms, has been shown to result in active phonological alternations via the process of phonologization. The stronger the phonetic tendency, the more frequent the resulting phonological alternation (Blevins 2004). Crucially, even if learnability differences that match the observed typology are experimentally confirmed, the typological distribution can nevertheless be explained within the historical bias approach. Experimental studies testing learnability only rarely target predictions that cannot be attributed to the historical bias (for one method, see Moreton 2008 and Yu 2011 who points to some of its shortcomings).

For example, high frequency of postnasal voicing (/ton topo/ → [tondopo]) and low frequency of postnasal devoicing (/χɔmbóná/ → [χɔmpóná]) can be explained by grammatical or learnability dispreferrences against a phonetically unmotivated process. Some proposals even argue that the grammar is incapable of accommodating phonetically unnatural processes (Kiparsky 2006, 2008). The same typological asymmetry can be explained under the historical bias approach: phonetic variation caused by mechanic factors is present in the first, but absent in the latter, which is why sound change can produce postnasal voicing, but not devoicing (unless a particular combination of minimally three sound changes conspire to produce the unnatural result; see Section 2.2).

This duplication of evidence goes even further: the complexity bias too has a duplicate historical explanation. Complex alternations such as PFOD (/mβona/ → [mpona]) can be rare because they are more difficult to learn, computationally more complex, or require a large perceptual distance between the target and result (all cognitive biases) or they can be rare because complex alternations require two sound changes and two historical events (sound changes) generally have lower probabilities than a single event (historical bias) (Bell 1970, 1971, Greenberg 1978, Cathcart 2015, Morley 2015, Beguš 2019).

While experimental results consistent with substantive bias have been reported in some studies (Wilson 2006, Carpenter 2010), many other experimental studies fail to find positive evidence for substantive bias (Pycha et al. 2003, Kuo 2009, Skoruppa & Peperkamp 2011 via Moreton & Pater 2012a,b and Seidl et al. 2007, Do et al. 2016, Glewwe 2017, Glewwe et al. 2017, Do & Havenhill 2021). As already mentioned, historical bias makes exactly the same predictions as the complexity and substantive bias approaches. Featureally complex alternations are predicted to be less frequent than simple alternations because they result from multiple sound changes (for a discussion, see Section 2.1). Phonetically motivated alternations are likewise predicted to be more frequent by the historical bias approach: they arise from a single sound change, while non-natural alternations require a combination of sound changes. This means that experimental results confirming learning biases for a given process almost always have an alternative historical explanation. This problem is called the ‘duplication problem’ henceforth.

Second, it is possible that the frequency of sound changes itself is crucially affected not only by

the robustness of phonetic variation (as assumed by the historical bias approach), but by learnability as well. In other words, what the historical bias approach assumes to be exclusively an emergent factor, i.e. frequency of a sound change based on historical factors such as articulatory or perceptual motivation of phonetic variation, can be influenced by learnability or other cognitive factors (Kiparsky 1995, 2006, 2008, Moreton 2008). For example, Moreton (2008) argues that cognitive factors determine when some phonetic variation will result in a sound change/synchronic alternation: phonetic variation of equal size can result in sound change or not. Cognitive biases are responsible for this asymmetry (for problematic aspects of the notion of equal size in phonetic variation/precursors, see Yu 2011).

Finally, elaborate models of how exactly learning influences the observed typology are lacking. The cognitive bias approach often uses first language (L1) acquisition to explain the link between learnability and typology, whereby learners fail to learn a process or restructure their phonological grammar based on learning biases. However, many experimental studies have shown that, given enough exposure, any alternation can be learned (Hayes et al. 2009, Coetzee & Pretorius 2010, Hayes & White 2013, White 2014, Avcu 2018). Additionally, human L1 learners get even more exposure to primary linguistic data compared to subjects in laboratory experiments and are able to reproduce language input with a high degree of faithfulness past some developmental stage (Kong et al. 2012a). Numerous studies have also confirmed the anti-alternation bias: Learners prefer no alternation to any alternation (Wilson 2006, Tessier 2012 and the literature therein). The fact that sound change gives rise to active phonological alternations means that learnability alone does not crucially affect the operation of contextually limited sound changes that result in alternations. Showing how learnability differences affect the typology is thus not trivial. The position that anything can be learned given enough exposure is supported by phonological data as well: unnatural alternations that operate in the exact opposite direction from universal phonetic tendencies that result from a combination of sound changes are attested as productive alternations in languages such as Tswana and Shekgalagari (postnasal devoicing; Coetzee & Pretorius 2010). This means that learners were able to learn a phonological grammar with an alternation as phonetically unnatural as postnasal devoicing. Second, while L1 acquisition is a potential source of sound change, it has been known that sound change operates within early adolescent and adult populations as well (Labov 1994, Sankoff & Blondeau 2007 and the literature therein). Deriving the typology with L1 learning thus needs to account for sound change in adult population. Third, that phonological processes in L1 acquisition are often different from sound change typology is often an overlooked and underresearched observation that poses a challenge to the L1 approach to sound change (Bybee 2001). Finally, some computational models suggest that learnability differences might not be sufficient in deriving surface typology (Rafferty et al. 2013).

In sum, it is nontrivial to show how cognitive biases result in phonological typology. To address this difficulty, we propose a model called CATALYSIS which outlines a possible mechanism on how synchronic cognitive bias can accelerate the operation of a sound change and consequently directly influence the typology. We test simulations of catalysis experimentally. Because we can argue that the historical bias makes the opposite predictions compared to cognitive bias based on the estimation of historical probabilities (Section 2.2) behind phonological processes, we can dissociate the two influences and test them against the observed typology.

## 1.2 Testing the hypotheses on mismatched predictions

This paper argues that phonology offers a unique test case for the debate on cognitive vs. historical influences that avoids the duplication problem outlined above if we adopt some assumptions proposed in Beguš 2019. Key to this discussion are the unnatural phonological processes, defined as those that operate in exactly the opposite direction from universal phonetic tendencies in Beguš 2019 (such as final voicing or postnasal devoicing). Beguš (2019) argues that sound change cannot operate in an unnatural direction and that unnatural alternations require at least three sound changes to arise (the so-called Minimal Sound Change Requirement; Section 2.1). This crucial condition makes unnatural alternations the best testing ground for distinguishing cognitive from historical influences on typology. The requirement that three historical events, i.e. sound changes, need to occur for an alternation to arise can be exclusively ascribed to the historical factor. Even if learnability influences the operation of individual sound changes, the requirement that three sound changes need to operate in a speech community in a given timeframe to produce an unnatural alternation is exclusively the influence of the historical factor.

Additionally, the historical and the cognitive bias approaches make opposing predictions regarding the relative frequencies of unnatural and complex alternations. These crucial mismatches allow us to test the influences of one approach while controlling for the other. A statistical technique proposed in Beguš 2020 for estimating the probabilities of sound changes based on diachronic factors facilitates the identification of mismatches in predictions between the two approaches to typology.

This paper presents experiments that simulate a development from a complex to an unnatural process, and thus, experimentally tests the mismatched predictions between the historical and cognitive approach. A historical trajectory required for an unnatural alternation, postnasal devoicing, is first identified and the historical probabilities of each stage in its development are estimated (based on Beguš 2020, 2019). The experiments test learning of the last two stages in the historical development of postnasal devoicing. The cognitive and the historical bias approaches make opposing predictions for the two stages. The historical bias approach predicts the unnatural stage (postnasal devoicing, e.g. /b/ → [p] / m\_) to be less frequent than the complex stage (postnasal fricative occlusion and devoicing, e.g. /v/ → [p] / m\_), while the cognitive bias approach predicts the opposite. If experimental results support the cognitive bias approach that matches the observed typology, the link between cognitive bias and typology is supported without the duplication problem – historical bias is likely not responsible for a typological distribution that operates against its predictions.

The experiments are conducted within the artificial grammar learning paradigm (Albright & Do 2019, overview in Moreton & Pater 2012a,b), but several diversifying factors are introduced. To diversify the design, online experiments targeting non-specific groups of speakers are combined with experimental fieldwork with a high number of subjects and a relatively high train-to-test ratio of stimuli. Supervision of the experiment by a research assistant at least partly addresses the concern of subjects' attention in online experiments. Additionally, the experiments are conducted on L1 speakers of English and Slovenian, two languages with different realizations of the feature [ $\pm$ voice] and different frequencies of segments tested in the experiment, which at least partially controls for interference from first language, a long-standing objection to the artificial grammar learning experimental paradigm.

The results suggest that subjects prefer the response consistent with the unnatural alternation

compared to the complex alternation when presented with ambiguous stimuli for at least a subset of places of articulation. These results are in line with the observed synchronic typology and operate against what is predicted by statistical modeling of the historical bias. This means that the experimentally confirmed link between cognitive bias and the observed typology is likely not due to the historical bias.

## 2. Background: Unnatural phonology

Unnatural phonological alternations are defined as those that operate against universal phonetic tendencies. Universal phonetic tendencies are defined as articulatory or perceptually motivated (Garrett & Johnson 2013) phonetic processes that ‘passively operate in speech production cross-linguistically and result in typologically common phonological processes’ (Beguš 2019:691).

Based on a typological and historical study, Beguš (2019) argues that unnatural segmental alternations always arise from a specific combination of three sound changes called the Blurring Process: (i) a sound change creates a complementary distribution, (ii) a sound change that targets a subset of segments in the complementary distribution, and (iii) a sound change that undoes the original complementary distribution (Beguš 2019) (for a schematic representation, see Table 1). The Blurring Process proposal argues that postnasal devoicing results from a combination of three sound changes in all surveyed cases. This is a crucial assumption that this paper adopts from Beguš 2019.<sup>3</sup> First, voiced stops [b, d, g] fricativize to [v, ð, y] except postnasally ( $D > Z / [-\text{nas}]_-$ ,<sup>4</sup> e.g. [bamba] > [vamba]). Then, voiced stops [b, d, g] devoice to [p, t, k] unconditionally ( $D > T$ , e.g. [vamba] > [vampa]), but because at this stage they only surface postnasally, the resulting devoicing appears to be limited to postnasal position.<sup>5</sup> Finally, voiced fricatives [v, ð, y] occlude back to stops [b, d, g] ( $Z > D$ , e.g. [vampa] > [bampa]), which results in a synchronic PND ( $D \rightarrow T / N_-$  or /bamba/ → [bampa]; Table 1).

### 2.1 Minimal Sound Change Requirement

Unnatural processes are crucial for identifying mismatched predictions between cognitive and historical biases because they have well-structured complex histories. Beguš (2019) provides a formal argument that at least three sound changes are required for an unnatural alternation to arise. In abstract terms, a set of segments represented by feature matrix A can alternate with a set of segments B in environment X. If B is phonetically universally preferred in X, then  $A \rightarrow B /$

---

<sup>3</sup>To be sure, it cannot be proven that PND cannot operate as a single sound change, but see Beguš 2019, 2018 for arguments that provide direct evidence for intermediate stages and for an example that strongly suggests the unnatural intervocalic devoicing cannot result from a single sound change (Beguš 2018).

<sup>4</sup>Feature nasal is abbreviated as [ $\pm$ nas] henceforth.

<sup>5</sup>While postnasal position facilitates voicing into closure, stop closure is nevertheless antagonistic to voicing, and speakers need to actively adjust for voicing even postnasally. For an extensive discussion on naturalness of unconditioned stop devoicing when stops appear in the postnasal position, see Beguš 2019. There exists phonetic evidence that supports this assumption: English voiced stops are relatively frequently realized as partially voiceless in the postnasal position (Davidson 2016, 2017, Beguš 2019) due to anti-voicing effects of closure.

Table 1: Blurring Cycle (schematic; left) yielding PND (right) (from Beguš 2020).

	Blurring Cycle	PND	Schematic example	
1.	$B > C / \neg X$	$D > Z / [-\text{nas}]_-$	[bamba]	$>$ [vamba]
2.	$B > A$	$D > T$	[vamba]	$>$ [vampa]
3.	$C > B$	$Z > D$	[vampa]	$>$ [bampa]
Result	$B \rightarrow A / X$	$D \rightarrow T / [+ \text{nas}]_-$	/bamba/	$\rightarrow$ [bampa]

$$\begin{array}{ccc}
 \text{Natural alternation} & & \text{Unnatural alternation} \\
 \begin{array}{ccccc}
 A & \rightarrow & B & / & X \\
 \left[ \begin{array}{c} \phi_1 \\ \vdots \\ \phi_{1+n} \end{array} \right] & \left[ \begin{array}{c} \alpha \\ \vdots \\ \beta \end{array} \right] & \rightarrow & \left[ \begin{array}{c} \neg\alpha \\ \vdots \\ \beta \end{array} \right] & / \quad X \\
 \end{array} & & \begin{array}{ccccc}
 B & \rightarrow & A & / & X \\
 \left[ \begin{array}{c} \phi_1 \\ \vdots \\ \phi_{1+n} \end{array} \right] & \left[ \begin{array}{c} \neg\alpha \\ \vdots \\ \beta \end{array} \right] & \rightarrow & \left[ \begin{array}{c} \alpha \\ \vdots \\ \beta \end{array} \right] & / \quad X
 \end{array}
 \end{array}$$

Figure 1: Natural and unnatural alternations.

$X$  is a natural alternation. The opposite process,  $B \rightarrow A / X$  is unnatural, where  $A$  is phonetically dispreferred in  $X$ . In feature matrix notation, a change of a feature value of  $\phi_1$  from  $\alpha$  to  $\neg\alpha$  in environment  $X$  given a constant set of other features  $\beta$ , where  $\neg\alpha$  is preferred in  $X$ , is natural. A change  $\neg\alpha$  to  $\alpha$  in  $X$  given  $\beta$ , where  $\alpha$  is dispreferred in  $X$  is unnatural (schematized in Figure 1).

Beguš (2019) assumes that sound change is always natural and cannot produce an unnatural alternation in a single step. It is also assumed (following the Minimality Principle in Picard 1994) that sound change almost always targets a single feature value (for a detailed discussion, see Beguš 2019).<sup>6</sup>  $\neg\alpha > \alpha$  can thus not operate under a constant set of feature values  $\beta$  in the matrix. To get  $\neg\alpha > \alpha$ ,  $\beta$  first needs to change to  $\neg\beta$ . Under this new condition,  $\neg\alpha > \alpha$  can become phonetically motivated and can operate as a sound change. To get the full unnatural alternation,  $\neg\beta$  then has to change back to  $\beta$ . Minimally three independent historical events, i.e. sound changes, are thus required for unnatural process to arise (Table 2). For details, see Beguš 2019.

## 2.2 Historical probabilities

The historical bias influences on phonological typology can be quantitatively estimated by combining the concept of alternations requiring a specific number of historical events (i.e. sound changes) with the estimation of individual probabilities of these historical events (Beguš 2019,

<sup>6</sup>Most of the predictions in Beguš 2019 still hold even if the Minimality Principle is a tendency rather than a hard rule.

Stage	1.	$>$	2.	$>$	3.	$>$	4.
$\left[ \begin{array}{c} \phi_1 \\ \vdots \\ \phi_{1+n} \end{array} \right]$	$\left[ \begin{array}{c} \neg\alpha \\ \vdots \\ \beta \end{array} \right]$		$\left[ \begin{array}{c} \neg\alpha \\ \vdots \\ \neg\beta \end{array} \right]$		$\left[ \begin{array}{c} \alpha \\ \vdots \\ \neg\beta \end{array} \right]$		$\left[ \begin{array}{c} \alpha \\ \vdots \\ \beta \end{array} \right]$

Table 2: Changes in feature values in a Blurring Process (from Beguš 2019).

2020). In other words, the HISTORICAL PROBABILITY ( $P_\chi$ ) of an alternation, i.e. the probability that an alternation arises based on historical factors, can be estimated from the number of sound changes the alternation requires and their respective probabilities (adopted from Beguš 2020).

The probability of each individual sound change  $S_i$  (in 1 below) is estimated based on historical typological surveys of sound changes (Kümmel 2007) from the number of occurrences (successes) of a sound change  $S_i$  and the number of languages surveyed (successes and failures). 95% confidence intervals adjusted for bias and skewness ( $BC_a$ ) are estimated using non-parametric bootstrap (Efron 1979, 1987). The historical probability of alternation  $A_j$  that requires more than a single sound change is estimated as a joint probability (a simple product of each individual change; for the assumption of independence, see Beguš 2020) of individual sound changes corrected for ordering of sound changes ( $n!$ ). The historical probability is again estimated with non-parametric bootstrap (in (2) below). For all details and underlying assumptions of the model as well as for a discussion on representativeness of samples, see Beguš 2020.

(1)

$$P_\chi(S_i) = \frac{\text{number of languages with sound change } S_i}{\text{number of languages surveyed}}$$

(2)

$$P_\chi(A_j) = \frac{\prod_{i=1}^n P_\chi(S_i)}{n!}$$

### 2.3 Mismatches

Each stage in the development of PND (Blurring Process; Table 1) has a historical probability that can be estimated using the bootstrapping technique (outlined in Section 2.2 above and in Beguš 2020). Table 3 lists counts of languages with the three sound changes that operate in the development of PND in the historical sample given in Kümmel 2007. The first sound change, D > Z / [-nas]/V\_(V), results in a synchronic alternation between voiced stops that surface postnasally and voiced fricatives that surface elsewhere (D → Z / [-nas]\_, e.g. /b/ → [v] / [-nas]\_). One phonological feature, [±continuant], is manipulated in this alternation. When the second sound change, unconditioned devoicing of voiced stops (D > T, e.g. [b] > [p]) occurs, the resulting alternation is PFOD: Voiced fricatives in the elsewhere condition alternate with voiceless stops postnasally (Z → T / N\_, e.g. /v/ → [p] / N\_). This alternation manipulates two phonological features, [±continuant] and [±voice], but the latter is automatic, because at that point, the system lacks voiced stops altogether. Finally, the third sound change, occlusion of fricatives (Z > D, e.g. [v] > [b]) results in the unnatural PND where underlying voiced stops surface as voiceless postnasally and as voiced elsewhere. PND manipulates one feature, [±voice]. Table 4 shows estimated Historical Probabilities ( $P_\chi$ ) (with 95%  $BC_a$  confidence intervals) of each of the three stages in the Blurring Process that result from operation of the first, the first and second, and all three sound changes.

A clear mismatch in predictions between the historical bias and the cognitive bias emerges (Beguš 2020) if it is assumed that a single sound change cannot operate in the phonetically unmotivated direction (e.g. PND is not a possible sound change; for an overview of the literature that holds this view and a detailed argumentation, see Beguš 2019). Given this assumption, historical

Table 3: Counts of languages that feature the sound changes required for PND to arise (Counts) and counts of languages surveyed (Surveyed) for each corresponding sound change (from Beguš 2020). Counts are based on a historical sample in Kümmel 2007.

Sound change	Example	Counts	Surveyed
	bamba		
D > Z / [-nas]/V_(V)	vamba	47	294
D > T	vampa	15	263
Z > D	bampa	17	216

Table 4: Estimated historical probabilities ( $P_\chi$  in %) with lower (Lo.) and upper (Up.) 95% BC<sub>a</sub> confidence intervals and the number of phonological features the resulting alternation manipulates. Arrows in the last two columns indicate changes in the predicted probability (compared to previous stage) of the historical bias ( $P_\chi$ ) and the complexity bias within the cognitive bias ( $P_{cplx}$ ); table from Beguš 2020.

Sound change	Alternation	$P_\chi$	Lo.	Up.	Features	$P_\chi$	$P_{cplx}$
	No alternation	83.5			0		
D > Z / [-nas]_	D → Z / [-nas]_	16.0	11.9	20.1	1	↓	↓
D > T	Z → T / [+nas]_	0.5	0.3	0.8	2	↓	↓
Z > D	PND	0.01	0.006	0.02	1	↓	↑

bias predicts that unnatural alternations are significantly less frequent than natural alternations: the probability of an alternation decreases with each additional sound change, regardless of the complexity of the resulting alternation. The probability of  $n + 1$  events is always lower than the probability of  $n$  events and this influence is exclusively the result of the historical bias, as sound changes are historical events operating in speech communities in time and space. On the other hand, complexity bias predicts featurally complex alternations to be typologically less frequent than featurally simple alternations (see Section 1). This prediction works regardless of what the underlying mechanism behind compexity bias are: structural (featural) complexity akin to concept learning (Moreton et al. 2017) or perceptual distance (in terms of P-map; Steriade 2001). Complexity avoidance where featurally or perceptually complex processes are tested against simple processes is experimentally confirmed in numerous studies (Moreton & Pater 2012a). The cognitive bias approach thus predicts PND to be more frequent than PFOD, because the first requires manipulation of one feature less than the second (or is perceptually less distant). Conversely, the historical bias approach predicts the unnatural PND to be less frequent than PFOD, because the first requires one sound change more (Beguš 2020).

Because no learning differences are observed between the natural postnasal voicing and the unnatural PND (Do et al. 2016, Do & Havenhill 2021), substantive bias alone is likely not responsible for the observed typological asymmetries. A clarification is warranted here (from Beguš 2020). The only type of learning differences between the natural and unnatural alternations (substantive bias) are those that involve articulatory effort: Segments that require greater articulatory effort are acquired later (Broselow 2018, Kong et al. 2012b, Clark & Bowerman 1986), which means that, for example, voiced stops word-finally are acquired later in L1 and L2 acquisition. Do & Havenhill

(2021) also find limited evidence that exposure to production increases natural responses in the postnasal position in adults, but it is not unexpected that an experiment with production (more exposure) improves learning in the variable condition.<sup>7</sup> The mechanism underlying this articulatory learning is likely different from phonological learning as observed, for example, as the complexity bias (Moreton & Pater 2012a,b) and, in fact, more consistent with the historical bias approach. Unnatural alternations require production of segments that require more articulatory effort. That learning of articulatorily more complex gestures is more difficult is not surprising. The very same mechanism is responsible for universal phonetic tendencies operating in the adult population: Articulatory effort of different gestures causes varying degrees of phonetic variation that results in phonological alternations via phonologization. This mechanism thus falls within the historical bias approach (Beguš 2020). Deriving typology from articulatory learning within the cognitive bias approach remains problematic, as children are able to replicate their linguistic input with high degree of faithfulness past some developmental age (Kong et al. 2012a). One of the main pieces of evidence against L1 articulatory effort influencing the typology comes from the observation that many articulatory adjustments in L1 acquisition do not result in sound changes if the variation is not present in the adult language as well (Bybee 2001).

In sum, the last sound change in the development of PND (Blurring Process) decreases the historical probability of the resulting alternation, but increases its synchronic preference, because it reduces its structural or perceptual complexity. This means that the historical bias approach predicts the unnatural alternation PND to be less frequent than PFOD, whereas the cognitive bias approach makes the opposite prediction.

These predictions can be directly tested against the observed typological distributions (Beguš 2020). Historical bias predicts PFOD to be approximately 50-times more frequent typologically compared to PND as estimated based on counts in Table 4 ( $P_x$  of PFOD is 0.5% [0.3%, 0.8%];  $P_x$  of PND is 0.01% [0.006%, 0.02%]). A typological survey in Beguš 2019 shows that PFOD is indeed more frequent than PND, but not as much as would be predicted by the historical bias. The survey in Kümmel 2007 and Beguš 2019 is complemented with the most comprehensive survey of phonological rules, the P-base (Mielke 2018). In only one system (Sie) is PFOD reported as a synchronic alternation in the P-base (Mielke 2018).<sup>8</sup> Approximately 3–6 languages are reported to feature PFOD in Beguš 2020. On the other hand, in two closely related languages (Tswana and Shekgalagari), PND is confirmed as a productive alternation. In one additional language PND is reported, but not yet confirmed (Buginese). In seven languages PND is reported as a sound correspondence resulting from a combination of sound changes. Exact counts are difficult to determine because productivity of a synchronic alternation needs to be experimentally confirmed and experimental work on many of these languages is lacking. Based on available typological data, however, PFOD seems to be substantially less than 50-times more frequent than PND (Beguš 2020).

<sup>7</sup>This is especially so because English postnasal voiceless stops have a considerable amount of voicing into closure (Davidson 2016, 2017). In other words, English has a gradual phonetic process of postnasal voicing which operates during the production and can thus affect the experimental results. Do & Havenhill (2021) report no differences in learning in the categorical condition. Moreover, test phase in Do & Havenhill 2021 is exclusively orthographic, which can introduce confounds.

<sup>8</sup>The query for searching of the P-base consisted of [−stop][+voiced] condition for search input and [+stop][−voiced] condition for search output.

Table 5: Contingency table and Fisher’s Exact Test of counts of occurrence (successes vs. failures) of sound change in an independent sample and in cases where sound change simplifies an alternation based on Kümmel 2007, Beguš 2020, and Mielke 2018.

	✓	✗	Total
Independent	44	172	216
Simplifies	6	4	10
Fisher’s Exact Test:	$p < 0.01$		

Another, more testable mismatch in predictions emerges from the proposed technique of estimating Historical Probabilities. The last sound change, occlusion of voiced fricatives to stops ( $Z > D$ , e.g. [v] > [b]), that reduces the complexity of the resulting alternation appears to operate more frequently than expected by only the historical bias approach (Beguš 2020). To test this observation, the historical probability of a sound change operating in an unconditioned sample is compared to the historical probability of the same sound change operating on languages that already undergo the first two sound changes in the Blurring Process (i.e. where the sound change simplifies an alternation). The probability of occlusion of fricatives ( $Z > D$ ), the last sound change in the Blurring Process that leads to PND, operating independently (44-times in a sample of 216 languages in Kümmel 2007) is compared to the probability of the same sound change when it operates as the last sound change in the Blurring Process (occlusion targets at least one place of articulation in six out of ten languages with PFOD in the sample in Beguš 2020, and P-base; Mielke 2018). Counts are given in Table 5; for details, see Beguš 2020. Occlusion of fricatives occurs significantly more frequently as the last sound change in the development of PND compared to the independent sample ( $p < 0.01$ , Fisher’s Exact Test; Beguš 2020). In other words, sound change that simplifies a featurally complex alternation and thereby learnability of the alternation operates significantly more frequently than predicted by the historical bias approach. While the sample in the condition group is small, the comparison suggests that cognitive bias influences frequency of sound change in this type of cases (Beguš 2020).

### 3. Catalysis

As mentioned in Section 1, elaborate models of how cognitive preferences and typology are connected are lacking and several objections have been raised against the existing models. Here we propose one possible mechanism for the direct link between cognitive bias and typology.

Phonetic variation resulting from universal articulatory or perceptual phonetic tendencies is the underlying condition for every non-analogical sound change (see Garrett & Johnson 2013, Moreton 2008, Yu 2011, and Section 1). For example, a sound change that operates in the development of PND is occlusion of fricatives ( $Z > D$ , e.g. [v] > [b]). It is a well-documented sound change (Kümmel 2007) with a relatively well-understood phonetic motivation: articulatory targets for fricatives require greater precision compared to the stops (Ladefoged & Maddieson 1996:137). Reducing the precision of articulatory targets can result in the occlusion of stops. Typologically, occlusion of fricatives (targeting at least a subset of places of articulation) is a relatively frequent sound change: It is attested in approximately 44 of 216 languages surveyed (see Table 5, Beguš 2020, and Kümmel 2007).

Occlusion of fricatives can thus operate as a passive phonetic tendency both in phonological systems that do not feature PFOD as well as in those that do. As described in Section 2, after the first two sound changes in the development of PND (Blurring Process) operate, the resulting synchronic alternation PFOD is complex and involves voiced fricatives surfacing in the elsewhere position and voiceless stops in the postnasal position. Example 3 illustrates this alternation.

(3) PFOD

$Z \rightarrow T / N_{\_}$ , e.g. [vo:na] : [ɔmpo:na]

At this point, deviation from articulatory targets described above can cause weak phonetic variation, where voiced fricatives are occasionally produced with occlusion (Ladefoged & Maddieson 1996:137) because occlusion requires less articulatory precision. These articulatory forces result in phonetic variation between voiced fricatives ( $Z$ , e.g. [v] or [z]) and voiced stops ( $D$ , e.g. [b] or [d]). In other words, the universal phonetic tendency of fricative occlusion causes variation in voiced fricative production across phonological systems and therefore also in those phonological systems that had already undergone the first two sound changes in the Blurring Process. At some stage, the complex alternation PFOD thus involves voiceless stops ( $T$ , e.g. [p] or [t]) in the postnasal position and voiced fricatives ( $Z$ , e.g. [v] or [z]) that are in phonetic variation with voiced stops ( $D$ , e.g. [b] or [d]) elsewhere due to the universal phonetic tendency – occlusion of fricatives. The variation in this latter case is schematized in 4.

(4) PFOD WITH VARIATION

$D \sim Z \rightarrow T / N_{\_}$ , e.g. [bo:na] ~ [vo:na] : [ɔmpo:na]

In the initial stages, this variation that cross-linguistically arises from automatic articulatory factors is expected to be highly skewed towards the faithful, in this case, fricative articulation ([v] or [z]). However, cognitive preference for simple (albeit unnatural) alternation that favors the variant with a stop ([b] or [d]), will be confirmed by experimental results in this paper. Despite the preference for the stop response in the experiment being weak, over time, this weak preference can result in an accelerated reversal of the skewed variation.

For example, a speaker of a language with PFOD is in the majority of inputs exposed to an alternation [vo:na] : [ɔmpo:na], but occasionally the speaker is also exposed to [bo:na] : [ɔmpo:na] due to the low-level phonetic process of fricative occlusion. Additionally, they will produce or perceive some of the target fricatives as stops. This is how a regular sound change would operate too: variation based on production and perception can result in a reversal of distribution (from [v] to [b] as the prevalent variant) and, via phonologization, in a sound change. Since this process appears to operate significantly more frequently when it simplifies an alternation, it is reasonable to assume that it is accelerated by a synchronic cognitive mechanism: speakers associate the variant [bo:na] with [ɔmpo:na] more readily when it is preferred by their grammar. This ‘catalyzes’ the operation (initiation) of sound change and results in a direct influence of the cognitive bias on typology. In other words, higher rate and frequency of those sound changes that simplify an alternation result in observed synchronic typology: alternations resulting from such sound changes accelerated by cognitive bias are more frequent than expected. In fact, because historical bias alone cannot explain the higher frequency of operation of the last sound change in those cases in which the sound change simplifies an alternation, the paper argues that this gradual preference for stop articulation operating on gradient phonetic variation is precisely what catalyzes

occlusion of voiced fricatives in the development of PND.

The proposed mechanism called CATALYSIS can be summarized as:

(5) CATALYSIS

- a. A subset of segments in an alternation is in passive phonetic variation.
- b. The less frequent variant in this variation is cognitively preferred.
- c. Subjects associate the variant preferred by cognitive factors with a given input more often than the variant that is less preferred (due to learnability, grammatical preferences, or perceptual distance).
- d. Distribution of variation initially skewed towards the faithful variant is reversed towards the variant that is cognitively preferred.

The advantage of catalysis is that it provides a plausible mechanism for how cognitive biases influence phonological typology that has an empirical basis in the present experiment (Section 4). Catalysis explains the higher rate of the last sound change in the Blurring Process that simplifies an alternation and its learning. The mechanism also offers a potential answer to the question of how cognitive factors influence the typology outside of the scope of L1 acquisition (see Bybee 2001 and Section 1.1) – the proposed mechanism can equally apply during the L1 acquisition and after it.

In this paper, we experimentally simulate conditions for catalysis in the second and third stage of the Blurring Process that leads to PND and test it in the artificial grammar learning paradigm. We train subjects on data two alternations: PFOD and PND. Subjects are then faced with ambiguous surface forms with voiceless stops postnasally (e.g. [ɔmpo:na]) that can go back to a variant with closure consistent with PND ([bo:na]) or a variant with a fricative consistent with PFOD ([vo:na]). According to catalysis, the first variant is the innovative variant that arises from low-level phonetic processes (e.g. fricative occlusion) and the second is the faithful variant. Analogous to the lab behavior in the real world are speakers who are faced with phonetic variation that results from universal phonetic tendency (fricative occlusion; see 4). They might generalize the variant with closure more frequently. Over time, this preference due to cognitive bias can result in a higher rate of the reverse distribution of variation (i.e. higher rate of sound change initiation) and consequently, this higher rate of operation results in PND being more frequent than predicted by the historical bias.

When native speakers of two languages with different phoneme frequencies and different voicing realizations are presented with equal amounts of evidence for PND and PFOD, they show a slight preference for the PND response over the PFOD response in the labial series. Subjects thus prefer the simple alternation to the complex one, even if the simple alternation is phonologically unnatural. The link between a synchronic preference for one type of alternation (cognitive bias) and typological rarity of complex alternations is thus experimentally confirmed and, crucially, this link cannot be interpreted as part of the historical bias.

## 4. Experiments

The Blurring Process assumes a stage with PFOD in the development of PND. Catalysis assumes that speakers of a system with PFOD are faced with PFOD and occasionally, with variants con-

sistent with PND. We expose the subjects to singular-plural nonce word pairs with these two alternations: half of the words are consistent with PFOD and half with PND, just as we assume happens in the development of PND.

The advantage of such a design is that the experiments resemble the proposed trajectory of historical changes as closely as possible. The experimental design also follows the approach proposed by Albright & Do (2019), where the same group of subjects is exposed to data consistent with multiple alternations and tested on ambiguous stimuli with both explicit and implicit tasks, which avoids the problem of heterogeneous subject groups in the two conditions and primes implicit rather than explicit learning. The experiment includes evidence for an explicit task, vowel harmony in feature [ $\pm$ front], and an implicit task, PND and PFOD.

Like any experiment in the artificial language learning paradigm, the experimental design in this paper does not completely replicate reality (for a discussion on why the artificial grammar learning paradigm is nevertheless valid, see Ettlinger et al. 2016). Unlike in initial stages of catalysis, subjects in the experiment are presented with equal amount of evidence for PND and PFOD. Such a design was chosen for several reasons.<sup>9</sup> First, it is assumed that in catalysis, initial evidence for PND (that results from low-level phonetic variation) is relatively small. It would be impractical to test synchronic preferences using such small proportions. In laboratory conditions, subjects can be exposed to only a few hundred stimuli. This means that if an experiment contained a small proportion of stimuli containing evidence for PND, subjects would be faced with a handful of PND examples in absolute terms. In the case of actual catalysis, even if the proportion of evidence for PND is small, speakers would be faced with substantially more PND forms in absolute terms. Moreover, in the laboratory, subjects are exposed to each unique item only once, whereas during the actual learning, speakers would be faced with the same form several times, thus amplifying evidence for PND.

The present experiment tests the existence of a synchronic preference for PND, all else being equal (i.e. when speakers are presented with equal evidence for the two alternations). If subjects choose disproportionately more PND than justified by data, it means there exist a synchronic preference that operates as a pressure every time the speaker needs to make a decision about whether to analyze [ɔm'po:na] as a prefixed ['vo:na] or ['bo:na]. There is no specific evidence suggesting such a preference would only operate when the subjects are faced with equal amount of evidence for PND and PFOD. We can assume that the preference that holds in the least conditioned case (when both are presented equally) also holds in case one variant is more or less frequently represented in the data. One reason to extend this assumption is that in reality, speakers would be faced with substantially more evidence for the PND variant in absolute terms and would have repeated exposure to individual forms consistent with PND.

Results of our experiments are, however, relevant even without the assumption that the preference for PND operates regardless of the proportions of input data. We saw that catalysis derives typology by explaining why the final sound change operates more frequently. Crucially, fricative occlusion is a required condition for catalysis, regardless of whether the novel stop variant is rare or frequent. It is likely that the preference towards PND starts operating when the evidence for stop-initial variant is very small – the synchronic preference for PND can operate every time a

---

<sup>9</sup>Future experiments can test preferences given lower proportions of PND-consistent evidence in the training data, but such an experiment will be more challenging to implement and statistically more difficult to model.

speaker needs to make a decision between two variants. But even if the synchronic preference starts operating only when the variation accumulates and reaches a threshold (due to general historical factors) such that the two variants are approximately equally represented (as is the case in the experimental design), the effect would be similar: an accelerated operation and an accelerated completion/phonologization of the final sound change. While in such a scenario, the phonetic variation would need to accumulate based on historical factors, the accelerated rate of sound change can still be influenced by catalysis: Catalysis predicts not only that the rate of sound change would be accelerated, but also that completion and phonologization would be higher because the speakers start analyzing [ɔm'po:na] as ['bo:na] due to the synchronic preference. Without this synchronic preference, the variation would lack the driving force towards its accelerated phonologization and would remain at the rates of the initial phonetic variation or at the rates of sound change from the unconditioned samples.

There are also limitations to the present approach. The experiments only indirectly test whether the observed effects are due to learnability or other synchronic factors: the forced choice approach in our experiments does not include an incorrect answer. Such a design was chosen for two reasons: not to overburden the subjects (three vs. two choices) and to resemble the assumed catalysis more closely (the assumed trajectory in catalysis only involves two variants – the stop and the fricative variant). Evidence for learning comes from the explicit task, where subjects choose between the correct and incorrect options. While the results suggest that subjects do learn vowel harmony in the explicit task, this does not necessarily entail that learning occurs in the implicit task too (PFOD vs. PND) and that asymmetries in experimental results can be explained by learnability. The preference for PND or PFOD in the implicit task can result from learnability differences (PND is easier to learn), perceptual similarities (e.g. P-map; Steriade 2001), or general markedness-driven avoidance of voiced fricatives. All three possibilities are relevant to catalysis and are considered part of the cognitive bias, because they operate in individual speakers in a lab setting and likely stem from an association of two variants into a phonological alternation (for historical arguments for this association, see Section 2). The one confounding factor that would make the results less informative – influence of subjects' native phonologies – is controlled for by conducting experiments with subjects of two languages that differ in the frequencies of voiced fricatives and stops.

## 4.1 Training

Subjects were trained on forming plural nouns from singular nouns in a made-up language called Martian. The plural prefixes were [ɔn-] (before coronals such as [t, d, s] and elsewhere) and [ɔm-] (before labials such as [p, b, f]), e.g. singular ['sanu]; plural [ɔn'sanu]. If the singular noun features a front first vowel ([ɛ, i] in English and [e, i:] in Slovenian), vowel harmony is triggered in the prefix, which then surfaces as [ɛn-] before coronals and [ɛm-] before labials, e.g. singular ['pʰimi]; plural [ɛm'pʰimi]. Subjects were explicitly instructed that the plural is formed with <en-> and <on-> prefixes presented auditorily and orthographically during the instruction phase. Subjects were explicitly asked to pay attention to how plural nouns are formed and were told that the final task would involve forming plural from singular nouns.

The data in the training phase also involves IMPLICIT evidence for two alternations: PFOD ( $Z > T / N_-$ ) and PND ( $D > T / N_-$ ). Because the learning of PND and PFOD is tested with an implicit task, experimental instructions never reference the two alternations. Stimuli for the im-

		[-voice]		[+voice]	
		[-cont]	[+cont]	[-cont]	[+cont]
#	T	S	D	Z	
N	T	S	T	T	

Table 6: Alternations in voiceless and voiced stops and fricatives according to position (implicit task).

		[-voice]		[+voice]			
		[-cont]	[+cont]	[-cont]	[+cont]		
		[LAB]	[COR]	[LAB]	[COR]	[LAB]	[COR]
Singular (#)	'p <sup>h</sup> ɔrɔ	t <sup>h</sup> aru	'furə	'sanu	'balu	'dɔru	'vɔnə
Plural (N)	ɔm'p <sup>h</sup> ɔrɔ	ɔn <sup>th</sup> aru	ɔm'furə	ɔn'sanu	ɔm'p <sup>h</sup> alu	ɔn't <sup>h</sup> oru	ɔn'p <sup>h</sup> ɔnə

Table 7: Examples of Martian words for the implicit task in the English experiment.

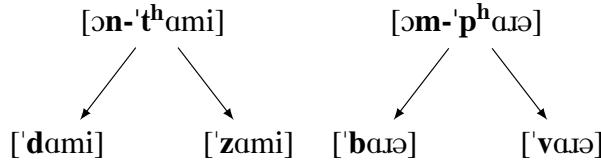
plicit task include items of the shape  $C_1V_2C_3V_4$  with an equal number of initial ( $C_1$ ) labial and coronal voiced stops ([b] and [d]) and initial labial and coronal voiced fricatives ([v] and [z]). In both cases, the plural prefix that ends in a nasal ([ən/m-] and [ən/m-]) causes the voiced stop to devoice (PND, e.g. ['balu] ~ [ɔm'p<sup>h</sup>alu]) and voiced fricative to occlude and devoice (PFOD, e.g. ['vɔnə] ~ [ɔm'p<sup>h</sup>ɔnə]). Plural forms are thus of the shape  $\{\epsilon/\circ\}\{n/m\}-C_1V_2C_3V_4$ , where  $C_1$  is the ‘devoiced’ voiceless stop [p] or [t]. In other words, subjects are exposed to the same amount of stimuli consistent with PND and PFOD. In order to prevent subjects from analyzing PFOD as a simple restriction on postnasal fricatives, subjects are also given evidence that voiceless stops ([p] and [t]) and voiceless fricatives ([f] and [s]) remain unchanged in postnasal position ([mp], [nt], [mf], and [ns], respectively). Table 6 schematically represents the alternations in the training phase; Table 7 lists some actual examples of the stimuli. One disadvantage of the present experiment is that L1 phonology can affect the complexity of the tested alternations. Both English and Slovenian feature voiced stops in the elsewhere condition, but at the PFOD stage, voiced stops are absent from the system. This would mean that [ $\pm$ voice] does not change automatically in the experiment, but it does so in the development of PND. We can still assume that changing two features is more complex compared to a single feature (even if one changes automatically). Under the perceptual distance hypothesis, this aspect is of course not problematic.<sup>10</sup> For all details of the experimental design and structure of the stimuli, see Supp. Materials Section 1.

During the training phase, the stimuli (as described above and summarized in Supp. Materials Section 1) were presented in a randomized order (randomized for each subject). A unique picture of a Martian creature was associated with each stimulus (also randomized for each subject).<sup>11</sup> Generally, no orthography was given with the training stimuli. To prompt subjects into focusing on the experiment, however, ten words in the training phase were chosen randomly (random-

<sup>10</sup>The data also involves place assimilation of the prefix nasal (e.g. [n] if coronal, [m] if labial). This assimilation is never tested experimentally; its primary purpose is articulatory ease during stimuli recording and to achieve conformity to subjects’ L1 phonologies, such that no attention is attracted to the distribution of nasals.

<sup>11</sup>Pictures of Martian creatures in the experiments were taken from van de Vijver & Baer-Henney 2014 with their permission.

Figure 2: Examples of ambiguous stimuli in plural form with a prefix and corresponding possible singular forms subjects were asked to choose from.



ized for each subject). Those ten items were presented orthographically as well as auditorily, and subjects had to enter a transcription of the plural form they had heard (the singular form was presented both in audio form and printed on the screen, whereas the plural form only appeared in audio form). This orthographic task was never chosen for words that started with voiced stops or fricatives (bearing evidence for PND and PFOD) in order to minimize the effect of orthography and the effect of increased memorization due to this task.

While PFOD is not an unnatural alternation according to the definition in Section 2 in the sense that postnasal frication and voicing (the opposite process to PFOD) are not universal phonetic tendencies at the same time, it is unnatural in the sense that when PFOD applies, stops surface as voiceless in postnasal position, where they are universally dispreferred (for an extensive discussion, see Beguš 2019). Unlike PND, however, PFOD allows subjects to construct alternative grammars which are not necessarily unnatural, especially if the two processes are treated separately in phonology. If no evidence existed that voiceless fricatives remain unchanged postnasally and voiced stops devoice postnasally, PFOD could be analyzed as a restriction against postnasal continuants and a global restriction against voiced stops (\*[+nasal][+cont] and \*[+voice, -cont] markedness constraints in Optimality Theoretic terms; Prince & Smolensky 1993/2004). This is exactly the synchronic system of stage 2 in the development of PND. Even if voiceless fricatives remain unchanged, PFOD can be analyzed as resulting from \*[+nasal][+cont] and \*[+voice, +cont] constraints with an additional higher ranked faithfulness constraint that preserves the identity of voiceless fricatives, although such an analysis is less likely. Because the experimental design requires evidence for voiced stops to remain unchanged except postnasally (as is also predicted by catalysis), it is possible that subjects analyze PFOD as complex and unnatural. Regardless of whether subjects treat PFOD as an isolated, complex and motivated process or unnatural and complex, the results have direct implications for modeling synchronic influences on typology (see Section 6).

## 4.2 Test

Learnability of the two alternations is tested with an implicit task because implicit learning might resemble L1 phonological acquisition more closely than explicit learning (for this question, see Moreton & Pertsova 2017 and Moreton et al. 2017, but further research is warranted). Subjects were not given any instructions on the implicit task. The alternations in the implicit task are summarized in Table 6, Table 7, and Figure 2.

After the training period, subjects were told that they would hear some Martian words that they had not heard before and that they would be asked to indicate the most likely way to say new words in Martian. In the test phase, subjects were given ambiguous stimuli in the plural

form of the shape  $\{\varepsilon/\circ\}\{n/m\}-C_1V_2C_3V_4$  with devoiced noun-initial stop ( $C_1 = [p]$  or  $[t]$ ) that can go back to either a voiced stop  $[b]$  or  $[d]$  (consistent with PND) or a voiced fricative  $[v]$  or  $[z]$  (consistent with PFOD). The subjects had to choose between two singular forms: one with a word-initial voiced labial or coronal fricative ( $ZV_2C_3V_4$ , consistent with PFOD) and one with a word-initial voiced labial or coronal stop ( $DV_2C_3V_4$ , consistent with PND). Figure 2 exemplifies this task. Preference for one option would suggest either that (a) one alternation is easier to learn than the other, (b) subjects choose the option that is typologically more frequent or/and articulatory easier and perceptually more salient (less marked in phonological terms), or (c) subjects use knowledge of phonemic distributions in their native language and match the phoneme frequencies of their native language to their experimental responses. Both (a) and (b) are informative for our purposes (see discussion in Sections 4.5 and 5.2 on how we control for c).

### 4.3 Procedure

The test phase consisted of two tasks, the explicit and implicit task. The order of items testing the explicit and implicit tasks was mixed and randomized for each subject. In the explicit task, the experiment tested whether or not the subjects learned the vowel harmony in plural prefixes. Six items in the singular were created for this task, three that trigger frontness harmony and three that do not: [rema], [liro], [leni] vs. [lonu], [ruro], and [lona]. Subjects were presented with the six items in the singular (both orthographically and auditorily) that were paired with pictures of a single Martian creature. After they heard the item, subjects were shown four pictures of the same creature and were given two orthographic stimuli to choose from: with the correct harmony and incorrect disharmony. In half of the six items, the correct response appeared on the left button and in half on the right button. The explicit task includes orthographic choices in order to avoid overburdening of subjects with auditory stimuli: the auditory-only stimuli are reserved for the more informative implicit task which always lacked any orthographic presentation. The button side ordering was kept constant for all subjects and items, but ordering of individual items was randomized.

For the implicit task, subjects were first presented with four identical pictures of a Martian creature accompanied by a recording of the test word in plural without any orthographic stimuli. The button that played the recording was embedded in a written sentence ‘These are [▷]’. After playing the sound, subjects were presented with a single picture of the same Martian creature. The subjects played the first stimulus presented in a sentence ‘Is this a [▷]?’ The second stimulus appeared afterwards, embedded in ‘or a [▷]?’. After the subjects heard both stimuli, they were asked ‘Which one is it?’<sup>12</sup> and were given two choices (‘1st’ and ‘2nd’). Stop-initial and fricative-initial singular forms were equally assigned the first or second position, so that the position of the answer would not influence the results. In other words, in half of the eight items, the stop-initial response appeared as the first option and in half of the items as the second option. The ordering was kept constant for all subjects and items, but ordering of individual items was randomized. That stimuli ordering does not affect the result is suggested by statistical analysis given in Section 5. None of the stimuli in the implicit task were presented orthographically.

The stimuli for the implicit task consisted of eight plural items of the shape *prefix-C<sub>1</sub>V<sub>2</sub>C<sub>3</sub>V<sub>4</sub>*, where prefix is [ɔn, εn, ɔm, εm] with the correct harmony based on V<sub>2</sub> (for an example, see Figure

---

<sup>12</sup>The question was followed by ‘(you can replay all three words).’

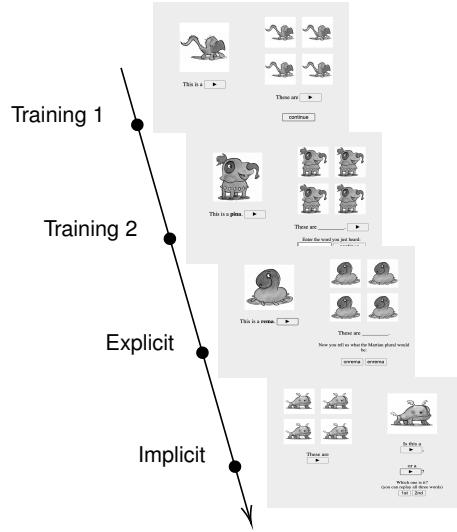


Figure 3: Examples from the Experigen (Becker & Levine 2013) interface as used in the experiment (converted from color to grayscale) for the training phase and test phase (explicit and implicit tasks).

2).  $C_1$  was a voiceless ('devoiced') labial stop in four item and a voiceless ('devoiced') coronal stop in four items. Four items included a noun with a front first vowel ( $V_2$ ) and four with a non-front first vowel (two for each place-of-articulation group).  $C_3$  consisted of [m, n, r, l, j, w]. Vowels  $V_2$  and  $V_4$  consisted of [a, ε, ɔ, i]. Test words were created such that there were no minimal pairs with the training words. Each plural form was matched with two corresponding singular forms of the shape  $C_1V_2C_3V_4$ , one with a voiced stop in  $C_1$  and one with a voiced fricative. A sample spectrogram of a stimulus [ɔmpara] and the two possible responses [bara] and [vara] are given in Figure 1 (Supp. Materials). Figure 3 illustrates the experimental interface.

#### 4.4 Slovenian

The Slovenian experiment is identical to the English experiment, except as described in this section and in the Supp. Materials Section 1.2. The Slovenian experiment was conducted in person in Slovenia with the supervision of research assistants. Subjects were native speakers of Slovenian, a south Slavic language with approximately 2.2 mil speakers (Simons & Fennig 2018). Subjects were recruited from the general public with the help of research assistants (via personal contacts and social media). Altogether 150 subjects participated in the Slovenian experiment.

Test items for the explicit tasks included six  $C_1V_2C_3V_4$  words with  $C_1 = [l, r]$  and three front  $V_2$ s (the other three are non-front). Subjects were asked to choose between the correct harmonic and incorrect disharmonic forms with prefixes [ɔn/m-] and [ɛn/m-]. Subjects were tested on the implicit task with twelve trials. First, subjects were presented with a plural form of the shape *prefix-C<sub>1</sub>V<sub>2</sub>C<sub>3</sub>C<sub>4</sub>* where  $C_1$  is a devoiced [p] or [t] that can go back to a voiced stop [b, d] consistent with PND or a voiced fricative [v, z] consistent with PFOD. In half (six) of these trials the 'devoiced' stop is a labial, and a coronal in the other six. In half of the items,  $V_2$  is non-front (therefore with prefix [ɔn/m-]); in the other half  $V_2$ , it is front (therefore with prefix [ɛn/m-]). All experimental

stimuli are given in Supp. Materials Section 1. After being exposed to the plural forms, subjects were asked to choose between the corresponding singular form with an initial voiced stop and one with a voiced fricative. Both options were presented to the subjects in each trial and in auditory form only without orthographic inputs. All other properties of the test phase are the same as described in the English experiment in Supp. Materials Section 1.1.1.

## 4.5 Why two languages

The possibility that a preference for one or the other response is due to the phonemic frequencies of subjects' L1 phonologies is uninformative, which is precisely why the experiment is conducted in two languages, English and Slovenian. In both, coronal /d/ is more frequent than /z/, but in Slovenian the labial fricative /v/ is more frequent than the labial stop /b/, and in English the distribution is the opposite or the difference is at least substantially smaller than in Slovenian. The difference in frequency between /b/ and /v/ in Slovenian is substantial, persists in word-initial position, and is established based on several lemmatized and non-lemmatized corpora of both written and spoken Slovenian (Suhadolc 2013, Marvin et al. 2018). Word-initially, /b/ has a relative frequency of 3.6%, while the relative frequency of /v/ is 5.3% (see Suhadolc 2013). In other corpora, the difference across positions is between 1.7–1.9% vs. 3.3–4.2% (see Marvin et al. 2018).<sup>13</sup> For English, on the other hand, different studies report slightly different results for the relative frequencies of /b/ and /v/. In most studies or corpora, /b/ is more frequent than /v/ (Wang & Crawford 1960, Mines et al. 1978, Kessler & Treiman 1997). This is true across positions as well as in the onset position (Kessler & Treiman 1997). One study shows the distribution of frequencies of /b/ and /v/ is the opposite (Hayden 1950; also in some corpora by Wang & Crawford (1960)), but differences here are minor. If subjects simply matched native phoneme frequencies in their experimental responses, we would expect a higher response rate for /v/-initial words in Slovenian and the opposite in English. For an even stronger piece of evidence that the L1 phonologies do not crucially affect the results based on bigram frequencies, see Section 5.2.

English and Slovenian are also suitable for testing the learning of alternations involving the feature [±voice] because phonetic realization of the phonological contrast differs substantially in the two languages. In fact, many analyses assume the phonological feature involved in the English contrast is [±spread glottis] rather than [±voice] (Iverson & Salmons 1995 and the literature therein). English voiceless stops are realized with a substantial period of aspiration, especially in the onset position of stressed syllables. Voiced stops are partially or fully voiced depending on the position. Utterance-initially, voicing is often lacking completely (Davidson 2016). Slovenian, on the other hand, is a 'true voicing' language in which voiced stops are fully voiced in all positions, including utterance-initially (with pre-voicing; see Supp. Materials Figure 2 and Toporišič 2004). The two languages also differ in exact realization of /v/. In English, labiodental /v/ is always analyzed as a fricative, while in Slovenian descriptions vary between a fricative /v/ (Toporišič 2004) and an approximant /v/ (Šuštaršič et al. 1995). The exact acoustic distinction between the

<sup>13</sup>Slovenian phoneme frequencies are calculated based on grapheme frequency in Suhadolc 2013, but because there is a strong tendency towards one-to-one correspondence of Slovenian consonant inventory and orthographic representation (especially of /b/ and /v/ in initial position), the results of the analysis based on grapheme can be extended to phonemes. Marvin et al. (2018) account for the grapheme-phoneme differences.

two is, however, challenging to establish. The nature of the bilabial fricative/approximant and its distribution, which is potentially dialectal in Slovenian, warrants a separate discussion, but our recordings indicate the presence of friction noise (an example in Supp. Materials Figure 3).

In addition to controlling for the voicing realization and phoneme frequencies, the experiments were designed with the goal of diversifying the subject pool, keeping the ratio of stimulus vs. test items as high as possible, and maximally balancing experimental stimuli. For these reasons, speakers of the two languages tested are recruited from the general public. The experiment with English speakers was conducted as an online experiment; the Slovenian experiment was conducted in person with the supervision of research assistants and with controlled experimental equipment. Subjects with prior linguistic experience have been excluded from the analysis (if they responded so). The ratio of training vs. test items is 32:8 for the explicit task in English and 32:12 in Slovenian in order to prevent the test phase itself from influencing the responses. Balancing experimental stimuli is described in detail in Supp. Materials Section 1.1.1 and 1.2.1.

## 4.6 Subjects

A total of 353 subjects participated in the two experiments. Of the 203 English participants, 198 finished the experiment in full and completed a final demographic questionnaire. Of the 150 Slovenian participants, 141 completed the test phase and finished the final questionnaire. Subjects who indicated that either English or Slovenian, respectively, was not their native language or who indicated that they have linguistic education (or took any classes in linguistics as part of their education) were excluded from the analysis (subjects with no responses on the two questions were included). Altogether, 170 subjects in the English group and 110 in the Slovenian group were analyzed on the implicit task. The 170 subjects in the English experiment provided 1346 analyzed responses on the implicit task; the 110 subjects in the Slovenian experiment provided 1317 responses. A total of 280 analyzed subjects is a relatively high number of participants compared to most other similar artificial grammar learning experiments, especially for the in-person experiment. For other details on the subjects and exclusion criteria, see Supp. Materials Section 1.4.

## 5. Results

Responses for the explicit (correct vs. incorrect) and implicit (PND vs. PFOD) tasks across the two experiments (English and Slovenian) and across the two places of articulation and vowel frontness are given in Tables 8 and 9. Raw counts reveal that the correct response (consistent with vowel harmony) was more frequent than the incorrect response for all groups in the explicit task.

Likewise, the raw counts suggest a response consistent with PND (stop response) was more frequent than the response consistent with PFOD (fricative response) for all four groups on the implicit task. This effect appears to be substantially stronger for the labial series of stops compared to the coronals. The preference for the correct response, consistent with vowel harmony, seems more robust than the preference for the stop response (PND), but the difference is not too substantial, especially given that the evidence for the first is explicit and categorical (occurring in every of the 58-60 items) without any ambiguous stimuli. On the other hand, the preference for the PND-consistent response emerges from the 32 items in the training phase that bear an equal amount of evidence both for PND or PFOD: devoiced postnasal stops can go back to either voiced

Table 8: Raw counts of number of correct (harmonic) and incorrect (disharmonic) responses on the explicit task across the two languages and vowel frontness.

	English			Slovenian		
	correct	incorrect	% correct	correct	incorrect	% correct
front	313	193	61.9%	226	104	68.5%
back	311	196	61.3%	207	122	62.9%

Table 9: Raw counts of the number of stop and fricative responses on the implicit task accross the two langauges and places of articulation.

	English			Slovenian		
	stop	fricative	% stop	stop	fricative	% stop
labial	388	286	57.6%	397	261	60.3%
coronal	354	319	52.6%	341	319	51.7%

stops or voiced fricatives. Additionally, the PND-consistent response emerges in an implicit task without any direct instructions to the subject.

Figure 4 plots counts of subjects according to how they performed on the implicit task (number of responses consistent with PND). Subjects who always selected the fricative (PFOD) response have a score of 0. Subjects with all stop responses (PND) have a score of 8 or 12, respectively (English vs. Slovenian). For reference, the plots also show predicted values if subjects responded randomly indicated by vertical lines with dots. These predicted values are calculated from a binomial distribution with  $n = 8$  or  $12$ , respectively,  $p = 0.5$ . For each  $k$ ,  $\Pr(k; n, p)$ , was calculated from a binomial distribution and multiplied by the number of subjects in each sample: 167 and 109, respectively.

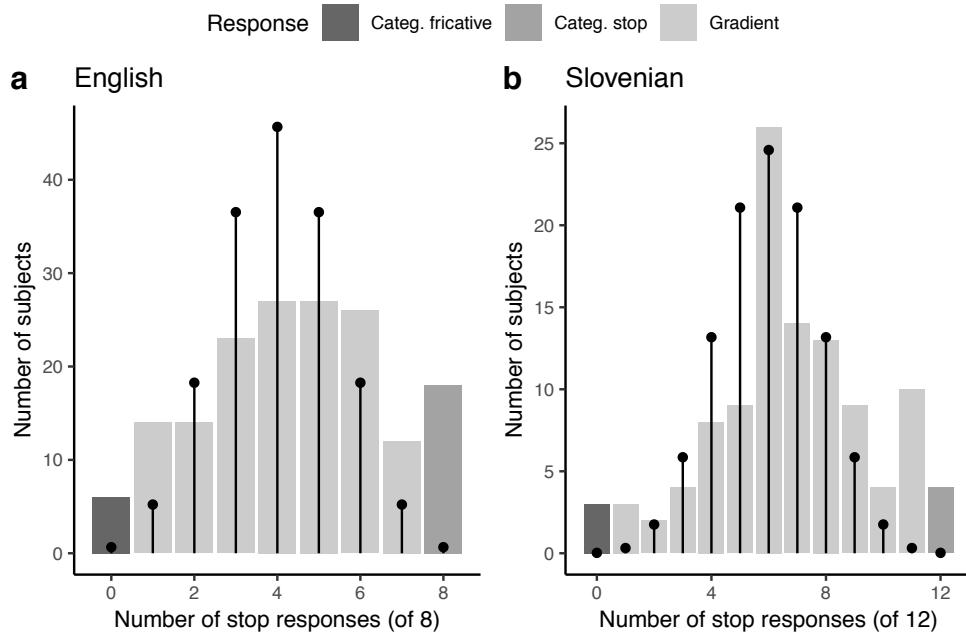
The distribution of subjects is substantially higher than expected on the two marginal ends of the stop-fricative opposition and is lower than expected in its middle (in the English experiment). There are more subjects that chose the stop response more often compared to the number of subjects that chose the fricative response more often (especially in the labial series). Especially in the English experiment, categorical responders that chose the PND-consistent response are notably more numerous.

## 5.1 Explicit task

To confirm that the subjects are learning the explicit alternation (vowel harmony), the results of the explicit task were fit to two logistic regression mixed effects linear models with harmonic (correct) responses as successes and disharmonic (incorrect) responses as failures, one for each experiment. For an exact description of the model, see Supp. Materials Section 1.5. Correct answers (consistent with the explicit vowel harmony) are significantly above chance level in the English experiment ( $\beta = 0.51$ ,  $z = 5.6$ ,  $p < 0.0001$ ) with 95% profile CIs at [57.8%, 67.2%]. The Slovenian experiment yields similar results. The harmonic responses that conform to the training data are significantly more frequent than the disharmonic responses ( $\beta = 0.70$ ,  $z = 6.72$ ,  $p < 0.0001$ ) with profile CI [61.9%, 71.9%].

The results suggest that subjects learned the explicit alternation above chance. Similar out-

Figure 4: Raw counts of subjects according to the number of stop responses in the English (A) and Slovenian (B) experiment.



comes in both experiments suggest that English speaking subjects recruited via Amazon MTurk who were not supervised during the experiment and were not given the same headphones and the DAC amplifier as in the Slovenian experiment (see Supp. Materials Section 1.2.1), did not perform substantially worse on the explicit task compared to in-person experiments. On average, the in-person Slovenian group performed slightly better than the English Amazon MTurk group.

## 5.2 Implicit task

To test the significance of the preference for the stop response (consistent with PND) vs. the fricative response (consistent with PFOD), the data were fit to two logistic regression mixed effects linear models, one for each experiment. Coefficients of the final models are given in Table 10; estimates with confidence intervals across places of articulation and across the two languages are given in Figure 5. Significance of the predictors of interest is the same in the full models including all interactions and random slopes and the final chosen models; here we report the non-full models because of interpretability and to signal which predictors are significant. For a detailed description of the models, see Supp. Materials Section 1.6.

Subjects choose the stop response (compared to the fricative response) in the English experiment significantly more frequently for the labial series of stops ( $\beta = 0.43, z = 2.02, p = 0.04$ ).<sup>14</sup> In the coronal series, no such difference is observed (based on confidence intervals, see Figure 5). Given that subjects are recruited via Amazon Mturk in the English experiment and potentially some subjects did not pay attention to the experiment, it is reasonable to also test responses in the data from which the lowest-scoring subjects on the explicit task are removed. If we include only

<sup>14</sup>Profile CIs cross the zero in the model not corrected for underdispersion.

subjects who scored 50% or better on the explicit task, the preference for the fricative response remains significant ( $\beta = 0.58$ ,  $z = 3.0$ ,  $p = 0.003$ ).

The same procedure is used to test significance in the Slovenian experiment. For a detailed description of the model, see Supp. Materials Section 1.6. The estimates are given in Table 10. Subjects prefer the stop response in the labial series ( $\beta = 0.62$ ,  $z = 2.78$ ,  $p = 0.005$ ), but again no such effect is observed in the coronal series. The same effect persists if we exclude non-learners (who scored less than 50%):  $\beta = 0.54$ ,  $z = 2.5$ ,  $p = 0.01$ .

English				
	Estimate	Std. Error	z value	Pr(> z )
(Intercept) = labial	0.43	0.21	2.02	0.04
coronal	-0.29	0.28	-1.03	0.30
Slovenian				
	Estimate	Std. Error	z value	Pr(> z )
(Intercept) = labial	0.62	0.22	2.78	0.01
coronal	-0.56	0.32	-1.79	0.07
HARMONY	-0.20	0.13	-1.49	0.14

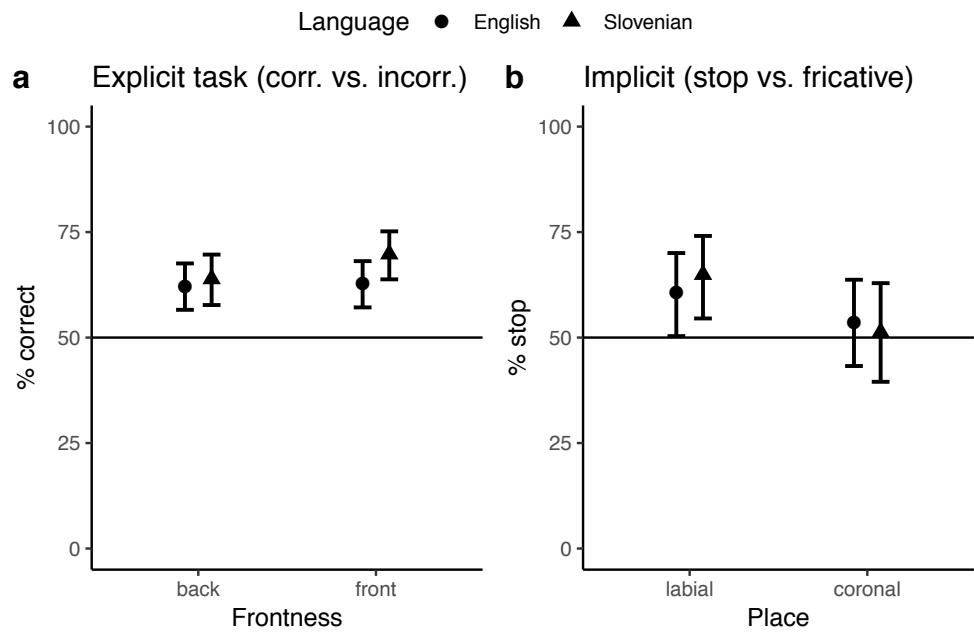
Table 10: Coefficients of the final model (English and Slovenian).

### 5.3 Influence of L1 phonology: bigrams in Slovenian

A strong piece of evidence against the possibility that subjects simply match frequencies from their L1 phonologies comes from bigram frequencies and absolute response rates in the experiments. Frequencies of the tested word-initial bigrams in lemmas (from the standard dictionary of Slovenian; Bajec et al. 2000) are not distributed equally. /b/ is more frequent than /v/ before /a/ (688 vs. 341 in counts) and before /u/ (267 vs. 32). /v/ is more frequent than /b/ before /i/ (476 vs. 388) and /e/<sup>15</sup> (782 vs. 489). To test non-lemmatized frequencies in a spoken corpus, word-initial sequences of /b/ and /v/ followed by vowels /a/, /i/, /e/, and /u/ were extracted from a non-lemmatized transcribed corpus of spoken Slovenian GOS (Korpus govorjene slovenščine; Verdonik & Zwitter Vitez 2011). Word initial /v/ is more frequent than /b/ before /a/ (3,542 vs. 1,353) and before /e/ (15,779 vs. 1,852), but /b/ is more frequent than /v/ before /i/ (13,074 vs. 5,059) and /u/ (693 vs. 221). Learnability of PND and PFOD is tested on items with initial /b/- and /v/- before six vowels: /i/ (twice), /e/, /a/ (twice), and /u/. The /b/-initial response is more frequent than /v/-initial response in five out of six items in the experiment. /b/ is a more frequent response before items with V<sub>2</sub> /a/ (two items), /e/, /i/, and /u/. In only one item with V<sub>2</sub> /i/ is the /vi/-initial response more frequent than the /bi/-initial response (56 vs. 53). In other words, subjects prefer the /b/-initial response even in those stimuli in which the /v/-initial bigram in the native phonology is more frequent than the /b/-initial bigram, estimated from a lemmatized dictionary and non-lemmatized spoken corpus. This means that phoneme frequency distribution in the native language likely does not influence the experimental outcomes. To the author’s knowledge, this experiment is one of the first to test learnability of an alternation on two languages with different relative frequencies of the tested phonemes.

<sup>15</sup>In this paragraph, /e/ stands both for /e:/ and /ɛ/, because the corpora do not distinguish them.

Figure 5: (a) Logistic regression mixed effects model with correct responses on the explicit task as the dependent variable and FRONTNESS as predictors with interactions (with 95% CI) for each of the two languages tested. (b) Logistic regression mixed effects model with % of stop response on the implicit task as the dependent variable and PLACE as the predictor (in the Slovenian model, HARMONY is also a predictor) (with 95% CI).



## 6. Discussion

When trained on data with an equal amount of evidence for PND and PFOD and tested on the ambiguous stimuli, subjects show a small, but significant preference for the stop response (consistent with PND) compared to the fricative response (consistent with PFOD) in the labial series. These results are likely not influenced by subjects' native phonological grammar, because English and Slovenian differ in phonemic frequencies of /b/ and /v/, how their bigram frequencies are distributed, and in the phonetic realization of the feature [±voice] (or [±spread glottis]). Preference for the stop response (PND) appears stronger in the Slovenian than in the English experiment; if phoneme frequencies affected the results, an opposite distribution would be expected (/v/ is more frequent than /b/ in Slovenian). Moreover, because the learning of the two alternations is tested with an implicit task, it is likely that the results resemble L1 phonological acquisition more closely (compared to the explicit task) (Moreton & Pertsova 2017, Moreton et al. 2017)). Additionally, recruitment for the experiment is done on the general public without prior linguistic experience in two experimental modes: online and in person.

While the preference for the PND response appears weak, comparing the results of the implicit task, in which both options are correct (and presented with equal frequencies in the training data), to the results of the explicit task in which there is a correct and an incorrect option reveal that the preference for a stop-initial response in absolute terms is comparable to the preference towards the correct response. Figure 5 illustrates that the proportion of correct responses on the vowel harmony pattern is comparable to the preference towards the stop-initial response.

The advantage of the proposed model of combining experimental data with statistical modeling of sound change is that the link between experimental responses and the typological distribution of phonological alternations cannot be attributed to historical bias. As discussed in Section 2.3, the historical bias approach alone does not derive the observed typological distributions: PFOD is less frequent than predicted by the historical bias approach, whereas PND is more frequent than expected. Additionally, the last sound change that leads to PND operates significantly more frequently when it simplifies the resulting alternation and consequently simplifies its learning. These non-experimental observations already suggest that cognitive biases directly influence the observed typology. Experimental data in this paper suggest that PND is favored to PFOD by the cognitive factors tested in the experiment. In other words, the typological distribution that cannot be explained within the historical bias approach, matches the experimental responses. The experimental results thus confirm the link between cognitive bias and typology. Experimental results suggest that the main manifestation of the cognitive factor is avoidance of complex alternations (in line with previous work; see Moreton & Pater 2012a,b) and that we can experimentally confirm this link even when the historical factor is controlled for. Additionally, the experiments simulate a historically plausible scenario for how a synchronic preference towards the stop response can directly result in the acceleration of sound change operation, which consequently results in typological distributions (the catalysis model).

Results also suggest that the preference for the PND response ([b] vs. [v] and [d] vs. [z]) is stronger for labials than for coronals. This asymmetry in experimental results, however, is not undesirable. A typological survey of PFOD and PND in Beguš 2019 reveals that the last sound change, occlusion of fricatives to stops, often does not include all places of articulation, but targets only a subset. It is thus possible that the asymmetry in the experimental results reflect the actual asymmetry in the observed typology. In other words, the experimental results on learnability are

consistent with the typology, where only a subset of places of articulation are targeted.

There are at least two possible explanations within the cognitive bias approach for why subjects prefer the response consistent with PND. Both possibilities are informative for modeling historical and cognitive influences on typology. First, it is possible that one alternation is easier to learn than the other. This possibility is consistent with a large body of research showing that complex alternations are more difficult to learn than simple alternations (i.e. complexity bias; see Moreton & Pater 2012a,b). Additionally, the higher than chance response in favor of the correct harmonic response (see Section 5.1) suggests that subjects indeed learned the alternation in the explicit task. It is possible that a similar learning mechanism underlies the preference for the PND response on the implicit task. The novel aspect of the present experiment (besides controlling for historical bias) is that the preference for simple alternation emerges even if the simple alternation is unnatural, i.e. operating in exactly the opposite direction from universal phonetic tendencies. To the author's knowledge, in none of the experiments testing complexity bias thus far is the alternative alternation unnatural.

The second possibility is that subjects prefer the variant that is articulatorily or perceptually less marked: either the variant that requires less articulatory effort when presented with ambiguous data or the variant that is perceptually closest to the stimulus. Voiced fricatives (such as [v, z]) are articulatorily dispreferred (Ladefoged & Maddieson 1996:137; Ohala 2006:688; Ohala 1983; Ohala 1997; Smith 1997) and typologically less common compared to voiced stops (such as [b, d]) (Moran et al. 2014). Such dispreferred segments have been labeled 'marked' in phonological theory (de Lacy 2006a).<sup>16</sup> It is possible that subjects use tacit universal phonological knowledge and choose the less marked option or the variant that requires less articulatory effort. This line of explanation faces some challenges. First, while voiced fricatives are generally more marked than voiced stops, it is unclear whether they are more marked for L1 speakers of languages in which voiced fricatives are more frequent than voiced stops. As discussed in Section 4.2, voiced labial fricatives are more frequent than voiced labial stops in Slovenian and some of the tested bigrams are more frequent if they involve an initial voiced fricative. According to the markedness hypothesis, subjects would have to disregard phonemic frequency and choose their responses based only on the general markedness or articulatory effort. Moreover, if markedness avoidance is the main mechanism behind subjects' experimental responses and if experimental responses are indicative of what happens in phonological development, we would expect the same rate of markedness avoidance and consequently the same rate of sound change occlusion of fricatives (/v/ > [b]) in all cases, regardless of whether the sound change simplifies an alternation (and therefore its learnability) or not (see Section 2.3). Because the rate is not equal in the two conditions, it is unlikely that articulatory markedness affects the results. Even if articulatory markedness affects experimental outcomes, the results are nevertheless informative for our purposes. The fact that higher rate of sound change occurs only when two segments are connected with a synchronic alternation provides evidence in favor of the link between the experimental results and typology (and therefore supports the catalysis model).

One of the more challenging questions in phonology is whether experimental evidence of the complexity bias is influenced by featural or perceptual complexity. It has been assumed that the more features an alternation manipulates, the more complex it is and therefore more difficult to learn. Featural complexity is, however, highly conflated with perceptual complexity. It is thus pos-

---

<sup>16</sup>For markedness and its diverse definitions, see de Lacy 2006a.

sible that the PND alternation is preferred and easier to learn compared to the PFOD alternation because perceptual distance between T and D ([p] vs. [b] or [t] vs. [d]) is smaller than between T and Z ([p] vs. [v] or [t] vs. [z]) (in line with the P-map hypothesis (Steriade 2001; see also Wilson 2006, White & Sundara 2014, White 2017).

To quantitatively evaluate the role of perceptual distance on the synchronic preference in the experimental results and potentially in the process called catalysis, we estimated  $d'$  values for the perceptual distance between the voiceless labial and coronal stop and corresponding voiced stops and voiced fricatives. We take consonant confusion matrices from listeners of four languages (including English) from Singh & Black 1966. We fit the data to a bias-reduced probit regression linear model coded such that it estimates  $d'$  values and performs significance testing on differences in  $d'$  values (for a detailed description of the data and models, see Supp. Materials Section 2). [p] and [b] are perceptually significantly more similar than [t] and [d] (the difference between  $d'$  of [p] ~ [b] vs. [t] ~ [d] is estimated at  $\beta(\Delta d') = 0.84$ ,  $z = 2.17$ ,  $p = 0.03$ ). To be sure, the difference in  $d'$  between [p] ~ [b] and [p] ~ [v] is significant:  $\beta(\Delta d') = 1.45$ ,  $z = 3.46$ ,  $p = 0.0005$ . In other words, perceptual distance between a voiceless and a voiced labial stop ([p] ~ [b]) is smaller than between a voiceless stop and a voiced fricative ([p] ~ [v]); perceptual distance between a voiceless and voiced labial stop ([p] ~ [b]) is smaller than between a voiceless and voiced coronal stop ([t] ~ [d]). This could explain the synchronic preference for stop response in the experiments: subjects prefer perceptually minimal alternations. The P-map hypothesis (Steriade 2001) similarly claims that the preference for perceptually minimal alternations is part of the synchronic grammar. It is possible that the synchronic preference in the experiment is strongest in the labial series precisely because [p] and [b] are perceptually most similar.

Conclusive tests of the two hypotheses (featural vs. perceptual complexity) are difficult to design. The results of the experiment presented here remain relevant even if perceptual complexity plays a role in experimental responses (see Section 3). Under this approach, the causing factor behind the typology-perception link still needs to be tied to subjects' internal preference for similarity of segments involved in an alternation. In other words, while perception plays a role both in cognitive and historical bias approaches (Section 1), it is likely the case that cognitive factors – relationship to a synchronic phonological alternation – accelerate the operation of sound change and consequently result in typology. If only perceptual difference, regardless of its association with a phonological alternation, influenced the typology, the same rate of application of fricative occlusion would be expected both where the sound change simplifies an alternation and where it does not (see Section 2 and 3). Regardless of its underlying driving forces, the synchronic preference towards the stop (simple) response has been confirmed in two experiments on two languages with 280 subjects in at least a subset of places of articulation. The observed preference is reflected in the typology and, crucially, is likely not influenced by the historical bias factors.

## 7. Conclusions and future work

Phonology offers a unique test case in the discussion of historical and cognitive influences on human behavior. Combining statistical modeling (Beguš 2020) and experimental work, this paper presents a framework for testing one approach while controlling for the other. The crucial aspect of this framework is identifying mismatched predictions between the cognitive and historical approaches and experimentally testing one approach while controlling for the other. The results

of one set of such experiments presented in this paper suggest that when subjects are trained on PFOD and PND and tested on ambiguous data, they show a weak but significant preference for the PND response. The results point to a causal link between cognitive bias and typology: The PND alternation that is synchronically preferred in a laboratory setting is also typologically more frequent than predicted by the historical bias alone. Crucially, the link between cognitive bias and typology does not have a competing historical explanation (unlike in many other experiments), because the historical bias approach makes the opposite prediction in these types of cases.

The results thus yield insights into those aspects of phonology that are primarily influenced by human cognition and are not emergent from language's transmission. The data suggests the cognitive part of phonology is responsible for avoiding (featurally or perceptually) complex alternations and keeping phonology structurally simple. Applying the framework to further alternations should provide a better understanding of which aspects of phonological grammar and typology are emergent from historical factors and which aspects are primarily influenced by cognition.

Some of the features in the experimental design that control for various non-linguistic variables are novel to the paradigm. The results are likely not influenced by subjects' L1 phonologies as Slovenian and English have different distributions of relative frequencies of both the tested phonemes and the tested word-initial bigrams. The experimental design aimed to diversify the subject pool by recruiting from the general public, both online and in person with controlled and uncontrolled auditory presentation conditions and excluding subjects with prior linguistic experience.

Finally, the presented experimental results combined with statistical models of sound change point to one potential mechanism for how cognitive bias directly influences the typology – catalysis. Universal phonetic tendencies operate cross-linguistically and cause variation in surface forms. Initially, the variation is heavily skewed towards the original, faithful variant. If the non-faithful variant simplifies learning, however, learnability preferences can skew the initial faithful distribution into a system that favors the featurally or perceptually less complex variant. This is precisely what likely underlies the higher rate of operation (initiation and phonologization) of fricative occlusion in the development of PND. Catalysis has several advantages – it explains the higher rate of operation of those sound changes that simplify an alternation, is able to derive sound change after L1 acquisition period, and provides an empirically based explanation for a question that is inherently difficult to answer – how biases in human cognition in general result in the observed distribution of patterns in human language. Exploring further implications of the proposed mechanism should yield a better understanding of the relationship between the historical transmission of language and the role that cognition plays in the phonological system of human language.

#### ACKNOWLEDGMENTS

The experiments were funded by the Mind Brain Behavior Interfaculty Initiative and the Department of Linguistics at Harvard University as well as by start-up funds for new faculty at the University of Washington. I would like to thank Kevin Ryan, Jay Jasanoff, Adam Albright, Donca Steriade, as well as the editors and reviewers at *Language* for useful comments on previous versions of this work. Special thanks go to the nine undergraduate research assistants at the University of Ljubljana who assisted with field experiments: Manca Uršič, Dafne Marko, Julija Kos, Karen Kavčič, Nina Kocjančič, Haris Agović, Denis Hacin, Ariela Herček, and Luka Horjak. I am grateful to Tatjana Marvin and Ana Mehle who helped coordinating the undergraduate research assistants and Samo Beguš who kindly allowed me to use the anechoic room at the Faculty of Electrotechnics, University of Ljubljana. All mistakes are my own.

## References

- ALBRIGHT, ADAM, and YOUNGAH DO. 2019. Three biases in learning phonological alternation. Ms., MIT and University of Hong Kong. Accessed on January 12, 2019. Online: <https://yd79.files.wordpress.com/2014/06/threebiases.pdf>.
- ALTARRIBA, JEANETTE. 1993. The influence of culture on cognitive processes. *Cognition and culture*, ed. by Jeanette Altarriba, *Advances in Psychology*, vol. 103, 379 – 384. North-Holland. Online: <http://www.sciencedirect.com/science/article/pii/S0166411508616738>.
- AVCU, ENES. 2018. Experimental investigation of the Subregular Hypothesis. *Proceedings of the 35th West Coast Conference on Formal Linguistics*, ed. by William G. Bennett, Lindsay Hracs, and Dennis Ryan Storoshenko, 77–86. Somerville, MA: Cascadilla.
- BAER-HENNEY, DINAH, and RUBEN VAN DE VIJVER. 2012. On the role of substance, locality, and amount of exposure in the acquisition of morphophonemic alternations. *Laboratory Phonology* 3.221–249.
- BAJEC, ANTON, ET AL. 2000. Slovar slovenskega knjižnega jezika. Online: <http://bos.zrc-sazu.si/sskj.html>.
- BECKER, MICHAEL; NIHAN KETREZ; and ANDREW NEVINS. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87.84–125.
- BECKER, MICHAEL, and JONATHAN LEVINE. 2013. Experigen – an online experiment platform. Online: <http://becker.phonologist.org/experigen>.
- BEGUŠ, GAŠPER. 2017. Effects of ejective stops on preceding vowel duration. *Journal of the Acoustical Society of America* 142.2168–2184.
- BEGUŠ, GAŠPER. 2018. *Unnatural phonology: A synchrony-diachrony interface approach*. Harvard University dissertation. Online: <http://nrs.harvard.edu/urn-3:HUL.InstRepos:40050094>.
- BEGUŠ, GAŠPER. 2019. Post-nasal devoicing and the blurring process. *Journal of Linguistics* 55.689–753.

- BEGUŠ, GAŠPER. 2020. Estimating historical probabilities of natural and unnatural processes. *Phonology* 37.515–549.
- BELL, ALAN. 1970. *A state-process approach to syllability and syllabic structure*. Stanford University dissertation.
- BELL, ALAN. 1971. Some patterns of the occurrence and formation of syllabic structure. *Working Papers on Language Universals* 6.23–138.
- BLEVINS, JULIETTE. 2004. *Evolutionary phonology*. Cambridge: Cambridge University Press.
- BROSELOW, ELLEN. 2018. Laryngeal contrasts in second language phonology. *Phonological typology*, ed. by Larry M. Hyman and Frans Plank, 312–340. Berlin: De Gruyter.
- BYBEE, JOAN. 2001. *Phonology and language use*. Cambridge Studies in Linguistics. Cambridge University Press.
- CARPENTER, ANGELA C. 2010. A naturalness bias in learning stress. *Phonology* 27.345–392.
- CATHCART, CHUNDRA A. 2015. A probabilistic model of Evolutionary Phonology. *Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society*, ed. by Thuy Bui and Deniz Özyıldız, 145–150. Amherst, MA: GLSA. Volume 1.
- CLARK, EVE V., and MELISSA BOWERMAN. 1986. On the acquisition of final voiced stops. *The fergusonian impact: In honor of charles a. ferguson on the occasion of his 65th birthday*, ed. by J. A. Fishman., vol. 1, 51–68. Berlin: De Gruyter.
- COETZEE, ANDRIES W., and RIGARDT PRETORIUS. 2010. Phonetically grounded phonology and sound change: The case of Tswana labial plosives. *Journal of Phonetics* 38.404–421.
- CROSS, IAN. 2012. Cognitive science and the cultural nature of music. *Topics in Cognitive Science* 4.668–677. Online: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1756-8765.2012.01216.x>.
- DAVIDSON, LISA. 2016. Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics* 54.35–50.
- DAVIDSON, LISA. 2017. Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association*, 1–26.
- DAVIDSON, LISA. 2018. Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association* 48.331–356.
- DICKENS, PATRICK J. 1984. The history of so-called strengthening in Tswana. *Journal of African Languages and Linguistics* 6.97–125.
- DO, YOUNGAH, and JONATHAN HAVENHILL. 2021. Production and substantive bias in phonological learning. *Proceedings of the 2020 Annual Meeting on Phonology*, ed. by Ryan Bennett, Richard Bibbs, Mykel Loren Brinkerhoff, Max J. Kaplan, Stephanie Rich, Amanda Rysling, Nicholas Van Handel, and Maya Wax Cavallaro, 101–114.

- DO, YOUNGAH; ELIZABETH ZSIGA; and JONATHAN HAVENHILL. 2016. Naturalness and frequency in implicit phonological learning. Talk presented at the 90th *Annual Meeting of the Linguistic Society of America*, Washington, DC, January 7-10, 2016.
- EFRON, BRADLEY. 1979. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* 7.1–26.
- EFRON, BRADLEY. 1987. Better bootstrap confidence intervals. *Journal of the American Statistical Association* 82.171–185.
- ETTLINGER, MARC; KARA MORGAN-SHORT; MANDY FARETTA-STUTENBERG; and PATRICK C.M. WONG. 2016. The relationship between artificial and second language learning. *Cognitive Science* 40.822–847. Online: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12257>.
- FERDINAND, VANESSA; SIMON KIRBY; and KENNY SMITH. 2019. The cognitive roots of regularization in language. *Cognition* 184.53 – 68. Online: <http://www.sciencedirect.com/science/article/pii/S0010027718303135>.
- FINLEY, SARA. 2012. Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes* 27.1550–1562. Online: <https://doi.org/10.1080/01690965.2012.660168>.
- FINLEY, SARA, and WILLIAM BADECKER. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language* 61.423 – 437. Online: <http://www.sciencedirect.com/science/article/pii/S0749596X09000564>.
- GARRETT, ANDREW, and KEITH JOHNSON. 2013. Phonetic bias in sound change. *Origins of sound change: Approaches to phonologization*, ed. by Alan C. L. Yu, 51–97. Oxford: Oxford University Press.
- GAUVAIN, MARY. 1995. Thinking in niches: Sociocultural influences on cognitive development. *Human Development* 38.25–45.
- GLEWWE, ELEANOR. 2017. Substantive bias in phonotactic learning: Positional extension of an obstruent voicing contrast. Talk presented at the 53rd meeting of *Chicago Linguistic Society*, Chicago, IL, May 25-27, 2017.
- GLEWWE, ELEANOR; JESSE ZYMET; JACOB ADAMS; RACHEL JACOBSON; ANTHONY YATES; ANN ZENG; and ROBERT DALAND. 2017. Substantive bias and word-final voiced obstruents: An artificial grammar learning study. Talk presented at the 92nd Annual Meeting of the Linguistic Society of America, Salt Lake City, UT, January 4-7, 2018.
- GREENBERG, JOSEPH H. 1978. Diachrony, synchrony, and language universals. *Universals of human language*, ed. by Joseph H. Greenberg, 61–92. Stanford, CA: Stanford University Press. Volume I: Method & Theory.
- GRIFFITHS, THOMAS L; MICHAEL L KALISH; and STEPHAN LEWANDOWSKY. 2008. Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363.3503–3514. Online: <https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2008.0146>.

- HANSSON, GUNNAR. 2008. Diachronic explanations of sound patterns. *Language and Linguistics Compass* 2.859–893.
- HAYDEN, REBECCA E. 1950. The relative frequency of phonemes in General-American English. *Word* 6.217–223.
- HAYES, BRUCE. 1999. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. *Functionalism and formalism in linguistics, volume i: General papers*, ed. by Michael Darnell and Edith Moravscik, 243–285. Amsterdam: John Benjamins.
- HAYES, BRUCE; PÉTER SIPTÁR; KIE ZURAW; and ZSUZSA LONDE. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85.822–863.
- HAYES, BRUCE, and DONCA STERIADE. 2004. Introduction: the phonetic bases of phonological markedness. *Phonetically based phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 1–33. Cambridge University Press.
- HAYES, BRUCE, and TANYA STIVERS. 2000. Postnasal voicing. Ms., University of California, Los Angeles. <http://linguistics.ucla.edu/people/hayes/Phonet/NCPhonet.pdf> (accessed 9 May 2018).
- HAYES, BRUCE, and JAMES WHITE. 2013. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44.45–75.
- HAYNIE, HANNAH J., and CLAIRE BOWERN. 2016. Phylogenetic approach to the evolution of color term systems. *Proceedings of the National Academy of Sciences* 113.13666–13671. Online: <https://www.pnas.org/content/113/48/13666>.
- HYMAN, LARRY M. 1975. *Phonology: Theory and analysis*. New York: Holt, Rinehart & Winston.
- HYMAN, LARRY M. 2001. The limits of phonetic determinism in phonology: \*NC revisited. *The role of speech perception in phonology*, ed. by Elizabeth Hume and Keith Johnson, 141–186. San Diego, CA: Academic Press.
- HYMAN, LARRY M. 2013. Enlarging the scope of phonologization. *Origins of sound change: Approaches to phonologization*, ed. by Alan C. L. Yu, 3–28. Oxford: Oxford University Press.
- IVERSON, GREGORY K., and JOSEPH C. SALMONS. 1995. Aspiration and laryngeal representation in germanic. *Phonology* 12.369–396. Online: <http://www.jstor.org/stable/4420084>.
- KESSLER, BRETT, and REBECCA TREIMAN. 1997. Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language* 37.295 – 311. Online: <http://www.sciencedirect.com/science/article/pii/S0749596X97925225>.
- KIPARSKY, PAUL. 1995. The phonological basis of sound change. *Handbook of phonological theory*, ed. by John Goldsmith, 640–670. Oxford: Blackwell, 1995.
- KIPARSKY, PAUL. 2006. Amphichronic program vs. Evolutionary Phonology. *Theoretical Linguistics* 32.217–236.

- KIPARSKY, PAUL. 2008. Universals constrain change, change results in typological generalizations. *Linguistic universals and language change*, ed. by Jeff Good, 23–53. Oxford: Oxford University Press.
- KIRBY, SIMON; MIKE DOWMAN; and THOMAS L. GRIFFITHS. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences* 104.5241–5245. Online: <https://www.pnas.org/content/104/12/5241>.
- KIRBY, SIMON; TOM GRIFFITHS; and KENNY SMITH. 2014. Iterated learning and the evolution of language. *Current Opinion in Neurobiology* 28.108 – 114, SI: Communication and language. Online: <http://www.sciencedirect.com/science/article/pii/S0959438814001421>.
- KLUENDER, KEITH R.; RANDY L. DIEHL; and BEVERLY A. WRIGHT. 1988. Vowel-length differences before voiced and voiceless consonants: an auditory explanation. *Journal of Phonetics* 16.153 – 169. Online: <http://www.sciencedirect.com/science/article/pii/S0095447019304802>.
- KONG, EUN JONG; MARY E. BECKMAN; and JAN EDWARDS. 2012a. Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: The cases of Greek and Japanese. *Journal of Phonetics* 40.725 – 744. Online: <http://www.sciencedirect.com/science/article/pii/S0095447012000526>.
- KONG, EUN JONG; MARY E. BECKMAN; and JAN EDWARDS. 2012b. Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: The cases of Greek and Japanese. *Journal of Phonetics* 40.725 – 744. Online: <http://www.sciencedirect.com/science/article/pii/S0095447012000526>.
- KÜMMEL, MARTIN. 2007. *Konsonantenwandel*. Wiesbaden: Reichert.
- KUO, LI-JEN. 2009. The role of natural class features in the acquisition of phonotactic regularities. *Journal of Psycholinguistic Research* 38.129–150. Online: <https://doi.org/10.1007/s10936-008-9090-2>.
- LABOV, WILLIAM. 1994. *Principles of linguistic change*. Oxford: Blackwell. Volumes 1 & 2.
- DE LACY, PAUL. 2006a. *Markedness: Reduction and preservation in phonology*. Cambridge Studies in Linguistics. Cambridge University Press.
- DE LACY, PAUL. 2006b. Transmissibility and the role of the phonological component: A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32.185–196.
- DE LACY, PAUL, and JOHN KINGSTON. 2013a. Synchronic explanation. *Natural Language & Linguistic Theory* 31.287–355. Online: <https://doi.org/10.1007/s11049-013-9191-y>.
- DE LACY, PAUL, and JOHN KINGSTON. 2013b. Synchronic explanation. *Natural Language and Linguistic Theory* 31.287–355.
- LADEFOGED, PETER, and IAN MADDIESON. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

MARVIN, TATJANA; JURE DERGANC; SAMO BEGUŠ; and SABA BATTELINO. 2018. Word selection in the Slovenian sentence matrix test for speech audiometry. *Zbornik konference jezikovne tehnologije in digitalna humanistika*, ed. by Darja Fišer and Andrej Pančur, 181–187. Ljubljana: Znanstvena založba Filozofske fakultete v Ljubljani. Online: [www.sjdt.si/wp/dogodki/konference/jtdh-2018/zbornik-jtdh-2018/](http://www.sjdt.si/wp/dogodki/konference/jtdh-2018/zbornik-jtdh-2018/).

MIELKE, JEFF. 2018. PBase: A database of phonological patterns. Online: <http://pbase.phon.chass.ncsu.edu>.

MINES, M. ARDUSSI; BARBARA F. HANSON; and JUNE E. SHOUP. 1978. Frequency of occurrence of phonemes in conversational English. *Language and speech* 21.221–241, Last updated - 2013-02-23. Online: <http://search.proquest.com.ezp-prod1.hul.harvard.edu/docview/1299110032?accountid=11311>.

MORAN, STEVEN; DANIEL McCLOY; and RICHARD WRIGHT (eds.) 2014. *Phoible online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: <https://phoible.org/>.

MORETON, ELLIOTT. 2008. Analytic bias and phonological typology. *Phonology* 25.83–127.

MORETON, ELLIOTT, and JOE PATER. 2012a. Structure and substance in artificial-phonology learning. Part I, Structure. *Language and Linguistics Compass* 6.686–701.

MORETON, ELLIOTT, and JOE PATER. 2012b. Structure and substance in artificial-phonology learning. Part II, Substance. *Language and Linguistics Compass* 6.702–718.

MORETON, ELLIOTT; JOE PATER; and KATYA PERTSOVA. 2017. Phonological concept learning. *Cognitive Science* 41.4–69. Online: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12319>.

MORETON, ELLIOTT, and KATYA PERTSOVA. 2017. Implicit and explicit processes in phonotactic learning. *Proceedings of the 40th boston university conference on language acquisition (BUCLD 40)*, ed. by Jennifer Scott and Deb Waugtal, 277–290.

MORLEY, REBECCA L. 2012. The emergence of epenthesis: An incremental model of grammar change. *Language Dynamics and Change* 2.59–97.

MORLEY, REBECCA L. 2015. Can phonological universals be emergent? Modeling the space of sound change, lexical distribution, and hypothesis selection. *Language* 91.e40–e70.

MYERS, SCOTT. 2002. Gaps in factorial typology: the case of voicing in consonant clusters. Ms., University of Texas, Austin. ROA.

NISBETT, RICHARD E., and ARA NORENZAYAN. Culture and cognition. *Stevens' handbook of experimental psychology*, 561–597. Wiley-Blackwell. Online: <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471214426.pas0213>.

OHALA, JOHN J. 1981. The listener as a source of sound change. *Papers from the parasession on language and behavior*, ed. by Carrie S. Masek, Roberta A. Hendrick, and Mary Frances Miller, 178–203. Chicago: Chicago Linguistic Society.

- OHALA, JOHN J. 1983. The origin of sound patterns in vocal tract constraints. *The production of speech*, ed. by Peter F. MacNeilage, 189–216. New York: Springer.
- OHALA, JOHN J. 1993. The phonetics of sound change. *Historical linguistics: Problems and perspectives*, ed. by Charles Jones, 237–278. London: Longman.
- OHALA, JOHN J. 1997. Aerodynamics of phonology. *Proceedings of the 4th Seoul International Conference on Linguistics (SICOL)*, 84–91. Seoul: Linguistic Society of Korea.
- OHALA, JOHN J. 2006. Speech aerodynamics. *Encyclopedia of language & linguistics*, ed. by Keith Brown, 684–689. Oxford: Elsevier. Second Edition.
- PATER, JOE. 2004. Austronesian nasal substitution and other NC effects. *Optimality theory in phonology: A reader*, ed. by John J. McCarthy, 271–289. Malden, MA: Blackwell.
- PICARD, MARC. 1994. *Principles and methods in historical phonology: From Proto-Algonkian to Arapaho*. Montreal: McGill-Queen's University Press.
- PRINCE, ALAN, and PAUL SMOLENSKY. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell. First published in 1993, Tech. Rep. 2, Rutgers University Center for Cognitive Science.
- PYCHA, ANNE; PAWEŁ NOWAK; EURIE SHIN; and RYAN SHOSTED. 2003. Phonological rule-learning and its implications for a theory of vowel harmony. *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, ed. by Gina Garding and Mimu Tsujimura, 101–114. Somerville, MA: Cascadilla Press.
- RAFFERTY, ANNA N.; THOMAS L. GRIFFITHS; and MARC ETTLINGER. 2013. Greater learnability is not sufficient to produce cultural universals. *Cognition* 129.70 – 87. Online: <http://www.sciencedirect.com/science/article/pii/S0010027713000966>.
- REALI, FLORENCIA, and THOMAS L. GRIFFITHS. 2009. The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition* 111.317 – 328. Online: <http://www.sciencedirect.com/science/article/pii/S0010027709000432>.
- SANKOFF, GILLIAN, and HÉLÈNE BLONDEAU. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83.560–588. Online: <http://www.jstor.org/stable/40070902>.
- SCHLOSS, KAREN B., and STEPHEN E. PALMER. 2011. Aesthetic response to color combinations: preference, harmony, and similarity. *Attention, Perception, & Psychophysics* 73.551–571. Online: <https://doi.org/10.3758/s13414-010-0027-0>.
- SEIDL, AMANDA; EUGENE BUCKLEY; and ALEJANDRINA CRISTIÀ. 2007. Complexity trumps naturalness. Talk presented at the 81st Annual Meeting of the Linguistic Society of America, Anaheim, CA, January 4–7, 2007.
- SIMONS, GARY F., and CHARLES D. FENNIG (eds.) 2018. *Ethnologue: Languages of the world*. 21st edn. Dallas, TX: SIL International. [www.ethnologue.com](http://www.ethnologue.com) (accessed 13 October 2018).

- SINGH, SADANAND, and JOHN W. BLACK. 1966. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *The Journal of the Acoustical Society of America* 39.372–387. Online: <https://doi.org/10.1121/1.1909899>.
- SKORUPPA, KATRIN, and SHARON PEPERKAMP. 2011. Adaptation to novel accents: Feature-based learning of context-sensitive phonological regularities. *Cognitive Science* 35.348–366. Online: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1551-6709.2010.01152.x>.
- SMITH, CAROLINE L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25.471 – 500. Online: <http://www.sciencedirect.com/science/article/pii/S009544709790053X>.
- SOLÉ, MARIA-JOSEP; LARRY M. HYMAN; and KEMMONYE C. MONAKA. 2010. More on post-nasal devoicing: The case of Shekgalagari. *Journal of Phonetics* 38.299–319.
- STERIADE, DONCA. 2001. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Ms., University of California, Los Angeles.
- SUHADOLC, BARBARA. 2013. *Statistična analiza slovenskih besedil [Statistic analysis of Slovenian texts]*. Master's thesis, University of Ljubljana. Online: [eprints.fri.uni-lj.si/2157/](https://eprints.fri.uni-lj.si/2157/).
- TESSIER, ANNE-MICHELLE. 2012. Testing for OO-faithfulness in the acquisition of consonant clusters. *Language Acquisition* 19.144–173. Online: <https://doi.org/10.1080/10489223.2012.660552>.
- TOPORIŠIČ, JOŽE. 2004. *Slovenska slovnica*. Maribor: Obzorja.
- VERDONIK, DARINKA, and ANA ZWITTER VITEZ. 2011. *Slovenski govorni korpus Gos*. Ljubljana: Trojina, zavod za uporabno slovenistiko. Online: [www.korpus-gos.net](http://www.korpus-gos.net).
- VAN DE VIJVER, RUBEN, and DINAH BAER-HENNEY. 2014. Developing biases. *Frontiers in Psychology* 5.634. Online: <https://www.frontiersin.org/article/10.3389/fpsyg.2014.00634>.
- ŠUŠTARŠIČ, RASTISLAV; SMILJANA KOMAR; and BOJAN PETEK. 1995. Slovene. *Journal of the International Phonetic Association* 25.86–90.
- WANG, WILLIAM S.-Y., and JOHN CRAWFORD. 1960. Frequency studies of English consonants. *Language & Speech* 3.131 – 139.
- WHITE, JAMES. 2014. Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130.96 – 115. Online: <http://www.sciencedirect.com/science/article/pii/S0010027713001923>.
- WHITE, JAMES. 2017. Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Language* 93.1–36.
- WHITE, JAMES, and MEGHA SUNDARA. 2014. Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition* 133.85 – 90. Online: <http://www.sciencedirect.com/science/article/pii/S0010027714001097>.

WILSON, COLIN. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30.945–982.

YU, ALAN C. L. 2011. On measuring phonetic precursor robustness : a response to Moreton. *Phonology* 28.491–518. Online: <http://www.jstor.org/stable/41475374>.

# Supplementary Materials: *Distinguishing cognitive from historical influences in phonology*

Gašper Beguš

*University of California, Berkeley*

## 1 Experiments

### 1.1 Experiment 1: English

Altogether, 203 subjects participated in the English experiment. Subjects were recruited using Amazon Mechanical Turk. Of the 203 participants, 198 finished the experiment in full and completed a final demographic questionnaire at the end of the experiment. Of the 170 English participants included in the final analysis (see Section 4.6 in the main paper for exclusion criteria), 166 filled out a demographic questionnaire. Age of the participants ranged from 19–64 years ( $M = 33.8$ ,  $SD = 10.1$ ).<sup>1</sup> 74 identified as female, 80 as male, and one as other.<sup>2</sup>

To diversify the subject pool, recruitment included three sets of 80 participants with different conditions and payment amounts. The last recruitment was stopped after 43 participants reported completion. A US IP-address was a requirement for all three recruitments. The first recruitment included no further requirements, the second required ‘HIT Approval Rate (%) for all Requesters’ HITs greater than 99’ and ‘Number of HITs Approved greater than 50’. The third recruitment required ‘HIT Approval Rate (%) for all Requesters’ HITs greater than 99’ and ‘Number of HITs Approved greater than 500’. Participants were paid \$3.30, \$2.00, and \$2.50 for their participation according to the three recruitment groups described above. Average time to completion across all three groups was approximately 22 min (28 min, 19 min, 20 min).

#### 1.1.1 Materials and procedure

The experiment was conducted using the Experigen experimental software (Becker & Levine, 2013), adjusted by Adam Albright (Albright & Do, 2019) and additionally adjusted by the author. Most of the experimental outline and instructions are as provided by Experigen and Albright’s modified version of it. Subjects were told that they would learn a made-up language called Martian and would be asked to provide some answers about this language. First, subjects were instructed to go to a quiet room and use headphones. To prompt subjects to do so, they were asked to transcribe a nonce-word in Martian. After this initial task, subjects had to agree to informed consent if they wished to proceed and were given some basic Martian spelling instructions with recorded examples and instructions on how to proceed with the experiment. Subjects were told that plural nouns in Martian are formed by either [ɔn-] or [ɛn-] (orthographically <on-> and <en->) and were given two preliminary examples. Subjects were encouraged to repeat words they would hear and were told that they would ‘be asked some questions about Martian singular and plural forms’ at the end of the experiment.<sup>3</sup>

<sup>1</sup>Estimated from self-reported year of birth.

<sup>2</sup>Nine participants did not answer, 2 refused to answer.

<sup>3</sup>Subjects were given the following instructions (adjusted from Adam Albright’s version of Experigen):

### 1.1.2 Stimuli

The training phase consisted of 58 singular-plural pairs of words. Twelve of the 58 items were filler words of the shape C<sub>1</sub>V<sub>2</sub>C<sub>3</sub>V<sub>4</sub> or C<sub>1</sub>V<sub>2</sub>C<sub>3</sub>. C<sub>1</sub>s in the filler group consisted of [j], [l], and [r] that did not change in the plural. Four items were created for each initial consonant C<sub>1</sub>, so that each group contained two mono- and two disyllabic words. Each group also contained two front first vowels ([ɛ, i]) and two non-front vowels ([a, ɔ, u]). C<sub>2</sub>s were taken from the following inventory: [s, r, l, n, m] and V<sub>4</sub> included vowels from the following inventory: [ɔ, u, i]. Table 1 lists all filler words used in the experiment.

Fourteen (out of 58) stimuli were created with initial voiceless stops and fricatives in order to train subjects on non-alternating voiceless fricatives (see Table 7 in the main paper). Because the English inventory lacks the voiced velar fricative, the experiment involves only labials ([p, f]) and coronals ([t, s]). All words were of the shape C<sub>1</sub>V<sub>2</sub>C<sub>3</sub>V<sub>4</sub>. Six words started with initial voiceless stops [p] and [t] (3 each). Four of the six words contained a front V<sub>2</sub> ([i, ɛ]) that triggered frontness harmony; two contained a back V<sub>2</sub> ([a, ɔ, u]). Eight items were created with initial voiceless fricatives [f] and [s], four for each group. Half of the words per group involved a front V<sub>2</sub> ([i, ɛ]). C<sub>3</sub> included consonants from the following inventory: [r, l, n, m]. V<sub>4</sub> included vowels from the following inventory: [i, ə, ɔ, u]. Table 2 lists all nonce words with voiceless fricatives or stops in initial position used in the training phase of the experiment.

The remaining 32 items of the training phase contain evidence for PND and PFOD and were created with initial voiced stops (16 words) and voiced fricatives (16 words). Because the test phase of the implicit task (PND vs. PFOD) involves a choice between stops and fricatives, the two groups needed to be maximally balanced. Half of the items involved a labial-initial ([b, v]) and a half a coronal-initial ([d, z]) obstruent (fully balanced). In each group, half of the items contained a front V<sub>2</sub> that triggered vowel harmony ([i, ɛ]) and half contained V<sub>2</sub> ([a, ɔ, u]). The design was thus as balanced as possible in order to minimize undesired effects that are not controlled for in the experiment. All words in this group were of the shape C<sub>1</sub>V<sub>2</sub>C<sub>3</sub>V<sub>4</sub>. In order to control for influences of the sequence -V<sub>2</sub>C<sub>3</sub>V<sub>4</sub> following the initial consonant that is of interest to the experiment, design was additionally balanced for vowel (V<sub>2</sub> and V<sub>3</sub>) and consonant identity (C<sub>3</sub>) across the stop vs. fricative groups. In other words, each of the two groups of items (stop vs. fricative) that are of primary interest in the experiment had approximately the same inventory and number of vowels and consonants in V<sub>2</sub>, C<sub>3</sub>, and V<sub>4</sub> in both the labial and coronal groups. To avoid too many minimal pairs, the order of vowels in V<sub>2</sub>, C<sub>3</sub>, and V<sub>4</sub> was mixed across the subgroups. For example, stimuli with the initial labial voiced stop ([b]) had the following inventory: V<sub>2</sub> = [i, i, ɛ, ɛ, ɔ, a, u, u]; C<sub>3</sub> = [l, l, l, r, r, r, n, m]; V<sub>4</sub> = [i, ɛ, ɛ, ɔ, ɔ, ə, ə, u]. Stimuli with initial voiced fricatives ([v]) had exactly the same inventories, but in different combinations to avoid too many minimal pairs.<sup>4</sup> Some minimal pairs were included in the training phase.<sup>5</sup> All items, including the non-filler and test items, were stressed on the first syllable of the root, which means that the first syllable is stressed in the singular (without the prefix) and the second syllable in the plural (with the prefix), e.g. singular [balu]; plural [om'pʰalu]. Table 3 lists all nonce words with voiced stops and fricatives in initial position used in the training phase of the experiment.

Stimuli were read by a phonetically trained native female speaker of American English who was un-

- 
- In order to make sure that you are paying attention to how Martian words sound, we will periodically ask you to type the word that you have just heard.
  - Don't despair if you don't see a pattern to the use of the en- and on! Just listen to the words, and by the end, you will probably have an opinion about which one 'sounds right'.
  - Don't try to memorize every word that you hear. You will not be asked to remember the meanings of the words.

<sup>4</sup>In the coronal series, C<sub>3</sub> of the stop-initial group consisted of one [l] more and one [n] less than the fricative-initial group. In the labial series, V<sub>4</sub> of the stop-initial group consisted of one [ɔ] more and one [ɔ̄] less than the fricative-initial group.

<sup>5</sup>The minimal pairs were [tels] and [dels], [dilɔ] and [bilɔ], [bunɛ] and [zunɛ], [dule] and [vule], [dema] and [bema] or [taru], [zaru], and [varu]; [bilɔ], [zilɔ], and [vilɔ].

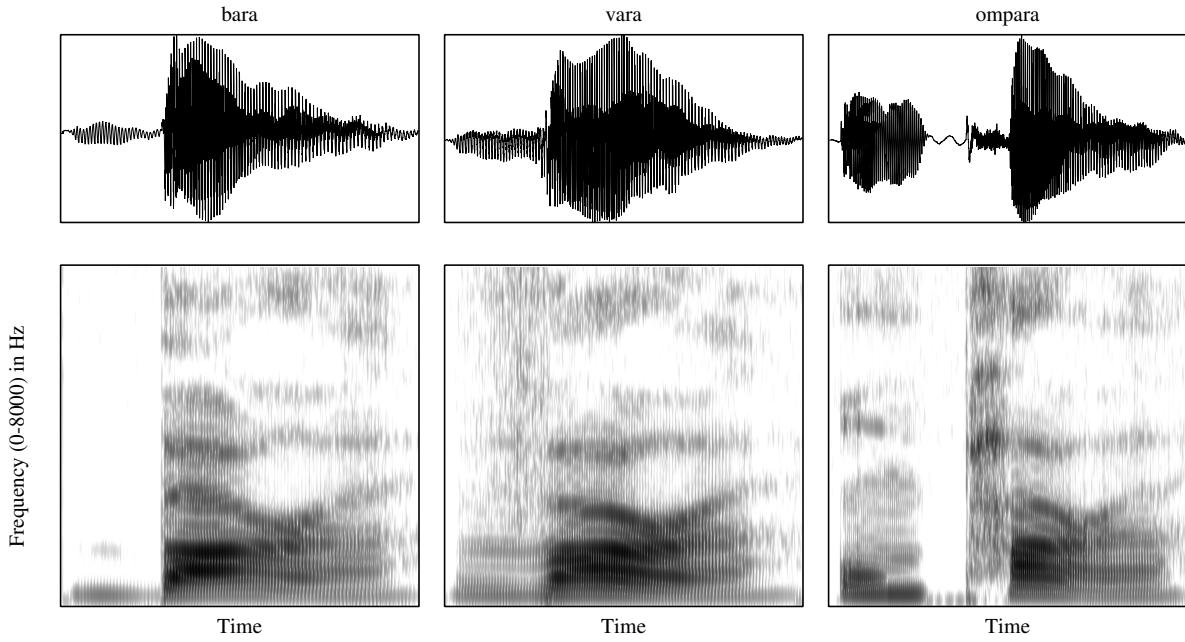


Figure 1: Waveforms and spectrograms

of [ˈbaɾə] (left), [ˈvaɾə] (center) and [ɔ̃mˈpaɾə] (left). Subjects were first played the plural form (the third in the diagram) and were asked to choose between the other two. Spectrograms illustrate a clear period of prevoicing. Frication is visible in realization of /v/. Note that the low frequency energy in the [ɔ̃mˈpaɾə] is environmental noise and not voicing.

aware of the task of the experiment. The speaker was compensated for participation. The speaker was asked to check if any word sounded familiar or English-like and reported no such words in the final inventory. Recordings were made in a sound-attenuated booth using Shure Beta 53 omnidirectional condenser head-mounted microphone and Sound Devices USB Pre2 pre-amplifier with Audacity recording software, sampled at 44.1 kHz (16-bit). To facilitate a constant rate and intensity across recorded words, the speaker was instructed to read the stimuli in a carrier phrase ‘Say [stimulus]’ and to make a long pause between the carrier word ‘Say’ and the stimulus. The speaker was also asked to produce voiced stops with prevoicing.

Stimuli were presented in a randomized order to the speaker on sheets of paper, but singular-plural pairs of the same noun were presented together. After the recording, every word-initial voiced stop was manually inspected for voicing. The speaker failed to prevoice voiced stops in some stimuli. Those were replaced by novel recordings, so that all word-initial voiced stops were prevoiced in the final inventory of recorded stimuli. To control for the effect of loudness, stimuli were RMS-equalized in Praat (Boersma & Weenink, 2015) using the RMS EQUALIZE Praat script written by Gabriel J. L. Beckers (Beckers, 2002). Recordings were converted to .mp3 format from .wav format using Audacity’s LAME mp3 encoder.

Figures 1 and 2 feature sample waveforms and spectrograms of experimental stimuli used in English and Slovenian experiments. Spectrograms illustrate different durations of aspiration between the two languages, clear pre-voicing of voiced stops in both languages, and frication during the production of Slovenian /v/.

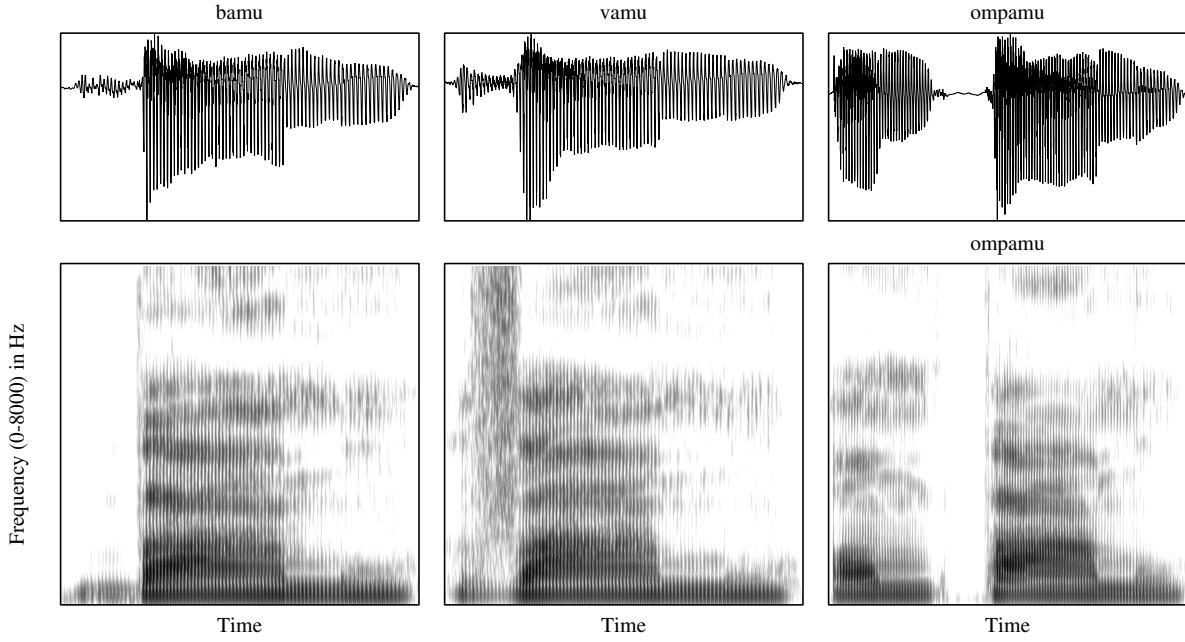


Figure 2: Waveforms and spectrograms of ['bamu] (left), ['vamu] (center) and [ɔm'pamu] (left). Subjects were first played the plural form (the third in the plot) and were asked to choose between the other two. Spectrograms illustrate a clear period of prevoicing. Frication is visible in realization of /v/.

## 1.2 Experiment 2: Slovenian

### 1.2.1 Materials and procedure

Of the 110 Slovenian participants included in the final analysis, 109 filled out the demographic questionnaire. Age ranged from 19–60 years ( $M = 29.9$ ,  $SD = 12.0$ ).<sup>6</sup> 57 identified as female, 47 as male, one as other.<sup>7</sup>

The training phase consisted of 60 singular-plural pairs. All stimuli were disyllabic words of the shape  $C_1V_2C_3C_4$  (singular) and *prefix*- $C_1V_2C_3C_4$  (plural) with a stressed and long  $V_2$  (stressed vowels are generally long in Slovenian). Twelve items were filler words, where  $C_1$  consists of non-alternating [j, l, r] and  $C_3 = [m, n, l, r]$ . In half of the filler items,  $V_2$  is a front vowel ([i:, e:]) with the other half being non-front [a:, o:, u:]. Subjects were trained on non-alternating voiceless stops and fricatives in sixteen items. The  $C_1$  in these sixteen items is a labial or a coronal voiceless stop or fricative [p, t, f, s] (4 each). Again, in half of these items,  $V_2$  is a front vowel (fully balanced).  $C_3$ s consist of [m, n, l, r]. Subjects were trained on PND and PFOD with 32 items.  $C_1$  in these 32 items are [b, d, v, z] (8 each) with half of them having a front  $V_2$  (fully balanced). To maximally balance the design,  $V_2$  consisted of an equal number of vowels [e:, i:, a:, o:, u:] across the fricative vs. stop-initial groups for each place of articulation (as described in Section 1.1.1). Likewise,  $C_1$  and  $V_4$  consisted of an equal number of consonants [m, n, l, r] and vowels [i, a, ɔ, u]. This maximally balanced design aims to control for unwanted influences of the  $-V_2C_3V_4$  sequence following the consonant of interest,  $C_1$ . To avoid real words, some deviations from the fully balanced design were necessary. A nasal [n] in  $C_3$  of the [b]-initial group was substituted with [l] in the [v]-initial group and one  $V_2$  [u] in the [b]-initial group was substituted with [e] in the [v]-initial group.<sup>8</sup>

In the test phase (see also Section 4.4 in the main text),  $C_3$  was a [j] in ten items, which did not appear as  $C_3$  in the training phase. In two items,  $C_1$  was a [m] and a [r], which are fully balanced consonants in the training phase, meaning that both voiced-stop- and fricative-initial items contain an equal number of [m] and [r] phonemes.

Stimuli for the experiment were read by a phonetically trained native female speaker of Slovenian with no prior knowledge of the experiment. The speaker volunteered her time. Recordings took place in anechoic chamber at the Faculty of Electrical Engineering at the University of Ljubljana, Slovenia using AKG C544L cardioid condenser head-mounted microphone connected via AKG MPAVL cable and Sound Devices USB Pre2 pre-amplifier (with 15 dB pad) using Praat recording software (Boersma & Weenink, 2015), sampled at 44.1 kHz (16-bit). The approximate distance from mouth to the microphone was 2 cm.

Nine undergraduate research assistants from the University of Ljubljana helped recruit the participants. The experimental interface, designed as described above in Experigen (Becker & Levine, 2013) was presented to subjects using laptop or desktop computers. To control for auditory presentation and experimental conditions, every subject wore the same Sony MDR7506 headphones that were connected to FiiO E10K USB digital-to-analog (DAC) converter and headphone amplifier. The volume level of the DAC converter was set at 3 for all subjects with low gain and low bass settings on the device. Undergraduate research assistants supervised each participant during the experiment, took informed consent, and filled out a short questionnaire describing each session. Subjects were not compensated for their time.

Stimuli were adjusted to conform to Slovenian phonology and to ensure none of the stimuli were actual words of Slovenian or resembled Slovenian words according to judgments by native-speaking research

<sup>6</sup>Estimated from self-reported year of birth.

<sup>7</sup>Four participants did not answer. One participant is additionally excluded because they only provided one answer on the implicit task and none on the explicit task, which means that their HARMONY score cannot be calculated.

<sup>8</sup>Three items are homophonous in the Slovenian experiment: [ɔnta:nu] that goes back to [za:nu] and [ta:nu]; [ɔntu:lɔ] that goes back to [du:lɔ] and [tulɔ], and [ɛmpi:ma] that goes back to [bi:ma] and [vi:ma]. The homophonous forms are thus equally distributed between the stop vs. fricative groups. [bima] and [vima] are the only minimal pairs within the voiced stop vs. fricative group. More minimal pairs were present across all initial consonants  $C_1$ .

Figure 3: A spectrogram  
of Slovenian /v/ in a made-up stimulus '/vamu/ (0–15000 Hz) with frication noise.

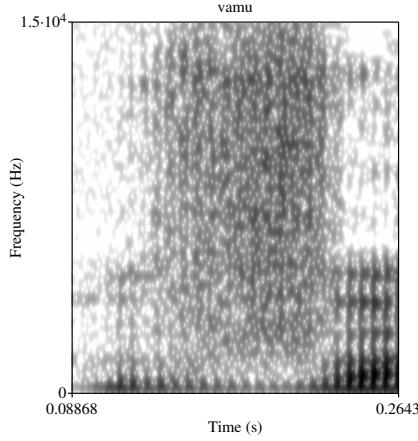


Table 1: Filler words in the English experiment.

Fillers					
#	Harm.	Sg.	Pl.	Orthography	
[l]	[+fr]	'len	en'len	len	enlen
	[l]	'lin'	en'lin	lino	enlino
	[-fr]	'lor	re'lōn	lor	onlor
[r]	[+fr]	'rel	er'el	rel	enrel
	[r]	'rinu	er'nu	rinu	enrinu
	[-fr]	'ras	sar'as	ras	onras
[j]	[+fr]	'yim	en'jim	yim	enyim
	[j]	'jeni	en'jeni	yeni	enyeni
	[-fr]	'jam	ən'jam	yam	onyam
		'jalu	ən'jalu	yalu	onyalu

assistants. All instructions were translated into Slovenian.

### 1.3 Stimuli

### 1.4 Subjects: exclusion criteria (additional information)

After the test phase, subjects were asked to fill out a short demographic questionnaire. They were asked about their year of birth, gender, place of abode, where other people think they are from (based on their speech), whether they are native speakers of English, whether they have taken any linguistics classes, and how did they choose their answers (based on a rule, intuition, or guesses).

Four subjects in the English experiment and two subjects in the Slovenian experiment started the test phase, but were interrupted probably due to technical issues and did not finish it. These subjects were nevertheless included in the analysis, except in the analysis for Figure 4 in the main paper and except if they have no recorded responses for one of the two categories (explicit/implicit task; one such subject per two experiments). The remaining subjects did not finish the test phase primarily due to technical issues.

Table 2: Nouns with initial voiceless stops and fricatives in the English training phase.

		Voiceless					
Place	#	Harm.	Sg.	Pl.		Orthography	
Labial	[-cont]	[+fr]	'hini	em'hini	pina	empina	
			'himi	em'himi	pimi	empimi	
		[-fr]	crch'	crch'd'me	poro	omporo	
	[+cont]	[+fr]	'fini	em'fini	fini	emfini	
			'fim	em'fim	fima	emfima	
		[-fr]	erfu'	erfu'me	fura	omfura	
Coronal	[-cont]	[+fr]	'tela	en'tela	telo	entelo	
			'tlini	en'tlini	tina	entina	
		[-fr]	'tha.ru	en'tha.ru	taru	ontaru	
	[+cont]	[+fr]	'cuso	cu'seo	seno	enso	
			'silis	elis'en	sila	ensila	
		[-fr]	'crs	crs'uc	soro	onsoro	
			'sanu	on'sanu	sanu	onsanu	

The analysis includes subjects who had indicated that they had linguistic education, but it was later revealed to one of the research assistants that they only took Slovenian language and literature classes as part of the regular high school curriculum that includes basic linguistic concepts. If a subject with the same Amazon MTurk Worker ID submitted their responses twice, the second set of responses were excluded from the analysis. Subjects with the same IP address but different Amazon MTurk Worker IDs were excluded if they had similar answers on their demographic questionnaire. One subject was excluded because they only submitted one test phase response.

### 1.5 Statistical analysis: Explicit task

The model was fit using the `glmer()` function from the `LME4` package (Bates et al., 2015) in the statistical software R (R Core Team, 2018). A full model was first fit with FRONTNESS of the vowel (back for [ɔ] and front for [ɛ]), crossed random slopes for SUBJECT and ITEM and by-subject and by-item random slopes for FRONTNESS. The non-converging models were excluded. The optimal model was selected based on Akaike Information Criterion (AIC) by first removing random slopes and then removing fixed effects predictors from the model. The final model includes only the intercept in the fixed structure and intercepts for SUBJECT and ITEM in the random structure.

In the Slovenian experiment, the model chosen with the same technique includes FRONTNESS in its fixed effects structure and random intercepts for SUBJECT and ITEM.

Mild underdispersion (estimated as described in Bolker 2019) (0.9 in the English and Slovenian experiments) should not crucially affect the results.

### 1.6 Statistical analysis: Implicit task

First, a full model was fit using the `glmer()` function (Bates et al., 2015) with three predictors: PLACE of articulation of C<sub>1</sub> (with two levels, labial and coronal), scaled and centered HARMONY scores, which represent the number of correct responses per subject on the explicit task, and the FRONTNESS of the prefix vowel. All interactions were included in the initial full model. The HARMONY scores were included in the model to test whether learning of harmony correlates with outcomes in the implicit task. In the English experiment,

Table 3: Nouns with initial voiced stops and fricatives in the English training phase. Orthographic stimuli were only presented to the phonetically trained speaker who recorded the stimuli and never to the subjects.

			Voiced			
	#_	Harm.	Sg.	Pl.	Orthography	
Labial	[+fr]		'bilə	ɛm'p <sup>h</sup> ilə	bila	empila
			'berə	ɛm'p <sup>h</sup> erə	bera	empera
			'bilo	ɛm'p <sup>h</sup> ilɔ	bilo	empilo
			'bemə	ɛm'p <sup>h</sup> emə	bema	empema
	[-cont]		'bulə	ɛm'p <sup>h</sup> ulə	bula	ompula
			'balu	ɛm'p <sup>h</sup> alu	balu	ompalu
			'bɔrə	ɛrɔd'mə	bora	ompora
			'bunə	ɛm'p <sup>h</sup> unə	bune	ompune
Coronal	[+fr]		'vilə	ɛm'p <sup>h</sup> ilə	vila	empila
			'vəmə	ɛm'p <sup>h</sup> emə	vemo	empemo
			'virə	ɛm'p <sup>h</sup> irə	vira	empira
			'velə	ɛm'p <sup>h</sup> elə	vela	empela
	[-fr]		'vulə	ɛm'p <sup>h</sup> ulɔ	vulo	ompulo
			'va.ru	ɛm'p <sup>h</sup> a.ru	varu	omparu
			'vɔnə	ɛm'p <sup>h</sup> ɔnə	vona	ompona
			'vule	ɛm'p <sup>h</sup> ulɛ	vule	ompule
Dental	[+fr]		'dilə	ɛn't <sup>h</sup> ilɔ	dilo	entilo
			'diri	ɛn't <sup>h</sup> iri	diri	entiri
			'dełɔ	ɛn't <sup>h</sup> elɔ	delo	entelo
			'demə	ɛn't <sup>h</sup> emə	dema	entema
	[-fr]		'dułe	ɛn't <sup>h</sup> ule	dule	ontule
			'doru	ɛn't <sup>h</sup> ɔru	doru	ontoru
			'dale	ɛn't <sup>h</sup> alɛ	dale	ontale
			'dunə	ɛn't <sup>h</sup> unə	duna	ontuna
Velar	[+fr]		'zilə	ɛn't <sup>h</sup> ilə	zila	entila
			'zirə	ɛn't <sup>h</sup> irə	zira	entira
			'zəmə	ɛn't <sup>h</sup> emə	zemo	entemo
			'zeni	ɛn't <sup>h</sup> eni	zeni	enteni
	[-fr]		'zulɔ	ɛn't <sup>h</sup> ulɔ	zulo	ontulo
			'zaru	ɛn't <sup>h</sup> a.ru	zaru	ontaru
			'zole	ɛn't <sup>h</sup> ɔlɛ	zole	ontole
			'zunə	ɛn't <sup>h</sup> unə	zune	ontune

Table 4: Nouns in the English test phase. Orthographic stimuli were only presented to the phonetically trained speaker who recorded the stimuli and never to the subjects.

		Test					
Place	Harm.	Pl.	Sg. 1	Sg. 2	Orthography		
Labial	[+fr]	em'p <sup>h</sup> ene	'bene	'vene	empene	bene	vene
		em'p <sup>h</sup> εjɔ	'bεjɔ	'vεjɔ	empeyo	beyo	veyo
	[-fr]	əm'p <sup>h</sup> aјε	'baјε	'vajε	ompaya	baya	vaya
		əm'p <sup>h</sup> ařař	'erbař	'ervař	ompara	bara	vara
Coronal	[+fr]	en't <sup>h</sup> ele	'dele	'zεle	entele	dele	zele
		en't <sup>h</sup> iwɔ	'diwɔ	'ziwɔ	entiwo	diwo	ziwo
	[-fr]	ən't <sup>h</sup> ami	'dami	'zami	ontami	dami	zami
		ən't <sup>h</sup> awɔ	'dawɔ	'zawɔ	ontawo	dawo	zawo

Table 5: Filler words in the Slovenian experiment.

		Fillers			
#_	Harm.	Sg.	Pl.	Orthography	
[l]	[+fr]	'le:mu	ən'le:mu	lemu	enlemu
		'li:ru	ən'li:ru	liru	enliru
	[-fr]	'lu:rε	ən'lu:rε	lure	onlure
[r]		'lu:ru	ən'lu:ru	luru	onluru
	[+fr]	'ri:ra	ən'ri:ra	rira	enrica
		're:la	ən're:la	rela	enrela
	[-fr]	'ru:lε	ən'ru:lε	rule	onrule
[j]		'ru:lɔ	ən'ru:lɔ	rulo	onrulo
	[+fr]	'ji:lε	ən'ji:lε	jile	enjile
		'je:nɔ	ən'je:nɔ	jeno	enjeno
	[-fr]	'ja:lu	ən'ja:lu	jalu	onjalu
		'jo:rε	ən'jo:rε	jore	onjore

Table 6: Nouns with initial voiceless stops and fricatives in the Slovenian training phase.

		Voiceless				
Place	#	Harm.	Sg.	Pl.	Orthography	
Labial	[-cont]	[+fr]	'pi:mɔ	ɛm'pi:mɔ	pimo	empimo
		[−fr]	'pe:ri	ɛm'pe:ri	peri	emperi
		[−fr]	'pu:ru	ɔm'pu:ru	puru	ompuru
	[+cont]	[+fr]	'pa:nɛ	ɔm'pa:nɛ	pane	ompane
		[+fr]	'fi:ma	ɛm'fi:ma	fima	emfima
		[−fr]	'fe:rɛ	ɛm'fe:rɛ	fera	emfera
Coronal	[-cont]	[−fr]	'fa:lɔ	ɔm'fa:lɔ	falo	omfalo
		[−fr]	'fu:na	ɔm'fu:na	funa	omfuna
		[+fr]	'ti:ɾɔ	ɛn'ti:ɾɔ	tiro	entiro
	[+cont]	[−fr]	'te:nɔ	ɛn'te:nɔ	teno	enteno
		[−fr]	'ta:nu	ɔn'ta:nu	tanu	ontanu
		[+fr]	'tu:lɔ	ɔn'tu:lɔ	tulo	ontulo
	[+cont]	[+fr]	'se:nɛ	ɛn'se:nɛ	sene	ensene
		[+fr]	'si:nɔ	ɛn'si:nɔ	sino	ensino
		[−fr]	'sa:ru	ɔn'sa:ru	saru	onsaru
	[-fr]	[−fr]	'su:li	ɔn'su:li	suli	onsuli

there is an insignificant but positive correlation between the performance on the explicit and implicit tasks; in the Slovenian experiment, the correlation is almost flat (slightly negative) for the labial series and negative in the coronal series. The random effect structure of the initial full model included a crossed random intercepts for SUBJECT and ITEM and by-subject random slopes for PLACE and by-item random slopes for PLACE and HARMONY (scaled and centered; with the interaction). First, random slopes were removed one by one from the model. The final random effects structure was chosen based on AIC and included crossed random intercepts for SUBJECT and ITEM and by-subject random slope for PLACE. The fixed effects structure was chosen based on this random effects structure by step-wise removal of higher order interaction terms based on AIC. All predictors are non-significant, but we keep the PLACE of articulation in the model.

To control for the effect of response button placement, a random intercept for placement (top or bottom) was added to the final model. Estimates of coefficients remain approximately the same (with equal significance levels), but because adding a random intercept for response side does not improve fit significantly (according to AIC), the final model excludes it. This suggests that placement of the stimuli in the test phase does not influence the results substantially.

The same technique is used in the Slovenian experiment. The best-fitting model (chosen based on the technique described in the English experiment) includes PLACE and HARMONY (scaled and centered) in the fixed effects structure (without the interaction) and random intercepts for subject and item with by-subject random slope for PLACE and by-item random slope for HARMONY. Significance of the predictors of interest is the same in the full model including all interactions and random slopes and the final chosen model.

Mild underdispersion was detected in both the English and Slovenian PND models: 0.7 in both models (estimated with *overdisp\_fun()* in Bolker 2019). Since underdispersion leads to conservative estimates, we leave the models as such. One potential correction using quasi-likelihood (by multiplying regression estimates with the square root of the dispersion ratio) yields slightly lower standard error values (as proposed in Bolker 2019):  $\beta(\text{Intercept}=\text{labial}) = 0.43, z = 2.5, p = 0.01$  for the English experiment and  $\beta(\text{Intercept}=\text{labial}) = 0.62, z = 3.3, p = 0.0009$  for the Slovenian experiment.

Table 7: Nouns with initial voiced stops and fricatives in the Slovenian training phase. Orthographic stimuli were only presented to the phonetically trained speaker who recorded the stimuli and never to the subjects.

			Voiced		
Place	#_	Harm.	Sg.	Pl.	Orthography
Labial	[-cont]	[+fr]	'bi:n <u>u</u>	ɛm'pi:n <u>u</u>	binu empinu
			'be:ru	ɛm'pe:ru	beru emperu
			'bi:no	ɛm'pi:no	bino empino
			'bi:ma	ɛm'pi:ma	bima empima
	[-fr]	[-fr]	'bu:lu	ɔm'pu:lu	bulu ompulu
			'ba:n <u>o</u>	ɔm'pa:n <u>o</u>	bano ompano
			'bo:mi	ɔm'po:mi	bomi ompomi
			'bu:nu	ɔm'pu:nu	bunu ompunu
	[+cont]	[+fr]	'vi:ma	ɛm'pi:ma	vima empima
			'vi:lu	ɛm'pi:lu	vilu empilu
			'vi:r <u>o</u>	ɛm'pi:r <u>o</u>	viro empiro
			've:lu	ɛm'pe:lu	velu empelu
Coronal	[-cont]	[+fr]	'vu:n <u>o</u>	ɔn'pu:m <u>o</u>	vuno ompuno
			'va:nu	ɔm'pa:nu	vanu opanu
			'vo:ni	ɔm'po:ni	toni omponi
			'vu:m <u>e</u>	ɔm'pu:m <u>e</u>	vume ompume
	[-fr]	[-fr]	'de:r <u>o</u>	ɛn'te:r <u>o</u>	dero entero
			'di:ni	ɛn'ti:ni	dini entini
			'de:ru	ɛn'te:ru	deru enteru
			'de:ma	ɛn'te:ma	dema entema
	[+cont]	[+fr]	'du:l <u>o</u>	ɔn'tu:l <u>o</u>	dulo ontulo
			'do:ru	ɔn'to:ru	doru ontoru
			'da:lu	ɔn'ta:lu	dalu ontalu
			'du:na	ɔn'tu:na	duna ontuna
	[+cont]	[+fr]	'zi:la	ɛn'ti:la	zila entila
			'ze:ra	ɛn'te:ra	zera entera
			'ze:m <u>o</u>	ɛn'te:m <u>o</u>	zemo entemo
			'ze:ri	ɛn'te:ri	zeri enteri
	[-fr]	[-fr]	'zu:r <u>o</u>	ɔn'tu:r <u>o</u>	zuro onturo
			'za:nu	ɔn'ta:nu	zaru ontanu
			'zo:lu	ɔn'to:lu	zolu ontolu
			'zunu	ɔn'tunu	zunu ontunu

Table 8: Nouns in the Slovenian test phase. Orthographic stimuli were only presented to the phonetically trained speaker who recorded the stimuli and never to the subjects.

Place	Harm.	Pl.	Test			Orthography	
			Sg. 1	Sg. 2			
Labial	[+fr]	ɛm'pe:mi	'be:mi	've:mi	empemi	bemi	vemi
		ɛm'pi:ju	'bi:ju	'vi:ju	empiju	biju	viju
		ɛm'pi:ja	'bi:ja	'vi:ja	empija	bija	vija
	[-fr]	ɔm'pa:mu	'ba:mu	'va:mu	ompamu	bamu	vamu
		ɔm'pa:rɔ	'ba:rɔ	'va:rɔ	omparo	baro	varo
		ɔm'pu:ji	'bu:ji	'vu:ji	ompuji	buji	vuji
Coronal	[+fr]	ɛn'ti:ju	'di:ju	'zi:ju	entiju	diju	ziju
		ɛn'te:ju	'de:ju	'ze:ju	enteju	deju	zeju
		ɛn'ti:je	'di:je	'zi:je	entije	dije	zije
	[-fr]	ɔn'to:ju	'do:ju	'zo:ju	ontoju	doju	zaju
		ɔn'tu:jɔ	'du:jɔ	'zu:jɔ	ontujo	dujo	zujo
		ɔn'tu:jɛ	'du:jɛ	'zu:jɛ	ontuje	duje	zuje

## 2 Probit regression for estimation of $d'$

The data for estimating  $d'$  were taken from consonant confusion matrices based on perceptual experiments in Singh & Black 1966. We extract counts for [p] and [b] responses (when presented with [p]) and vice versa to get hits, rejections, misses, and false alarms ([p] and [b] responses when presented with [b]) (the stop condition). We also extract [p] and [v] responses (when presented with [p]) and vice versa to get hits, rejections, misses, and false alarms ([p] and [v] responses when presented with [v]) in the fricative condition. This resembles the experimental design (Slovenian and English), where subjects are faced with NT and have to choose between D and Z. The same information is extracted for the coronal series too. We only extract information for experiments in Singh & Black 1966 in which the stimuli are from native speakers (and listeners too are native speakers) of Hindi, English, Arabic, and Japanese. For counts, see Table 9.

Due to complete separation in the data, the model was fit to a bias reduced probit linear regression models using the *brglm()* function from Kosmidis & Firth 2020. The model has counts as described above as the dependent variables and TARGET (voiceless vs. voiced), CONDITION (stop vs. fricative, treatment-coded with stop as the reference level), and PLACE as predictors. In order to estimate  $d'$  values, the two levels in the TARGET predictor (voiced and voiceless) are coded as -0.5 and 0.5 (or vice versa). Predictor PLACE has two levels (labial, coronal); the labial is the reference level (treatment-coded). The model also features LANGUAGE as predictor (sum-coded with four levels: Hindi, English, Japanese, Arabic). The model includes all interactions. Table 10 gives estimates of the model.

language	place	condition	target	voiceless	voiced
Hindi	labial	stop	voiceless	59	11
Hindi	labial	stop	voiced	4	66
Hindi	labial	fricative	voiceless	59	0
Hindi	labial	fricative	voiced	3	63
Hindi	coronal	stop	voiceless	55	13
Hindi	coronal	stop	voiced	0	44
Hindi	coronal	fricative	voiceless	55	0
Hindi	coronal	fricative	voiced	0	53
English	labial	stop	voiceless	72	0
English	labial	stop	voiced	0	72
English	labial	fricative	voiceless	72	0
English	labial	fricative	voiced	0	64
English	coronal	stop	voiceless	62	1
English	coronal	stop	voiced	0	70
English	coronal	fricative	voiceless	62	0
English	coronal	fricative	voiced	0	67
Arabic	labial	stop	voiceless	56	14
Arabic	labial	stop	voiced	8	59
Arabic	labial	fricative	voiceless	56	0
Arabic	labial	fricative	voiced	0	61
Arabic	coronal	stop	voiceless	63	0
Arabic	coronal	stop	voiced	0	61
Arabic	coronal	fricative	voiceless	63	1
Arabic	coronal	fricative	voiced	0	53
Japanese	labial	stop	voiceless	61	11
Japanese	labial	stop	voiced	1	66
Japanese	labial	fricative	voiceless	61	0
Japanese	labial	fricative	voiced	1	63
Japanese	coronal	stop	voiceless	61	6
Japanese	coronal	stop	voiced	1	60
Japanese	coronal	fricative	voiceless	61	0
Japanese	coronal	fricative	voiced	0	51

Table 9: Counts of outcomes given target and response (voiceless vs. voiced) in the two conditions (stop vs. fricative) for two places of articulation from Hindi, English, Arabic, and Japanese perceptual experiments (consonant confusion matrices) with native stimuli and listeners in Singh & Black 1966.

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.2367	0.1167	-2.03	0.0425
target1	3.1402	0.2334	13.45	0.0000
conditionfricative	0.3929	0.2099	1.87	0.0612
placecoronal	-0.0921	0.1930	-0.48	0.6334
language1	0.0738	0.1486	0.50	0.6192
language2	0.2367	0.2902	0.82	0.4146
language3	-0.2748	0.1806	-1.52	0.1281
target1:conditionfricative	1.4518	0.4198	3.46	0.0005
target1:placecoronal	0.8371	0.3861	2.17	0.0301
conditionfricative:placecoronal	-0.1063	0.3199	-0.33	0.7398
target1:language1	-1.1553	0.2972	-3.89	0.0001
target1:language2	1.8706	0.5803	3.22	0.0013
target1:language3	-0.1004	0.3612	-0.28	0.7810
conditionfricative:language1	-0.2449	0.3562	-0.69	0.4917
conditionfricative:language2	-0.3726	0.4316	-0.86	0.3879
conditionfricative:language3	0.3364	0.3392	0.99	0.3213
placecoronal:language1	0.2605	0.3450	0.76	0.4501
placecoronal:language2	-0.1527	0.3983	-0.38	0.7014
placecoronal:language3	0.2638	0.2781	0.95	0.3428
target1:conditionfricative:placecoronal	-0.6813	0.6398	-1.06	0.2870
target1:conditionfricative:language1	1.4299	0.7125	2.01	0.0448
target1:conditionfricative:language2	-1.4923	0.8632	-1.73	0.0838
target1:conditionfricative:language3	-0.0307	0.6784	-0.05	0.9639
target1:placecoronal:language1	2.0857	0.6899	3.02	0.0025
target1:placecoronal:language2	-1.3459	0.7967	-1.69	0.0911
target1:placecoronal:language3	-0.5706	0.5562	-1.03	0.3049
conditionfricative:placecoronal:language1	-0.2405	0.5584	-0.43	0.6668
conditionfricative:placecoronal:language2	0.3174	0.6066	0.52	0.6008
conditionfricative:placecoronal:language3	-0.2518	0.5194	-0.48	0.6278
target1:conditionfricative:placecoronal:language1	-2.6964	1.1169	-2.41	0.0158
target1:conditionfricative:placecoronal:language2	1.1545	1.2132	0.95	0.3413
target1:conditionfricative:placecoronal:language3	0.7876	1.0388	0.76	0.4483

Table 10: Estimates of the probit logistic regression model based on data from Table 9 from Singh & Black 1966.

## References

- ALBRIGHT, ADAM, and YOUNGAH DO. 2019. Three biases in learning phonological alternation. Ms., MIT and University of Hong Kong. Accessed on January 12, 2019. Online: <https://yd79.files.wordpress.com/2014/06/threebiases.pdf>.
- BATES, DOUGLAS; MARTIN MÄCHLER; BEN BOLKER; and STEVE WALKER. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1–48.
- BECKER, MICHAEL, and JONATHAN LEVINE. 2013. Experigen – an online experiment platform. Online: <http://becker.phonologist.org/experigen>.
- BECKERS, GABRIEL J. L. 2002. rms equalize. praat script. Online: [http://www.gbeckers.nl/pages/praat\\_scripts/rms\\_equalize.praat\\_script](http://www.gbeckers.nl/pages/praat_scripts/rms_equalize.praat_script).
- BOERSMA, PAUL, and DAVID WEENINK. 2015. Praat: doing phonetics by computer [computer program]. version 5.4.06. Retrieved 21 February 2015 from <http://www.praat.org/>.
- BOLKER, BEN. 2019. Glmm faq. Retrieved on June 22, 2022. Online: [bbolker.github.io/mixedmodels-misc/glmmFAQ.html](https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html).
- KOSMIDIS, IOANNIS, and DAVID FIRTH. 2020. Jeffreys-prior penalty, finiteness and shrinkage in binomial-response generalized linear models. *Biometrika*. Online: <https://doi.org/10.1093/biomet/asaa052>.
- R CORE TEAM. 2018. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Online: <https://www.R-project.org/>.
- SINGH, SADANAND, and JOHN W. BLACK. 1966. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *The Journal of the Acoustical Society of America* 39.372–387. Online: <https://doi.org/10.1121/1.1909899>.