



# When statistics collide: The use of transitional and phonotactic probability cues to word boundaries

Rodrigo Dal Ben<sup>1,2</sup> · Débora de Hollanda Souza<sup>1</sup> · Jessica F. Hay<sup>3</sup>

Accepted: 1 March 2021  
© The Psychonomic Society, Inc. 2021

## Abstract

Statistical regularities in linguistic input, such as transitional probability and phonotactic probability, have been shown to promote speech segmentation. It remains unclear, however, whether or how the combination of transitional probabilities and subtle phonotactic probabilities influence segmentation. The present study provides a fine-grained investigation of the effects of such combined statistics. Adults ( $N = 81$ ) were tested in one of two conditions. In the Anchor condition, they heard a continuous stream of words with small differences in phonotactic probabilities. In the Uniform condition, all words had comparable phonotactic probabilities. In both conditions, transitional probability was stronger in words than in part-words. Only participants from the Anchor condition preferred words at test, indicating that the combination of transitional probabilities and subtle phonotactic probabilities may facilitate speech segmentation. We discuss the methodological implications of our findings, which demonstrate that even small phonotactic variations should be accounted for when investigating statistical speech segmentation.

**Keywords** Language development · Phonotactics · Speech segmentation · Statistical learning

Learning a new language involves discovering how sound sequences combine to make words. For a novice learner, this can be challenging, as words are not typically produced in isolation (Brent & Siskind, 2001), and unlike written language, spoken language does not have reliable pauses between words. Fortunately, there are a number of probabilistic cues, including transitional probability (TP) and phonotactic probability (PP) that may work together to help learners overcome the segmentation problem (Johnson, 2016; Jusczyk, 2002; Onnis et al., 2005). The current study further investigates the combined effects of TP and PP on speech segmentation and highlights the methodological implications of overlooking PPs when studying statistical speech

segmentation, which is often the case in both classical and more recent studies.

Transitional probabilities (TPs) are a type of sequential statistic found between words' segments (e.g., syllables) that can be used to predict the occurrence of the next or the previous segment (Aslin et al., 1998; Hay et al., 2011). For instance, given sufficient experience with English, one might learn that in the sequence *pretty#baby*, the chance of *pre* being followed by *ty* is greater than the chance of *ty* being followed by *ba*, signaling a word boundary. Infants and adults are able to track differences in TPs to find word boundaries in continuous speech (Saffran, Aslin, et al., 1996a; Saffran et al., 1997). Phonotactic probabilities (PPs), on the other hand, are positional statistics based on the frequency of phonological segments in given positions within words of a language (Vitevitch & Luce, 2004). For example, in the same sequence *pretty#baby*, the English PP of the word *pretty* ( $\approx 0.0440$ ) is higher than the PP of the word *baby* ( $\approx 0.0050$ ), which, in turn, is higher than the PP of the part-word *ty#ba* ( $\approx 0.0022$ ). Differences in PPs between words and/or part-words have also been shown to signal word boundaries and promote speech segmentation for both infants and adults (Jusczyk et al., 1994; Mattys & Jusczyk, 2001; Mattys et al., 1999).

To study statistical speech segmentation, researchers usually combine made-up words to create continuous speech streams. The very nature of made-up words implies that they

---

✉ Rodrigo Dal Ben  
dalbenwork2@gmail.com

<sup>1</sup> Department of Psychology, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil

<sup>2</sup> Programa de Pós-Graduação em Psicologia, Centro de Ciências Humanas, Universidade Federal de São Carlos, Via Washington Luís, Km 235–Caixa Postal 676, São Carlos, SP 13565-905, Brazil

<sup>3</sup> Department of Psychology, University of Tennessee, Knoxville, TN, USA

have TPs = 0 in the participants' native language. For instance, the syllables from the sequence *pretty#baby* could be rearranged to create the made-up sequence *tyba#preby*. Native TP knowledge will not aid segmentation of this made-up sequence, but given sufficient exposure, participants can track the new TPs and use it to find word boundaries. Artificial languages have been successful in precisely controlling TP information in their made-up words; however, this level of strict control is not automatically extended to PP. For instance, in the same made-up sequence (*tyba#preby*), the PP of the word *preby* is much higher ( $\approx 0.0391$ ) than the PP of the part-words *ba#pre* ( $\approx 0.0110$ ), which, in turn, is higher than the word *tyba* ( $\approx 0.0022$ ). If an artificial language has systematic PP differences among its' words and/or part-words, participants' native PP knowledge may affect segmentation, regardless of TP information (Finn & Hudson Kam, 2008; Mersad & Nazzi, 2011). Made-up words in statistical learning studies usually have legal phonotactics (e.g., Karaman & Hay, 2018; Mirman et al., 2008; Pelucchi et al., 2009). However, variations in phonotactic probabilities are seldom controlled. We expand this point and discuss its' implications, with data, in the Discussion.

At least two studies have focused on the combined effects of TP and PP on speech segmentation. Finn and Hudson Kam (2008) familiarized English-speaking adults to a continuous stream of speech formed by disyllabic words that always had strong TPs, but varying PPs. In one condition, words began with legal consonant clusters (e.g., /kr/, /pl/as in, "kraft," "plural"). In the other condition, words began with illegal clusters (e.g., /tf/, /km/, /bt/). Despite TPs being stronger within words (ranging from 0.25 to 1.0) than part-words (ranging from 0.035 to 0.143) in both conditions, only participants who heard the speech with legal clusters preferred words more than part-words at test. These results show that adults combine TP information learned in the task with native PP knowledge when segmenting novel speech. It also shows that segmentation may be impaired if these statistics collide (i.e., strong TP, but illegal PP).

In a more graded approach, Mersad and Nazzi (2011) showed that probabilistic differences between words' PPs may serve as cues to word segmentation. French-speaking adults were familiarized to either a uniform speech stream, with all words having medium-low PPs, or to a non-uniform speech stream, with some words having high and others having medium-low PPs. For both speech streams, TP was always stronger within words (1.0) than within part-words (0.5). Interestingly, only participants exposed to the non-uniform speech stream preferred words to part-words during test. These results suggest that the PP contrasts can be combined with TPs to promote segmentation on continuous speech. One explanation offered by the authors is that words with stronger PPs could have functioned as anchors that highlighted word boundaries.

Support for this idea comes from research showing that infants as young as 6 months old can use their knowledge of highly familiar words, such as their own name and caregivers' name, as anchors to find unfamiliar word boundaries in continuous speech streams (Bortfeld et al., 2005). In addition, by taking advantage of anchors, infants can segment speech streams with more challenging structures, such as ones with varying word length (Mersad & Nazzi, 2012). In the adult literature, the presence of recently learned words in the speech stream facilitates speech segmentation (Cunillera et al., 2010) and creates an anticipatory expectation for the words surrounding the anchor (Cunillera et al., 2016).

The findings from Finn and Hudson Kam (2008) and Mersad and Nazzi (2011) provide compelling evidence that transitional and phonotactic statistics can affect speech segmentation. However, both studies used very distinct phonotactics values (i.e., legal vs. illegal; high vs. medium-low). On one hand, given the implicit nature of PP knowledge, large PP variations may be necessary for the anchoring effect to be observed. On the other hand, PP learning starts early in life. Even before segmenting speech into words, 5-month-olds are already sensitive to PP information from their native language (Sundara & Breiss, 2020). In addition, both infants (Chambers et al., 2003) and adults (Adriaans & Kager, 2017; Onishi et al., 2002) can learn PPs from brief exposures to continuous speech in the lab. Further, PPs have been shown to affect a range of psychological phenomena, from memory to speech production (Apel et al., 2006; Gathercole et al., 1999; Graf Estes et al., 2016; MacKenzie et al., 2012; Zamuner, 2009). Together, this evidence suggests that PP knowledge may be robustly encoded in a manner that impacts speech segmentation (Onnis et al., 2005).

The present study aims to provide a more fine-grained test of the combined effects of PP and TP on speech segmentation. We conducted a conceptual replication of Mersad and Nazzi (2011), but using smaller differences in words' PPs. Again, adults heard either a stream of speech with PP anchors or a uniform stream of speech without anchors. If participants rely on TP alone, we expect both speech streams to be segmented. If participants rely on the combination of TP and PP anchors, we expect those in the Anchor condition to demonstrate better segmentation than those in the Uniform condition. By investigating the integration of transitional and phonotactic statistics, our results will provide a more fine-grained understanding of statistical language learning (Saffran, 2020).

## Method

### Participants

Eighty-one students (53 females,  $M_{\text{age}} = 22.59 \pm 5.28$  SD) from Universidade Federal de São Carlos were recruited to

participate through online postings. All participants were native speakers of Brazilian-Portuguese and had typical hearing, vision, and motor control, according to self-report. The research was approved by the Institutional Review Board of the Universidade Federal de São Carlos (#1.484.847). Participants signed the informed consent form and were randomly assigned to a condition. No compensation was provided for participation.

## Stimuli

To ensure that PPs were tightly controlled, we focused on PP information for initial stimulus selection. To do so, we used a PP calculator especially designed for this study (Estivalet & Dal Ben, *n.d.*). The calculator had two parts: the database and the search engine (algorithm based on Vitevitch & Luce, 2004). The database was built from a comprehensive Brazilian-Portuguese corpus (Estivalet & Meunier, 2015) in five steps. First, grapheme-phone transcriptions were performed automatically following the conventions set by Barbosa et al. (2003). Second, all words from the corpus were split into biphones. Third, the log (base 10) value of the frequencies of each biphone by position was summed (e.g., sum of /mæ/ frequency in the first position, in the second, and so forth). Fourth, the log value (base 10) of total word frequencies by number of biphones was summed (i.e., the total frequency of words with one, two, three biphones, and so forth). Finally, the biphones' sums, which were calculated on the second step, were divided by the words' sums, calculated on the fourth step. The quotient was the log-transformed PP of biphones in Brazilian-Portuguese. Log transformations were used because they better correlate with performance in linguistic experiments than raw frequency (Vitevitch & Luce, 2004). The search engine worked in four steps—the reverse of Vitevitch and Luce (2004). First, it took as input user-specified log-transformed PPs (e.g., exactly 0.002 in the first position, 0.003 in the second, and so forth), or PP ranges (e.g., 0.002 or higher in the first position), for each desired biphone position. Second, it searched the database for those PPs and created a list of biphones by position. Third, it recombined overlapping biphones by position (e.g., /pe/ in the first position would be recombined with /em/ in the second position, and with /mi/ in the third position to make /pemi/). Finally, it returned the combinations that matched the user-specified statistics.

Using the calculator, we selected six disyllabic pseudo-words with consonant–vowel structure (CVCV) and with the highest possible PPs before becoming words in Brazilian-Portuguese. We labeled these six pseudo-words PP+. Then, we recombined the biphones from PP+ pseudo-words to create a second set of six pseudo-words that had slightly less probable, but still high, PPs. We labeled these PP– (see Table 1).

**Table 1** Phonetic transcription (IPA), Phonotactic Probabilities (PP), and Phonetic Levenshtein Distances (PLD20) of the pseudowords with highest possible phonotactic probabilities (PP+) and their recombinations (PP–)

PP+ <sup>a</sup>			PP– <sup>b</sup>				
	IPA	PP	PLD20		IPA	PP	PLD20
dini	[d ini]	0.0090	1.4	nipe	[nipe]	0.0066	1.65
deta	[deta]	0.0085	1	tadi	[tad i]	0.0074	1.8
pemi	[pemi]	0.0082	1.6	mide	[mide]	0.0075	1
sute	[sute]	0.0084	1.2	teba	[teba]	0.0074	1.15
viko	[viko]	0.0080	1.35	kosu	[kosu]	0.0078	1.95
bara	[bara]	0.0090	1	ravi	[ravi]	0.0073	1.55
<b>Mean</b>		<b>0.0085</b>	<b>1.26</b>	<b>Mean</b>		<b>0.0073</b>	<b>1.52</b>

<sup>a</sup> Items with the highest possible PPs (before becoming words) in Brazilian Portuguese.

<sup>b</sup> Items with slightly lower PPs, but that still had relatively high PP.

To ensure that both sets were very similar to actual words from Brazilian-Portuguese, we calculated the mean number of insertions, deletions, and substitutions needed to transform PP+ and PP– words into their closest 20 phonetic neighbors (i.e., Levenshtein Distance; Yarkoni et al., 2008). We used the package *vwr* (Keuleers, 2013) for R (R Core Team, 2017) and the same Brazilian-Portuguese corpus in these calculations (Estivalet & Meunier, 2015; see Table 1). The analysis indicated that both sets were very close to actual Brazilian-Portuguese words (1.26 for PP+, 1.52 for PP–; the smaller the number of operations, the closer the words were to Brazilian-Portuguese) and to each other—difference of 0.26 mean operations.

We used disyllabic words, instead of trisyllabic ones, because it avoids possible confounding variations that arise from recombining final + initial + middle syllables or middle + final + initial syllables of trisyllabic words (Mersad & Nazzi, 2011). We chose to use a CVCV structure because it is one of the most common word structures in Brazilian-Portuguese (Estivalet, 2018). Finally, we decided to use PPs over raw biphone positional frequency (Mersad & Nazzi, 2011) because it provides a comparable measure across all biphones for each stimulus (i.e., CV, VC, CV, in the first, second and third position, respectively), it is a standard measure in psycholinguistic research (Vitevitch & Luce, 2004), and it allows for cross-linguistic comparisons (e.g., see Table 3).

When recombining these words to create continuous speeches, we aimed to mimic, as closely as possible, the type of artificial language used in many of the classic speech segmentation experiments (Aslin et al., 1998; Saffran, Aslin, et al., 1996a; Saffran et al., 1997). To rule out frequency of repetition of test items as a confound, we created four frequency-balanced languages, two for each condition (see

Aslin et al., 1998 for advantages of using a frequency-balanced language). All languages had six words (TP = 1.0). Half of the words were presented 300 times (i.e., high-frequency words) and half of the words were presented 150 times (i.e., low-frequency words). Words were concatenated in a quasi-random order following two constraints. First, a given word was never repeated in succession (e.g., *nipenipe...*). Second, high-frequency words were arranged so that their part-words (TP = 0.5) were matched in frequency with low-frequency words (i.e., they occurred 150 times each). During both familiarization and test, each word lasted for 696 ms, had a mean F0 of 220 Hz, and a mean intensity of 77 dB. Words were synthesized and concatenated, with co-articulation, using the MBROLA synthesizer (Dutoit et al., 1996) with the female Brazilian-Portuguese database br4. Each language had a total duration of approximately 15 min 40 s (stimuli are available at: <https://osf.io/s9thk/>).

In the Anchor condition, both languages had words with varying PPs in order to generate anchor effects. High-frequency words from one of the languages (Language 2) came from the PP+ set and had the highest possible PPs ( $M = 0.0085$ , range from 0.0090 to 0.0080). In the other language (Language 1), the high-frequency words came from the PP− set and had slightly lower PPs ( $M = 0.0072$ , range from 0.0075 to 0.0066; see Table 2). The opposite was true for low frequency words. Low frequency words in Language 2 came from the PP− set and had slightly lower PPs ( $M = 0.0072$ ) than they did in Language 1 ( $M = 0.0085$ ; PP+). Thus, the PP difference between the highest and the lowest PP was 0.0024 during familiarization, for both languages. Furthermore, the low-frequency words and the part-words used at test had similar PPs in both languages. It was slightly higher for Language 1 ( $M = 0.0085$  and 0.0086, respectively) and slightly lower for Language 2 ( $M = 0.0072$  and 0.0075, respectively).

In the Uniform condition, PPs were flattened for both languages during familiarization (i.e., no anchors were present). In one of the languages (Language 3), all words (high and low frequency) had slightly lower (but still high) PPs ( $M = 0.0072$ , range from 0.0066 to 0.0078; PP− set). In Language 4, all words had the highest possible PP values ( $M = 0.0085$ , range from 0.0080 to 0.0090; PP+ set; see Table 2). Thus, the PP difference between the highest and the lowest PP during familiarization was 0.0012 for Language 3 and 0.0010 for Language 4. As a consequence of the flattened PP distribution during familiarization, part-words used at test had the highest possible PPs in Language 3 ( $M = 0.0085$ ), and slightly lower in Language 4 ( $M = 0.0072$ ).

During warm-up trials, before the test phase, four existing words in Brazilian-Portuguese were contrasted with four novel words (not used in the experiment)—that is, *bola* (ball), *tela* (screen), *cabo* (cable), *pato* (duck) versus *sibu*, *bafi*, *guvi*, *tibo*, respectively.

## Procedure

The experiment consisted of three phases: familiarization, test, and self-evaluation. The task was computer administrated using PsychoPy2 (Peirce et al., 2019) and stimuli were played on high-definition neutral headphones (AKG 740 powered by Fiio E10K dac/amp) in a sound-attenuated room. At the beginning of the experiment, before the familiarization phase, music was played through the headphones at the same intensity as the experimental stimuli ( $\approx 77$  dB) and participants were instructed to adjust the volume to a comfortable level. The familiarization language was then played at the individually selected volume. In the familiarization phase, participants were told that they would hear a new language and were provided with puzzles to play with while they were listening (e.g., wood puzzles, slide puzzles, dexterity puzzle; Saffran et al., 1997). They were not instructed about any aspect of the language, nor were they instructed to pay attention to it.

After familiarization, participants were presented with four warm-up trials during which they were instructed to select the existing word in Brazilian-Portuguese compared with a novel word, by pressing either 1 or 2 on an adapted keyboard. There was a 500-ms pause between the two stimuli. In the following test phase, participants were presented with 18 two-alternative forced-choice trials that had the same structure as the warm-up trials. There were six trials for each of the three words. During each trial, participants heard one word (TP = 1) paired with a part-word (TP = 0.5)—order of presentation was counterbalanced across trials—and were instructed to select the stimulus that sounded more like the language they just heard.

Finally, during the self-evaluation phase, participants were asked to estimate the percentage of the words they were able to detect (between 0–25%, 25–50%, 50–75%, 75–100%), and to inform how focused they were on the puzzles while listening to the new language (very focused, focused, poorly focused, not focused). After the experimental task, participants completed a questionnaire including questions about their age, native language, and whether they had any experience with foreign languages.

## Data analysis

The main dependent variable was the proportion of words (versus part-words) selected during the test phase. To test whether participants relied on TP alone or on a combination of TP and PP anchors to segment the languages, the proportion of words selected in each condition was tested against chance (0.5) using classic and Bayesian  $t$  tests. All tests against chance were one-sample, two-tailed, and Bayesian tests had a default Cauchy prior (0, 0.707; Wagenmakers et al., 2018). To test whether performance differed between conditions, we used an independent-samples  $t$  test. We also

**Table 2** Phonotactic Probabilities (PP), Frequency (Freq), and Transitional Probabilities (TP) of familiarization and test stimuli of Languages (L) from Anchor and Uniform conditions (Cond)

Cond	L	Familiarization			Test								
		Words	PP	Freq	Words	PP	TP	Part-words	PP	TP			
Anchor	1	nipe	[nipe]	–	300	sute	[sute]	+	1.0	dini	[d ini]	+	0.5
		tadi	[tad i]	–	300	viko	[viko]	+	1.0	deta	[deta]	+	0.5
		mide	[mide]	–	300	bara	[bara]	+	1.0	pemi	[pemi]	+	0.5
		sute	[sute]	+	150								
		viko	[viko]	+	150								
	2	sute	[sute]	+	300	nipe	[nipe]	–	1.0	teba	[teba]	–	0.5
		viko	[viko]	+	300	tadi	[tad i]	–	1.0	kosu	[kosu]	–	0.5
		bara	[bara]	+	300	mide	[mide]	–	1.0	ravi	[ravi]	–	0.5
		nipe	[nipe]	–	150								
tadi		[tad i]	–	150									
Uniform	3	teba	[teba]	–	300	nipe	[nipe]	–	1.0	sute	[sute]	+	0.5
		kosu	[kosu]	–	300	tadi	[tad i]	–	1.0	viko	[viko]	+	0.5
		ravi	[ravi]	–	300	mide	[mide]	–	1.0	bara	[bara]	+	0.5
		nipe	[nipe]	–	150								
		tadi	[tad i]	–	150								
		mide	[mide]	–	150								
	4	dini	[d ini]	+	300	sute	[sute]	+	1.0	nipe	[nipe]	–	0.5
		deta	[deta]	+	300	viko	[viko]	+	1.0	tadi	[tad i]	–	0.5
		pemi	[pemi]	+	300	bara	[bara]	+	1.0	mide	[mide]	–	0.5
		sute	[sute]	+	150								
		viko	[viko]	+	150								
		bara	[bara]	+	150								

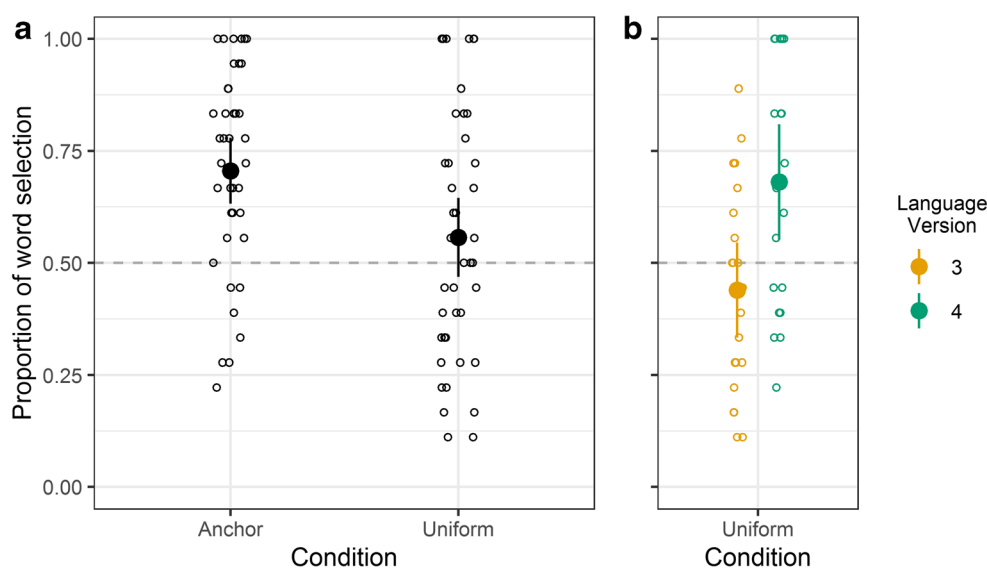
tested correlations between the proportion of words selected, self-evaluations of performance and focus on the task, and experience with foreign languages. All analyses were performed on JASP (JASP Team, 2018) and are available at OSF (<https://osf.io/s9thk/>). A post hoc power analysis of the Anchor condition, the one with results significantly different than chance, revealed 99% of power to detect a true effect (Faul et al., 2007).

## Results

We found a significant difference in word selection across conditions,  $t(79) = 2.613$ ,  $p = 0.011$ , Cohen's  $d = 0.581$ , 95% CI [0.035, 0.262], suggesting that information from TP and PP anchors were combined during word segmentation (see Fig. 1a). In the Anchor condition, words were selected ( $M = 0.70$ ,  $SD = 0.22$ ) significantly above chance (0.5),  $t(39) = 5.661$ ,  $p < .001$ ,  $d = 0.895$  [0.523, 1.259], with extreme evidence in favor of a combined effect of TP and PP on word selection,  $BF_{\text{alternative}} = 11,446$ . Furthermore, there was no

significant difference between performance across language versions (1 and 2),  $t(38) = 0.378$ ,  $p = .707$ ,  $d = -0.120$  [–0.176, 0.121]. Participants were equally successful when tested on target words with the highest phonotactics (PP+; Language 1) or with slightly lower phonotactics (PP–; Language 2). In addition, no significant correlations were found between the proportion of words selected, self-evaluation of performance and focus, or experience with foreign languages.

In contrast, in the Uniform condition, words were not selected ( $M = 0.55$ ,  $SD = 0.28$ ) above chance,  $t(40) = 1.305$ ,  $p = .199$ ,  $d = 0.204$  [–0.107, 0.512], and no evidence was found for the effects of TP alone on word segmentation,  $BF_{\text{alternative}} = 0.370$ . Nonetheless, a significant difference was found between language versions (3 and 4; see Fig. 1b),  $t(39) = -3.038$ ,  $p = .004$ ,  $d = -0.949$  [–1.519, –0.296]. Participants who heard Language 3 did not select words above chance ( $M = 0.43$ ,  $SD = 0.23$ ),  $t(20) = -1.197$ ,  $p = .245$ ,  $d = -0.261$  [–0.693, 0.177],  $BF_{\text{alternative}} = 0.427$ . In contrast, participants that listened to Language 4 selected words above chance ( $M = 0.68$ ,  $SD = 0.27$ ),  $t(19) = 2.936$ ,  $p = .008$ ,  $d = 0.656$  [0.165, 1.134],



**Fig. 1** **a** Proportion of word selection across conditions (Anchor and Uniform). **b** Proportion of word selection across Languages 3 and 4, from the Uniform condition. Empty dots represent individual

performance, and filled dots represent the mean performance for each Condition (**a**) or Language (**b**). Point ranges indicate 95% confidence intervals. Dashed lines indicate chance level (0.5)

$BF_{\text{alternative}} = 5.875$ . In addition, a positive correlation was found between the proportion of words selected and self-evaluation of performance ( $r_s = 0.634$ ,  $p = 0.003$ ).

The overall successful segmentation in the Anchor condition but not in the Uniform condition suggests that adults combined information from TP and PP anchors to segment the continuous speech. It is noteworthy that the flat PP distribution in the Uniform condition created test items with unbalanced PPs (see Table 2). In this condition, phonotactic information present during test may have led to word preferences in one language version (Language 4), but not in the other (Language 3). We return to these points in the Discussion.

## Discussion

The present study investigated the combined effects of transitional probabilities and phonotactic probabilities on speech segmentation. Across two conditions, participants were exposed either to artificial languages in which all statistically defined words (TP = 1) had similar phonotactic probabilities (i.e., Uniform condition) or to artificial languages in which words had varying degrees of phonotactic probabilities (i.e., Anchor condition). If participants relied on TP alone, we expected similar segmentation performance for both conditions, given that TP was always stronger in words (TP = 1) in comparison with part-words (TP = 0.5). On the other hand, if participants relied on a combination of TP and PP, we expected segmentation to happen only in the Anchor condition.

The successful segmentation found in the Anchor condition, but not in the Uniform condition, suggests a combined effect of TP and PP anchors on speech segmentation. These

results were found using very small PP differences between words in the Anchor condition. Thus, we replicated Mersad and Nazzi's (2011) findings under more challenging circumstances and these results add support to the argument that anchors play an important role in speech segmentation (Bortfeld et al., 2005; Cunillera et al., 2010; Cunillera et al., 2016; Mersad & Nazzi, 2011, 2012). Nonetheless, the separate analysis of each speech stream from the Uniform condition indicates that PP effects may have extended to the test phase. In this condition, PP distributions were flat during familiarization for both versions of the familiarization language. Thus, participants could not have relied on PP anchors to find word boundaries. However, as a side effect of having flat PPs during the familiarization, words and part-words used during test had unbalanced PPs. For one language (Language 3), part-words had higher PPs than did words; for the other language, words had higher PPs (Language 4). Thus, even if participants did not segment the speech stream during familiarization, unbalanced PP information during test could still affect performance. Based solely on the PP information during test, participants exposed to Language 4 should prefer words (higher PP on words), but participants exposed to Language 3 should not (lower PP on words), which is in accordance with our findings. Furthermore, the positive correlation between performance and self-evaluation found only in Language 4 might be a further indication that participants were using PP information recently presented in the test phase—a recency effect. We did not expect these results because we assumed that the small PP differences in our stimuli would only drive learning after repeated exposure during the familiarization phase (at least 150 repetitions of each word). However, it seems that the limited experience with only six repetitions of each word

and part-word during the test was sufficient to drive participants' preferences. Although unexpected, this is in line with evidence showing that abrupt distinction between legal versus illegal phonotactics at test affect preference, regardless of TP information presented during familiarization (Finn & Hudson Kam, 2008). Future research can prevent PP effects during test by using the same set of test items across conditions (Adriaans & Kager, 2017; Reber & Perruchet, 2003). It is worth noting, however, that the higher the stimuli PPs, as in the current study, the harder it is to balance PPs at both familiarization and test—especially when differences in PP, as small as 0.0024, might affect performance, a point we return in the next paragraphs.

Our findings also provide a note of caution for speech segmentation researchers, who may overlook the impact that subtle phonotactic differences in their experimental stimuli might have on segmentation. Typically, statistical speech segmentation researchers use counterbalanced speech streams to control for any “general preferences for certain syllable strings” (Saffran, Aslin, et al., 1996a, p 1928) across participants. Thus, it is assumed that all observed learning reflects listeners sensitivity to TP information presented in the exposure stream rather than participants' previous knowledge of their native language (such as phonotactics). However, counterbalancing does not necessarily control for differences in PP that, combined with TP, could transform words into anchors.

The results from the present study suggest that PP differences as small as 0.0024 (highest PP minus the lowest [0.0090 – 0.0066], Anchor condition) between stimuli can be combined with TP information to create anchors that promote segmentation. Using this difference as a threshold, we conducted a brief analysis of PPs from words, part-words, and nonwords used in artificial languages from a number of well-known statistical segmentation studies. Based on our knowledge of the literature, we analyzed highly cited seminal studies with infants and adults (Aslin et al., 1998; Saffran, Aslin, et al., 1996a; Saffran, Newport, et al., 1996b; 1209, 5590, and 1389 citations, respectively<sup>1</sup>); one study that used disyllabic stimuli and a combination of TP and stress cues to segmentation (Thiessen & Saffran, 2003; 607 citations); one study that combined speech segmentation and word learning (Estes et al., 2007; 491 citations); and one study published by one of us (Hay & Saffran, 2012; 62 citations). For each study, we first calculated PPs for each reported stimulus using Vitevitch and Luce (2004) phonotactic calculator. Then, we subtracted the lowest PP from the highest PP for each word from the speech stream and between test items (words, part-words, nonwords). All phonetic transcriptions, PPs, and calculations are openly available at OSF (<https://osf.io/s9thk/>).

<sup>1</sup> All number of citations were retrieved from Google Scholar on December 12, 2020.

Overall, we found an average difference of 0.0082 across studies' speech streams (range from 0.0042 to 0.0199; see Table 3), which is 3.4 higher than the 0.0024 difference from our stimuli (Anchor condition). Even the smallest difference, 0.0042 (Estes et al., 2007; Experiment 2) is still almost twice as high as the difference in our stimuli. This shows that, regardless of counterbalancing, PP differences persist in speech segmentation tasks. Such differences, combined with TP information, may have created PP anchors that highlighted word boundaries and facilitated speech segmentation during familiarization—as in our Anchor condition. Furthermore, with the exception of the first experiment from Saffran, Aslin, et al. (1996a), all test items also had systematic PP differences. The overall mean difference was 0.0055 across studies' test phases (range from 0.0020 to 0.0090), 2.3 times higher than the PP difference in the test phase of the Uniform condition of our study (0.0024). Such differences could have also driven word preferences during test.

Moreover, the effects of PP anchors in segmentation may not be limited to artificial languages. With this in mind, we also calculated PPs for one of the Italian languages used in a number of statistical learning studies by Hay et al. (2011, Corpus 2A; e.g., Karaman & Hay, 2018; Pelucchi et al., 2009; Shoaib et al., 2018). This natural language controls for critical TP information, but it is composed of meaningful and grammatically correct phrases in Italian and is naturally produced by a native Italian-speaker (Hay et al., 2011). Again, we used the PP calculator by Vitevitch and Luce (2004) to calculate the English PPs for the Italian words, as participants were native English-speakers. All phonetic transcriptions, PPs, and calculations are openly available at OSF (<https://osf.io/s9thk/>).

We found an overall difference between PPs of 0.0120 (range from 0.0001 to 0.0617), which is 5 times higher than the 0.0024 difference from our own stimuli. In addition, the PPs follow a skewed distribution similar to word frequencies distributions in natural languages (Fig. 2a; Kurumada et al., 2013; Zipf, 1965). Most of the words had low PPs and few words had very high PPs. When combined, they generate phrases with highly variable PP information (see Fig. 2b), which, in turn, may generate PP anchors that are likely to support speech segmentation. Future corpus analyses could be very informative on this matter.

The presence of PP differences in all the analyzed languages (artificial and natural) raises the possibility that differences in words' PPs may have created anchors that helped to promote speech segmentation in these studies. Thus, it is possible that learning that has been previously assumed to depend solely on TPs may be a result, instead, of a combination of TP and PP information. However, at least two important counterpoints should be noted.

First, our data, as in Mersad and Nazzi (2011) and Finn and Hudson-Kam (2008), come from adult participants and may not reflect linguistic sensitivities in infants. On one hand,

**Table 3** Differences between highest and lowest phonotactic probabilities and their range for each experiment of the analyzed studies

Study	Exp <sup>a</sup>	Familiarization				Test	
		Language A		Language B		Phonotactic difference	
		Phonotactic difference		Phonotactic difference			
		High – Low	Range	High – Low	Range	High – Low	Range
Present study	1	0.0024	0.0066, 0.0090	0.0024	0.0066, 0.0090	0.0011 <sup>b</sup>	*
	2	0.0012	0.0066, 0.0078	0.0010	0.0080, 0.0090	0.0024 <sup>c</sup>	*
Saffran, Aslin, et al. (1996a)	1	0.0098	0.0064, 0.0162	0.0059	0.0046, 0.0105	0.0020	0.0066, 0.0086
	2	0.0199	0.0032, 0.0231	0.0056	0.0046, 0.0102	0.0090	0.0032, 0.0122
Saffran, Newport, et al. (1996b)	1	0.0069	0.0053, 0.0122	**	**	***	***
	2	0.0100	0.0032, 0.0132	**	**	***	***
Saffran et al. (1997)	1	0.0069	0.0053, 0.0122	**	**	0.0056	0.0045, 0.0101
Aslin et al. (1998)	1	0.0090	0.0032, 0.0122	0.0056	0.0046, 0.0102	0.0090	0.0032, 0.0122
Saffran & Thiessen (2003)	1	0.0063	0.0020, 0.0083	**	**	0.0063	0.0020, 0.0083
Graf Estes, Evans, Alibali, & Saffran (2007)	1	0.0165	0.0029, 0.0194	0.0080	0.0041, 0.0121	0.0038	0.0029, 0.0067
	2	0.0042	0.0029, 0.0071	0.0057	0.0012, 0.0069	0.0036	0.0012, 0.0048
Hay & Saffran (2012)	1 & 3	0.0063	0.0020, 0.0083	0.0045	0.0066, 0.0021	0.0045	0.0066, 0.0021

<sup>a</sup> For the Present study, Exp signals conditions (1 = Anchor, 2 = Uniform).

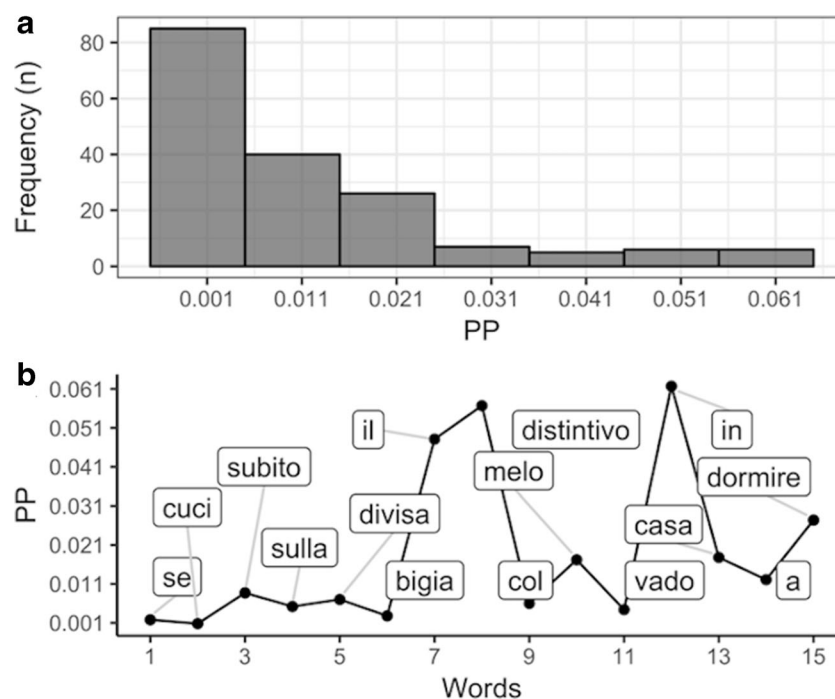
<sup>b</sup> Mean PP difference across test items from Languages 1 and 2 (A and B).

<sup>c</sup> Mean PP difference across test items from Languages 3 and 4 (A and B).

\*The range varies according to the Language being tested (see Table 2).

\*\*Not used.

\*\*\*Not described.



**Fig. 2** **a** Distribution of phonotactic probabilities (PP) for words of Italian speech from Hay et al. (2011, Language 2A). **b** Phonotactic probabilities (PP) of words from one sentence (8th) of the same speech



adults' extensive experience with their native language most likely generates a richer phonotactic knowledge when compared with infants'. On the other hand, phonotactic learning begins very early in life, as early as 5 months of age (Sundara & Breiss, 2020), and adults and infants may rely on similar strategies to segment speech (Cunillera et al., 2010). Future studies should extend this investigation to infants, which could provide more insights into the role of PP anchors during language development.

Second, PP differences may only function as anchors when combined with strong TPs. For example, after being familiarized with the Italian languages, infants readily segment words with high TPs (1.0), but not words with low TPs (0.33; Hay et al., 2011; Pelucchi et al., 2009); despite the fact that anchors could have highlighted both types of words. In natural languages, TP values are usually much smaller than the absolute 1.0 used in most segmentation studies (Yang, 2004). Future corpus analyses, experimental studies, and neurophysiological measures (Cunillera et al., 2016) may provide better resolution on the role of varying degrees of TP and PP combination on word segmentation.

Furthermore, although we have focused on PP differences in previously published studies, it may be even more informative for future analyses to explore whether differences in PPs can account for the fact that numerous studies have failed to replicate the original statistical learning findings and thus were not published—file drawer problem (Black & Bergmann, 2017; Rosenthal, 1979).

In sum, we believe that our findings, together with those from Mersad and Nazzi (2011), and the aforementioned observations about PPs in previous segmentation studies, make a reasonable argument for the effects of PP anchors on speech segmentation. In addition, the present study is, to our knowledge, the first to investigate statistical learning with Brazilian-Portuguese speakers, thus providing additional support for the generality of statistical learning research.

Language learning is a complex process. Several statistical cues have now been shown to play a role in speech segmentation. Here, we present evidence that small variations in phonotactic probability can be combined with transitional probability information to impact speech segmentation. Our findings provide a more nuanced understanding of the role PP anchors play in speech segmentation and suggest that, in future studies, we need to consider even subtle differences in PP when selecting stimuli for speech segmentation research.

**Open practices statement** The data and materials for all experiments are available at <https://osf.io/s9thk/>. None of the experiments was preregistered.

**CRedit statement** Rodrigo Dal Ben: Conceptualization, Methodology, Software, Investigation, Writing—Original draft preparation. Débora de Holanda Souza and Jessica F. Hay:

Conceptualization, Resources, Writing—Reviewing and Editing, Supervision.

## Declarations

**Declarations of interest** None.

## References

- Adriaans, F., & Kager, R. (2017). Learning novel phonotactics from exposure to continuous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 8(1), 12. <https://doi.org/10.5334/labphon.20>
- Apel, K., Wolter, J., & Masterson, J. (2006). Effects of phonotactic and orthotactic probabilities during fast mapping on 5-year-olds' learning to spell. *Developmental Neuropsychology*, 29(December 2015), 21–42. <https://doi.org/10.1207/s15326942dn2901>
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological Science*, 9(4), 321–324. <https://doi.org/10.1111/1467-9280.00063>
- Barbosa, F., Pinto, G., Resende, F. G., Gonçalves, C. A., Monserrat, R., & Rosa, M. C. (2003). Grapheme-phone transcription algorithm for a Brazilian Portuguese TTS. *Computational Processing of the Portuguese Language*, 2721, 23–30. [https://doi.org/10.1007/3-540-45011-4\\_4](https://doi.org/10.1007/3-540-45011-4_4)
- Black, A., & Bergmann, C. (2017). Quantifying infants' statistical word segmentation: A meta-analysis. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 124–129). Cognitive Science Society. <https://pdfs.semanticscholar.org/0807/41051b6e2b74d2a1fc2e568c3dd11224984b.pdf>
- Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me. *Psychological Science*, 16(4), 298–304. <https://doi.org/10.1111/j.0956-7976.2005.01531.x>
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33–B44. [https://doi.org/10.1016/S0010-0277\(01\)00122-6](https://doi.org/10.1016/S0010-0277(01)00122-6)
- Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, 87(2), B69–B77. [https://doi.org/10.1016/s0010-0277\(02\)00233-0](https://doi.org/10.1016/s0010-0277(02)00233-0)
- Cunillera, T., Càmarà, E., Laine, M., & Rodríguez-Fornells, A. (2010). Words as anchors: Known words facilitate statistical learning. *Experimental Psychology*, 57(2), 134–141. <https://doi.org/10.1027/1618-3169/a000017>
- Cunillera, T., Laine, M., & Antoni, R.-F. (2016). Headstart for speech segmentation: a neural signature for the anchor word effect. *Neuropsychologia*, 82, 189–199. <https://doi.org/10.1016/j.neuropsychologia.2016.01.011>
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & van der Vrecken, O. (1996). The MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes. *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP*, 3, 1393–1396. <https://doi.org/10.1109/ICSLP.1996.607874>
- Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? *Psychological Science*, 18(3), 254–260. <https://doi.org/10.1111/j.1467-9280.2007.01885.x>
- Estivalet, G. L. (2018, November). Algoritmo de silabação e tonicidade: inventário e frequências das sílabas e estruturas silábicas do português brasileiro [Syllabation and tonicity algorithm: Inventory

- and frequencies of syllables and syllable structures in Brazilian Portuguese]. *XXVII Jornada Internacional Do Grupo de Estudos Linguísticos Do Nordeste—GELNE*. <http://www.jornada27.gelne.com.br/>
- Estivalet, G. L., & Dal Ben, R. (n.d.). *An online calculator to compute phonotactic and orthotactic probability, and phonologic and orthographic neighborhood densities for words and nonwords in Brazilian-Portuguese*. Unpublished instrument.
- Estivalet, G. L., & Meunier, F. (2015). The Brazilian Portuguese Lexicon: An instrument for psycholinguistic research. *PLOS ONE*, *10*(12), 1–24. <https://doi.org/10.1371/journal.pone.0144016>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, *108*(2), 477–499. <https://doi.org/10.1016/j.cognition.2008.04.002>
- Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 84–95. <https://doi.org/10.1037/0278-7393.25.1.84>
- GrafEstes, K., Gluck, S. C.-W., & Grimm, K. J. (2016). Finding patterns and learning words: Infant phonotactic knowledge is associated with vocabulary size. *Journal of Experimental Child Psychology*, *146*, 34–49. <https://doi.org/10.1016/j.jecp.2016.01.012>
- Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*(2), 93–106. <https://doi.org/10.1016/j.cogpsych.2011.06.002>
- Hay, J. F., & Saffran, J. R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, *17*(6), 610–641. <https://doi.org/10.1111/j.1532-7078.2011.00110.x>
- JASP Team. (2018). JASP (Version 0.9) [Computer software]. <https://jasp-stats.org/>
- Johnson, E. K. (2016). Constructing a proto-lexicon: An integrative view of infant language development. *Annual Review of Linguistics*, *2*(1), 391–412. <https://doi.org/10.1146/annurev-linguistics-011415-040616>
- Jusczyk, P. W. (2002). How infants adapt speech-processing capacities to native-language structure. *Current Directions in Psychological Science*, *11*(1), 15–18. <https://doi.org/10.1111/1467-8721.00159>
- Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*(5), 630–645. <https://doi.org/10.1006/jmla.1994.1030>
- Karaman, F., & Hay, J. F. (2018). The longevity of statistical learning: When infant memory decays, isolated words come to the rescue. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(2), 221–232. <https://doi.org/10.1037/xlm0000448>
- Keuleers, E. (2013). *vwr: Useful functions for visual word recognition research*. <https://CRAN.R-project.org/package=vwr> 41.
- Kurumada, C., Meylan, S. C., & Frank, M. C. (2013). Zipfian frequency distributions facilitate word segmentation in context. *Cognition*, *127*(3), 439–453. <https://doi.org/10.1016/j.cognition.2013.02.002>
- MacKenzie, H., Curtin, S., & Graham, S. A. (2012). 12-month-olds' phonotactic knowledge guides their word-object mappings. *Child Development*, *83*(4), 1129–1136. <https://doi.org/10.1111/j.1467-8624.2012.01764.x>
- Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 644–655. <https://doi.org/10.1037/0096-1523.27.3.644>
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*(4), 465–494. <https://doi.org/10.1006/cogp.1999.0721>
- Mersad, K., & Nazzi, T. (2011). Transitional probabilities and positional frequency phonotactics in a hierarchical model of speech segmentation. *Memory & Cognition*, *39*(6), 1085–1093. <https://doi.org/10.3758/s13421-011-0074-3>
- Mersad, K., & Nazzi, T. (2012). When mommy comes to the rescue of statistics: Infants combine top-down and bottom-up cues to segment speech. *Language Learning and Development*, *8*(3), 303–315. <https://doi.org/10.1080/15475441.2011.609106>
- Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*(1), 271–280. <https://doi.org/10.1016/j.cognition.2008.02.003>
- Onishi, K., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), B13–B23. [https://doi.org/10.1016/S0010-0277\(01\)00165-2](https://doi.org/10.1016/S0010-0277(01)00165-2)
- Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, *53*(2), 225–237. <https://doi.org/10.1016/j.jml.2005.02.011>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*(3), 674–685. <https://doi.org/10.1111/j.1467-8624.2009.01290.x>
- R Core Team. (2017). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Reber, R., & Perruchet, P. (2003). The use of control groups in artificial grammar learning. *The Quarterly Journal of Experimental Psychology Section A*, *56*(1), 97–115. <https://doi.org/10.1080/02724980244000297>
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Saffran, J. R. (2020). Statistical Language Learning in Infancy. *Child Development Perspectives*, *14*(1), 49–54. <https://doi.org/10.1111/cdep.12355>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language*, *35*(4), 606–621. <https://doi.org/10.1006/jmla.1996.0032>
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*(2), 101–105. <https://doi.org/10.1111/j.1467-9280.1997.tb00690.x>
- Shoib, A., Wang, T., Hay, J. F., & Lany, J. (2018). Do infants learn words from statistics? Evidence from English-learning infants hearing Italian. *Cognitive Science*, *42*(8), 3083–3099. <https://doi.org/10.1111/cogs.12673>
- Sundara, M., & Breiss, C. (2020). 5-month-olds are sensitive to phonotactic patterns in their native language. *45th Annual Boston University Conference on Language Development*. <https://labphon.org/labphon17/searchable-programme>
- Thiessen, E. D., & 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*(4), 706–716. <https://doi.org/10.1037/0012-1649.39.4.706>
- Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English.

- Behavior Research Methods, Instruments, & Computers*, 36(3), 481–487. <https://doi.org/10.3758/BF03195594>
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmira, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 25(1), 35–57. <https://doi.org/10.3758/s13423-017-1343-3>
- Yang, C. D. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences*, 8(10), 451–456. <https://doi.org/10.1016/j.tics.2004.08.006>
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15(5), 971–979. <https://doi.org/10.3758/PBR.15.5.971>
- Zamuner, T. S. (2009). Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech Language and Hearing Research*, 52(1), 49. [https://doi.org/10.1044/1092-4388\(2008/07-0138\)](https://doi.org/10.1044/1092-4388(2008/07-0138))
- Zipf, G. K. (1965). *Human behavior and the principle of least effort: An introduction to human ecology*. Hafner.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.