

JOURNAL ARTICLE

# Learning novel phonotactics from exposure to continuous speech

Frans Adriaans and René Kager

Utrecht Institute of Linguistics OTS, Utrecht University, NL

Corresponding author: Frans Adriaans (f.w.adriaans@uu.nl)

Knowledge of phonotactics is commonly assumed to derive from the lexicon. However, computational studies have suggested that phonotactic constraints might arise before the lexicon is in place, in particular from co-occurrences in continuous speech. The current study presents two artificial language learning experiments aimed at testing whether phonotactic learning can take place in the absence of words. Dutch participants were presented with novel consonant constraints embedded in continuous artificial languages. Vowels occurred at random, which resulted in an absence of recurring word forms in the speech stream. In Experiment 1 participants with different training languages showed significantly different preferences on a set of novel test items. However, only one of the two languages resulted in preferences that were above chance-level performance. In Experiment 2 participants were exposed to a control language without novel statistical cues. Participants did not develop a preference for either phonotactic structure in the test items. An analysis of Dutch phonotactics indicated that the failure to induce novel phonotactics in one condition might have been due to interference from the native language. Our findings suggest that novel phonotactics can be learned from continuous speech, but participants have difficulty learning novel patterns that go against the native language.

**Keywords:** phonological acquisition; phonotactics; statistical learning; constraint induction; speech segmentation

## 1 Introduction

One of the major challenges in early language acquisition is to learn the phonology that defines the words of the native language. Phonological acquisition starts in the first year of life, and includes the learning of stress patterns (Jusczyk et al., 1993a; Jusczyk et al., 1999b; Mehler et al., 1988; Morgan & Saffran, 1995; Nazzi et al., 1998), co-articulation (e.g., Fowler et al., 1990; Johnson & Jusczyk, 2001), allophony (e.g., Jusczyk et al., 1999a), and phonotactics (e.g., Friederici & Wessels, 1993; Jusczyk et al., 1993b; Jusczyk et al., 1994).

Experimental studies addressing the learning of phonotactics show that both adults and infants induce constraints from exposure to items designed to meet those constraints. Onishi et al. (2002) examined the learning of constraints which confined consonants to either word-initial or word-final position (e.g., /bæp/, not \*/pæb/), and constraints which linked specific consonant-vowel sequences (e.g., /bæp/ or /pib/, but not \*/pæb/ or \*/bip/). Adult participants were faster at repeating novel test words which obeyed the experiment-induced constraints than test words which violated those constraints. Similar findings have been reported in studies on phonotactic learning through speech production (e.g., Dell et al., 2000; Goldrick & Larson, 2008), and in studies with infants (Chambers et al., 2003).

Various studies have aimed to determine which phonological properties drive phonotactic learning. For example, it has been shown that learners induce constraints on consonants across different vowel contexts, suggesting that consonants and vowels are processed independently (Bonatti et al., 2005; Chambers et al., 2010). In addition, several studies indicate that patterns of natural classes are easier to learn than patterns of arbitrary classes of segments (Cristià & Seidl, 2008; Finley & Badecker, 2009; Saffran & Thiessen, 2003; Seidl & Buckley, 2005).

Recent studies have started to investigate the source of input that is used in phonotactic learning. In particular, the question has been whether learning is primarily driven by exposure to word types or by exposure to word tokens. Experiments with adults and children indicate that phonotactic learning is primarily driven by phonotactic probabilities in word types (Richtsmeier, 2011), although token variability due to speaker variation also has a positive impact on learning (Richtsmeier et al., 2011). These findings are in line with computational work showing that phonotactic constraints can be modeled as abstractions over co-occurrence patterns in the lexicon (e.g., Albright, 2009; Hayes & Wilson, 2008), and that models trained on word types typically perform better than models trained on word tokens (Albright, 2009; Bailey & Hahn, 2001; Hay et al., 2004; Pierrehumbert, 2003).

There is, however, another source of input that may affect phonotactic learning. Computational work has shown that phonotactic constraints are learnable from transcriptions of continuous speech, and that such constraints have a positive effect on word segmentation (Adriaans & Kager, 2010; Blanchard et al., 2010; Brent & Cartwright, 1996; Cairns et al., 1997; Daland & Pierrehumbert, 2011). Phonotactic constraints might thus arise before the lexicon is in place, in particular from co-occurrence probabilities in continuous speech. This view is supported by experimental evidence showing that nine-month-old infants use phonotactic probabilities to segment words from speech (Mattys et al., 1999; Mattys & Jusczyk, 2001).

Additional support for the learning of phonotactics from continuous speech comes from studies that focus on the role of segment transitional probabilities (TPs) in word segmentation (Bonatti et al., 2005; Newport & Aslin, 2004; Saffran et al., 1996; Toro et al., 2008). Bonatti et al. (2005) found that learners exploit consonant probabilities (but not vowel probabilities) for segmentation. Importantly, the study showed that learners generalize the phonotactic structure of familiarization words to novel test items (i.e., consonant frames matched with a novel vowel structure). These findings were taken as evidence that learners pick up consonantal word roots from the speech stream, rather than complete word forms (which consist of consonants and vowels). This could explain generalization to novel vowels, since no vowel information is extracted from the speech stream.

However, there is no direct evidence that human learners can induce novel phonotactic constraints directly from the speech stream, without the mediation of a lexicon of word forms. The artificial language in Bonatti et al. (2005) consisted of nine different words which occurred repeatedly in the speech stream. A possible explanation of their findings is therefore that learners induced phonotactics from a statistically learned lexicon. That is, it may have been the case that phonotactic learning was preceded by word segmentation, in which participants learned the artificial lexicon (containing nine trisyllabic word forms) from the speech stream, and subsequently judged test items based on their similarity to the stored word forms. This explanation is in line with the view that statistical learning provides a basis for an initial lexicon from which further linguistic generalizations can be derived (e.g., Swingley, 2005; Thiessen & Saffran, 2003).

The idea of constructing generalizations from unsegmented speech also appears to be in conflict with studies examining the learning of syllable class dependencies (e.g., Endress & Bonatti, 2007; Gómez, 2002; Gómez & Lakusta, 2004; Gómez & Maye, 2005;

Marcus et al., 1999; Pena et al., 2002). In such studies, learners are presented with a similar challenge of detecting dependencies across varying intervening material. For example, Gómez (2002) found that adults and infants can learn dependencies between artificial words (e.g., *pel* and *jic*) across different intervening words (e.g., *pel-wadim-jic*, *pel-kicey-jic*). Work by Pena et al. (2002) and Endress and Bonatti (2007) has addressed the important question of whether such dependencies might be learnable from continuous speech. They found that the learning of generalizations can only take place if the speech signal contains subliminal acoustic cues (i.e., short silences) to word boundaries. These findings cast doubt on whether generalizations can be learned directly from speech, without the mediation of word forms, and without acoustic cues that signal word boundaries to the learner.

The current study provides a direct test for the possibility that learners induce phonotactics from continuous speech. In two experiments with artificial languages, we investigate whether phonotactic learning can take place in the absence of recurring word forms in the speech stream.

### 1.1 The current study

Throughout this study, we will use the term ‘word’ to refer to CVCVCV sequences, and ‘frame’ to refer to C\_C\_C\_ structures (similar to lexical roots in Semitic languages, e.g., Bonatti et al., 2005; McCarthy, 1988). Any change in vowel thus represents a new word, but not a new frame.

The artificial languages used in the current experiments are relatively complex, compared to earlier studies. First, vowel slots are filled at random, so unlike in previous studies, there are no recurring words in the speech stream. This rules out the possibility that participants learn a small artificial lexicon from which generalizations are derived. Second, the consonant frames in our study are probabilistic. That is, while most earlier studies on consonant co-occurrence probabilities use a transitional probability (TP) of 1.0 within words (e.g., Bonatti et al., 2005; Newport & Aslin, 2004), our languages exhibit relatively subtle differences between ‘within-word’ and ‘between-word’ consonant probabilities ( $TP_{within} = 0.5$ ,  $TP_{between} = 0.33$ ). This is in line with a large body of work that suggests that phonotactic constraints used in speech processing are gradient rather than absolute (e.g., Bailey & Hahn, 2001; Vitevitch & Luce, 1998, 1999, 2005). Our languages thus exhibit some of the probabilistic variation that is found in natural language. The only other study we are aware of that uses relatively subtle differences in consonant TP is Toro et al. (2008), who use a consonant TP of 0.7 within words, and 0.16 between words.

Our study employs different training languages, and one single set of test items. This was done to ensure that participants’ responses are not solely driven by native language preferences or properties of test items (Reber & Perruchet, 2003; Redington & Chater, 1996). If participants base their decisions on knowledge from their native language, then they would display the same preferences in the test phase, regardless of their training condition. Conversely, if there is a significant difference between training conditions, then this can be ascribed to the structure of different familiarization languages.

## 2 Experiment 1

The experiment focuses on consonant triplets that occur across intervening vowels in a continuous speech stream. Two artificial languages are defined by different statistical structures: An ‘ABC’ language and a ‘BCA’ language. Test items in the experiment are novel combinations of consonant frames and vowel fillers. The experiment thus tests whether knowledge that is acquired by participants generalizes to novel words that have the same consonantal structure as the sequences in the continuous familiarization

language. If participants learn the phonotactic structure of the language they are exposed to, then they should show a preference for novel words that conform to the pattern of their training language.

## 2.1 Method

### 2.1.1 Participants

Forty native speakers of Dutch (33 female, 7 male) were recruited from the Utrecht Institute of Linguistics OTS subject pool (mean age: 21.4, range: 18–39). Participants received five euros for participation. Participants were assigned randomly to either the ABC or BCA condition.

### 2.1.2 Materials

Artificial languages were created using six consonant frames. Consonants were taken from natural classes of obstruents. Class A consisted of voiceless /p/, /t/, /s/. Class B consisted of their voiced counterparts: /b/, /d/, /z/. Class C had three dorsals with mixed voicing: /k/, /g/, and /x/. Consonants were concatenated to create one language consisting of ABC frames, and one language consisting of BCA frames. The materials are given in **Table 1**.

A continuous stream of consonants was generated for each language by concatenating 600 randomly selected frames (about 100 occurrences per frame), resulting in a stream of  $600 \times 3 = 1800$  consonants. ‘Within-word’ probabilities were 0.5 (consonants had 2 possible successors within frames), while ‘between-word’ probabilities were 0.33 (consonants had 3 possible successors between frames). The two languages differed with respect to the locations of high and low probabilities in the speech stream:

- (1) ...ABC.ABC.ABC.ABC.ABC.ABC... (*ABC language*)
- (2) ...A.BCA.BCA.BCA.BCA.BCA.BC... (*BCA language*)

where ‘.’ indicates a boundary as predicted by a low transitional probability between consonants. A set of six vowels (/a/, /e/, /o/, /i/, /u/, /y/) was used to fill the vowel slots between consonants. A vowel was chosen at random for each position. As a result, vowel TPs were 0.17 within words as well as between words. Due to the random insertion of vowels, each syllable bigram on average occurred only 2.7 times. Syllable TPs were thus unreliable and took on a wide range of values. Importantly, within-word syllable probabilities were not systematically higher than between-word probabilities.

Audio streams were generated using MBROLA (Dutoit et al., 1996), using the Dutch ‘nl2’ voice. Streams were synthesized with flat intonation and had a total duration of 7 minutes (with an average syllable duration of 232 ms).

**Table 1:** Artificial languages. (A = voiceless obstruents, B = voiced obstruents, C = dorsal obstruents)

ABC language		BCA language	
Consonant frames (C1_C2_C3_)	Vowel fillers (random)	Consonant frames (C2_C3_C1_)	Vowel fillers (random)
p_d_g_	[_a, _e, _o, _i, _u, _y]	d_g_p_	[_a, _e, _o, _i, _u, _y]
p_z_k_		z_k_p_	
t_b_x_		b_x_t_	
t_z_g_		z_g_t_	
s_b_k_		b_k_s_	
s_d_x_		d_x_s_	

Thirty-six novel test items were created for each language by combining consonant frames with novel vowel structures. There were no vowel repetitions within items (i.e., each word had 3 different vowels). For each test trial (e.g., /tibaxo/ – /dugopa/) a counterpart was created in which vowel frames had been switched (/tuboxa/ – /digapo/). Vowel frames were thus distributed evenly between ABC and BCA items, so that any learning effect must be due to consonant structure. Test items were chosen with a minimal difference in cohort density (i.e., the number of Dutch words starting with the initial CV syllable) between two items in a trial. All items were checked to make sure that the exact combination of consonants and vowels had not occurred in either of the two familiarization languages. The complete set of trials is given in the Appendix.

### 2.1.3 Procedure

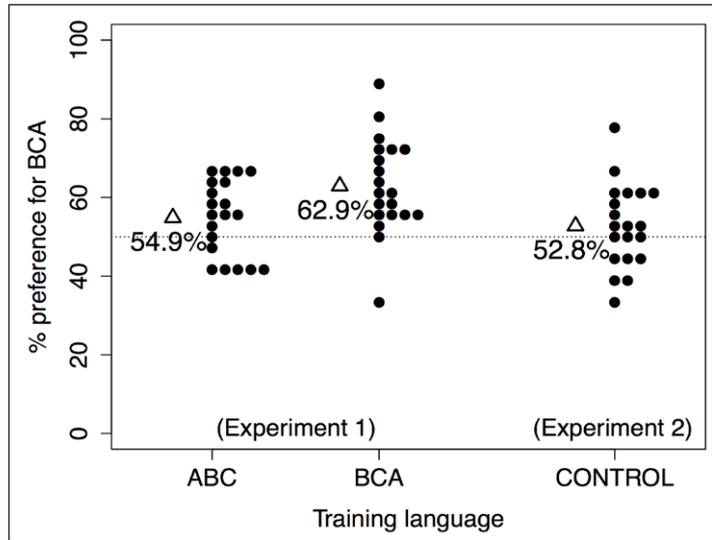
Participants were tested in a sound-attenuated booth. Participants received written instructions which were explained to them by the experimenter. Audio was presented over a pair of headphones. Participants responded by selecting one out of two response options (indicated visually with ‘1’ and ‘2’) with a mouse on the screen. The instruction given to participants was that they would hear a novel (‘Martian’) language, and that their task was to discover the words of this language. Before starting the actual experiment, participants were given a short pre-test in which they had to indicate whether a particular syllable had occurred in first or in second position in a trial. This was done to familiarize participants with the setup of the experiment.

The 7-minute familiarization stream was presented twice, with a 2-minute silence between presentations of the stream. As a consequence, total familiarization time was 14 minutes. The stream started with a 5-second fade-in, and ended with a 5-second fade-out. There were no indications of word endings or word beginnings in the speech stream. After familiarization, participants were given a two-alternative forced-choice (2AFC) task in which they had to indicate for each trial which out of two words sounded more like the Martian language they had just heard. The order of test trials was randomized, and the order of presentation of ABC and BCA items within trials was balanced within and across participants.

## 2.2 Results and discussion

**Figure 1** (left half) shows the mean preferences of individual participants in the ABC and BCA conditions, coded as the percent preference for BCA items (which would be expected to be lower than 50% in the ABC condition, and higher than 50% in the BCA condition). Participants’ responses were analyzed with mixed-effects logistic regression with subjects and items as random factors.<sup>1</sup> We first tested responses in each training condition against chance (**Table 2**). Participants in the ABC condition showed no significant preference for either ABC or BCA items ( $p = 0.114$ ). In contrast, participants in the BCA condition significantly preferred BCA words over ABC words ( $p < 0.001$ ). We then tested whether there was a significant effect of training condition on participants’ responses (**Table 3**). Participants exposed to the BCA language chose BCA items significantly more often than participants exposed to the ABC language ( $p = 0.0168$ ). The odds of choosing a BCA item for participants in the BCA condition was  $e^{0.3628} = 1.44$  times higher than for participants in the ABC condition. The significant difference between groups indicates that participants’ preferences were affected by the phonotactic structure of the continuous language to which they had been exposed.

<sup>1</sup> The ‘item’ random factor refers to Trial IDs (as listed in the Appendix). Including this random factor increased the model fit. Random slopes were tested but did not improve model fit, and were therefore not included in the reported model.



**Figure 1:** Results for the ABC and BCA (Experiment 1) and control (Experiment 2) training languages. Circles indicate mean preferences for BCA words for individual participants. Triangles indicate the mean score for each condition.

**Table 2:** Experiment 1: Summary of mixed logit model for individual conditions against chance.

Predictor	Coefficient	SE	Pr(> z )
Training language = ABC	0.2136	0.1353	0.114
Training language = BCA	0.5764	0.1368	<0.001***

**Table 3:** Experiment 1: Summary of mixed logit model testing for difference between conditions (learning effect).

Predictor	Coefficient	SE	Pr(> z )
Intercept	0.2136	0.1353	0.1144
Training language	0.3628	0.1517	0.0168*

While we can conclude from the significant difference between conditions that phonotactic learning took place in the experiment, Experiment 1 leaves open the question of which condition resulted in the learning of novel constraints. That is, even though performance in the ABC condition was not significant from chance, this does not necessarily imply that no learning took place in the ABC condition. Several studies have shown that artificial language learning from continuous speech is affected by phonological patterns in the native language (Boll-Avetisyan & Kager, 2014; Finn & Hudson Kam, 2008; Mersad & Nazzi, 2011; Onnis et al., 2005). If participants would somehow be biased towards BCA patterns, then ABC participants could have un-learned this bias (resulting in chance-level performance), and BCA participants could have developed a preference for BCA patterns while ignoring statistical cues in the artificial language. To test this possibility, we designed a control experiment aimed at testing whether Dutch participants develop a preference for BCA words in the absence of statistical cues to novel phonotactic structure.

### 3 Experiment 2

One way to assess the effects of native language phonotactics on artificial language learning is to create a continuous artificial language that is neutral in terms of transitional probability (Boll-Avetisyan & Kager, 2014). By removing the probabilistic cue from the artificial language, any preference for test items must be due to factors other than

phonotactic learning. If exposure to a language with probabilities that are constant throughout the familiarization stream leads to a preference for BCA items, then this can be taken as evidence for a native language bias.

### 3.1 Method

#### 3.1.1 Participants

Twenty native speakers of Dutch (18 female, 2 male) were recruited from the Utrecht Institute of Linguistics OTS subject pool (mean age: 24.3, range: 19–35). Participants received five euros for participation.

#### 3.1.2 Materials

A new familiarization language was constructed using the same consonant classes as in Experiment 1. Rather than imposing phonotactic constraints (i.e., specific ABC and BCA frames) on the speech stream, a sequence of consonants was generated by selecting one of the three consonants for each class at random. That is, every consonant from Class A could be followed by any of the consonants from Class B, etc. As a result, there was no difference between ‘within-word’ and ‘between-word’ probabilities. Consonant probabilities were 0.33 throughout the speech stream, providing no cue to the structure of the language:

(3) ...ABCABCABCABCABC... (*control language*)

As in Experiment 1, the resulting stream contained 1800 consonants, and six different vowels were used to fill the vowel slots at random. The method for generating the speech stream was identical to Experiment 1. The same test trials were used as in Experiment 1.

#### 3.1.3 Procedure

The procedure was the same as in Experiment 1.

### 3.2 Results and discussion

The mean preferences of participants in the control condition are shown in **Figure 1** (right half). Participants’ responses were analyzed using mixed-effects logistic regression with subjects and items as random factors. Participants in the control condition showed no significant preference for either ABC or BCA items ( $p = 0.3283$ ). We then compared performance in the control condition to performance in the ABC and BCA conditions of Experiment 1. **Table 4** shows that performance in the ABC condition was not significantly different from the control condition ( $p = 0.5400$ ). In contrast, there was a significant difference between the BCA condition and the control condition ( $p = 0.0023$ ). The odds of choosing a BCA item for participants in the BCA condition was  $e^{0.4434} = 1.56$  times higher than for participants in the control condition.

In the absence of statistical cues in the training language, participants did not develop a preference for either ABC or BCA. This result allows us to further interpret the results of Experiment 1. The significant difference between the BCA language and the control

**Table 4:** Experiment 2: Summary of mixed logit model for the control condition against chance and against the conditions from Experiment 1 (ABC and BCA).

Predictor	Coefficient	SE	Pr(> z )
Intercept (Training language = <i>control</i> )	0.1191	0.1219	0.3283
Training language = ABC	0.0883	0.1441	0.5400
Training language = BCA	0.4434	0.1454	0.0023**

language indicates that participants in the BCA condition induced novel phonotactic constraints from continuous speech. However, participants in the ABC condition failed to induce the structure of their training language.

### 3.3 Similarity to Dutch phonotactics

Why did ABC participants fail to learn the phonotactics of their training language? Earlier work by Finn and Hudson Kam (2008) found that adult learners were unable to use novel statistical information in an artificial language task when this information conflicted with their native language phonotactics. This raises the possibility that native language phonotactics is not actively guiding participants' preferences in the absence of statistical cues in the artificial language, but it is interfering with the learning of novel statistical patterns that go against a more common native language pattern. We assessed this possibility by examining ABC versus BCA patterns in Dutch words in CELEX (Baayen et al., 1995).

Although only a few words in CELEX followed the exact ABC or BCA patterns used in our experiment, we did find more words that followed the BCA pattern (6 types, 128 tokens) than the ABC pattern (2 types, 51 tokens).<sup>2</sup> Since our languages were defined in terms of natural classes of obstruents, we also looked at how common *voiceless-voiced* (AB), *voiced-dorsal* (BC), and *dorsal-voiceless* (CA) patterns are in Dutch monomorphemic words. The likelihood of these natural class patterns was calculated using the observed/expected (O/E) ratio, a measure of co-occurrence probability commonly used to quantify gradient phonotactic patterns (e.g., Adriaans & Kager, 2010; Frisch et al., 2004; Kager & Shatzman, 2007). An O/E ratio smaller than 1 indicates that a particular pattern is underrepresented (relatively uncommon) in the language, whereas patterns with a value larger than 1 are overrepresented (relatively common).

**Table 5** shows that in word-initial position, *voiceless-voiced* patterns are somewhat underrepresented in Dutch (O/E = 0.74). In contrast, *voiced-dorsal* occurrences appear to be neutral (i.e., the O/E is close to 1). In word-final position, only a few relevant sequences were found, resulting in low observed and expected frequencies. In this case *voiced-dorsal* sequences are somewhat more common than *dorsal-voiceless* sequences. To get an overall estimate of the similarity of ABC and BCA natural class patterns to Dutch phonotactics, we summed the initial and final observed and expected values for each pattern. The net result is that BCA is a relatively neutral pattern in the Dutch lexicon (O/E = 1.04), whereas ABC is somewhat underrepresented (O/E = 0.76). The fact that underrepresentation occurs

**Table 5:** Observed/expected ratios for natural class patterns in Dutch monomorphemic words (CELEX).

Position	Pattern	Observed	Expected	Observed/Expected
Initial	<i>voiceless-voiced</i> (AB)	180	243.9	0.74
Initial	<i>voiced-dorsal</i> (BC)	81	78.6	1.03
Final	<i>voiced-dorsal</i> (BC)	8	4.5	1.77
Final	<i>dorsal-voiceless</i> (CA)	10	9.1	1.10
Initial + final	<i>voiceless-voiced-dorsal</i> (ABC)	188	248.5	0.76
Initial + final	<i>voiced-dorsal-voiceless</i> (BCA)	91	87.7	1.04

<sup>2</sup> Dutch ABC words: *tudekə* (*toedekken*, 'to cover'), *tedukə* (*theedoeken*, 'tea towels'); Dutch BCA words: *bəkopə* (*bekopen*, 'to pay for'), *buketə* (*boeketten*, 'bouquets'), *dykatə* (*ducaten*, 'ducats'), *bigətə* (*bigotte*, 'bigoted'), *beikopə* (*bijkopen*, 'to buy more'), *zekutə* (*zeekoeten*, 'guillemots').

in *voiceless-voiced* sequences might be due to the involvement of two different values of a single feature, 'voice.' The other patterns involve two different features, 'voice' and 'place,' which are not *a priori* expected to interact.

In sum, we found that ABC words are relatively unlikely in terms of Dutch phonotactics. This gradient phonotactic restriction may have made it more difficult for participants in the ABC condition to learn the structure of their training language. In contrast, BCA words appear to be neutral with respect to Dutch phonotactics. The neutrality of BCA patterns could have enabled participants to induce novel statistical patterns from the BCA artificial language, while at the same time keeping neutral preferences in the absence of novel statistical cues in the random control condition. This interpretation of the results is in line with earlier findings that participants have difficulty with artificial language learning when the language being learned conflicts with native language phonotactics (Finn & Hudson Kam, 2008).

#### 4 General discussion

Human learners have been shown to be capable of learning novel phonotactic constraints from exposure to isolated words. The current study tested another possibility, which is that learners might be able to induce phonotactics from continuous speech input. Using artificial languages that had no recurring word forms, we were able to investigate whether participants induce phonotactics directly from speech. We found that novel phonotactic constraints can emerge as a result from exposure to consonant co-occurrence patterns in continuous speech. However, learning was not symmetrical between two conditions with opposite statistical patterns. The asymmetry of our results is in line with studies showing that artificial language learning is particularly difficult when the novel statistical cues conflict with native language patterns (Finn & Hudson Kam, 2008). Performance in the learning condition was nevertheless impressive given the complexity of our languages, which had random vowels, and relatively subtle statistical manipulations. Despite difficult learning conditions, participants were able to learn novel phonotactics, and generalized these structures to novel test words.

Our study is related to work on the learning of structural generalizations (e.g., Endress & Bonatti, 2007; Gómez, 2002; Gómez & Maye, 2005; Pena et al., 2002). In such studies, participants learn long-distance dependencies across intervening syllables (e.g., *pel X jic*, Gómez, 2002). In our study, learners face a similar task of abstracting consonant patterns across intervening vowels (e.g., *pXdYgZ*). The aim of our study was to see whether such generalizations can be learned from continuous speech. Earlier work by Bonatti et al. (2005) found that learners can induce C\_C\_C\_ frames and apply them to novel vowel structures. However, their design allowed for the possibility that generalizations were derived from a small, statistically learned lexicon of CVCVCV word forms, rather than directly from the speech stream. To rule out the possibility that word learning preceded phonotactic learning in our experiment, we avoided having recurring CVCVCV word forms in the speech stream. This design makes it likely that participants derived generalizations directly from continuous speech, without the mediation of an artificial lexicon.

Our findings appear to be in contrast with studies showing that generalization can only take place if the speech signal contains acoustic cues to word boundaries. Work by Pena et al. (2002) and Endress and Bonatti (2007) has shown that the learning of long-distance syllable dependencies relies on the presence of subliminal acoustic word boundaries in the speech stream. In contrast, our results suggest that the induction of phonotactic generalizations might not depend on the presence of such cues. Perhaps the difference in findings could be explained by the different nature of the generalizations involved. That is, it is conceivable that the generalization of consonant patterns can operate directly on the

speech stream, whereas the generalization of long-distance syllable patterns might require a more established segmentation. Further research will have to determine the extent to which different types of generalizations operate on different types of input.

In earlier computational work we have argued that young infants might be able to induce phonotactics from continuous speech, and use the resulting phonotactic constraints for the detection of word boundaries (Adriaans & Kager, 2010). The ability to induce phonotactics from continuous speech could serve as a way of bootstrapping into lexical acquisition. The experiments presented here provide partial support for this view by showing that under certain conditions learners are indeed able to induce novel phonotactics from continuous speech. However, more work is needed to fully understand how artificial language learning interacts with the participants' native language.

Our study has broader implications for models of phonotactic learning, which typically operate on word types in the lexicon (e.g., Hayes & Wilson, 2008). Given our finding that learners' preferences for novel words can be affected by co-occurrence patterns in connected speech, and given the vast amounts of connected speech that listeners hear on a daily basis, it seems that models of phonotactic learning should allow for contextual effects to occur. Phonotactic learning might be affected by multiple sources of input, including word types, word tokens, and continuous speech.

### Additional File

The additional file for this article can be found as follows:

- **Appendix.** Experimental items. DOI: <https://doi.org/10.5334/labphon.20.s1>

### Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research (NWO grant 277-70-001 to R.K., and NWO grant 446-10-027 to F.A.). We would like to thank Tom Lentz, Natalie Boll-Avetisyan, Elise de Bree, Annemarie Kerkhoff, and two anonymous reviewers for helpful discussion and comments. We also thank Jorinde Timmer and Natalia Davidson for assistance with running the experiments, and Arthur Dirksen for providing the speech synthesis script.

### Competing Interests

The authors have no competing interests to declare.

### References

- Adriaans, F. and Kager, R. 2010. Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language*, 62, 311–331. DOI: <https://doi.org/10.1016/j.jml.2009.11.007>
- Albright, A. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology*, 26, 9–41. DOI: <https://doi.org/10.1017/S0952675709001705>
- Baayen, H. R., Piepenbrock, R. and Gulikers, L. 1995. *The CELEX lexical database*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Bailey, T. M. and Hahn, U. 2001. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44, 568–591. DOI: <https://doi.org/10.1006/jmla.2000.2756>
- Blanchard, D., Heinz, J. and Golinkoff, R. 2010. Modeling the contribution of phonotactic cues to the problem of word segmentation. *Journal of Child Language*, 37, 487–511. DOI: <https://doi.org/10.1017/S030500090999050X>
- Boll-Avetisyan, N. and Kager, R. 2014. OCP-Place in speech segmentation. *Language and Speech*, 57, 394–421. DOI: <https://doi.org/10.1177/0023830913508074>

- Bonatti, L. L., Pena, M., Nespor, M. and Mehler, J. 2005. Linguistic constraints on statistical computations. *Psychological Science*, 16, 451–459. DOI: <https://doi.org/10.1111/j.0956-7976.2005.01556.x>
- Brent, M. R. and Cartwright, T. A. 1996. Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61, 93–125. DOI: [https://doi.org/10.1016/S0010-0277\(96\)00719-6](https://doi.org/10.1016/S0010-0277(96)00719-6)
- Cairns, P., Shillcock, R., Chater, N. and Levy, J. 1997. Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology*, 33, 111–153. DOI: <https://doi.org/10.1006/cogp.1997.0649>
- Chambers, K. E., Onishi, K. H. and Fisher, C. 2003. Infants learn phonotactic regularities from brief auditory experience. *Cognition*, 87, B69–B77. DOI: [https://doi.org/10.1016/S0010-0277\(02\)00233-0](https://doi.org/10.1016/S0010-0277(02)00233-0)
- Chambers, K. E., Onishi, K. H. and Fisher, C. 2010. A vowel is a vowel: Generalizing newly learned phonotactic constraints to new contexts. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36, 821–828. DOI: <https://doi.org/10.1037/a0018991>
- Cristià, A. and Seidl, A. 2008. Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development*, 4, 203–227. DOI: <https://doi.org/10.1080/15475440802143109>
- Daland, R. and Pierrehumbert, J. B. 2011. Learning diphone-based segmentation. *Cognitive Science*, 35, 119–155. DOI: <https://doi.org/10.1111/j.1551-6709.2010.01160.x>
- Dell, G. S., Reed, K. D., Adams, D. R. and Meyer, A. S. 2000. Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1355–1367. DOI: <https://doi.org/10.1037/0278-7393.26.6.1355>
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F. and van der Vrecken, O. 1996. The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes. In: *Proceedings of the ICSLP'96* (p. 1393–1396). Philadelphia, PA. DOI: <https://doi.org/10.1109/icslp.1996.607874>
- Endress, A. D. and Bonatti, L. L. 2007. Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, 105, 247–299. DOI: <https://doi.org/10.1016/j.cognition.2006.09.010>
- Finley, S. and Badecker, W. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language*, 61, 423–437. DOI: <https://doi.org/10.1016/j.jml.2009.05.002>
- Finn, A. S. and Hudson Kam, C. L. 2008. The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, 108, 477–499. DOI: <https://doi.org/10.1016/j.cognition.2008.04.002>
- Fowler, C., Best, C. and McRoberts, G. 1990. Young infants' perception of liquid coarticulatory influences on following stop consonants. *Perception & Psychophysics*, 48, 559–570. DOI: <https://doi.org/10.3758/BF03211602>
- Friederici, A. D. and Wessels, J. M. I. 1993. Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception & Psychophysics*, 54, 287–295. DOI: <https://doi.org/10.3758/BF03205263>
- Frisch, S. A., Pierrehumbert, J. B. and Broe, M. B. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, 22, 179–228. DOI: <https://doi.org/10.1023/B:NALA.0000005557.78535.3c>
- Goldrick, M. and Larson, M. 2008. Phonotactic probability influences speech production. *Cognition*, 107, 1155–1164. DOI: <https://doi.org/10.1016/j.cognition.2007.11.009>
- Gómez, R. L. 2002. Variability and detection of invariant structure. *Psychological Science*, 13, 431–436. DOI: <https://doi.org/10.1111/1467-9280.00476>

- Gómez, R. L. and Lakusta, L. 2004. A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, 7, 567–580. DOI: <https://doi.org/10.1111/j.1467-7687.2004.00381.x>
- Gómez, R. L. and Maye, J. 2005. The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7, 183–206. DOI: [https://doi.org/10.1207/s15327078in0702\\_4](https://doi.org/10.1207/s15327078in0702_4)
- Hay, J., Pierrehumbert, J. and Beckman, M. 2004. Speech perception, well-formedness, and the statistics of the lexicon. In: *Papers in Laboratory Phonology VI* (p. 58–74). Cambridge, UK: Cambridge University Press.
- Hayes, B. and Wilson, C. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39, 379–440. DOI: <https://doi.org/10.1162/ling.2008.39.3.379>
- Johnson, E. K. and Jusczyk, P. W. 2001. Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 548–567. DOI: <https://doi.org/10.1006/jmla.2000.2755>
- Jusczyk, P. W., Cutler, A. and Redanz, N. J. 1993a. Infants' preference for the predominant stress patterns of English words. *Child Development*, 64, 675–687. DOI: <https://doi.org/10.1111/j.1467-8624.1993.tb02935.x>
- Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y. and Jusczyk, A. M. 1993b. Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402–420. DOI: <https://doi.org/10.1006/jmla.1993.1022>
- Jusczyk, P. W., Hohne, E. and Bauman, A. 1999a. Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, 61, 1465–1476. DOI: <https://doi.org/10.3758/BF03213111>
- Jusczyk, P. W., Houston, D. M. and Newsome, M. 1999b. The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159–207. DOI: <https://doi.org/10.1006/cogp.1999.0716>
- Jusczyk, P. W., Luce, P. A. and Charles-Luce, J. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630–645. DOI: <https://doi.org/10.1006/jmla.1994.1030>
- Kager, R. and Shatzman, K. 2007. Phonological constraints in speech processing. In: Los, B. and van Koppen, M. (eds.), *Linguistics in the Netherlands 2007* (p. 100–111). John Benjamins.
- Marcus, G., Vijayan, S., Rao, S. B. and Vishton, P. 1999. Rule learning by seven-month old infants. *Science*, 283, 77–80. DOI: <https://doi.org/10.1126/science.283.5398.77>
- Mattys, S. L. and Jusczyk, P. W. 2001. Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78, 91–121. DOI: [https://doi.org/10.1016/S0010-0277\(00\)00109-8](https://doi.org/10.1016/S0010-0277(00)00109-8)
- Mattys, S. L., Jusczyk, P. W., Luce, P. A. and Morgan, J. L. 1999. Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465–494. DOI: <https://doi.org/10.1006/cogp.1999.0721>
- McCarthy, J. J. 1988. Feature geometry and dependency: A review. *Phonetica*, 43, 84–108. DOI: <https://doi.org/10.1159/000261820>
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J. and Amiel-Tison, C. 1988. A precursor of language acquisition in young infants. *Cognition*, 29, 143–178. DOI: [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- Mersad, K. and Nazzi, T. 2011. Transitional probabilities and positional frequency phonotactics in a hierarchical model of speech segmentation. *Memory and Cognition*, 39, 1085–1093. DOI: <https://doi.org/10.3758/s13421-011-0074-3>
- Morgan, J. L. and Saffran, J. R. 1995. Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, 66, 911–936. DOI: <https://doi.org/10.1111/j.1467-8624.1995.tb00913.x>

- Nazzi, T., Bertoncini, J. and Mehler, J. 1998. Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 756–766. DOI: <https://doi.org/10.1037/0096-1523.24.3.756>
- Newport, E. L. and Aslin, R. N. 2004. Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48, 127–162. DOI: [https://doi.org/10.1016/S0010-0285\(03\)00128-2](https://doi.org/10.1016/S0010-0285(03)00128-2)
- Onishi, K. H., Chambers, K. E. and Fisher, C. 2002. Learning phonotactic constraints from brief auditory experience. *Cognition*, 83, B13–B23. DOI: [https://doi.org/10.1016/S0010-0277\(01\)00165-2](https://doi.org/10.1016/S0010-0277(01)00165-2)
- Onnis, L., Monaghan, P., Richmond, K. and Chater, N. 2005. Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, 53(2), 225–237. DOI: <https://doi.org/10.1016/j.jml.2005.02.011>
- Pena, M., Bonatti, L. L., Nespors, M. and Mehler, J. 2002. Signal-driven computations in speech processing. *Science*, 298, 604–607. DOI: <https://doi.org/10.1126/science.1072901>
- Pierrehumbert, J. B. 2003. Probabilistic phonology: Discrimination and robustness. In: Bod, R., Hay, J. and Jannedy, S. (eds.), *Probabilistic linguistics*. Cambridge, MA: The MIT Press.
- Reber, R. and Perruchet, P. 2003. The use of control groups in artificial grammar learning. *The Quarterly Journal of Experimental Psychology*, 56A, 97–115. DOI: <https://doi.org/10.1080/02724980244000297>
- Redington, M. and Chater, N. 1996. Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology: General*, 125, 123–138. DOI: <https://doi.org/10.1037/0096-3445.125.2.123>
- Richtsmeier, P. T. 2011. Word-types, not word-tokens, facilitate extraction of phonotactic sequences by adults. *Laboratory Phonology*, 2, 157–183. DOI: <https://doi.org/10.1515/labphon.2011.005>
- Richtsmeier, P. T., Gerken, L. and Ohala, D. 2011. Contributions of phonetic token variability and word-type frequency to phonological representations. *Journal of Child Language*, 38, 951–978. DOI: <https://doi.org/10.1017/S0305000910000371>
- Saffran, J. R., Newport, E. L. and Aslin, R. N. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621. DOI: <https://doi.org/10.1006/jmla.1996.0032>
- Saffran, J. R. and Thiessen, E. D. 2003. Pattern induction by infant language learners. *Developmental Psychology*, 39(3), 484–494. DOI: <https://doi.org/10.1037/0012-1649.39.3.484>
- Seidl, A. and Buckley, E. 2005. On the learning of arbitrary phonological rules. *Language Learning and Development*, 1, 289–316. DOI: <https://doi.org/10.1080/15475441.2005.9671950>
- Swingle, D. 2005. Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology*, 50, 86–132. DOI: <https://doi.org/10.1016/j.cogpsych.2004.06.001>
- Thiessen, E. D. and Saffran, J. R. 2003. When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39, 706–716. DOI: <https://doi.org/10.1037/0012-1649.39.4.706>
- Toro, J. M., Nespors, M., Mehler, J. and Bonatti, L. L. 2008. Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*, 19, 137–144. DOI: <https://doi.org/10.1111/j.1467-9280.2008.02059.x>
- Vitevitch, M. S. and Luce, P. A. 1998. When words compete: Levels of processing in perception of spoken words. *Psychological Science*, 9, 325–329. DOI: <https://doi.org/10.1111/1467-9280.00064>

- Vitevitch, M. S. and Luce, P. A. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374–408. DOI: <https://doi.org/10.1006/jmla.1998.2618>
- Vitevitch, M. S. and Luce, P. A. 2005. Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, 52, 193–204. DOI: <https://doi.org/10.1016/j.jml.2004.10.003>

**How to cite this article:** Adriaans, F and Kager, R 2017 Learning novel phonotactics from exposure to continuous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1): 12, pp. 1–14, DOI: <https://doi.org/10.5334/labphon.20>

**Published:** 18 May 2017

**Copyright:** © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** The Open Access logo, which is a stylized 'a' inside a circle.