The Impact of Similar Place Avoidance on Novel Word Learning in Adults

Annie Holtz (aholtz@ed.ac.uk)
Centre for Language Evolution, University of Edinburgh, 3 Charles Street
Edinburgh EH8 9AD, UK

Kenny Smith (kenny.smith@ed.ac.uk)
Centre for Language Evolution, University of Edinburgh, 3 Charles Street
Edinburgh EH8 9AD, UK

Mitsuhiko Ota (mota@ed.ac.uk)
School of Philosophy, Psychology and Language Sciences, University of Edinburgh, 3 Charles Street
Edinburgh EH8 9AD, UK

Abstract

Similar Place Avoidance (SPA) is the cross-linguistic tendency whereby languages avoid transvocalic consonants with the same place of articulation within a word. In this study, we examine if SPA is the result of learning biases against words where the consonants share a place of articulation. In two experiments we examine whether adults show a learning difference between place-disagreeing novel words (e.g. [tip]) and place-agreeing novel words (e.g. [tid], where [t] and [d] are coronal). Participants are taught novel words and are then tested in an object-mapping or lexical decision task. We measure participants' learning performance based on accuracy and reaction times. Results indicate that, while accuracy is comparable for place-agreeing and place-disagreeing words in both tasks, participants' lexical decision responses are generally slower for place-agreeing words. These results suggest that participants experience processing difficulties when accessing newly-formed representations of place-agreeing words, which may contribute to the existence of SPA.

Keywords: phonotactics; word learning; typology; similar place avoidance; phonology

Introduction

Different languages have different restrictions on what sound patterns are permissible in their phonology, i.e. they have different *phonotactics*. For example, no word in Modern English may begin with the sound sequence [kn], but this sequence is acceptable in languages like German and Swedish. Other phonotactic patterns, such as Similar Place Avoidance, are shared across many languages. Similar Place Avoidance (hereafter SPA) is the tendency whereby consonants separated by vowels are unlikely to share the same major place of articulation (Greenberg, 1950; Frisch, Pierrehumbert, & Broe, 2004). This pattern is a wide-spread statistical tendency and research has consistently found that words which violate SPA (i.e. place-agreeing words), like *geek* [gik], and *date* [deɪt] are under-represented in the lexicons of a large number of languages, including Arabic (Greenberg, 1950), Dutch (Shatzman & Kager, 2007), English (Berkley, 2000), French (Plénat & Roché, 2001), Hebrew (Berent & Shimron, 2003), Japanese (Kawahara, Ono, & Sudo, 2006), Javanese (Yip, 1989), Latin (Berkley, 2000), Muna (Coetzee & Pater, 2008), Russian (Padgett, 1991), Maori (Rácz, Hay, Needle, King, & Pierrehumbert, 2016), and Niger-Congo languages like Fula, Wolof and Balanta (Pozdniakov & Segerer, 2007). In contrast, no languages have been documented to exhibit the opposite phonotactic pattern, preferring tranvocalic consonants to share a major place of articulation.

One explanation for such asymmetries in natural language typology is an intrinsic bias against learning certain types of linguistic patterns. Biases of this kind have been proposed as an explanation for a number of linguistic phenomena across phonology (e.g. Wilson, 2006; A. Martin & White, 2021; Finley, 2015), syntax (e.g. Culbertson, Smolensky, & Legendre, 2012; A. Martin, Holtz, Abels, Adger, & Culbertson, 2020), and semantics (e.g. Maldonado & Culbertson, 2022). Extending this approach, we might hypothesise that SPA is due to an inductive bias against a configuration that violates a place-specific version of the more general prohibition of adjacent identical segments given by the so-called Obligatory Contour Principle (McCarthy, 1986).

Alternatively, distributional asymmetries such as SPA may be rooted in mislearning of word forms due to a channel bias, such as the tendency for multiple occurrences of the same feature to be misperceived as a single occurrence of that feature (Ohala, 1981). If word forms containing place-sharing consonants are more likely to be misremembered, they become underrepresented in the lexicon over time. Such skewed transmission of word forms can give rise to typological asymmetries in phonology (A. T. Martin, 2007). To our knowledge, there is currently no research that speaks directly to this possibility. Some studies have demonstrated that words containing consonant repetitions are more, not less, accurately learned than those without (Ota, San José, & Smith, 2021; Basnak & Ota, 2024), but these are about *identical* consonants, which are known to be exempted from SPA in some languages (Gordon, 2016; Gallagher, 2013). If words with place-agreeing *non-identical* consonants were found to be harder to learn, that would provide evidence that a word learning bias underlies the typological tendency towards SPA.

To test this prediction, we carried out two word learning experiments, in order to examine whether adults' ability to learn words is affected by the presence of place-agreeing (non-identical) consonants in the target word forms. Learning was assessed using a word-object mapping task (Experiment 1), and a lexical decision task (Experiment 2), for which we measured accuracy and reaction times. The object mapping task requires recalling associations between objects and novel words, whereas the lexical decision task requires identifying novel words as either having been part of the training data or not. The materials and procedure of both experiments were designed to be compatible with future studies on infant/child participants. [1]

Experiment 1

Experiment 1 was an artificial word learning experiment in which adult participants were trained on associations between novel words and novel objects. They were then tested on their recollection of the words and their associated objects using a forced-choice mapping task.

Methods

Participants A total of 74 participants were recruited through the online crowdsourcing platform Prolific (31 female and 43 male). Eligible participants were those who reported English as their first language, had at least a 98% previous task approval rate, and had not completed any of our previous experiments/pilots. Participants were paid £3.47. Two participants were excluded due to indicating at debrief that they had taken notes, and data from 27 individual testing trials were excluded as participants responded too quickly or too slowly.[2] The remaining data (1,701 trials) came from 72 participants, with two participants per stimuli set.

Materials Novel CVC words with unique initial and final consonants were created by combining the English stops /p, b, t, d, g, k/ with one of 9 vowels /i, e, ɛ, æ, ʌ, u, ɪ, aʊ, aɪ/. From this list we created 36 stimuli sets of 12 words (six place-disagreeing and six place-agreeing) by sampling among the available words based on the schematic in Table 1. Sets had an equal number of dorsal, coronal, and labial consonants across the place-agreeing and place-disagreeing words, and an equal number of voiced and voiceless consonants.

The 36 sets consisted of 108 unique novel words, and the number of times each words was used across sets varied. We calculated the mean neighbourhood density, mean log frequency of neighbours (based on CELEX (Baayen, Piepenbrock, & Gulikers, 1995)) and bigram probability (based on IPhOD Vaden, Halpin, & Hickok, 2009) for all instances of place-agreeing and place-disagreeing words used across the 36 sets. The word types were balanced for mean log frequency of neighbours and bigram frequency (i.e. difference fell below the pre-registered tolerance threshold of 0.2 SD for each value across the 36 sets, see Table 2). The difference between place-agreeing and place-disagreeing words for neighbourhood density fell just outside our tolerance window (0.2 SD), and so we accounted for this variation by including neighbourhood density as a fixed effect in our statistical models.[3]

The experiments also used 7 filler words (/faɪʊv/, /havəŋ/, /ɪʊnəv/, /ʃɔləm/, /wɛəlʊn/, /valən/, /nʊʃəf/) and 5 real English CVC words for catch trials (*kitten* /kɪtən/, *rabbit* /ræbɪt/, *parrot* /pærət/, *chicken* /ʧɪkən/, *pigeon* /pɪʤən/). During the training phase, these words were embedded in three carrier-phrases, namely "*Look! A ...*", "*Wow! A ...*", and "*Oh! A ...*".

All novel words and indefinite articles were synthesised using the IPA-to-Speech tool (Vasetenkov, 2025), set to the English (United Kingdom) pronunciation and recorded using Dipper (Existential Audio, Version 1.9). The phrases were synthesised using Amazon POLLY (UK voice Amy).

The visual stimuli were images of 16 novel objects, taken from the Novel Object and Unusual Name (NOUN) database (Horst & Hout, 2016). They are a subset of the 18 images used in Ota et al. (2021), and were selected based on their

Table 1: Conditional coding for selection procedure of stimuli sets. The Vs in place-agreeing (PA) words can be any of the 9 vowels. Vs in place-disagreeing (PD) words can be any vowel except the one which appears in the paired place-agreeing word. $C_{!x}$ denotes any consonant except x (i.e. excluding the final consonant in the paired place-agreeing word and the initial consonant in the place-disagreeing word).

Pair	PA words	PD words
1	p V b	p V $C_{!b\|p}$
2	t V d	t V $C_{!d\|t}$
3	k V g	k V $C_{!g\|k}$
4	b V p	b V $C_{!p\|b}$
5	d V t	d V $C_{!t\|d}$
6	g V k	g V $C_{!g\|k}$

Table 2: Mean neighbourhood density (ND), mean bigram frequency (Bigram), and mean log frequency (Frequency) of neighbours for all instances of place-agreeing (PA) and place-disagreeing (PD) words used across the 36 stimuli sets.

Agreement status	ND	Bigram	Frequency
PD words	22.171	0.001	4.055
PA words	19.680	0.001	3.960

[1] All experimental design and analysis for both experiments were pre-registered on the Open Science Framework prior to data collection and can be viewed at https://osf.io/8y9vq/?view_only=3670ec814bfc4d20b709e8a69e93d899

[2] We define responses as being too fast if they are faster than than 220 ms from audio onset. This value was derived from the audio tracks used in the testing trial such that a response is not valid if participants did not listen to at least 100 ms of the word audio (i.e. 200 ms is average initial silence + 100 ms). Responses over 10 seconds were classified as too slow. All exclusions applied in this study were preregistered.

[3] Note that this matching procedure differs from the one declared in the preregistration. This is due to an initial miscalculation of the number of neighbours.

Table 3: Sample set of novel place-agreeing (PA) and place-disagreeing (PD) words, as well as fillers, assigned to a participant in Experiment 1 and 2. In Experiment 1, participants only encountered the “heard” words. In Experiment 2, all words were used, with “heard” words appearing in training and test, whereas “unheard” words were only used at test.

Familiarity	PA words	PD words	Filler
Heard	/bæp/	/pug/	/faɹʊv/
	/taʊd/	/daɪp/	/wɛəlʊn/
	/gɪk/	/kit/	/nʊʃəf/
Unheard	/paɪb/	/beɪg/	/havəŋ/
	/dut/	/tip/	/ɹʊnəv/
	/kɪg/	/gʌd/	/ʃɔləm/

low familiarity scores (i.e., low percentages of adults who indicated they had seen the object before) and low nameability scores (i.e., low percentages of adults who spontaneously came up with the same name for the object).

The experiment itself was developed to run in participants’ web browsers using the JavaScript library jsPsych (version 7) (de Leeuw, Gilbert, & Luchterhandt, 2023).

Procedure Each participant was pseudo-randomly assigned one of the 36 stimuli sets (two participants per set). From the assigned set, three place-agreeing words were pseudo-randomly selected so that one had dorsal agreement, one coronal, and one labial agreement. The three place-disagreeing words which were *not* paired with the selected place-agreeing words were then used as the place-disagreeing words. The set also included three randomly selected filler words. All of the words were randomly assigned an object from the 16 novel items.

Participants were exposed to each novel word in their training set (six critical words and three filler words, see Table 3 for sample stimuli set), presented auditorily while simultaneously seeing an image of its associated object. The novel word was heard three times per trial, once for each carrier phrase in random order. The experiment automatically progressed between training trials. Each word appeared in three training trials, meaning that participants heard every word nine times during training. In total, there were 27 training trials, 18 for critical items, and 9 trials for filler words.

In the testing phase, participants were tested on how well they remembered the *objects* associated with the words they had learned. Participants were tested on the 9 novel words they heard during training, plus two real English words which served as catch trials. Testing trials consisted of images of four items, organised in an inverted trapezoid layout (see Figure 1), that appeared together on the screen, without any accompanying sound, for 800 ms. One of these items was the object associated with the target word (a novel object, or a picture of the real-world object for English words used in catch trials), and the three distractors were novel objects drawn randomly from the stimuli set. These images remained on the screen as participants heard the target word spoken once. Their task was to press the key associated with the correct object, either Q, Z, O or M (corresponding to the top left, bottom left, top right or bottom right object respectively). Participants’ accuracy and reaction times were measured. Participants were tested on each word four times, making a total of 44 testing trials.

After completing the testing phase, participants went through a debrief where they answered a short questionnaire about their language background and their experience of the experiment.

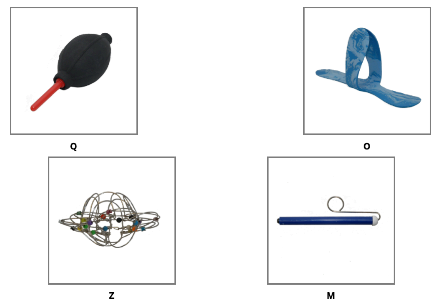

Figure 1: Sample test trial from Experiment 1.

Results

For Experiment 1 we predicted a main effect of condition, such that there would be a difference in the speed and accuracy with which participants select the objects for words based on them being place-agreeing or place-disagreeing.

Accuracy The left side of Figure 2 shows the proportion of accurate responses to critical trials. Participants performed well above chance (25% for 4-alternative forced choice) for both place-agreeing and place-disagreeing words, but there is no obvious difference in accuracy based on SPA status. We analysed this data using a mixed effects logistic regression model, run using the glmer function in the lme4 (Bates, Mächler, Bolker, & Walker, 2015) package in R, predicting accuracy (1 for correct object, 0 for incorrect), including a fixed effect for word type (place-agreeing = 1, place-disagreeing = -1), and a fixed effect for neighbourhood density (centred raw counts). The model also included a by-participant random intercept and random slope for word type, and a by-item (word form) random intercept. The model intercept was above chance, indicating successful learning of word-object associations (b = 2.75, *SE.* = 0.30, z = 12.73, $p < 0.001$; chance log odds were adjusted to account for 0.25 chance level in the forced-choice task), but there was no main effect of word type (b = -0.13, *SE.* = 0.19, z = -0.68, $p = 0.49$) or neighbourhood density (b = -0.02, *SE.* = 0.02, z = -0.91, $p = 0.37$). There is therefore no evidence that the presence or absence of word-internal consonants with the same place of articulation affects learning accuracy in an object mapping task.

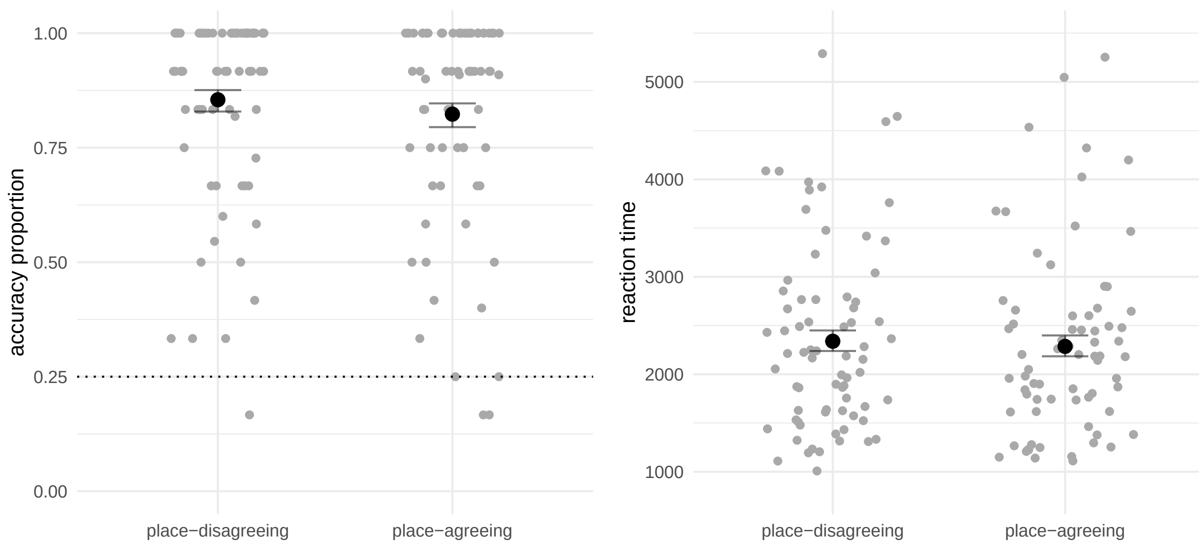

Figure 2: Left: Proportion of accurate responses split by SPA status. In both plots, small dots represent individual participant means, large dots represent group-level means, dotted line indicates chance performance and error bars show bootstrapped 95% CIs. Participants performed above chance for both word types, and there was no difference in accuracy based on SPA status. Right: Mean raw RTs for correct trials split by SPA status. Participants responded at similar speeds for both word types. NB. RTs above 5500 ms are not depicted, but are included in all analyses.

Reaction time The mean reaction times (correct responses only) can be seen on the right side of Figure 2. Visual inspection suggests that there was no difference in reaction time between the two word types. We used the lmer function to run a mixed effects linear regression model, predicting log-transformed RTs (for correct trials) based on fixed effect for word type and neighbourhood density with the same coding of fixed effects and random effect structure as in the previous analysis. The model shows no main effect of word type (b = -0.0001, *SE.* = 0.020, t = -0.01, $p = 0.94$), nor of neighbourhood density (b = 0.001, *SE.* = 0.002, t = 0.71, $p = 0.48$). These results again suggest that the presence or absence of word-internal transvocalic consonants with the same place of articulation does not affect learnability, as measured via the speed of object selection in an object mapping task.

Experiment 2

Experiment 2 was another artificial word learning experiment, featuring the same training procedure as Experiment 1 but a final test on word recollection using a lexical decision task.

Methods

Participants A total of 79 participants (35 female and 44 male) were recruited using the same platform and criteria as Experiment 1 and paid the same amount (£3.47). We excluded data from participants where technical difficulties meant we recorded partial or incorrect data (2 participants), and those who responded incorrectly on more than 1 catch trial (5 participants). Data from 28 individual trials were excluded due to responses being too fast or too slow. The remaining data (1,724 trials) came from 72 participants, with two participants per stimuli set.

Materials We used the same stimuli as in Experiment 1, and the experiment was also developed using jsPsych.

Procedure Participants were assigned a set of stimuli in the same way as in Experiment 1, and the training phase was identical. Experiment 2 differed from Experiment 1 in the composition of the test set and the testing procedure. In Experiment 1 participants were tested on the 9 novel words they were trained on (plus 2 English words in catch trials). In Experiment 2, participants were also tested on the remaining six novel words in their assigned stimulus set, not encountered during training, and an additional three fillers, also not encountered during training; these 9 additional novel items constitute *unheard* words during testing. The total test set for each participant therefore contained 18 novel words (9 heard in training, 9 unheard) and 2 English words (used in catch trials). The full stimuli set for a participant in Experiment 2 can be seen in Table 3.

Unlike testing trials in Experiment 1, testing trials in Experiment 2 did not include object images. Instead, on each trial, participants heard a single spoken word. This word was either a heard (trained) or unheard non-word, or an English word in catch trials and was presented without an indefinite article or carrier phrase. Participants were instructed to give a "yes" response if they had heard the word in training, or if it was a real English word (i.e. it was a familiar word), and to give a "no" response if they had not heard the word in training and it was not a real English word (i.e. it was an unfamiliar word). For each participant, "yes" and "no" was randomly assigned to the E and I keys. Participants were tested on all 20 words in their test set twice, in random order, for a total of 40 test trials. Accuracy and reaction time data were gathered for each trial.

Results

For Experiment 2 we again tested for effects of word type, such that there will be a difference in the speed and accuracy of responses depending on whether the word was place-agreeing vs. place-disagreeing.

Accuracy The left side of Figure 3 shows the proportion of accurate responses given by participants to non-catch trials. Mean accuracies were well above chance for both place-agreeing and place-disagreeing words, and higher for heard words compared to unheard words. We analysed this data using a mixed effects logistic regression model, again predicting accuracy (1 for correct response, 0 for incorrect response) based on fixed effects of word type (place-agreeing = 1, place-disagreeing = -1) and familiarity (1 for heard words, -1 for unheard words). The model also had a fixed effect of neighbourhood density (centred raw counts). We included by-participant random intercepts and random slopes for word type, familiarity, and their interaction, and by-item random intercepts and slopes for familiarity. The intercept was above chance, indicating successful learning (b = 2.38, *SE.* = 0.25, z = 9.46, $p < 0.001$), and there was a significant positive effect of familiarity (b = 0.95, *SE.* = 0.27, $z = 3.55$, $p < 0.001$) showing that participants were more accurate for heard words. There was no main effect of word type (b = 0.04, *SE.* = 0.13, $z = 0.29$, $p = 0.77$), nor an interaction between word type and familiarity (b = -0.003, *SE.* = 0.13, z = -0.03, $p = 0.98$), and no effect of neighbourhood density (b = 0.02, *SE.* = 0.013, z = 1.35, $p = 0.18$). These results show that, overall, participants were able to correctly recognise words they had heard during training and differentiate those from similar-sounding unheard words, and that they did this more successfully for heard words. However, SPA status did not influence their accuracy.

Reaction time The mean reaction times of participants' correct responses to non-catch trials can be seen on the right side of Figure 3. Participants' means for correct responses were shorter for heard words (-362 ms), and longer for place-agreeing words (+88 ms). We analysed this data (correct responses only) using a mixed effects linear regression, predicting log-transformed RTs based on fixed effects of word type, familiarity and neighbourhood density (model structure and effects coding were the same as in the accuracy model). There was a positive effect of word type (b = 0.03, *SE.* = 0.01, t = 2.09, $p = 0.04$), showing that participants gave slightly slower responses (i..e higher RTs) for place-agreeing words. Familiarity was significant and negative (b = -0.11, *SE.* = 0.02, t = -5.75, $p < 0.001$), showing that responses were faster when responding to heard words. There was no interaction between word type and familiarity (b = 0.001, *SE.* = 0.012, t = 0.11, $p = 0.92$), and no effect of neighbourhood density (b = -0.001, *SE.* = 0.001, t = 0.69, $p = 0.49$). These results show that both familiarity and SPA status influenced participants' RTs during word recognition: participants' responses were faster for heard words, and slower for place-agreeing words.

General Discussion

In this study we examined if place-agreeing words affected adults' word-learning. No effects of place-agreeing consonants on word learning were found in terms of accuracy. However, the fact that there was an effect of SPA status for the reaction time data in the lexical decision experiment (Experiment 2) indicates that there is some difference in participants' treatment of these word types, with the typologically-dispreferred place-agreeing words proving more challenging in lexical decision. Our interpretation is that slower RTs could indicate that the representations that participants form of place-agreeing words are not as robust as their representations of place-disagreeing words, leading to slightly slower responses when deciding whether the given word had been heard before or not. This slight processing disadvantage could give rise to the under-representation of place-agreeing words in the lexicon as previous research has shown that weak synchronic biases can have cumulatively large effects on language structure over time (A. T. Martin, 2011; Kirby, Cornish, & Smith, 2008; Reali & Griffiths, 2009). If this is the right interpretation, it also means that typological skews in phonotactics can result not only from inductive biases for certain types of generalisation, but also from token learning biases for word forms with certain configurations.

Why does this effect on RTs not show up in Experiment 1? It may be that the additional time required to motor plan the responses in a forced-choice task with four options may have masked this effect. More importantly, the word-object mapping task did not require participants to build and retain very fine-grained phonological representations of the novel words. If the word /kig/ was mapped onto object X, they only needed to differentiate it from the other 8 labels. Those 8 words included the fillers, which had completely different syllable structures from the critical words, and the remaining 5 critical words differed from /kig/ by at least two segments. The participant could select the correct object as long as they remembered that the label for X was something roughly like /kig/. In contrast, the lexical decision task was designed such that the participants had to make precise judgments about the phonological forms of the novel words. If the participant was unsure about fine details such as whether the last segment in /kig/ was indeed /g/ or something else, that would have caused their reaction time to slow down.

Why does the effect show up in RTs in Experiment 2 but not in accuracy? Both of our tasks were quite easy for most participants, with many participants performing at 100% accuracy, which may have obscured this effect — if so, a replication with even briefer exposure or a larger lexicon might show a similar penalty for place-agreeing words in less sen-

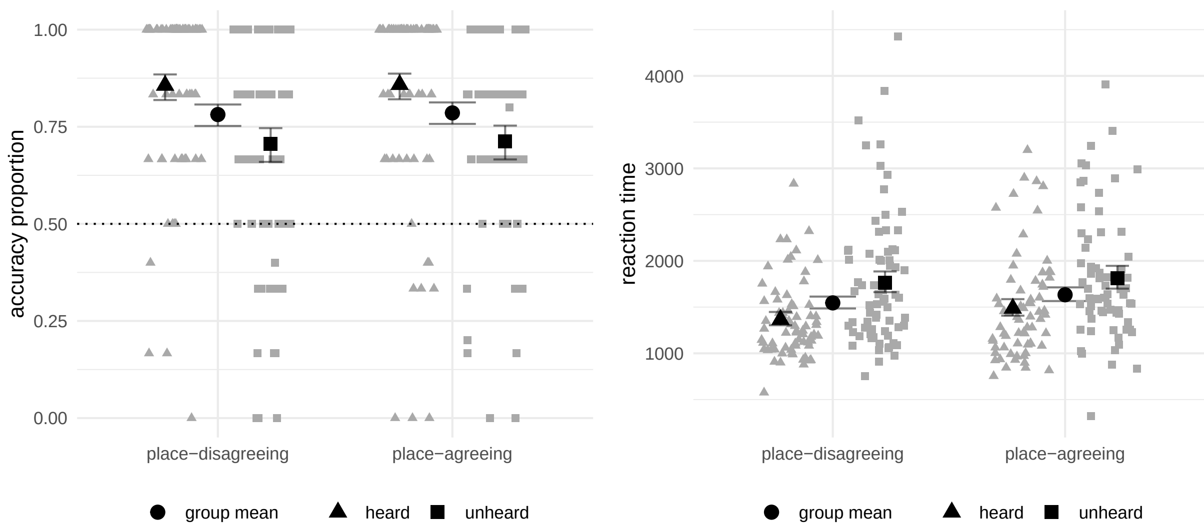

Figure 3: Left: Proportion of accurate responses split by SPA status. In both figures small grey shapes represent individual participant means, larger black shapes represent group-level mean for heard (triangle) and unheard (square) words, and black dots show group-level means for place-disagreeing and place-agreeing words. Dotted line indicates chance performance and error bars show bootstrapped 95% CIs. Participants performed above chance for both word types and accuracy was higher for heard compared to unheard words. Right: Mean raw RTs for correct trials split by SPA status. Participants responded faster for heard words, and slower for place-agreeing words. NB. RTs above 4500 ms are not depicted, but are included in all analyses.

sitive measures such as accuracy in an object mapping task.

It is worth noting that our reaction time results in Experiment 2 contrast with previous experimental work showing that learning is *facilitated* by word-internal repetition (Basnak & Ota, 2024; Ota & Skarabela, 2016; Ota et al., 2021). However, these studies have mainly examined repetition of identical elements, either full syllable repetitions (Ota & Skarabela, 2016), or repetition of whole segments (Basnak & Ota, 2024; Ota et al., 2021); as reviewed in the introduction, there are typological reasons for suspecting that identical repetition, rather than similarity due to place sharing, may behave differently, with identical repetition being more acceptable.

Our finding that participants were *slower* to reject place-agreeing words is the opposite pattern to that found in studies where participants have to differentiate between familiar (i.e. natural language) and novel words. For example, Berent, Everett, and Shimron (2001) found that Hebrew speakers were *faster* at rejecting novel OCP-violating words, compared to their speed at rejecting novel OCP-compliant words. Similarly, speakers of Dutch rejected novel place-agreeing words more quickly than they rejected novel place-disagreeing words (Shatzman & Kager, 2007). However, these studies do not look at the impact of SPA status on word *learning*; our findings suggest that our word learning task taps into a different feature of the representations of newly-learned words than that measured in these tasks. Additionally, the study with Hebrew speakers examined gemination patterns, so repetition of identical segments, rather than SPA effects (Berent et al., 2001), and the study with Dutch speakers compared place-disagreeing words to words with full segment repetition and words with place-agreeing consonants but different manners of articulation (Shatzman & Kager, 2007). The delay associated with the place-agreeing words used in this study could therefore be due to processing costs that are specific for words with place-agreeing consonants, and matched manners of articulation. Future studies will explore the effect that varying the manner of articulation has on participants' performance in these word-learning tasks.

Given that SPA is a widespread typological pattern, (Pozdniakov & Segerer, 2007; Frisch et al., 2004) and exists in the native language of our participants, namely English (Berkley, 2000)), this means that our participants have less experience of learning words with place-agreeing consonants. To ensure that the processing cost we have identified for place-agreeing words is not just a reflection of this experience, but rather an inductive bias against such words, we are planning to extend the current study to infant participants. Testing the word learning of infants allows us to minimise the impact of prior expectations regarding the shape of valid words that may have influenced our adult participants, who have a mature lexicon and an implicit understanding of the probabilities of encountering place-agreeing and place-disagreeing words. Examining how infants learn these types of words may also provide details about which population of learners contributes to creating the typological tendency for SPA. This extension would contribute to the active debate regarding children's role in shaping the structure of language (Cournane, 2019; Kam & Newport, 2009).

Acknowledgments

This project was supported by funding from the Economic and Social Research Council (project reference ES/X014312/1). We wish to thank the anonymous reviewers for their comments, as well as Johanna Basnak and the audience at the Centre for Language Evolution talk series for their helpful feedback and discussion.

References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *CELEX2 LDC96L14*. Linguistic Data Consortium. doi: 10.35111/gs6s-gm48

Basnak, J., & Ota, M. (2024). Learnability advantage of segmental repetitions in word learning. *Language and Speech*, *67*(4), 1-28. doi: 10.1177/00238309231223909

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

Berent, I., Everett, D. L., & Shimron, J. (2001). Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive psychology*, *42*(1), 1–60. doi: 10.1006/cogp.2000.0742

Berent, I., & Shimron, J. (2003). Co-occurrence restrictions on identical consonants in the Hebrew lexicon: Are they due to similarity? *Journal of Linguistics*, *39*(1), 31–55. doi: 10.1017/S0022226702001949

Berkley, D. M. (2000). *Gradient obligatory contour principle effects*. Doctoral dissertation, Northwestern University.

Coetzee, A. W., & Pater, J. (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language & Linguistic Theory*, *26*, 289–337. doi: 10.1007/s11049-008-9039-z

Cournane, A. (2019). A developmental view on incrementation in language change. *Theoretical Linguistics*, *45*(3-4), 127–150. doi: 10.1515/tl-2019-0010

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*(3), 306–329. doi: 10.1016/j.cognition.2011.10.017

de Leeuw, J. R., Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software*, *8*(85), 5351. doi: 10.21105/joss.05351

Existential Audio. (2025). *Dipper [Computer software]*. Retrieved from `https://dipper.audio/` (Version 1.9)

Finley, S. (2015). Learning nonadjacent dependencies in phonology: Transparent vowels in vowel harmony. *Language*, *91*(1), 48-72. doi: 10.1353/lan.2015.0010

Frisch, S. A., Pierrehumbert, J. B., & Broe, M. B. (2004). Similarity avoidance and the OCP. *Natural language & linguistic theory*, *22*(1), 179–228. doi: 10.1023/B:NALA.0000005557.78535.3c

Gallagher, G. (2013). Learning the identity effect as an artificial language: Bias and generalisation. *Phonology*, *30*(2), 253–295. doi: 10.1017/S0952675713000134

Gordon, M. K. (2016). *Phonological typology* (Vol. 1). Oxford University Press.

Greenberg, J. H. (1950). The patterning of root morphemes in Semitic. *Word*, *6*(2), 162–181. doi: 10.1515/9781503623217-019

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior research methods*, *48*, 1393–1409. doi: 10.3758/s13428-015-0647-3

Kam, C. L. H., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive psychology*, *59*(1), 30–66. doi: 10.1016/j.cogpsych.2009.01.001

Kawahara, S., Ono, H., & Sudo, K. (2006). Consonant co-occurrence restrictions in Yamato Japanese. *Japanese/Korean Linguistics*, *14*, 27–38.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686. doi: 10.1073/pnas.0707835105

Maldonado, M., & Culbertson, J. (2022). Person of interest: Experimental investigations into the learnability of person systems. *Linguistic Inquiry*, *53*(2), 295–336. doi: 10.1162/ling_a_00406

Martin, A., Holtz, A., Abels, K., Adger, D., & Culbertson, J. (2020). Experimental evidence for the influence of structure and meaning on linear order in the noun phrase. *Glossa: a journal of general linguistics*, *5*(1), 1–21. doi: 10.5334/gjgl.1085

Martin, A., & White, J. (2021). Vowel harmony and disharmony are not equivalent in learning. *Linguistic Inquiry*, *52*(1), 227–239. doi: 10.1162/ling_a_00375

Martin, A. T. (2007). *The evolving lexicon*. Unpublished doctoral dissertation, University of California, Los Angeles.

Martin, A. T. (2011). Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language*, *87*(4), 751–770.

McCarthy, J. J. (1986). OCP effects: Gemination and antigemination. *Linguistic inquiry*, *17*(2), 207–263.

Ohala, J. J. (1981). The listener as a source of sound change, dutch, and javanese. In C. S. Masek, R. A. Hendrick, & M. F. Miller (Eds.), *Papers from the parasession on language and behavior* (pp. 178–203).

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

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

Padgett, J. E. (1991). *Stricture in feature geometry*. Doctoral dissertation, University of Massachusetts Amherst.

Plénat, M., & Roché, M. (2001). Prosodic constraints on suffixation in french. In *Topics in morphology. selected papers from the third mediterranean morphology meeting* (pp. 285–299).

Pozdniakov, K., & Segerer, G. (2007). Similar place avoidance: A statistical universal. *Linguistic Typology*, *11*(2), 307–348. doi: 10.1515/LINGTY.2007.025

Rácz, P., Hay, J., Needle, J., King, J., & Pierrehumbert, J. B. (2016). Gradient Maori phonotactics. *Reo, Te*, *59*, 3–21. doi: 10.3316/informit.371049747671257

Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, *111*(3), 317–328. doi: 10.1016/j.cognition.2009.02.012

Shatzman, K., & Kager, R. (2007). A role for phonotactic constraints in speech perception. In *Proceedings of the 16th international congress of phonetic sciences* (pp. 1409–1412).

Vaden, K. I., Halpin, H. R., & Hickok, G. S. (2009). *Irvine phonotactic online dictionary[Data file].* Retrieved from `http://www.iphod.com.` (Version 2.0.)

Vasetenkov, A. (2025). *IPA to Speech (IPA Reader)[Computer software].* Retrieved from `https://www.antvaset.com/ipa-to-speech`

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive science*, *30*(5), 945–982. doi: 10.1207/s15516709cog0000_89

Yip, M. (1989). Contour tones. *Phonology*, *6*(1), 149–174. doi: 10.1017/S095267570000097X[Opensinanewwindow]