Short Research Article

Rapid Serial Auditory Presentation

A New Measure of Statistical Learning in Speech Segmentation

Ana Franco,[1,2] Julia Eberlen,[1] Arnaud Destrebecqz,[1] Axel Cleeremans,[1] and Julie Bertels[1,3]

[1]Cognition, Consciousness and Computation Group, Université Libre de Bruxelles, Belgium, [2]Fonds de la Recherche Luxembourg – FNR, Luxembourg, [3]Fonds de la Recherche Scientifique – FNRS, Belgium

Abstract. The Rapid Serial Visual Presentation procedure is a method widely used in visual perception research. In this paper we propose an adaptation of this method which can be used with auditory material and enables assessment of statistical learning in speech segmentation. Adult participants were exposed to an artificial speech stream composed of statistically defined trisyllabic nonsense words. They were subsequently instructed to perform a detection task in a Rapid Serial Auditory Presentation (RSAP) stream in which they had to detect a syllable in a short speech stream. Results showed that reaction times varied as a function of the statistical predictability of the syllable: second and third syllables of each word were responded to faster than first syllables. This result suggests that the RSAP procedure provides a reliable and sensitive indirect measure of auditory statistical learning.

Keywords: speech segmentation, statistical learning, indirect measures, RSAP

When learning a foreign language, the first challenging step of the process consists of segmenting the continuous speech. Although we can easily extract the words in written language because there are blank spaces between them, spoken language is arranged in a continuous speech flow and words are not separated by clear cues (Klatt, 1980). Because the pauses in the speech flow are not informative enough to allow its segmentation, one needs to find other speech cues to rely on.

An important speech segmentation cue consists of the processing of the statistical information in the language. These statistical cues exist at several levels of the language; at the basic level, transitional probabilities (TPs) between adjacent syllables may be used (Jusczyk, Houston, & Newsome, 1999). TPs are defined as the probability of Y given X. This is computed by taking into account the frequency of XY/frequency of X in a given language. It can be shown that the probability of one syllable following another is higher within words than between words. For example, in the sentence "pretty baby," the transitional probabilities between "pre" and "tty" and between "ba" and "by" are higher than the transitional probability between "tty" and "ba" (Saffran, Aslin, & Newport, 1996).

It is well established that human beings tend to associate the syllables with high transitional probabilities and separate those with low probabilities, creating word candidates from the continuous flow of speech. Saffran, Newport, Aslin, Tunick, and Barrueco (1997) demonstrated this phenomenon in adults and 6- to 7-year-old children. In their study, six nonsense words were created by randomly assembling three of a set of 12 syllables to form each word. The words were concatenated into a continuous speech produced by a speech synthesizer, so that there were no pauses or any other acoustic or prosodic cues of word boundaries. Only the transitional probabilities between syllable pairs were informative: they were higher within words (ranging between 0.3 and 1.0) than across word boundaries (ranging between 0.1 and 0.2). After 20 min of exposure to the continuous speech, the knowledge of the participants was assessed. On each trial, two sounds were presented: one was a word from the language they were exposed to, while the other was composed by the same set of syllables but had never occurred during the exposure phase. Participants were asked to judge which one of the sounds resembled more to what they had heard. Results show that both adults and children performed above chance level in this task, indicating that they were able to extract words from the continuous speech using statistical information. These results were also observed in very young children (Pelucchi, Hay, & Saffran, 2009; Saffran et al., 1996; Thiessen & Saffran, 2003) and widely replicated with adults under various other conditions (Romberg & Saffran, 2010 for a review).

In most of the studies, statistical learning of an artificial language is measured using a two alternative forced choice (2AFC) task, as in Saffran et al. (1997), or a simple recognition task (Abla, Katahira, & Okanoya, 2008). Yet, these tasks are direct measures of learning, namely the instructions explicitly require that participants use acquired

DOI: 10.1027/1618-3169/a000295

knowledge. However, in some cases, direct tasks might not be sensitive enough to show the acquisition of knowledge (e.g., Shanks & St. John, 1994), especially in incidental situations such as language acquisition. Moreover, in both types of tasks, the (non)words are usually presented in isolation, and participants are asked to decide on their regularity. However, these presentation conditions differ from those in which these words were potentially learned (i.e., the conditions of the exposure phase). In addition, these presentation conditions might seem "unecological" given that words are generally presented in continuous speech (i.e., in a succession of words) rather than in isolation. The goal of the present study is to present an alternative measure of statistical learning in speech segmentation situations, which we believe to be more sensitive toward learning and more ecological.

More precisely, we propose to use rapid serial auditory presentation (RSAP) as an indirect measure of statistical learning. This was recently done by a number of studies (Bertels, Boursain, Destrebecqz, & Gaillard, 2014; Bertels, Demoulin, Franco, & Destrebecqz, 2013; Bertels, Franco, & Destrebecqz, 2012; Kim, Seitz, Feenstra, & Shams, 2009; Turk-Browne, Jungé, & Scholl, 2005; Turk-Browne, Scholl, Johnson, & Chun, 2010) in the visual modality. For example, in Bertels et al. (2012), participants were exposed to a continuous stream consisting of visual shapes made of the repeated presentation of four triplets (i.e., sequences of three shapes presented successively in a fixed order). Afterwards, participants were asked to detect a target shape on each trial in a rapid serial visual presentation (RSVP) stream consisting of one presentation of the four triplets they were exposed to, one shape at time. The rationale was that if participants learned the regularities, they would detect the second and third item of each triplet faster than the first one because these were reliably predicted by the preceding item. Coherently, results revealed shorter reaction times for the predictable items (i.e., second and third ones) compared to the unpredictable ones (i.e., first item of each triplet).

Similarly, in an artificial language stream, such as the ones used in statistical learning paradigms, the first syllable of each word predicts the second and the third. The idea of the present study is that if participants are able to extract the words from the continuous speech they were exposed to, we should observe faster reaction times for the second and third syllables of each word when compared to the first syllable in a subsequent RSAP task.

Method

Participants

Thirty monolingual French-speaking undergraduate psychology students (21 women; mean age = 21.13, *SD* = 5.45) were included in this study and received course credits for their participation. None reported hearing problems. The experiment was approved by the ethics committee of the Université Libre de Bruxelles.

Material

Two artificial speech streams were generated using the MBROLA speech synthesizer (Dutoit, Pagel, Pierret, Bataille, & Van der Vreken, 1996) with the French male diphone database fr1, with a sampling frequency of 16 KHz. Each stream contained four nonsense words based on the same corpus of syllables. Language A contained *bamoli*, *kobite*, *vemachu*, and *tichalu*. Language B contained *techabi*, *komati*, *lumoba*, and *velichu*. Four other nonsense words were created in order to be presented as nonwords during the test phase: *motecha*, *baluti*, *liveko*, and *chubima*. These nonwords were recombinations of the same syllables presented during exposure. These recombinations were built so that the transitional probabilities between syllables was 0. Each syllable lasted 200 ms. There was a 30 ms pause between each syllable, regardless of being within a word or between two words. Each participant was randomly assigned to either Language A or B. Stimulus presentation, timing and data collection were controlled using Psyscope X software and the Psyscope USB button box (Cohen, MacWhinney, Flatt, & Provost, 1993) in combination with a Mac mini 2.4 GHz Intel Core 2 Duo.

Procedure

Up to three participants were tested simultaneously in the same dimly lit room. Each participant sat in an experimental booth in front of a computer screen. They performed the experiment wearing soundproof headsets in order to be isolated from external noise. The experiment started by an auditory presentation of all 12 syllables presented in a randomized order (none of the words or nonwords was presented here). The purpose of this presentation was to familiarize participants with the different syllables composing the artificial language. Then, during the exposure phase, they listened to the language. Participants were instructed to pay attention to the speech stream spoken in an "unknown language" and to extract the words from the speech. The speech stream of the exposure phase consisted of 100 presentations of every word in a pseudorandom order: the same word never occurred twice in succession. There were no other speech cues than the transitional probabilities between syllables, 1.0 for the within-word transitions and 0.33 for the between-word transitions. The exposure phase lasted for 5 min and was immediately followed by the RSAP test. Participants were instructed to detect a target syllable in a nonsense speech stream. The target syllable, one of the 12 syllables presented during the exposure phase, was presented once and was followed by the stream of syllables. This stream consisted of a random presentation of the four words, at the same rate as during the exposure phase. Participants were asked to press a key on the button box as soon as they heard the target. The RSAP was then stopped, and the next target was presented. Each target syllable was presented six times – in the first, second, or third position of the second or third triplet in the RSAP stream – resulting in 72 trials. Finally, participants performed the 2AFC task. For each test item, participants heard two

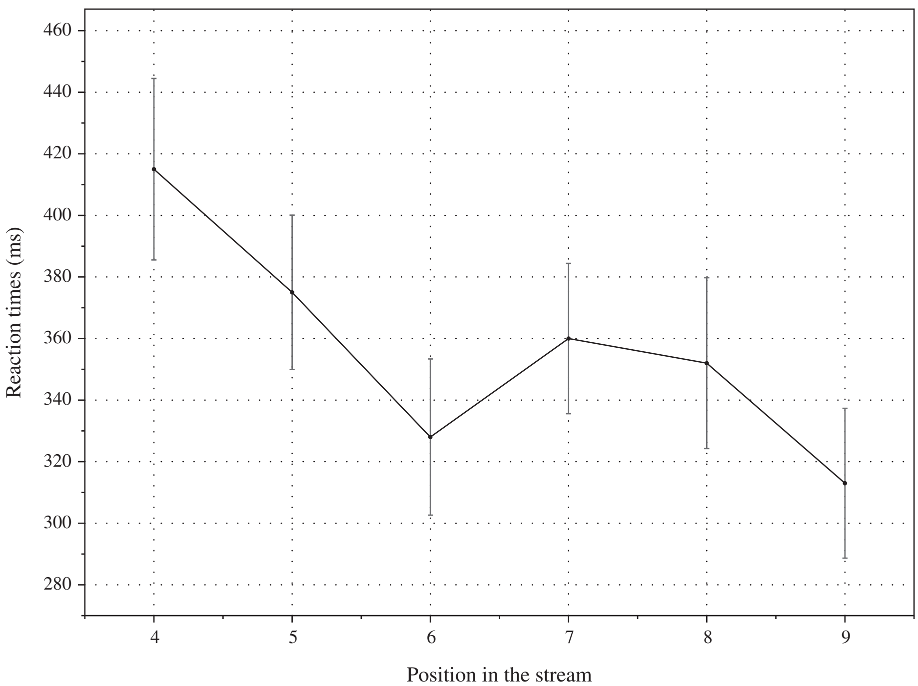

Figure 1. Mean detection latencies for the three item positions (Positions 1, 2, and 3) in the rapid serial auditory presentation (RSAP) task. Error bars represent 95% confidence intervals.

trisyllabic strings, separated by 500 ms of silence. One of the strings was a word from the nonsense language, while the other was not. Participants were instructed to report which one of the stimuli sounded more like the language they had previously heard. The test items were constructed by pairing the four words of the language with each one of the four nonwords. Each word was paired exhaustively with each nonword in any order, rendering 32 trials presented in a pseudo-random order, and the same word was never presented twice in succession. The 2AFC task served two purposes. First, it verified that any absence of effect in the RSAP task truly reflects an absence of learning, in which case performance in the 2AFC task would be at chance. Second, it tested the possibility that the RSAP task was not sensitive enough to detect participants' acquired knowledge.

Results

Analyses were only performed on correct reaction times (RTs). Trials without response (on average 2.5%) and RTs longer than 1,000 ms or shorter than 100 ms (on average 5.5%) were excluded from the analyses. A repeated measures analysis of variance (ANOVA) was applied on RTs, with Triplet (2 levels: the target could be presented in the second or third triplet of the RSAP stream) and Position (3 levels: the target could be the first, second, or third syllable of a word) as within-subject factors and Language (2 levels: A, B) as a between-subjects factor. Figure 1 shows the average RTs for each position by triplet.

We observed a significant effect of Triplet, $F(1, 56) = 68.675$, $p < .001$, $\eta_p^2 = .710$, indicating that mean RTs were shorter when the target was presented in the third triplet (350 ms) than in the second triplet (373 ms). Most likely, this speeding up mirrors participants' expectancy. The more they wait for the appearance of the target, the more they expect that it will occur next (Perruchet, Cleeremans, & Destrebecqz, 2006).

A significant effect of Position was also found, $F(2, 56) = 67.957$, $p < .001$, $\eta_p^2 = .708$. In accordance with our predictions, Bonferroni adjusted comparisons revealed that RTs in Positions 2 and 3 (i.e., in predictable positions, 362 and 320 ms, respectively) were significantly shorter than RTs in Position 1 (388 ms), both $p < .001$. Also, RTs in Position 3 were significantly shorter than RTs in Position 2, $p < .001$.

The interaction between Triplet and Position, $F(2, 56) = 9.055$, $p < .001$, $\eta_p^2 = .244$, indicated that RTs for the three positions differed significantly in the second triplet (RT_{1_2}: $t(29) = 6.373$, $p < .001$; RT_{1_3}: $t(29) = 10.471$, $p < .001$ and RT_{2_3}: $t(29) = 6.655$, $p < .001$). In the third triplet, the differences between Positions 1 and 3 as well as between 2 and 3 were significant (RT_{1_3}: $t(29) = 7.220$, $p < .001$ and RT_{2_3}: $t(29) = 5.085$, $p < .001$, respectively). The difference between the Positions 1 and 2 was not significant, although in the predicted direction, $p > .1$. The speeding up of participants' responses along the course of the stream may have diminished the effect of the syllable predictability.

Since the statistical predictability of the syllables is confounded with their position in the stream (i.e., more predictable syllables are presented later in the stream), one might wonder whether the speeding up of participants' RTs along the course of the stream (both intra- and inter-triplets) might simply reflect participants' position-based expectancies. However, if this were true, RTs would decrease

Table 1. Means and standard deviations for RT differences between Positions 1 and 3 across the RSAP trials

	Part 1	Part 2	Part 3	Part 4
Mean	66.27	54.09	62.36	95.72
SD	11.55	6.99	9.68	11.45

linearly from one stream position to the next. However, we observed a significant RT increase between the sixth (i.e., the last position of the first triplet in which the target can be presented, thus being highly predictable) and the seventh position (i.e., the first position of the second triplet in which the target can be presented, thus being less predictable although in a later position in the stream), $t(29) = -4.371$, $p < .001$ (see Figure 1).

Neither the effect of Language nor its interaction with the other two variables were significant, all $p > .1$ and $F < 1$. Consequently, in the following analyses, results from both languages were pooled together. Moreover, because the RT difference between Positions 1 and 3 is the most robust, in the following analyses we will focus on this facilitation effect as a measure of learning in the RSAP task.

Since the RSAP test was composed of 72 trials, we cannot exclude that participants pursued to learn the triplets during the task, especially because the RSAP trials began and ended with the presentation of a word unit, although the latter only occurred in case participants did not respond to a trial (see Method). This strong nonstatistical segmentation cue at the beginning of each RSAP trial could therefore allow participants to increase their learning of the words during the task. In order to assess the evolution of the facilitation effect and to ensure that it was already present at the early RSAP trials, we divided the trials in four parts (each part composed of 18 trials). Means and standard deviations of the facilitation effects (i.e., RT differences between Positions 1 and 3) across the RSAP trials (i.e., across the four parts of the task) are presented in Table 1.

A repeated measures ANOVA on the RT difference between Positions 1 and 3 with Part (1, 2, 3, and 4) as a within-subject factor revealed a significant effect of this factor, $F(3, 87) = 3.639$, $p < .05$, $\eta_p^2 = .111$. Facilitation effects were larger in Part 4 than in Part 2, $t(29) = 3.328$, $p < .005$, supporting the idea that participants pursued learning of the triplets across the RSAP trials. None of the other comparisons reached significance, $p > .1$.

Nevertheless, in order to ensure that participants learned the triplets during exposure, namely that the facilitation effect (RT_{1_3}) was already present in the first trials of the RSAP test, the first part was analyzed in isolation. A one-sample *t*-test showed that mean RT_{1_3} was significantly higher than zero, $t(29) = 5.743$, $p < .001$.

These results thus support that participants extracted the words from the continuous speech stream during the exposure phase. Unsurprisingly, results from the 2AFC task support this statement: the overall recognition performance was 74.70%, which differed significantly from a chance-level of 50%, $t(29) = 8.224$, $p < .001$.

Since both the RSAP and the 2AFC tasks aimed at measuring any knowledge of the triplets, the two measures might be correlated. In order to check for such a relationship, we ran correlational analyses of participants' performance on both tasks. The general accuracy of the 2AFC task and the RT_{1_3} of the RSAP task were not correlated ($r(30) = .247$; $p = .188$). This result was somehow expected. Extra exposure to the artificial language – combined with the presentation of a nonstatistical segmentation cue – is given in the RSAP task. It is thus very likely that accuracy on the 2AFC task does not exclusively reflect learning during exposure, but also during the RSAP test itself. Moreover, and despite possible additional learning during the RSAP test, it should be noted that both tasks measure related but different aspects of the same material. While the RSAP reflects sensitivity to the transitional probabilities present in the speech stream, success on the 2AFC task requires knowledge of the artificial language units. This will be discussed further in the following section.

Discussion

Identifying word boundaries in continuous speech seems to be an easy task for native adult listeners. However, the problem of word segmentation can be very challenging for infants or for adults listening to a foreign language. In a series of studies, Saffran and colleagues (Saffran et al., 1996; Saffran, Newport, & Aslin, 1996) reported that when exposed to an artificial speech stream, both infants and adults are able to extract words by tracking the transitional probabilities between syllables. We thus seem to be able to use statistical information present in continuous speech. Later studies showed that such statistical computations are modality general (Kirkham, Slemmer, & Johnson, 2002).

Despite the attention that statistical learning has received, most studies assessed speech segmentation using the standard 2AFC task or recognition tests, namely direct measures of learning. One exception is a recent study (Gómez, Bion, & Mehler, 2011; but see Franco, Gaillard, Cleeremans, & Destrebecqz, 2014) that proposed an adaptation of the click detection paradigm (Fodor & Bever, 1965) as an online measure of statistical computations. In the visual domain, although direct measures have also been widely used, recent studies explored statistical learning using an indirect measure, the rapid serial visual presentation task (Bertels et al., 2012; Kim et al., 2009). This measure has the advantage to be (1) more sensitive to any acquired knowledge than a direct measure of learning (e.g., Shanks & St. John, 1994), and (2) more ecological, since it displays the test items in the same conditions as those they were learned in and as those they were usually heard in. Thus, one might want to explore whether this technique could be adapted to the auditory modality. This was the purpose of the present study.

Mirroring Bertels et al.'s (2012) experimental design, we first exposed participants to an artificial language made of the repeated presentation of four trisyllabic words and assessed learning of these words by asking participants to detect specific syllables in RSAP streams. The rationale was that the statistical relationships between syllables (high vs. low transitional probabilities) cause each syllable to become either highly predictable (i.e., the second and third syllables of each word) or less predictable (i.e., the first syllable of each word). Thus, if participants can track transitional probabilities and extract the words from the continuous speech, the detection of the predictable syllables should be faster than the detection of the unpredictable ones. This is indeed what we observed, thereby showing that the statistical information present in the continuous speech was processed. Replicating results from the visual modality, our results thus demonstrate that the RSAP is an interesting alternative to assess auditory statistical learning.

As mentioned above, when comparing the discrimination tasks usually used, the RSAP has the advantage of being more ecologically valid. It presents the words in a familiar context, like in real language. The fact that the syllables have to be detected in a continuous speech identical to the exposure phase creates a situation more comparable to the learning phase but also to the situation of natural language learning.

Another benefit of the RSAP method is that unlike the conventional methods used to measure statistical learning, there is no need to present nongrammatical elements which could reduce evidence of learning. This advantage however is linked to the main weakness of this method: Because the streams presented during the RSAP task were composed of words from artificial language – followed by each stream's beginning with the presentation of the first syllable of a word – participants are provided with a strong nonstatistical segmentation cue that allows them to increase their learning of the words during the task. Results show that even though the facilitation effect for the predictable syllables is already present at the beginning of the RSAP task, it tends to increase gradually along the blocks. This limitation of the RSAP in its current form could however be easily solved. In future studies, the RSAP streams should start at random positions of the words (i.e., first, second, or third syllable), rather than systematically at the first syllable. This will avoid providing a strong nonstatistical segmentation cue that clearly has an impact on learning.

Consistent with Kim et al. (2009)'s study that used a comparable experimental design, we did not observe any significant correlation between performance in the RSAP and in the 2AFC tasks. Given that both tasks aim to measure statistical knowledge learned in a speech segmentation situation, one might expect these measures to be linked to each other. However, while the performance in the 2AFC task can be interpreted in gradual terms – greater acquired knowledge results in a better performance – RT differences in the RSAP task should be interpreted in a dichotomic way: a positive RT_{1_3} difference reflects a sensitivity to the statistical information, while a negative or null RT_{1_3} difference should be interpreted as an insensitivity to the statistical information. In other words, unlike the 2AFC task, a large RT_{1_3} difference would not necessarily be associated with greater knowledge about the artificial language. In light of these considerations, the lack of correlation between the RSAP and the 2AFC tasks in the present study and in Kim et al. (2009)'s study is not surprising. Following Shanks and Perruchet (2002), this pattern of results perfectly fits with a model in which performance in both tasks depends on exactly the same underlying knowledge source, with the dissociations between tasks reflecting subtle differences between the retrieval processes they recruit. As mentioned above, it is very likely that these tasks reflect different outcomes of learning (see Karuza, Emberson, & Aslin, 2013 for a theoretical review). While the RSAP task captures the extent to which participants are sensitive to the statistical information, performing above chance in the 2AFC task also requires the actual extraction and memorization of word units.

Finally, and most importantly, this method allows an indirect assessment of learning. The terms "direct" and "indirect" are used to characterize two types of memory measure. Direct tasks are defined as those in which the instructions of the test make reference to the relevant discrimination. In contrast, indirect measures do not make any reference to the relevant discrimination. Crucially, indirect measures would be more sensitive to any acquired knowledge, even allowing detection of the acquisition of knowledge about which the participant lacks metaknowledge (i.e., he/she does not know that he/she possesses this knowledge). Indeed, although direct and indirect measures should exhibit equal sensitivity to consciously held task-relevant information, direct measures would not be sensitive enough when the information is not fully conscious (Reingold & Merikle, 1988). Yet, a growing body of research tends to show that statistical learning involves both conscious and unconscious knowledge (Bertels et al., 2012; Franco, Cleeremans, & Destrebecqz, 2011; Kim et al., 2009; Perruchet & Pacton, 2006). Therefore, we believe that an indirect measure is more appropriate to assess statistical learning. This is particularly relevant when the acquisition of conscious knowledge is made difficult by training conditions or task demands, for example when the exposure is brief or when the regularities are difficult to perceive, so that a quantitative dissociation between direct and indirect measures will be observed.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at http://dx.doi.org/10.1027/1618-3169/a000295

ESM 1. Raw data (sav).
Raw data of the Experiment.

References

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

Bertels, J., Boursain, E., Destrebecqz, A., & Gaillard, V. (2014). Visual statistical learning in children and young adults: How implicit? *Frontiers in Psychology, 5*, 1541. doi: 10.3389/fpsyg.2014.01541

Bertels, J., Demoulin, C., Franco, A., & Destrebecqz, A. (2013). Side Effects of Being Blue: Influence of Sad Mood on Visual Statistical Learning. *PLoS One, 8*, e59832. doi: 10.1371/journal.pone.0059832.t003

Bertels, J., Franco, A., & Destrebecqz, A. (2012). How implicit is visual statistical learning? *Journal of Experimental Psychology: Learning, Memory and Cognition, 38*, 1425–1431.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using MacIntosh computers. *Behavior Research Methods, Instruments, & Computers, 25*, 257–271.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Van der Vreken, O. (1996). The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. In *Proc. ICSLP'96* (Vol. 3, pp. 1393–1396). Philadelphia, PA: IEEE.

Fodor, J. A., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4*, 414–420.

Franco, A., Cleeremans, A., & Destrebecqz, A. (2011). Statistical learning of two artificial languages presented successively: How conscious? *Frontiers in Psychology, 2*, 229. doi: 10.3389/fpsyg.2011.00229

Franco, A., Gaillard, V., Cleeremans, A., & Destrebecqz, A. (2014). Assessing segmentation processes by click detection: Online measure of statistical learning, or simple interference? *Behavior Research Methods*. Advance online publication. doi: 10.3758/s13428-014-0548-x

Gómez, D. M., Bion, R. A. H., & Mehler, J. (2011). The word segmentation process as revealed by click detection. *Language and Cognitive Processes, 26*, 212–223.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159–207.

Karuza, E. A., Emberson, L. L., & Aslin, R. N. (2013). Combining fMRI and behavioral measures to examine the process of human learning. *Neurobiology of Learning and Memory, 109*, 193–206.

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: Is it implicit and long-term? *Neuroscience Letters, 461*, 145–149. doi: 10.1016/j.neulet.2009.06.030

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical leaning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*, B35–B42.

Klatt, D. H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 243–288). Hillsdale, NJ: Erlbaum.

Perruchet, P., Cleeremans, A., & Destrebecqz, A. (2006). Dissociating the effects of automatic activation and explicit expectancy on reaction times in a simple associative learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 955–965.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: Two approaches, one phenomenon. *Trends in Cognitive Sciences, 10*, 233–238.

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.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception and Psychophysics, 44*, 563–575.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*. doi: 10.1002/wcs.78

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science, 274*, 1926–1928.

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.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning. *Psychological Science, 8*, 101–105.

Shanks, D. R., & Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin & Review, 9*, 362–367. doi: 10.3758/BF03196294

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*, 367–447.

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*, 706–716.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005, May). *Attention and automaticity in visual statistical learning*. Talk presented at Vision Sciences Society Conference, Sarasota, FL. Abstract published in *Journal of Vision, 5*, 1067a.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience, 30*, 11177–11187.

Received August 6, 2014
Revision received January 24, 2015
Accepted January 27, 2015
Published online May 7, 2015

Ana Franco

Consciousness, Cognition, and Computation Group
Université Libre de Bruxelles CP 191
50 avenue F.-D. Roosevelt
1050 Bruxelles
Belgium
Tel. +32 2 650-2633
Fax +32 2 650-2209
E-mail afranco@ulb.ac.be

