Naturalness bias in palatalization: an experimental study

Joanna Zaleska and Matías Guzmán Naranjo
University of Leipzig

Abstract

In the present study, we report on an artificial language learning experiment aiming to test the idea that it is easier to learn palatalization before a front vowel than it is to learn depalatalization in the same context. The motivation for the study comes from recent work by Czaplicki (2013), who provides a detailed analysis of palatalization-related effects in Polish, showing that they have no phonological basis. The conclusion he reaches is that 'phonological naturalness does not play a role in linguistic computation' (Czaplicki 2013:32). It is nevertheless the case that palatalization is cross-linguistically much more common than depalatalization. If naturalness plays no role in computation, the typological asymmetry must arise from elsewhere, for example, from biases in learning difficulty. Our results provide provisory evidence that there is a small but statistically significant advantage for learning palatalization over depalatalization for adult Hungarian speakers.

1. Introduction [1]

In his 2013 article on palatalization effects in Polish, Czaplicki argues that these are phonologically arbitrary and cannot be predicted from the phonological make-up of the segments involved. While a subset of those effects can be viewed as 'natural', in the sense that they can be expressed in terms of feature spreading or sharing, Czaplicki discusses a whole range of processes that cannot be accounted for in these terms. He shows that the vowels /ɛ/ and /a/, as well as the consonant /n/, can (i) trigger diverse types of palatalization, (ii) fail to trigger palatalization and (iii) cause depalatalization.

To give a specific example, the vowel /ɛ/ in the locative singular suffix *-e* results in alveolars changing into prepalatals, (1a–c), and in velars turning into (post)alveolars, (1d–e). At the same time, [ɛ] in the diminutive suffix *-ek* fails to palatalize alveolar consonants (2a), turns velars into palato-alveolars (2b), and causes the depalatalization of some (2c) but not all (2d) pre-palatals.

(1) a. *psot+a* [psɔ**t**+a] 'prank' vs. *psoci-e* [psɔ**ʨ**+ɛ]
  b. *wod+a* [vɔ**d**+a] 'water' vs. *wodzi-e* [vɔ**ʥ**+ɛ]
  c. *stron+a* [strɔ**n**+a] 'side' vs. *stroni-e* [strɔ**ɲ**+ɛ]
  d. *ręk+a* [rɛŋ**k**+a] 'hand' vs. *ręc+e* [rɛn**ʦ**+ɛ]
  e. *much+a* [mu**x**+a] 'fly' vs. *musz-e* [mu**ʃ**+ɛ]

(2) a. *świat* [ɕfja**t**] 'world' vs. *świat-ek* [ɕfja**t**+ɛk]
  b. *krok* [krɔ**k**] 'step' vs. *krocz+ek* [krɔ**ʧ**+ɛk]
  c. *liść* [lʲi**ɕʨ**] 'leaf' vs. *list-ek* [lʲi**st**+ɛk]
  d. *miś* [mʲi**ɕ**] 'bear' vs. *misi-ek* [mʲi**ɕ**+ɛk]

The change illustrated in (1a–c) can be said to be phonetically grounded, as it consists in a non-palatalized ([+back]) segment becoming palatalized ([−back]) in the context of a front ([−back]) vowel. This is less true of the change in (1d–e) and (2b): the resulting segments are normally classified as ([+back]) because, unlike prepalatals, they pattern with other 'hard' consonants in Polish. Furthermore, if the palatalization in (1a) is a natural process, then its lack in (2a) is unexpected and the reverse change in (2c) is simply

[1] We would like to thank Ildikó Emese Szabó for her indispensable research assistance, to Sven Grawunder, Ludger Paschen and Andrew Murphy for help at various stages of the experiment and to the participants of the Colloquium *Neuere Arbeiten zur Grammatiktheorie* ('Theory of grammar') at the University of Leipzig as well as to Bartłomiej Czaplicki and Brandon Prickett, for their helpful comments and criticism. We would also like to thank the anonymous Nordlyd reviewers whose valuable comments and suggestions substantially improved the quality of this paper. Any errors or omissions are our own. This work was completed as part of the DFG-funded graduate school *Interaktion Grammatischer Bausteine* 'Interaction of Grammatical Building Blocks' (IGRA).

unnatural. On the basis of these and similar examples, Czaplicki argues that the type and applicability of palatalization in Polish cannot be effectively explained in terms of spreading or assimilation and that it must be governed by some mechanism that does not rely on phonological universals. He concludes that naturalness plays no role in the computation of phonological patterns.

While Czaplicki's analysis is both coherent and appealing, it is nevertheless true that examples of palatalization abound (see, e.g. Bateman 2011 for a recent review), whereas depalatalization is cross-linguistically much less common. Czaplicki addresses this point by observing that typological asymmetries are the result of phonetically conditioned sound change (as suggested by Hyman 1976, Ohala 1993 and Blevins 2004, among others). He rejects the competing view, which attributes the cross-linguistic scarcity of some phonological patterns and the pervasiveness of other superficially similar ones to *inductive biases* (Moreton and Pater 2012a;b), that is, to a human learner's cognitive predispositions that reduce the learnability of the former and enhance the learnability of the latter. In support, he cites Seidl and Buckley's (2005) experiment on infant subjects, as well as Pycha et al.'s (2003) study with adult learners, both of which find that phonologically arbitrary and phonologically motivated patterns are equally learnable.

Nonetheless, scholars do not unequivocally dismiss the possibility of inductive biases as the source of typological asymmetries. For one thing, experimental literature offers abundant evidence for the existence of *structural biases* (e.g. Pycha et al. 2003, Peperkamp et al. 2006, McMullin 2013), related to the preference for simple patterns over complex ones (Pater and Moreton 2012). While research into *naturalness* or *substantive biases* (Wilson 2006), based on the acoustic, auditory and perceptual properties of speech sounds, has produced less consistent results, some studies (e.g. Schane et al. 1974, Carpenter 2010, Wilson 2006) have indeed found that phonetically motivated patterns are easier to learn than phonetically unmotivated ones. One explanation for the inconsistency of experimental results could be related to Schane et al.'s (1974) observation that learning biases are only detectable at early stages of training and can be overridden given enough exposure to the unnatural pattern. Additionally, as noted by Finley (2012), it could be the case that some phonological patterns are more susceptible to naturalness biases than others.

The aim of the experiment described below, then, is to test whether palatalization and depalatalization are subject to naturalness constraints. To do this, we compare the acquisition of two artificial languages that are identical except for the palatalization pattern, with one language exhibiting palatalization, and the other one depalatalization. The results indicate that although both languages are learnable, speakers find the language with palatalization easier than the one with depalatalization.

2. Method

The experiment combines and adapts the methodology used by Carpenter (2010), who investigates the naturalness bias in learning stress patters as well as by Ettlinger (2008), who studies the learning of phonologically opaque forms. In contrast to recent experimental studies on palatalization (e.g. Wilson 2006 and Kapatsinski 2010), we do not focus on velar palatalization. Rather, we look at the more cross-linguistically common alternations (cf. Chen 1973, Bateman 2007) involving alveolar stops /t, d/ and palato-alveolar affricates /ʧ, ʤ/. The experiment was carried out with native speakers of Hungarian, as the language contains all the relevant consonants and does not contain a rule of palatalization or depalatalization before front vowels.

We constructed two miniature languages that were identical except for the palatalization pattern. The experiment began with a familiarization phase, in which the participants were exposed to words from one of the two artificial languages, paired with images of animals, food and objects related to nature. This was followed by a test phase, split into two parts: a two-alternative forced choice task, in which the participants had to select the correct form between two forms provided, and a lexical decision task, in which they had to judge whether an item belonged to the language they had learnt or not. By comparing the level of attainment between the two groups of participants, we could see whether learners exhibit a bias towards learning palatalization or depalatalization.

2.1. Conditions

We constructed two artificial languages for the purpose of the study, inspired by the Polish pattern, but vastly simplified. The two mini languages contain words built up from identical stems and suffixes. The only difference is the effect that one of the suffixes has on the final consonant of some stems. In the 'palatalizing' language, stem-final alveolar stops become palato-alveolar affricates before the suffix that begins with the front mid vowel, as shown in (3a). In the 'depalatalizing' language, the same context changes stem-final palato-alveolars to plain alveolars, as in (3b).

(3) a. *Natural pattern: palatalization*
t, d → ʧ, ʤ /__ ɛ
b. *Unnatural pattern: depalatalization*
ʧ, ʤ→ t, d /__ ɛ

We follow Czaplicki (2013) in assuming that palatalization before front vowels is a natural process, whereas depalatalization is not. This assumption is also supported by the typological distribution of the two patterns. While most articulatory and perception studies investigating the naturalness of palatalization have focused on the change from the *velar* to palato-alveolar place of articulation (see, eg. Keating and Lahiri 1993, Guion 1996; 1998), (3a) could be said to favour phonological naturalness in that the resulting segments involve the fronted tongue body position (Flemming 2003:338), becoming more similar in this respect to a front vowel than plain alveolars. In this light, the reverse change in (3b) may be deemed unnatural, since it results in greater phonetic distance between the consonant and the conditioning vowel.

2.2. Stimuli

The consonant inventory of both languages consists of the sounds /p t k b d g s z ʧ ʤ m n l/, while the vowels are drawn from the set /a, ɛ, i, ɔ, u/. The "words" in the languages are built up of consonant-final stems followed by an optional diminutive suffix and an obligatory inflectional suffix. The inflectional suffix marks number, with /-a/, indicating singular and /-u/, indicating plural. The diminutive suffix has the form /-ɛn/. The suffixes can combine freely, which means that for every stem, four different forms are possible: a singular one, a plural one, a singular diminutive one and a plural diminutive one.

The stems are built up from one or two CV(C) syllables. For the first two stages of the experiment (familiarization and the forced choice task), we created 30 stems. 10 of those end with alveolar stops, /t/ and /d/, which alternate with /ʧ,ʤ/ in the 'palatalizing' language and do not exhibit any alternations in the 'depalatalizing' language, thus acting as fillers. A sample paradigm for a *-t -d* stem in both conditions is shown in (4a). 10 further stems end with palato-alveolar affricates, /ʧ/ and /ʤ/, which alternate with /t,d/ in the 'depalatalizing' language and do not alternate at all in the the 'palatalizing' language, see (4b). Finally, 10 stems end in a nonalternating consonant, drawn from the set /p b m n l/. An example is given in (4c). These stems act as fillers in both groups.

(4) Sample stimuli

a. *-t*, *-d* stems

	Palatalizing language	Depalatalizing language
Singular	lɛnsa**d**-a	lɛnsad-a
Plural	lɛnsa**d**-u	lɛnsad-u
Diminutive	lɛnsa**ʤ**-ɛn-a	lɛnsad-ɛn-a
Plural diminutive	lɛnsa**ʤ**-ɛn-u	lɛnsad-ɛn-u

b. *-ʧ*, *-ʤ* stems

	Palatalizing language	Depalatalizing language
Singular	nɔʧ-a	nɔ**ʧ**-a
Plural	nɔʧ-u	nɔ**ʧ**-u
Diminutive	nɔʧ-ɛn-a	nɔ**t**-ɛn-a
Plural diminutive	nɔʧ-ɛn-u	nɔ**t**-ɛn-u

c. Neutral stems

	Palatalizing language	Depalatalizing language
Singular	pɛʤɔm-a	pɛʤɔm-a
Plural	pɛʤɔm-u	pɛʤɔm-u
Diminutive	pɛʤɔm-ɛn-a	pɛʤɔm-ɛn-a
Plural diminutive	pɛʤɔm-ɛn-u	pɛʤɔm-ɛn-u

Some additional constraints were imposed on the forms in both languages. First, we made sure they did not resemble any actual Hungarian words. Second, the coda postion could only be filled with a sonorant (/l, m/ or /n/). Finally, stem-internal /t d ʧ ʤ/ could only occur before the vowels /ɔ, u/ and /a/, and not not before the front vowels, /ɛ/ and /i/. Thus, for example, /pɛʤɔm-/ is a licit stem, whereas */pɛʤɛm-/ is not. This was done to ensure that the alternations before the diminutive suffix can be the only source of information about the distribution of the alveolar and palato-alveolar sounds before front vowels.[2,3]

Full lists of items used in the familiarization phase and in the first part of the test phase can be found in Appendix A and Appendix B, respectively. For the purposes of the second part of the test phase, we created 48 more forms, which are listed in Appendix C.

Each stem was a assigned a meaning, illustrated with one of the images from the BOSS database, (Brodeur et al. 2010). The visual stimuli for the diminutive forms were created by resizing the respective images, and those for the plural forms were created by copying the object in the image 3 to 9 times. Sample visual stimuli for the item meaning 'starfish' are shown in (5).

[2] The reason for including /i/ in this set is related to the implicational universal for palatalization triggers, discussed by Bateman (2007:75) and Chen (1973:177): if lower mid front vowels trigger palatalization, so do high front vowels. Including /ɛ/ as a palatalization trigger to the exclusion of /i/ would make the pattern cross-linguistically unattested.

[3] An anonymous reviewer notes that given Hungarian front-back vowel harmony, it is not clear how Hungarian speakers process suffixes that do not harmonize with the stem. Our motivation for including stems that do not harmonize was twofold. First, we intend to run the experiment with speakers of languages that do not have vowel harmony, like Colombian Spanish. Making the artificial languages less 'foreign' to Hungarian speakers would then make them more 'foreign' to Spanish speakers. Although the lack of harmony does presumably make the task more difficult to Hungarian speakers, the results of the experiment show that they nevertheless managed to learn a substantial portion of the language. Secondly, and perhaps more importantly, if the experimental items included either front vowels or back vowels only, the speakers might make generalizations about the distribution of (palato-)alveolars on the basis on the quality of the preceding, rather than following vowel.

(5) Sample visual stimuli for the item meaning 'starfish'

Singular	Plural	Diminutive singular	Diminutive plural
			

All the items were digitally recorded with stress on the initial syllable in a sound-attenuated booth by a phonetically-trained female native speaker of Polish, using using sounddevice T788 and a head-set mic Audix HT-5, with 48kHz/16bit. The items in Appendix B, used in the first test phase, were recorded by a phonetically-trained male native speaker of German, using the same equipment and settings.[4] All items were recorded as whole words, rather than being created via concatenation. They were cut into separate files using Audacity (Audacity Team 2013). Apart from minor corrections, no normalization or adjustments were made.

2.3. Procedure

The experiment was coded using *Experigen* (Becker and Levine 2014). The files for the two conditions were hosted in different locations on the server and the conditions were accessed via separate URLs. Participants were recruited trhough a variety of methods including word of mouth and requests for participation on social media sites. Some of the initial participants shared the link to the experiment with others, which resulted in a snowball sample and uneven distribution of participants across the two conditions.

After accessing the experiment's URL, a welcome screen with instructions appeared. These were given in the form of a scenario in which the participants are asked to imagine that they are survivors of a shipwreck who end up living with a native tribe on a remote island. To maintain good relations, they need to try to learn the language spoken by the tribe, called Zaguna. The participants were told they would not have to memorise all the words but they should try to figure out the rules that govern the language. No explicit reference to any sound changes was made. An English translation of the instructions is reported in Appendix D.

In the familiarization stage, the participants were shown a subset of the stimuli (listed in Appendix A), accompanied by the corresponding photographic images. Each sound/image combination was presented for 1.8 seconds. The same combination was randomly repeated 4 times throughout the familiarization stage. For 4 stems, including two experimental ones, the entire paradigm (singular, plural, diminutive singular, diminutive plural) was shown. For other stems, parts of the paradigm were withheld from exposure, making it possible to test participants on known stems with new inflectional forms. The stage lasted about 8 minutes and included 240 trials.

The first test stage was a two-alternative forced choice task, in which the participants were presented with familiar stems (but some unfamiliar forms), as well as an image, and were asked to select the correct form of the two forms provided. The stage consisted of 42 trials, 12 of which involved the alternating stems (*-t -d* stems in the palatalizing condition and *-ʧ -ʤ* stems in the depalatalizing condition). In these

[4]Two speakers were used in keeping with the scenario laid out in the instructions to the experiment, where the participants are told that they meet a different learner of the language (see subsection 2.3 and Appendix D). An anonymous reviewer expresses a concern that since the voiced palato-alveolar affricate occurs in German only marginally, a German speaker might nativize the words containing /ʤ/ by devoicing it. While the choice of speakers was dictated by practical limitations, both of them were phonetically trained and were able to produce the items, provided in IPA transcription, accurately. Additional auditory inspection of the items containing the /ʤ/ affricate recorded by the German speaker did not reveal any devoicing.

trials, two forms with the same suffix were given, one with an alternation and another without. The image always corresponded to the meaning of the stems. A sample question is given in (6a). In 6 of the items, the form with the alternation was heard before the form without it, and in the remaining 6 items the order was reversed. The remaining trials included two different forms created on the basis of the same stem and an image corresponding to one of them. These items tested whether the participants had learned the meaning of the derivational and inflectional suffixes. An example is given in (6b). The full list of test 1 items can be found in Appendix B.

(6) Sample test 1 questions (palatalizing condition)

a. Experimental item

Visual:	four small avocados
Audio:	/mɛldɛnu/ ... /mɛlʤɛnu/
Correct answer:	2

b. Filler item

Visual:	nine small clownfish
Audio:	/mundulɛnu/ ... /mundulɛna/
Correct answer:	1

The second test stage was a lexical decision task, in which participants were presented with forms built from unfamiliar stems and were asked to assess whether they belonged to the language they had learnt or not. This stage consisted of 48 trials. 12 of those included a word in which /t,d/ appeared before the diminutive suffix. An example is provided in (7a). These items were crucial in the palatalizing condition: if the speakers accepted them, it would mean that they have not learned the palatalization rule. 12 further items, relevant for the depalatalizing condition, were words in which /ʧ, ʤ/ appeared before the diminutive suffix. Finally, 24 items were formed on the basis of 'neutral' stems, ending in /p,b,m,n,l/. Half of those were paired with an appropriate image and another half were paired with an incorrect image. An example of a filler experimental item is shown in (7b) and a full list of items used in the second test stage is given in Appendix C.

(7) Sample test 2 questions (palatalizing condition)

a. Experimental item

Visual:	one small butterfly
Audio:	/pupɔldɛna/
Correct answer:	NO

b. Filler item

Visual:	four (big) feathers
Audio:	/ʤɔmlu/
Correct answer:	YES

The last part of the experiment consisted of demographic questions regarding age, gender, level of education, place of origin and known foreign languages. Information about speech and hearing deficits was not collected.

2.4. Native language of the participants

To minimize the possibility of the participants' native language skewing the results, it was necessary to carry out the experiment with speakers of a language that meets two conditions. First, the inventory of the native languages of the participants should contain the consonants /t, d, ʧ, ʤ/, so that the participants do not have to learn new sound categories in addition to learning the pattern. Second, the language could not contain rules of palatalization or depalatalization of the type shown in (3), as this might give the learners in one of the groups an unfair advantage over the other one.

A language that seems to meet the criteria described above is Hungarian. As shown in Table 1, Hungarian has a rich consonant inventory that contains both alveolar and palatal consonants. Although the status of /ʤ/ as an underlying segment in Hungarian is disputed, we follow Siptár and Törkenczy (2007:89) in assuming that, unlike its dental counterpart [ʣ], [ʤ] is not derived from an underlying sequence of a stop + fricative.[5]

	Stops	Fricatives	Affricates	Nasals	Liquids
Labial	p/b	f/v		m	
Dental /Alveolar	**t/d**	s/z	ʦ/	n	l/r
Palatal	c/ɟ	ʃ/ʒ	**ʧ/ʤ**	ɲ	j
Velar	k/g	x			

Table 1: Hungarian consonant inventory (Siptár and Törkenczy 2007:75–76, Ladefoged 2001:147–148)

Hungarian has no process of (de)palatalization before front vowels. It must, however, be noted, that it does have a process of coalescence, whereby dental /t,d,n,l/ fuse with the following /j/, becoming /c,ɟ,j,ɲ/, as well as an optional process of assimilation, which turns /t,d,n/ into /c,ɟ,ɲ/ before /c,ɟ,j,ɲ/ (see Siptár and Törkenczy 2007:177). Additionally, the stem final /t/ in verbs becomes /ʃ/ or /ʧ/ before the imperative affix /j/ (Siptár and Törkenczy 2007:183). As pointed out by Bartłomiej Czaplicki (personal communication), this might result in a general bias towards palatalization before front segments. However, as observed by Bateman (2007:75) and Chen (1973:177), while palatalization before mid front vowels entails palatalization before high vowels and glides cross-linguistically, the reverse does not hold. The same asymmetry was found in a study by Wilson (2006), where adults exposed to velar palatalization before [i] did not generalize to [e] but those exposed to palatalization before [e] did generalize to [i]. Taken together, these findings suggest that speakers of Hungarian do not necessarily generalize the palatalization patterns of their own language (before the front glide) to the artificial language (before the mid front vowel). However, we admit that the issue is a potential confound. Nevertheless, given the paucity of easily accessible languages that fully conform to the criteria described above, we opted to use Hungarian speakers irrespective of this complication. In section 4, we suggest a further study relevant to this concern.

2.5. *Participants*

All experiment participants were native speakers of Hungarian, aged between 17 and 49 years old (age average 28). There were 42 participants in the 'palatalizing' group and 27 in the 'depalatalizing' group. The majority of participants (41 in the 'palatalizing' group and 24 in the 'depalatalizing' group) reported knowing English to some level. Many reported familiarity with additional languages (German: 26 participants altogether, Italian: 7 participants, French: 6 participants, Spanish: 5 participants, Romanian: 4 participants, Latin: 3 participants, Russian: 3 participants, Swedish: 2 participants, Finnish: 2 participants, Polish: 2 participants, Arabic: 1 participant, Estonian 1 participant). Subjects were not individually compensated for their participation but an Amazon voucher worth 10 Euro was raffled among the participants of each condition. 5 participants from the 'depalatalizing' group had to be dropped because they did not complete the experiment.

[5] Even if the participants do treat /ʤ/ differently than /ʧ/, the results of our experiment did not reflect that, revealing no statistically significant difference between items with voiced and voiceless alternating sounds.

3. Results and discussion

If learners do have a cognitive bias towards natural patterns, then the language in which alveolars palatalize before a front vowel should turn out easier to learn than the language in which palato-alveolars depalatalize in the same environment. If there is no naturalness bias, both languages should be learnable with comparable ease.[6]

3.1. Data

Let us first inspect the overall picture for all participants in both conditions by looking at the the results for control and filler items plotted in Figure 1.[7] The first thing that stands out here is the substantial variation across participants: one speaker managed to produce a correct answer in every trial, while others answered less than one third of the questions correctly. The second thing to note is that there seems to be a slight difference between the patalalizing language and the depalatalizing one.

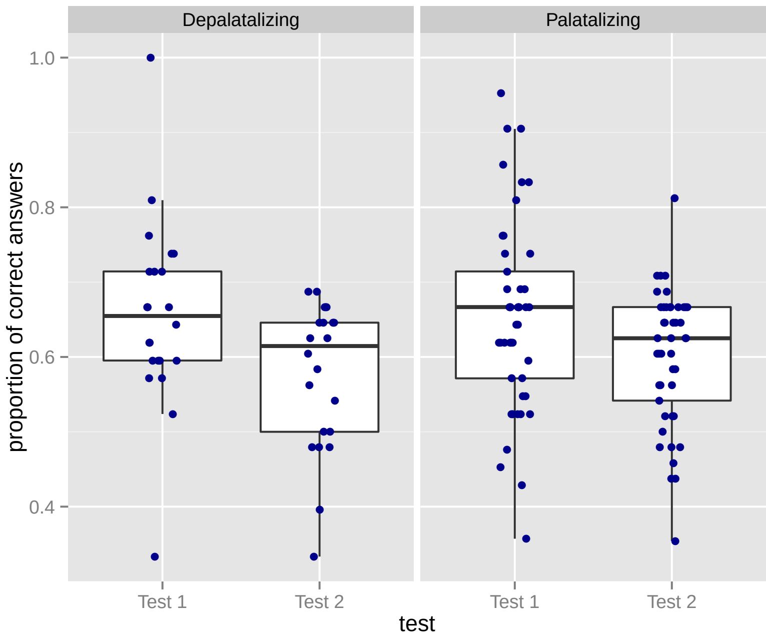

Figure 1: Results for all participants for all items in all conditions (including fillers)

Next, let us take a look at the performance of the participants on those questions that did not test palatalization/depalatalization, shown in Figure 2. What immediately strikes the eye in this plot is that the scores for most participants are now much better than when they include the results for the cricital items. This suggests that the majority of participants did manage to learn a good portion of the language during the training phase, and that a big part of the difficulties was related to the palatalization/depalatalization items.

[6]Data processing, statistical tests and plots in this section were made using R statistical software, version 3.2.2 (R Core Team 2013).

[7]Here and below, each dot in the plot represents a participant. The boxes range from 25% to 75% of the data, and the middle lines indicate the median. Whiskers extend from the hinges to 1.5*IQR, where IQR is the inter-quartile range (the distance between the first and third quartile). Observations have been jittered along the x-axis to reduce overlap.

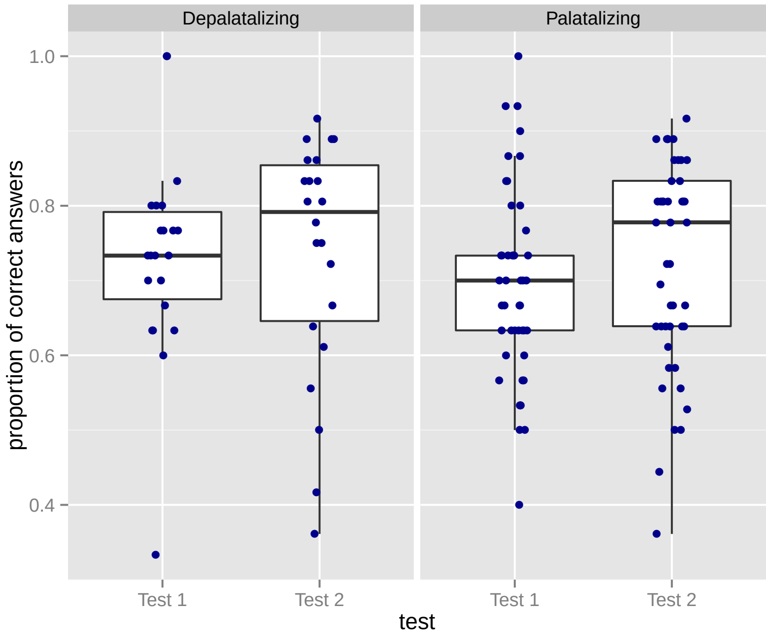

Figure 2: Results for all participants for the nonalternating items in all conditions

This can be seen more clearly in Figure 3, which shows the results for all participants only for the critical items, which test palatalization or depalatalization. The figure indicates that most participants perfomed rather poorly on those items, especially in the second test for the depalatalizing language.

To sum up, it seems that most participants dispreferred both palatalization and depalatalization over all. We attribute this to the fact that speakers tend to dislike stem alternations. Seen from this perspective, participants who failed to palatalize or depalatalize, simply refused to make generalizations about stem alternations and opted for a completely regular paradigm.

3.2. *Mixed effects models*

To test whether the differences we observed in the previous section are meaningful or not, we performed mixed effect logistic regression on the subset of the data testing for palatalization/depalatalization. For the models, we do not use the proportion of correct responses but rather we try to predict directly whether the participant gave a correct or incorrect response (coded as 1 and 0, respectively). To do this, we use the following formula: `response ∼ language + test + number + (1|item) + (1+test+number|participant)`.[8] This formula takes as independent variables the language (palatalizing or depalatalizing), the test (Test 1 or Test 2), and whether the observed item was in the singular or plural.[9] Additionally it controls for participant variation, together with participant variation across the two tests, as well as for inter item variation (to control for items that might have been harder to learn or to generalize to). The coefficients of this model are shown in Table 2.

In logistic regression, the estimates of the fixed effects express a log odds of the response being

[8]We used a logit link function for all models. Initially we tested a model with an interaction between `test` and `language` but this was not significant.

[9]A factor that turned out not to be relevant (small effect size and large p value) was the voicing specification of the alternating sound. We therefore do not include it in the model. Similarly, the trial number turned out not to be a significant factor in the model.

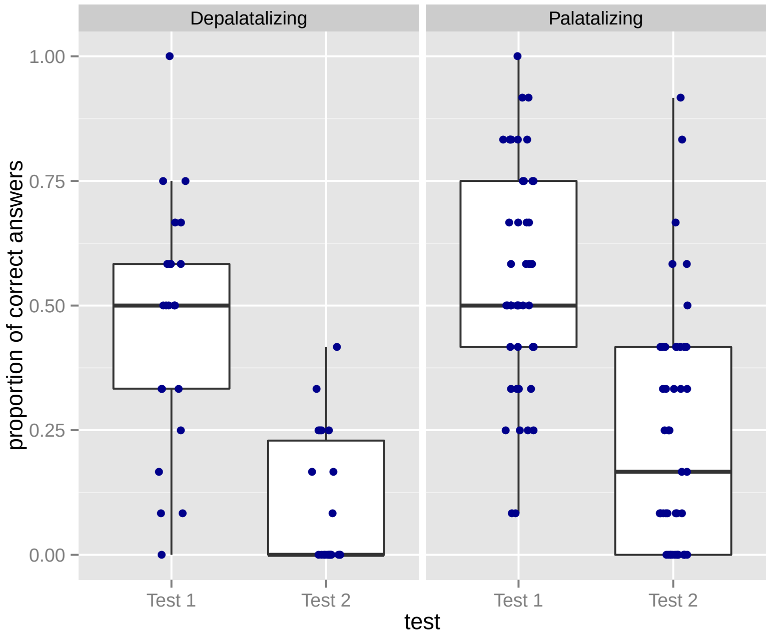

Figure 3: Results for all participants for palatalizing/depalatalizing items in all conditions

Random effects:					
Groups	Name	Variance	Std.Dev.	Corr	
participant	(Intercept)	1.01963	1.0098		
	test Test 2	3.67919	1.9181	-0.50	
	number sg	0.01522	0.1234	-0.90	0.09
item	(Intercept)	0.18024	0.4246		
Number of obs: 1608, groups: participant, 67; item, 48					
Fixed effects:					
	Estimate	Std. Error	z value	Pr(< $\|t\|$)	
(Intercept)	0.5699	0.2262	2.519	0.0118 *	
language Depalatalizing	-0.7285	0.2974	-2.450	0.0143 *	
test Test 2	-2.3171	0.3351	-6.915	4.68e-12 ***	
number singular	-0.3877	0.1879	-2.064	0.0390 *	

Table 2: Coefficients for the logistic regression model

correct (1). The intercept has the default levels of the factors, while each additional level that comes after the intercept expresses a change in the log odds. A positive estimate means that the change in level increases the odds of the response being 1 and a negative estimate means that the change in level decreases the odds. This can be understood as a level 'favouring' or 'disfavouring' a particular response. The greater the magnitude, the greater the effect. In this case, the intercept has the factors: *palatalizing, test 1, plural*. We see in the model that the strongest factor was, as would be expected, `test`. Speakers performed much worse in the second test (which included unfamiliar stems), than in the first test. The next strongest effect is that of `language` (palatalizing/depalatalizing). We see that speakers performed worse in the depalatalizing language, which confirms our initial hypothesis that palatalization before front vowels should be easier to learn. Finally, the last fixed effect `number`, shows that speakers found both palatalization and depalatalization harder to learn (or apply) for singular items than for plural ones (it should be mentioned that there was no relevant interaction between this predictor and `test`, which means they found them equally hard for recollection and generalization to new items).[10] A possible explanation of this effect is that Hungarian has a regular vowel lengthening process in stems in plural formation:

(8) alma [ɔlmɔ] 'apple' – almák [ɔlmaːk] 'apples'

It is therefore likely that Hungarian speakers found it easier to relate a stem alternation to the notion of plural than to that of singular. This would explain the observed effect.

From the random effects we can see that there was a very large degree of variation between different users, but this variation was even more pronounced during the second test. This means that the difference in performance between speakers was considerably more pronounced when the task involved generalizing to new items rather than just recalling already seen items. Two things could be at work here. The first possibility is that some of the speakers performing poorly during the second test just did not learn the generalization well enough to apply it to the new items. An alternative would be that speakers who actively dispreferred the stem alternation did manage to learn 'exceptions' in the form of seen items, but considered the new items as fully regular (from the point of view of their generalization).

Next we need to assess model performance. The corresponding confusion matrix for the regression model with accuracy and C score is presented in Table 3.

Confusion Matrix		
	Prediction	
Reference	0	1
0	890	207
1	140	371
Accuracy: 0.7842%		
C score: 0.7687		

Table 3: Confusion Matrix for the regression model

The model explains a relatively large amount of variation, and has fairly good accuracy and C scores, especially considering the type of task. This suggests that the factors considered were very relevant to the performance of the participants.

Having compared the items in the palatalizing and depalatalizing groups, let us now test the claim that speakers found it much more difficult to learn the target items that the fillers. To do this, we look at the whole data-set and fit a model with the following formula: `response ~ test + type * language + (1+type|item) + (1+test|participant)`.[11] The main difference between this and the previous

[10] We also tested `form` (diminutive vs non-diminutive) but this did not present a significant effect.

[11] To facilitate interpretation we created an interaction term with `interaction(type, language)`, instead of providing the interaction in the formula. We also removed `number` because it did not show any significant effect once the whole data set was considered. We also added an interaction between `item` and `type`, because there is a high correlation between these two variables

model is that we now include the variable `type`, which codes whether a particular item was a critical item or not. The results can be seen in Table 4.

Random effects:					
Groups	Name	Variance	Std.Dev.	Corr	
item	(Intercept)	1.4338	1.1974		
	type PalDepal	2.5463	1.5957	-0.97	
participant	(Intercept)	0.4054	0.6367		
	test Test 2	0.3034	0.5509	-0.71	
Number of obs: 6030, groups: userCode, 67; item, 95					
Fixed effects:					
	Estimate	Std. Error	z value	Pr(> $\|z\|$)	
(Intercept)	1.7318	0.2027	8.543	< 2e-16	***
testTest 2	-1.0552	0.1837	-5.745	9.17e-09	***
typeLang Critical.Palatalizing	-1.6633	0.2014	-8.260	< 2e-16	***
typeLang Filler.Depalatalizing	0.2592	0.1525	1.700	0.0891	.
typeLang Critical.Depalatalizing	-2.2660	0.2600	-8.716	< 2e-16	***

Table 4: Coefficients for the logistic regression model with the whole dataset

Table 4 again shows that the second test was much harder than the first one. What is new in this table, is that now the largest coefficient is that of `type`, which tells us that speakers performed much worse on the critical items. Moreover, `language` on its own reveals no significant effect. However, its effect becomes significant once the interaction wiht `type` is considered. This tells us that us that participants in the depalatalizing group fared worse than those in the palatalizing group only when learning the critical (alternating) items.

A post-hoc comparison using the `glht` function provides confirmation. The critical items were consistently more difficult than the fillers, and the palatalizing and depalatalizing languages only differ in the critical items.

	Estimate	Std. Error	z value	Pr(> $\|z\|$)	
Critical.Palatalizing - Filler.Palatalizing == 0	-1.6633	0.2014	-8.260	<0.001	***
Filler.Depalatalizing - Filler.Palatalizing == 0	0.2592	0.1525	1.700	0.3009	
Critical.Depalatalizing - Filler.Palatalizing == 0	-2.2660	0.2600	-8.716	<0.001	***
Filler.Depalatalizing - Critical.Palatalizing == 0	1.9224	0.2435	7.896	<0.001	***
Critical.Depalatalizing - Critical.Palatalizing == 0	-0.6027	0.2203	-2.736	0.0289	*
Critical.Depalatalizing - Filler.Depalatalizing == 0	-2.5252	0.2316	-10.901	<0.001	***

Table 5: Multiple comparisons for the interaction term

4. Summary and outlook

To sum up, the experiment yielded two noteworthy results. First, the participants in both groups managed to learn the inflectional and derivational suffixes of the artificial language, but struggled to learn the consonantal alternations. Second, the results reveal a small but statistically significant advantage for learning

as can be seen in the random effects.

palatalization over depalatalization for adult speakers of Hungarian. The first finding lends support to Kapatsinski's (2013) observation that learners have a tendency to avoid stem changes. The second one should be interpreted with greater caution, since potential confounding factors have not been thoroughly explored.

First of all, as noted by Moreton (2014), an important assumption that underlies artificial language experiments is that they simulate second-language learning rather than first-language acquisition. L2 learning may be subject to transfer effects coming from either Universal Grammar, or the participants' native language. To permit any conclusions about the UG, the design of the experiment must exclude L1 transfer as a source of the discovered biases. With respect to the present study, it is necessary to make sure that Hungarian speakers do not have the relevant linguistic experience. As discussed in section 2.4, this is not necessarily the case here. Additionally, although it is often assumed that if the pattern under investigation is absent from the native language of the participants, L1 influence should not be expected, it has recently been argued (Baer-Henney et al. 2015, Guzmán Naranjo and Zaleska to appear) that the native language of the participants can influence the results of artificial language learning experiments in unpredictable ways, even when it does not contain the relevant structures. To control for this, it would be necessary to compare the results obtained here with a similar experiment run with speakers of a different language. One potential candidate is Colombian Spanish. Like Hungarian, North-Western Spanish dialects of Colombia contain both /t,d/ and /ʧ,ʤ/ (although the latter is a positional realization of /ɟ/, occurring after a pause, a nasal consonant or /l/, see Quilis 2010:59) and have no productive rule of palatalization before front vowels.

Secondly, as noted in section 2.5, the majority of experiment participants reported knowing at least one foreign language. Some of those languages do contain palatalization before front vowels (e.g. Polish). It could therefore be the case that the effects we discovered are not the result of a cognitive bias but rather are related to the participants' pre-existing knowledge. To exclude this possibility, it would be necessary to compare the results with a control group, in which the participants complete the same test procedure either without any exposure to the artificial languages or with exposure to items that give no evidence for palatalization or depalatalization.

Another limitation of our study refers to the nature of the bias we have discovered. While the patterns look like palatalization and depalatalization before the front vowel /ɛ/, it could be the case that the learners viewed them, for example, as an arbitrary change before a derivational suffix. To investigate this, it would be necessary to run a similar study in which the suffix includes a back vowel like /u/ or /a/, which is not normally a palatalization trigger.

Finally, even if the confounding factors described above are discounted, the results are consistent not only with the with the hypothesis that ascribes the cross-linguistic asymmetry between palatalization and depalatalization to a cognitive bias but also with the competing view which offers a diachronic explanation of this asymmetry, attributing it to listener misperception, or *channel bias* (Moreton 2008). As observed by Yu (2011), channel biases and cognitive biases are extremely difficult to isolate. Indeed, a recent study by Greenwood (2014) reports on a follow-up experiment to a study comparing a 'natural' and 'unnatural' artificial language. Its results suggest that the learnability difference discovered in the initial study is related to different perceptability of the stimuli in the two conditions. A similar follow-up perception experiment using the stimuli used here, or a more general confusability study involving plain alveolars and palato-alveolars before front vowels, might help to tease apart the two alternative explanations of our results.

To sum up, the results obtained in the present study represent a first step towards explaining the typological patterns of palatalization and depalatalization. Although further research is required, these results can provide a preliminary contribution to the investigation of naturalness biases in learning palatalization.

References

Audacity Team. 2013. Audacity®. version 2.0.5. Audio editor and recorder. Available from: http://audacityteam.org/.

Baer-Henney, Dinah, Frank Kügler, and Ruben van de Vijver. 2015. The interaction of language specific and universal factors during the acquisition of morphophonemic alternations with exceptions. *Cognitive Science* 39 7: 1537–1569.

Bateman, Nicoletta. 2007. *A crosslinguistic investigation of palatalization*. Ph.D. thesis, University of California, San Diego.

Bateman, Nicoletta. 2011. On the typology of palatalization. *Language and Linguistics Compass* 5/8: 588–602.

Becker, Michael and Jonathan Levine. 2014. Experigen – an online experiment platform. Available at http://becker.phonologist.org/experigen.

Blevins, Juliette. 2004. *Evolutionary Phonology: the emergence of sound patterns*. Cambridge University Press, Cambridge.

Brodeur, Mathieu B., Emmanuelle Dionne-Dostie, Tina Montreuil, and Martin Lepage. 2010. The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE* 5 5: e10773.

Carpenter, Angela. 2010. A naturalness bias in learning stress. *Phonology* 27: 345–392.

Chen, Matthew. 1973. Predictive power in phonological description. *Lingua* 32: 173–191.

Czaplicki, Bartłomiej. 2013. Arbitrariness in grammar: Palatalization effects in Polish. *Lingua* 123: 31–57.

Ettlinger, Marc. 2008. *Input-driven opacity*. Doctoral dissertation, University of California, Berkeley.

Finley, Sara. 2012. Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes* 27 10: 1550–1562.

Flemming, Edward. 2003. The relationship between coronal place and vowel backness. *Phonology* 20 3: 335 –373.

Greenwood, Anna. 2014. Unpacking the effects of naturalness and simplicity biases on stress pattern learning. Talk presented at the 45th Meeting of the North East Linguistic Society, Cambridge, MA.

Guion, Susan G. 1996. *Velar palatalization: Coarticulation, perception and sound change*. Doctoral dissertation, University of Texas, Austin.

Guion, Susan G. 1998. The role of perception in the sound change of velar palatalization. *Phonetica* 55: 18–52.

Guzmán Naranjo, Matías and Joanna Zaleska. to appear. Probing the naturalness bias in learning stress. a cross-linguistic study. In *ConSOLE XXIII: Proceedings of the 23rd Conference of the Student Organization of Linguistics in Europe (7-9 January 2015, Paris)*, edited by Kate Bellamy, Elena Karvovskaya, Martin Kohlberger, and George Saad. Leiden University Centre for Linguistics, Leiden.

Hyman, Larry. 1976. Phonologization. In *Linguistic studies presented to Joseph H. Greenberg*, edited by Alphonse Juilland, pp. 407–418. Anma Libri, Saratoga.

Kapatsinski, Vsevolod. 2010. Velar palatalization in Russian and artificial grammar: Constraints on models of morphophonology. *Journal of Laboratory Phonology* 1 2: 361–393.

Kapatsinski, Vsevolod. 2013. Conspiring to mean: Experimental and computational evidence for a usage-based harmonic approach to morphophonology. *Language* 89 1: 110–148.

Keating, Patricia and Aditi Lahiri. 1993. Fronted velars, palatalized velars, and palatals. *Phonetica* 50: 73–101.

Ladefoged, Peter. 2001. *Vowels and consonants*. Blackwell, Massachusetts.

McMullin, Kevin. 2013. Learning consonant harmony in artificial languages. In *Actes du Congrès de l'ACL 2013 / 2013 CLA Conference Proceedings*. Association canadienne de linguistique / Canadian Linguistic Association.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25: 83–127.

Moreton, Elliott. 2014. Phonological concept learning: Experiments and models. Talk presented at LAGB Workshop on Learning Biases in Natural and Artificial Language Acquisition, Oxford, UK.

Moreton, Elliott and Joe Pater. 2012a. Structure and substance in artificial-phonology learning. Part I: Structure. *Language & Linguistics Compass* 6: 686–701.

Moreton, Elliott and Joe Pater. 2012b. Structure and substance in artificial-phonology learning. Part II: Substance. *Language & Linguistics Compass* 6: 702–718.

Ohala, John. 1993. The phonetics of sound change. In *Historical linguistics: Problems and perspectives*, edited by Charles Jones, pp. 237–278. Longman Academic.

Pater, Joe and Elliott Moreton. 2012. Structurally biased phonology: complexity in learning and typology. *Journal of the English and Foreign Languages University* 3: 1–44.

Peperkamp, Sharon, Skoruppa Katrin, and Emmanuel Dupoux. 2006. The role of phonetic naturalness in phonological rule acquisition. In *Proceedings of the 30th Boston University Conference on Language Development*, edited by David Bamman, Tatiana Magnitskaia, and Colleen Zaller, pp. 464–475. Cascadilla Press, Somerville, MA.

Pycha, Anne, Pawel Nowak, Eurie Shin, and Ryan Shosted. 2003. Phonological rule-learning and its implications for a theory of vowel harmony. In *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL) 22*, edited by Mimu Tsujimura and Gina Arding, pp. 423–435. Cascadilla Press, Somerville, MA.

Quilis, Antonio. 2010. *Principios de fonología y fonética españolas*. Arco Libros, S.L., Madrid.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schane, Sanford A., Bernard Tranel, and Harlan Lane. 1974. On the psychological reality of a natural rule of syllable structure. *Cognition* 3 4: 351–358.

Seidl, Amanda and Eugene Buckley. 2005. On the learning of arbitrary phonological rules. *Language Learning and Development* 1: 289–316.

Siptár, Péter and Miklós Törkenczy. 2007. *The phonology of Hungarian*. Oxford University Press.

Wilson, Colin. 2006. Learning phonology with substantive bias: An experimental and computational investigation of velar palatalization. *Cognitive Science* 30: 945–982.
Yu, Alan C. L. 2011. On measuring phonetic precursor robustness: a response to Moreton. *Phonology* 28: 419–518.

A. Stimuli for the familiarization phase

Singular	Plural	Diminutive singular		Diminutive plural		Meaning
		Palatalizing	Depalatalizing	Palatalizing	Depalatalizing	
-t, *-d* stems						
zɔlt-a	zɔlt-u	zɔltʧ-ɛn-a	zɔlt-ɛn-a	zɔltʧ-ɛn-u	zɔlt-ɛn-u	'ear'
lɛnsad-a	lɛnsad-u	lɛnsaʤ-ɛn-a	lɛnsad-ɛn-a	lɛnsaʤ-ɛn-u	lɛnsad-ɛn-u	'blackberry'
mungut-a		munguʧ-ɛn-a	mungut-ɛn-a			'chimpanzee'
	tɔnd-u			tɔnʤ-ɛn-u	tɔnd-ɛn-u	'ant'
	bit-u					'starfish'
ʤulamd-a						'flamingo'
zukɔnt-a	zukɔnt-u					'cloud'
bad-a	bad-u					'arrow'
zaldut-a						'seaturtle'
	mɛld-u					'avocado'
-ʧ, *-ʤ* stems						
nɔʧ-a	nɔʧ-u	nɔʧ-ɛn-a	nɔt-ɛn-a	nɔʧ-ɛn-u	nɔt-ɛn-u	'eye'
sɔtɔlʤ-a	sɔtɔlʤ-u	sɔtɔlʤ-ɛn-a	sɔtɔld-ɛn-a	sɔtɔlʤ-ɛn-u	sɔtɔld-ɛn-u	'blueberry'
zɔlpɔʧ-a		zɔlpɔʧ-ɛn-a	zɔlpɔt-ɛn-a			'chipmunk'
	tamʤ-u			tamʤ-ɛn-u	tamd-ɛn-u	'cockroach'
	kɔnʧ-u					'jellyfish'
ʧumuʤ-a						'heron'
gadulʧ-a	gadulʧ-u					'rock'
pɔʤ-a	pɔʤ-u					'bow'
nɛmʧ-a						'mussels'
	guntuʤ-u					'pecan'
Neutral stems						
nɛmb-a	nɛmb-u					'nose'
	bun-u	bun-ɛn-a				'egg'
		kum-ɛn-a		kum-ɛn-u		'antelope'
ʧalb-a				ʧalb-ɛn-u		'spider'
mundul-a		mundul-u				'clownfish'
	zɔnkap-u	zɔnkap-ɛn-a				'parrot'
		pɛʤɔm-ɛn-a		pɛʤɔm-ɛn-u		'tree'
tugalp-a				tugalp-ɛn-u		'spear'
bem-a	bem-u					'fingernail'
	gamdal-a	gamdal-ɛn-a				'banana'

B. Items for the first part of the test phase

Boldface font indicates the correct answer

B.1. ‘Palatalizing’ language

First form	Second form	Image
-t, *-d* stems		
bitʃɛna	bitɛna	starfish (dim.sg.)
zukɔntɛna	**zukɔntʃɛna**	cloud (dim.sg.)
badɛna	**badʒɛna**	arrow (dim.sg.)
dʒulamdʒɛna	dʒulamdɛna	flamingo (dim.sg.)
zaldutɛna	**zaldutʃɛna**	seaturtle (dim.sg.)
mɛldʒɛna	mɛldɛna	avocado (dim.sg.)
bitɛnu	**bitʃɛnu**	starfish (dim.pl.)
zukɔntʃɛnu	zukɔntɛnu	cloud (dim.pl.)
badʒɛnu	badɛnu	arrow (dim.pl.)
dʒulamdɛnu	**dʒulamdʒɛnu**	flamingo (dim.pl.)
zaldutʃɛnu	zaldutɛnu	seaturtle (dim.pl.)
mɛldɛnu	**mɛldʒɛnu**	avocado (dim.pl.)
-tʃ, *-dʒ* stems		
nɔtʃa	nɔtʃu	eye (sg.)
tʃumudʒu	tʃumudʒa	heron (pl.)
gadultʃɛna	**gadultʃa**	rock (sg.)
tamdʒa	**tamdʒɛna**	cockroach (dim.sg.)
kɔntʃɛnu	**kɔntʃɛna**	jellyfish (dim.sg.)
sɔtɔldʒɛna	**sɔtɔldʒɛnu**	blueberry (dim.pl)
zɔlpɔtʃu	zɔlpɔtʃɛnu	chipmunk (pl.)
pɔdʒɛnu	pɔdʒu	bow (dim.pl.)
nɛmtʃɛnu	**nɛmtʃu**	mussels (pl.)
guntudʒa	guntudʒɛna	pecan (sg.)

B.2. ‘Depalatalizing’ language

First form	Second form	Image
-tʃ, *-dʒ* stems		
kɔntɛna	kɔntʃɛna	jellyfish (dim.sg.)
gadultʃɛna	**gadultɛna**	rock (dim.sg.)
pɔdʒɛna	**pɔdɛna**	bow (dim.sg.)
tʃumudɛna	tʃumudʒɛna	heron (dim.sg.)
nɛmtʃɛna	**nɛmtɛna**	mussels (dim.sg.)
guntudɛna	guntudʒɛna	pecan (dim.sg.)
kɔntʃɛnu	**kɔntɛnu**	jellyfish (dim.pl.)
gadultɛnu	gadultʃɛnu	rock (dim.pl.)
pɔdɛnu	pɔdʒɛnu	bow (dim.pl.)
tʃumudʒɛnu	**tʃumudɛnu**	heron (dim.pl.)
nɛmtɛnu	nɛmtʃɛnu	mussels (dim.pl.)
guntudʒɛnu	**gundtudɛnu**	pecan (dim.pl.)
-t, *-d* stems		
zɔlta	zɔltu	ear (sg.)
dʒulamdu	dʒulamda	flamingo (pl.)
zukɔntɛna	**zukɔnta**	cloud (sg.)
tɔnda	**tɔndɛna**	ant (dim.sg.)
bitɛnu	**bitɛna**	starfish (dim.sg.)
lɛnsadɛna	**lɛnsadɛnu**	blackberry (dim.pl)
mungutu	mungutɛnu	chimpanzee (pl.)
badɛnu	badu	arrow (dim.pl.)
zaldutɛnu	**zaldutu**	seaturtle (pl.)
mɛlda	mɛldɛna	avocado (sg.)

B.3. Both languages

First form	Second form	Image	First form	Second form	Image
Neutral stems					
nɛmba	nɛmbu	nose (sg.)	zɔnkapɛnu	**zɔnkapɛna**	parrot (dim.sg.)
pɛdʒɔmu	**pɛdʒɔma**	tree (sg.)	**bunɛna**	bunɛnu	egg (dim.sg.)
tugalpu	tugalpa	spear (pl.)	kumɛna	**kumɛnu**	antelope (dim.pl.)
tʃalba	**tʃalbu**	spider (pl.)	**mundulɛnu**	mundulɛna	clownfish (dim.pl.)
zɔnkapa	zɔnkapɛna	parrot (sg.)	nɛmbɛnu	**nɛmbu**	nose (pl.)
kumɛna	**kuma**	antelope (sg.)	**tugalpu**	tugalpɛnu	spear (pl.)
buna	**bunɛna**	egg (dim.sg.)	**pɛdʒɔmɛnu**	pɛdʒɔmu	tree (dim.pl.)
mundulɛna	mundula	clownfish (dim.sg.)	tʃalbu	**tʃalbɛnu**	spider (dim.pl.)
bɛmɛna	bɛmɛnu	fingernail (dim.sg.)	gamdalu	**gamdalɛnu**	banana (dim.pl.)
gamdalu	**gamdala**	banana (dim.sg.)	**bɛmɛna**	bɛma	fingernail (dim.sg)

C. Items for the second part of the test phase (identical for both languages)

C.1. Experimental items

Form	Image	Correct? ('palatalizing')	Correct? ('depalatalizing')
kutʃɛna	foot (dim.sg.)	YES	**NO**
ʤumʤɛna	lemon (dim.sg.)	YES	**NO**
dazɔntʃɛna	lizard (dim.sg.)	YES	**NO**
tukɔmʤɛna	honeybee (dim.sg.)	YES	**NO**
tutʃɛna	dolphin (dim.sg.)	YES	**NO**
lunʤɛna	owl (dim.sg.)	YES	**NO**
ʤɔltɛna	hand (dim.sg.)	**NO**	YES
tʃadɛna	orange (dim.sg.)	**NO**	YES
pigultɛna	scorpion (dim.sg.)	**NO**	YES
pupɔldɛna	butterfly (dim.sg.)	**NO**	YES
gumbɛtɛna	branch (dim.sg.)	**NO**	YES
gɔmpɛdɛna	vase (dim.sg.)	**NO**	YES
tumtʃɛnu	shark (dim.pl.)	YES	**NO**
miʤɛnu	toucan (dim.pl.)	YES	**NO**
gumbɛtʃɛnu	branch (dim.pl.)	YES	**NO**
gɔmpɛʤɛnu	vase (dim.pl.)	YES	**NO**
ʤɔltʃɛnu	hand (dim.pl.)	YES	**NO**
tʃaʤɛnu	orange (dim.pl.)	YES	**NO**
tutɛnu	dolphin (dim.pl.)	**NO**	YES
lundɛnu	owl (dim.pl.)	**NO**	YES
lazɔmtɛnu	leaf (dim.pl.)	**NO**	YES
pikumdɛnu	pot (dim.pl.)	**NO**	YES
dazɔntɛnu	lizard (dim.pl.)	**NO**	YES
tukɔmdɛnu	honeybee (dim.pl.)	**NO**	YES

C.2. Fillers

Form	Image	Correct?	Form	Image	Correct?
nunla	leg (sg.)	YES	ʤɔmlu	feather (pl.)	YES
munsupu	pineapple (sg.)	NO	tʃandalɛnu	crab (pl.)	NO
ʤalsumu	alligator (pl.)	YES	nubamnɛnu	statuette (dim.pl.)	YES
taba	dragonfly (pl.)	NO	lɛbu	seashell (dim.pl.)	NO
tukɔmla	shrimp (sg.)	YES	malbu	malbu (sg.)	NO
pibɛna	bird (sg.)	NO	zɔʤulma	kiwi (pl.)	NO
punlɛna	flower (dim.sg.)	YES	lɛba	seashell (sg.)	YES
tʃalimpa	mask (dim.sg.)	NO	tʃandalɛna	crab (dim.sg.)	YES
zɔʤulmɛna	kiwi (dim.sg)	YES	ʤalsumɛnu	alligator (dim.sg.)	NO
tamɛnu	arm (dim.sg.)	NO	nunlɛna	leg (dim.pl)	NO
malbɛnu	ladybug (dim.pl)	YES	tʃalimpu	mask (pl.)	YES
pɔntʃɔpɛna	chameleon (dim.pl)	NO	pibɛnu	bird (dim.pl)	YES

D. Experiment instructions (translation from Hungarian to English)

D.1. Introduction

Welcome! In this experiment you are going to hear words in a language that you've never heard before, called Zaguna. First, we will teach you some Zaguna words. Then we will ask you some questions about Zaguna. The whole experiment will last about 15 minutes. Each participant has a chance to win one of two Amazon vouchers worth 10 EUR. To be entered in the prize draw, please provide your email address at the end of the experiment.

For this experiment you will need headphones. Please put your headphones on now and listen to this made up word. [. . .] Okay, we're ready to go!

D.2. Familiarization

Imagine you are the only survivor of a massive shipwreck. You are stranded on a small island inhabited by friendly and peaceful natives, the Zaguna tribe. The Zaguna allow you to live with them but they speak a language that you don't understand. To maintain good relations with them and to improve you position in the tribe, you try to learn their language.

You will now hear some Zaguna words. You don't have to memorize all of them, but try to figure out the rules that govern the language. Sometimes, instead of hearing a word, you will hear 'mmm'. When this happens, press the !!! button at the bottom. Press 'continue' to begin learning Zaguna.

D.3. Test 1

After a month of living with the Zaguna, you meet another survivor of the shipwreck. He also begins to live with the tribe and tries to learn the language. After two weeks, he has learnt some basic words, but still struggles with some aspects of the language. He asks you to help him by clarifying some of his doubts.

You will now see an image and you will hear two words. If you think the first word is the correct Zaguna word for what you see in the picture, press 1. If you think the second word is the right one, press 2. If you can't tell, make your best guess. You will then see another picture and hear another pair of words.

Listen to the following words... Which one is the correct Zaguna word?

D.4. Test 2

One day, you wander off into the bush and can't find your way back to the Zaguna village. After several hours of walking through the forest, you finally hear some voices. Before you approach the people, you need to make sure that they are indeed the Zaguna and not the Tinana. The Tinana is a cannibal tribe and you wouldn't like to fall into their hands. The languages of the the Zaguna and the Tinana are very similar but there are some differences. Unfortunately, the people you hear in the forest are using words that you haven't learned yet. You have to use the knowledge of the words that you HAVE learnt to decide if the words you hear could belong to Zaguna.

You will now see images you haven't seen before and you will hear a word. Your task is to decide whether the word you hear could be a Zaguna word for the thing you see in the picture. Press YES if you think the word could belong to the language and NO if you think if it couldn't. If you can't tell, make your best guess. The computer will then play the next word, until you have finished the test.

Listen to the following word... Could it be a Zaguna word for what you see in the picture?