BRIEF REPORT

Regularity of unit length boosts statistical learning in verbal and nonverbal artificial languages

L. Hoch · M. D. Tyler · B. Tillmann

Abstract Humans have remarkable statistical learning abilities for verbal speech-like materials and for nonverbal music-like materials. Statistical learning has been shown with artificial languages (AL) that consist of the concatenation of nonsense word-like units into a continuous stream. These ALs contain no cues to unit boundaries other than the transitional probabilities between events, which are high within a unit and low between units. Most AL studies have used units of regular lengths. In the present study, the ALs were based on the same statistical structures but differed in unit length regularity (i.e., whether they were made out of units of regular vs. irregular lengths) and in materials (i.e., syllables vs. musical timbres), to allow us to investigate the influence of unit length regularity on domain-general statistical learning. In addition to better performance for verbal than for nonverbal materials, the findings revealed an effect of unit length regularity, with better performance for languages with regular- (vs. irregular-) length units. This unit length regularity effect suggests the influence of dynamic attentional processes (as proposed by the dynamic attending theory; Large & Jones (Psychological Review 106: 119–159, 1999)) on domain-general statistical learning.

L. Hoch · B. Tillmann (✉)
Lyon Neuroscience Research Center Team Auditory Cognition and
Psychoacoustics CNRS UMR5292, INSERM U1028,
Université Claude Bernard–Lyon I,
50 Av. Tony Garnier,
69366 Lyon Cedex 07, France
e-mail: btillmann@olfac.univ-lyon1.fr

L. Hoch · M. D. Tyler · B. Tillmann
Marcs Institute and School of Social Sciences and Psychology,
University of Western Sydney,
Locked Bag 1797,
Penrith, New South Wales 2751, Australia

Language and music are complex, structured systems of which at least some aspects can be learned via mere exposure (Saffran, 2003; Tillmann, 2005). This implicit learning ability has been investigated in the auditory modality with artificial languages (ALs) that used syllables (Saffran, Newport, & Aslin, 1996) and nonverbal events (Creel, Newport, & Aslin, 2004; Saffran, Johnson, Aslin, & Newport, 1999; Tillmann & McAdams, 2004), and it has been also confirmed in the visual modality (e.g., Fiser & Aslin, 2002). One characteristic under focus is the contribution of transitional probabilities (TPs) to indicate the boundaries of word-like units. The ALs designed by Saffran and collaborators concatenated three-syllable nonsense words into a continuous stream: TPs between two syllables within a word were higher than those between two syllables spanning word boundaries. After exposure to the continuous verbal stream, adults and infants were able to discriminate between words (referred to as *statistical units* hereafter) and nonwords (or part-words), thus suggesting statistical learning abilities (see Saffran, 2003, for a review). These learning abilities have also been shown for music-like ALs using tones (Saffran et al., 1999), sung syllables (Schön, Boyer, Moreno, Besson, Peretz and Kolinsky 2008), and musical timbres (Tillmann & McAdams, 2004), suggesting domain-general implicit learning mechanisms for verbal and nonverbal materials.

Perruchet and Vinter (2002) proposed that statistical learning results from the interaction between the structural regularities of the AL, listeners' attentional focus, and general principles of memory (repetition, influence of interfering events, and forgetting). A continuous stimulus stream is

perceived as a succession of small, disjunctive chunks. The size of the perceived chunks depends on listeners' attentional focus and the stream's acoustic characteristics (e.g., pauses or prosodic cues such as variations in pitch, duration, and intensity). While the memory trace of a previously perceived chunk decreases over time (i.e., simulating interference and forgetting), its trace is reinforced when it reoccurs, and the repetition of chunks thus contributes to progressively shaping the continuous stream into language-relevant units. In addition to repetition, acoustic cues and similarities as well as previously perceived chunks (e.g., Perruchet & Tillmann 2010) influence the emergence of perceptual units (i.e., with the contained information being processed together). These emerging perceptual units guide listeners' attention over time, leading to results that suggest statistical learning.

Event-related potential (ERP) studies further suggested that auditory temporal attention is guided toward the onsets of emerging perceptual units. Initial syllables (i.e., the onsets) of words (Astheimer & Sanders, 2009) and artificial words without acoustic segmental cues (statistical units; Sanders, Newport, & Neville, 2002) presented in a continuous stream elicited an enhanced centro-frontal negativity at around 100 ms (N100) in comparison to within-unit events, suggesting the recruitment of more "orienting" components (see Alcaini, Giard, Thevenet, & Pernier, 1994). This early-onset negativity emerges with exposure and has been reported for verbal (Sanders et al., 2002) and nonverbal (Alba, Katahira, & Okanoya, 2008; Sanders, Ameral, & Sayles, 2009) materials, indicating that it can be elicited by domain-general segmentation mechanisms. Listeners' attention is enhanced in the time window containing unit onsets because (1) a preceding unit just reached closure; (2) the first event (i.e., the onset) of the next unit is rather unpredictable (given the low TPs between the final event [or offset] of a unit and the onset of the following unit); and (3) unit onsets are particularly informative, allowing for the prediction of the next event within the unit (due to high within-unit TPs). When available, listeners' attention can be directed temporally by stress patterns (Toro-Soto & Rodríguez-Fornells, 2007; Tyler & Cutler, 2009), by previously acquired lexical information in natural language (Astheimer & Sanders, 2009), and by chunks that emerge earlier because of some "initial word-likeliness" (e.g., Perruchet & Tillmann, 2010). However, in most AL studies, TPs are the only cue to segmentation and the statistical units are of equal length, reinforced by syllables of equal duration (unlike in natural language). Temporal regularity might thus reinforce the allocation of auditory temporal selective attention in AL learning.

The importance of temporally regular structures for guiding attention over time has been developed in Jones's (1976) theory of dynamic attending. This theoretical framework was developed initially for music processing, but it applies also to general auditory sequencing (Large & Jones, 1999). The dynamic attending theory (DAT) proposes that auditory attention is not equally and continuously distributed over time, but develops with attentional cycles. When listening to an event sequence, external rhythmic cues direct attention periodically and allow listeners to develop temporal expectations about the occurrence of future events, thus facilitating auditory sequencing. Beyond explaining temporal expectancy in music (e.g., Jones & Boltz, 1989), the DAT offers a framework for speech perception: Regular timing between stressed syllables enhances phoneme detection (Quené & Port, 2005) and syntactic processing (Schmidt-Kassow & Kotz, 2009). When applied to ALs with no acoustic cues as to unit boundaries, the DAT—together with the observation that unit onsets benefit from increased attentional resources (Astheimer & Sanders, 2009)—leads us to hypothesize that regular unit onsets guide attention over time and allow for the development of temporal expectations about the next unit onset, thus boosting (and bootstrapping) learning.

The ALs developed by Saffran and collaborators oversimplified natural languages by concatenating units of equal lengths (see also Johnson & Jusczyk, 2003). These units of equal lengths lead to regular unit onsets, while natural languages vary in word length, leading to irregular word onsets. Only a few studies have used units of varying length without additional cues to word boundaries. For infants, the absence of learning with irregular-length units (Johnson & Tyler, 2010) suggests that the use of regular-length units boosts learning.[1] Furthermore, preexposure to units of lengths matching the to-be-segmented units contributes to infants' speech segmentation (Lew-Williams & Saffran, 2012). For adults, learning has been reported for irregular-length units (Tyler & Cutler, 2009), but to our knowledge, no adult study has made direct comparisons of AL learning with units of regular or irregular lengths.

On the basis of the DAT and the absence of learning with irregular units in infants, our study tested the prediction of more effective statistical learning in adults when the units are of regular lengths rather than irregular lengths (i.e., a unit length regularity effect). In addition, we investigated the domain specificity of the influence of temporally driven attentional mechanisms on statistical learning. The previous studies showing statistical learning for music-like materials have all used units of regular lengths (three tones or timbres).

In our study, we manipulated the length regularity of units in verbal speech-like (synthesized spoken syllables)

---

[1] Children do learn ALs of irregular-length units when additional cues are available, such as prosodic cues in infant-directed speech (Thiessen, Hill, & Saffran, 2005) or natural language (Pelucchi, Hay, & Saffran, 2009).

and nonverbal music-like (musical timbres) ALs. For both verbal and nonverbal materials, two ALs were tested: One AL concatenated units of regular lengths, and the other concatenated units of irregular lengths. In all four conditions, the ALs used the same statistical structure: TPs were high between events (i.e., syllables or timbres) within a statistical unit, and low between events spanning unit boundaries. In addition, there were no acoustic cues as to the unit boundaries. Participants were first exposed to one of the four AL conditions and then tested in a two-alternative forced choice task in which they were asked to select statistical units over partial units. Length regularity effects were expected for verbal and nonverbal materials: Statistical learning should be better when the AL is composed of regular-length units rather than irregular-length units. In addition, on the basis of the strong metrical structure of music (e.g., Jones & Boltz, 1989), nonverbal AL learning might be more sensitive to unit length regularity than is verbal AL learning.

Method

Participants

A group of 96 introductory psychology students at the University of Western Sydney (native English speakers) participated for course credit, with 24 in each experimental condition: verbal (V) languages with regular- and irregular-length units (i.e., V-regular and V-irregular) and nonverbal (NV) languages with regular- and irregular-length units (i.e., NV-regular and NV-irregular).

Materials

Verbal ALs Six consonants (/m/, /n/, /f/, /s/, /b/, and /d/) and three vowels (/a/, /i/, and /u/) were combined exhaustively to create 18 consonant–vowel (CV) syllables (e.g., /ma/, /fi/, or /du/). Two AL sets were constructed: one with units of regular length (the V-regular language) and one with units of irregular lengths (the V-irregular language). The V-regular language combined six units of three syllables (using all 18 CV syllables), and the V-irregular language combined three units of two syllables and three units of three syllables (using 15 of the 18 CV syllables). These units were concatenated in a pseudorandom order with the constraint that the same unit never occurred twice in a row. For regular and irregular languages, the units were indicated by high TPs (equal to 1) between syllables within a unit and low TPs (equal to .20) between syllables spanning unit boundaries.

The stream of concatenated units was generated with the MBROLA diphone synthesizer in its totality, without any postediting of the stimuli, at a fundamental frequency of

220 Hz (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996) and with consonant and vowel lengths set to 100 ms and 300 ms,[2] respectively. As in Tyler and Cutler (2009), a female French speaker (fr1) was used because MBROLA's voices do not include Australian English, and the stimulus quality was higher than the US English voices. No acoustic cues as to the unit boundaries (e.g., stress, duration, or pauses) were added, and each phoneme was coarticulated with the following phoneme. No word-like units matched any English word, and no words could be formed across concatenated units in the stream.

The exposure phase contained five blocks of 2 min 24 s apiece, each separated by a 10-s break, for a total duration of about 12 min. In each of the five blocks, each of the six units was presented 20 times, resulting in 120 units per block and 100 presentations of each unit across the exposure phase. A 5-s fade-in and a 5-s fade-out were applied to each block to avoid giving participants access to unit boundaries from the beginning and the end of the stream.

The test phase consisted of a two-alternative forced choice task with 24 pairs. One item of the pair was a statistical unit from the exposed language, and the other item was a partial unit—a sequence of syllables that occurred in the stream across a unit boundary. For irregular languages, the test pairs consisted only of units and partial units of the same length.

To control for general perceptual biases, (1) each condition included four different ALs (L1–L4) with the same statistical structure (i.e., units made out of the same events and presented without direct repetition of a unit in the exposure stream), aiming to introduce a control proposed by Reber and Perruchet (2003), and (2) those ALs were constructed by pair (see Saffran et al., 1999), so that the statistical unit of one AL was the partial unit of the other, and vice versa (this was counterbalanced across participants). That is, for L1 and L2 (i.e., the first AL pair), participants completed the same test phase after exposure to either AL, but the correct response for each test pair changed as a function of the exposure language. The same procedure was applied to L3 and L4 (i.e., the second AL pair). Participants were randomly allocated to one of the four languages ($n = 6$). Across the test phase, each statistical unit was tested against two different partial units, resulting in 12 test pairs. These pairs were repeated twice to counterbalance the presentation order of statistical and partial units. The resulting 24 test pairs were presented in pseudorandom order, such that the statistical unit was not in the same

---

[2] Although previous AL studies have used shorter syllables (less than 300 ms), we used longer syllables to match the duration of the musical timbres of the nonverbal AL.

position in the pair more than four items in a row. For each test pair, the two items were separated by an interstimulus interval of 250 ms.

Nonverbal ALs The nonverbal regular and irregular ALs were based on the same construction as the verbal ALs, but each of the CV syllables was replaced with one synthesized musical timbre. The musical timbres (previously used in Tillmann & McAdams, 2004, and originally from McAdams, Winzberg, Donnadieu, deSoete, & Krimphoff, 1995) were played at 311 Hz (i.e., Eb4). To match the timbres and syllables in duration, these timbres were shortened to 400 ms by removing 100 ms from the steady-state part of the timbre (using Praat software), while the attack, resonance, and timbre envelope of the timbre were preserved, thus resulting in no audible alteration (http://olfac.univ-lyon1.fr/bt-sound.html for sound examples). Following Tillmann and McAdams (2004, language S3), the units of L1, L2, L3, and L4 were constructed in such a way that the timbral distances within a unit and across unit boundaries did not differ (in terms of the average similarity ratings of McAdams et al., 1995). Thus, timbral distances between timbres could not serve as a cue to unit boundaries.

Procedure

For the exposure phase, participants were asked to pay attention to the continuous stream (verbal or nonverbal). They were informed that they would answer questions about the stream after listening to it. After exposure, participants were told that, within the continuous stream, there were word-like units of syllables (i.e., nonsense words) or timbres. Participants who had been exposed to a V-regular (or NV-regular) language were informed that the word-like units contained three syllables (or timbres), while participants who had been exposed to a V-irregular (or NV-irregular) language were informed that the word-like units could contain either two or three syllables (or timbres).

In the test phase, participants were asked to decide which one of the two items (i.e., one statistical unit and one partial unit) was a unit that they had previously heard (also introduced as a "word-like unit") in the stimulus stream. They answered by pressing "1" for the first item or "2" for the second item on the computer keyboard. If they were unsure, they were encouraged to guess. The test phase started with one practice trial consisting of two novel items made out of combinations of syllables (or timbres) that did not occur in the exposure phase. Participants were told that there was no correct answer for the practice trial, as it only aimed to show the organization of a trial. The experiment lasted for about 20 min.

Results

Choosing the statistical unit over the partial unit was scored as a correct response. Accuracy (Fig. 1) was analyzed using a 2 × 2 ANOVA with Unit Length Regularity (regular vs. irregular) and Materials (verbal vs. nonverbal) as between-participants factors. The main effect of unit length regularity was significant, $F(1, 92) = 7.33$, $p = .008$, $MSE = 131.47$, $\eta_p^2 = .074$: Languages combining units of regular lengths were learned better than languages combining units of irregular lengths. The main effect of materials was also significant, $F(1, 92) = 6.17$, $p = .015$, $MSE = 131.47$, $\eta_p^2 = .063$: Verbal languages were learned better than nonverbal languages. The two-way interaction between unit length regularity and materials was not significant ($p = .684$). Performance did not differ for two-versus three-event units for syllables (55 % vs. 55 %; $p = .946$) or timbres (51 % vs. 49 %; $p = .670$).

Performance was significantly above chance (50 %, one-tailed t tests) for V-regular, $t(23) = 4.84$, $p < .001$, NV-regular, $t(23) = 2.59$, $p = .008$, and V-irregular languages, $t(23) = 1.88$, $p = .036$, but not different from chance for NV-irregular ones, $t(23) = 0.07$, $p = .474$.

As there was a smaller number of events used for the irregular than for the regular condition, we ran a new irregular-length condition with two-, three- and four-event units (two of each) for syllables and timbres, thus reaching an average unit length of three events and using all 18 events (as for the regular condition). This new material was tested (same procedure as described above) with another 48 introductory psychology students (24 for syllables and timbres, respectively), who scored 58 % for syllables and 47 % for timbres. We ran a 2 × 2 ANOVA, with unit length regularity and materials as between-participants factors, comparing the original regular-length condition to the new irregular-length condition. This analysis confirmed the results here above, notably a main effect of unit length regularity, $F(1, 92) = 7.82$, $p = .006$, $MSE = 119.91$, $\eta_p^2 = .078$ (better performance for regular-length than for irregular-length units), a main effect of
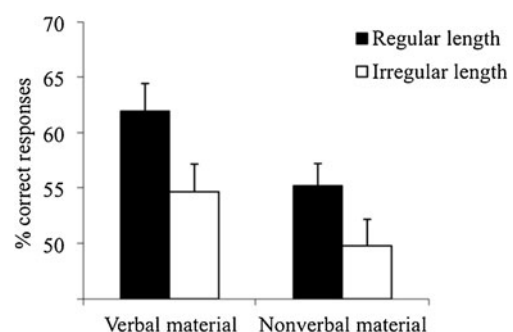

Fig. 1 Percent correct responses presented as a function of unit length regularity (regular vs. irregular) and materials (verbal vs. nonverbal). Error bars represent between-participants standard errors of the means

materials, $F(1, 92) = 16.95$, $p < .001$, $MSE = 119.91$, $\eta_p^2 = .156$, and no interaction $(p = .28)$.[3]

Discussion

In our study, we investigated a unit length regularity effect on statistical learning of verbal and nonverbal ALs. The results revealed better learning for languages with regular-length units than for languages with irregular-length units and better learning for verbal than for nonverbal materials.[4] The observed unit length regularity effect did not interact with materials.

Without being informed about the AL, participants were able to learn statistical regularities of the verbal and nonverbal ALs with regular-length units. This finding replicates the statistical learning of ALs with regular-length units in adults that has previously been observed for verbal (Saffran et al., 1996) and nonverbal (Tillmann & McAdams, 2004) materials. While previous studies using nonverbal materials pitted statistical units against nonunits (new event sequences; Tillmann & McAdams, 2004, Exp. 1) or partial units that did not occur in the stream (Saffran et al., 1999; Tillmann & McAdams, 2004, Exp. 2), our study pitted statistical units against partial units that occurred in the stream across unit boundaries, as was previously done for verbal materials (e.g., Tyler & Cutler, 2009). Hence, the above-chance performance for nonverbal languages with regular-length units indicates more refined nonverbal statistical learning than has previously been shown, as well as supporting the hypothesis of general implicit-learning mechanisms.

Most importantly, our study showed increased learning for languages with regular-length as compared to irregular-length units, even though all languages shared the same statistical properties regarding TPs. These findings are difficult to explain from a purely computational approach, but rather support a chunk-based approach, which includes attentional mechanisms (Perruchet & Pacton, 2006; Perruchet & Tillmann, 2010). On the basis of the DAT, we predicted that learning artificial languages of regular-length units might benefit from listeners' attentional cycles. In our study,

the attentional cues were not based on acoustic changes (like stress in speech; Quené & Port, 2005; Schmidt-Kassow & Kotz, 2009), but emerged during the learning process, as suggested by the onset negativity that has been observed in ERP studies (Alba et al., 2008; Sanders et al., 2009). Over AL exposure, listeners progressively learn the high TPs within units and can anticipate when a unit will end and when a new one will start. Even though the identity of the next onset is not predictable, its processing benefits from the enhanced attentional resources in this temporal window. Together with the special status of unit onsets in segmentation (e.g., Sanders et al., 2002), emerging attentional cues might be reinforced by the regular-length units (leading to regularly timed onsets/offsets). As predicted by the DAT, the regular units created temporal regularities that may have guided attention over time and allowed for the development of expectations about the temporal occurrence of the next onset (Large & Jones, 1999), thus reinforcing the formation of perceptual units (Perruchet & Vinter, 2002). In ALs with regular-length units, the dynamic attentional mechanisms, together with the emergence of perceptual units corresponding to statistical units, might thus boost statistical learning. In contrast, in ALs with units of irregular lengths, no temporal regularities drive attention over time, and basic statistical learning abilities cannot benefit from dynamic attentional mechanisms, thus leading to lower performance. It is interesting to note that in AL studies, the onset-based regularity is reinforced by the use of events of equal duration. Our study points out that to investigate statistical learning without the potential influence of attentional boosts linked to regularity, future studies should use ALs composed of temporally irregular unit onsets that are based on irregular unit lengths (as in the irregular language of this study) and/or of events of varying duration.

---

[3] To investigate the potential random effects of participants and items, we ran binomial probit mixed-model analyses with additive random item and participant effects. These analyses confirmed the two main effects (ps < .001) and the nonsignificant interaction (ps > .29), for the findings obtained with both types of irregular languages in comparison to the regular languages.

[4] The main effect of materials might be due to (1) more familiarity with syllables than with synthetic musical timbres (similarly, Gebhart, Newport, & Aslin, 2009, reported more difficult learning with nonlinguistic noise); (2) mental rehearsal for syllables that is not available or is less efficient for timbres (requiring additional labeling); and/or (3) an interfering influence of perceptual grouping other than the ones controlled here (Creel et al., 2004; Gebhart et al., 2009).

References

Alba, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience, 20,* 952–964.

Alcaini, M., Giard, M. H., Thevenet, M., & Pernier, J. (1994). Two separate frontal components in the N1 wave of the human auditory evoked response. *Psychophysiology, 31,* 611–615.

Astheimer, L. B., & Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological Psychology, 80,* 23–34.

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of non-adjacent dependencies in tone

sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1119–1130.

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. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 1393–1396). Philadelphia: ICSLP.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99,* 15822–15826.

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among non-linguistic sounds. *Psychonomic Bulletin & Review, 16,* 486–490. doi:10.3758/PBR.16.3.486

Johnson, E. K., & Jusczyk, P. W. (2003). Exploring statistical learning by 8-month-olds: The role of complexity and variation. In *Jusczyk Lab Final Report*. Retrieved from http://hincapie.psych.purdue.edu/Jusczyk

Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science, 13,* 339–345.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention and memory. *Psychological Review, 83,* 323–355.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106,* 119–159.

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition, 122,* 241–246.

McAdams, S., Winzberg, S., Donnadieu, S., deSoete, G., & Krimphoff, J. (1995). Perceptual scaling of synthetized musical timbres: Common dimensions, specificities and latent subject classes. *Psychological Research, 58,* 177–192.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development, 80,* 674–685. doi:10.1111/j.1467-8624.2009.01290.x

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233–238. doi:10.1016/j.tics.2006.03.006

Perruchet, P., & Tillmann, B. (2010). Exploiting multiple sources of information in learning an artificial language: Human data and modeling. *Cognitive Science, 34,* 255–285. doi:10.1111/j.1551-6709.2009.01074.x

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *The Behavioral and Brain Sciences, 25,* 297–388.

Quené, H., & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica, 621,* 1–13.

Reber, R., & Perruchet, P. (2003). The use of control groups in artificial grammar learning. *Quarterly Journal of Experimental Psychology, 56A,* 97–115.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12,* 110–114.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70,* 27–52.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35,* 606–621.

Sanders, L. D., Ameral, V., & Sayles, K. (2009). Event-related potentials index segmentation of nonsense sounds. *Neuropsychologia, 47,* 1183–1186.

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: An event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience, 5,* 700–703.

Schmidt-Kassow, M., & Kotz, S. A. (2009). Attention and perceptual regularity in speech. *NeuroReport, 20,* 1643–1647.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition, 106,* 975–983.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7,* 53–71. doi:10.1207/s15327078in0701_5

Tillmann, B. (2005). Implicit investigations of tonal knowledge in nonmusician listeners. *Annals of the New York Academy of Sciences, 1060,* 100–110.

Tillmann, B., & McAdams, S. (2004). Implicit learning of musical timbre sequences: Statistical regularities confronted with acoustical (dis)similarities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1131–1142.

Toro-Soto, J. M., & Rodríguez-Fornells, A. (2007). Stress placement and word segmentation by Spanish speakers. *Psicológica, 28,* 167–176.

Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America, 126,* 367–376.