Frequency-based organization of speech sequences in a nonhuman animal

Juan M. Toro[1], Marina Nespor[2], Judit Gervain[3,4]

1. ICREA - Universitat Pompeu Fabra, Barcelona, Spain

2. ERC PASCAL - SISSA, Trieste, Italy

3. Laboratoire Psychologie de la Perception, Université Paris Descartes, Sorbonne Paris Cité, Paris, France

4. Laboratoire Psychologie de la Perception, CNRS, Paris France

Word count: 5170

Corresponding author:

Juan M. Toro

Universitat Pompeu Fabra

C. Roc Boronat, 138

CP. 08018, Barcelona

mail: juanmanuel.toro@upf.edu

Phone number: +34 935422629

Fax: +34 935422517

ABSTRACT

A recurrent question regarding language acquisition is the extent to which the mechanisms human infants use to discover patterns over the linguistic signal are highly specialized and uniquely human, or are the result of more general mechanisms present in other species. Research with very young infants suggests that they are able to use the relative frequency of elements in a linguistic sequence to infer word order. Here we ask if this ability could emerge from grouping biases present in nonhuman mammals. We show that animals discover differences in the frequency of elements in a sequence and can learn the relative order of frequent and infrequent elements. Nevertheless, in animals, relative frequency does not appear to be overridden by other cues that have been shown to be important to human infants, such as prosody. Our results demonstrate that the basic mechanism that allows listeners to extract ordering relations based on frequency is shared across species.

Keywords: Word frequency, comparative cognition, function words, prosody

1. Introduction

Despite important advances in the study of language acquisition, we still do not have a complete understanding of how very young infants manage to learn the complexities of language. In the last years exciting discoveries have been made regarding infants' abilities to use cues present in the speech signal to learn linguistic regularities. For example, infants can compute statistics over elements in speech to discover word-like units (Saffran, Aslin & Nesport, 1996), extract token-independent patterns even at very young ages (Marcus, Vijayan, Bandi Rao & Vishton, 1999; Gervain Macagno, Cogoi, Peña & Mehler, 2008), and use prosodic cues to establish long-distance relations between categories of items (Endress & Bonatti, 2007; Marchetto & Bonatti, 2013; Peña, Bonatti, Nespor & Mehler, 2002;). All these abilities play a central role in the infant's quest to discover the elements and the structure of their language of exposure.

Recent research has suggested that infants not only extract these regularities from the signal, but they also map them onto abstract linguistic categories, e.g. functors and content words. In a series of experiments, Gervain and collaborators showed that infants use the relative frequency distribution of elements in a sequence to group them following the main word order of their native language (Gervain, Nespor, Mazuka, Horie & Mehler, 2008). In their study, the authors presented 8-month-old infants with sequences of alternating frequent and infrequent words, and then tested how the infants grouped the elements in the sequences. They found that eight-month-old infants growing in an Italian-speaking environment tended to group the words putting the frequent words first and the infrequent words second. On the contrary, infants growing in a Japanese-speaking environment tended to group the words differently, putting the infrequent words first and the frequent words second. These two patterns nicely mirrored those observed for functors and content words in Italian and Japanese, respectively. Across languages, functors are frequent grammatical words such as articles, pronouns, prepositions (as *in the, is, he, on, to*), while content words are infrequent and tend to carry lexical meaning such as nouns, verbs, adjectives (as *table, eat, pretty*; for differences in frequency between functor and content words in English see Kucera & Francis, 1967, for differences in frequency across Italian, Japanese see Gervain et al. 2008, and for differences across Farsi and Turkish see Gervain & Werker, 2013). Italian is a functor-initial language, in which functors are in the great majority of cases located in phrase-initial positions, i.e. before content words. In contrast, Japanese is a functor-final language, which tends to locate functors in phrase-final positions, i.e. after content words. Thus, infants in the Gervain et al. (2008) experiments were not only using relative frequency of words to group them, but more importantly, they were mapping such differences in relative frequency to the word order of their native language. As a result, infants placed frequent words in the position assigned to functors in their native language, while infrequent words were placed in the position assigned to content words. These results were confirmed and extended in later studies showing that infants of the same age are even able to use these frequency cues together with prosodic information to derive language-appropriate representations of word order pre-lexically, i.e. before they could rely on their understanding of word meanings to support grammar learning (Bernard & Gervain, 2012; Gervain & Werker, 2013).

Once infants enter the word-learning stage, they associate these two frequency-based grammatical categories with the expected functions. Thus, 17-month-olds expect infrequent (i.e. content words), but not frequent words (i.e.

functors), to be object labels using both natural (Hochmann, Endress & Mehler, 2010) and artificial linguistic stimuli (Hochmann, 2013). Importantly, the links infants establish between relative frequency and function/content words is not appropriately explained by simpler accounts involving mutual exclusivity or cross-situational statistics (Hochmann et al. 2010). That is, results converge on the idea that by this age, infants reliably infer the functional properties of words from relative distribution information.

A fundamental issue regarding these early language acquisition mechanisms is whether they are highly specialized for language and uniquely-human or, on the contrary, they are manifestations of more general abilities already present in other species. For example, categorical perception is a signature of how we perceive phonemes. However, instead of being language- or even human-specific, it appears to be the result of how humans as well as other species organize relevant acoustic stimuli (e.g. Kuhl & Miller, 1975). Similarly, it has been proposed that humans use the prosodic grouping principles described by the Iambic-Trochaic Law (according to which sequences alternating in intensity and pitch tend to be grouped into prominence-initial units, i.e. trochees, while sequences alternating in duration tend to be grouped into prominence-final units, i.e. iambs) to infer syntactic properties of their language, such as the head direction parameter (Nespor, Shukla, van de Vijver, Avesani, Schraudolf & Donati, 2008; Bion, Benavides-Varela & Nespor, 2011; Yoshida et al. 2010). Recent research suggests that the trochaic grouping bias is common to other perceptual domains (Peña, Bion & Nespor, 2011), and is also shared with other mammals (de la Mora, Nespor & Toro, 2013). In parallel, several sources of evidence suggest that the iambic grouping bias emerges as a consequence of language-specific experience (e.g. Hay & Saffran, 2012; Iversen, Patel, & Ohgushi, 2008, Bion et al., 2011). For example, Object-Verb languages such as Japanese, have a trochaic pattern at the phrasal level (e.g. Gervain & Werker, 2013; Nespor et al. 2008). It has been suggested that this might modulate how native Japanese speakers group duration-varying sequences (e.g. Iversen et al. 2008). Interestingly, experience seems also to play a role in the development of iambic grouping biases in non-human animals (de la Mora et al. 2013). Thus, at least some features characterizing how humans process the linguistic input appear to derive from more general abilities present in other animals, and therefore, are neither language-, nor human-specific.

It remains currently unexplored whether the frequency-based categorization and organization of the speech input, essential to human infants' early acquisition of their native language (Gervain et al. 2008; Gervain & Werker, 2013), is unique to how humans process the speech signal, or whether it is an ability shared with other animals. That is, it has so far not been explored whether the use of frequency as a grouping cue by infants might be based on a priori grouping biases that could be observed in other species. In the present study we address this question by presenting non-human animals with sequences of alternating frequent and infrequent elements and then testing how they group them. Our hypothesis is that if the way humans use relative frequency to discover word order can be traced to more general grouping biases present in non-human animals, we should find some precursors of this bias in our experiments. On the contrary, differences should be observed if the way humans map the relative frequency of elements in a sequence to syntactic categories is the result of how humans uniquely process speech.

2. Experiment 1. Frequency as an organizing cue

In the present experiment we tested whether non-human animals use the frequency of items in a sequence as a grouping cue, just as human infants as well as adults do (Gervain et al. 2008; Gervain & Werker, 2013). We presented rats with auditory sequences of alternating frequent and infrequent syllables and then tested how they group them.

2.1. Methods

2.1.1. Subjects

Subjects were 12 Long-Evans rats (six males) of 4 months of age. They were food-deprived until they reached 90% of their free-feeding weight. They had access to water ad libitum. Food was administered after each training session.

2.1.2. Stimuli

We used the same material as in Gervain and Werker (2013), except for the control random sequences. It comprised an artificial language composed of frequent (A, B) and infrequent (X, Y) CV syllables. Frequent syllables corresponded to a single token (A: fi; B: ge). Infrequent syllables corresponded to nine tokens each (X: ba, de, du, fo, pe, pa, ra, ru, to; Y: bo, bi, do, ka, ku, mu, na, ri, ro). Thus frequent syllables appeared nine times more often than the infrequent syllables. To create the sequences, syllables were arranged following an alternating pattern between frequent and infrequent syllables, as in AXBYAXBY (e.g. fibagebofifogeku; Alternating Sequences, AS; see table 1). Each of the sequences presented during training contained 36 syllables and lasted 9 secs. Half of the sequences started with a frequent syllable and half started with an infrequent one. Each sequence was faded in and out for the duration of 1 sec at its onset and offset, respectively, to avoid animals simply focusing on the first or the last element of the sequence. As a control condition, we also created sequences that consisted of exactly the same syllables as the alternating sequences, but with their elements organized at random (thus, not following an alternating pattern between frequent and infrequent items; e.g. AXYBAYXB...). Nine new control sequences were created by randomly combining the same syllables composing each sequence (that is, without any alternation between frequent and infrequent syllables; e.g. XBAYXXAYB…; Random Sequences, RS). Test stimuli consisted in eight new sequences composed of four syllables each. Each one had an alternation of frequent and infrequent syllables. Importantly, four of these sequences started with a frequent syllable (e.g. AXBY), while the other four started with an infrequent syllable (e.g. XAYB). We expected animals to respond more to test sequences that followed whatever grouping pattern they had parsed during training (for a similar design, see de la Mora et al., 2013). All sequences were synthesized using MBROLA software with the fr4 female database with a fundamental frequency of 200 Hz, and syllable duration set to 240 ms.

Frequent - A: fi; **B**: ge **Infrequent - X**: ba, de, du, fo, pe, pa, ra, ru, to; **Y**: bo, bi, do, ka, ku, mu, na, ri, ro			
	Training		**Test**
Experiment 1	Alternating ..AXBYAXBY..	..fibagebofifogeku..	fibagebo - bafiboge
Experiment 2	Random ..XBAYYAXB..	..bagefibokufifoge..	fifi - baba
Experiment 4	Alternating ..AXBYAXBY.. Random ..XBAYYAXB..	..fibagebofifogeku.. ..bagefibokufifoge..	fibagebo - bafiboge

Table 1. Sequences and test stimuli used in Experiments 1 to 4. In Experiments 1, and 2 the same training sequences were used. However, test stimuli differed. In Experiment 1 animals were presented with test sequences alternating frequent and infrequent items that could either start with a frequent (AXBY) or with an

infrequent (XAYB) element. In Experiment 2, the animals were presented during test with frequent-frequent (XX or YY) or with infrequent-infrequent (AA or BB) items. In Experiment 3, random (not alternating) sequences were reinforced. Finally, in Experiment 4 pitch was increased in the infrequent items during training (underlined items). The test was the same as in Experiment 1.

2.1.3. Apparatus

We used Letica L830-C Skinner boxes (Panlab S. L., Barcelona, Spain). A custom made program (RatboxCBC) presented stimuli, recorded the lever-press responses and provided reinforcement during the experiment. Syllable sequences were acoustically presented using E. V. (s-40) speakers located besides the boxes.

2.1.4. Procedure

Before the experiment started, rats were trained to press a lever to obtain sucrose pellets. Once rats learned the target response, the experiment started. The experiment consisted of a discrimination training phase and a test phase. During discrimination training, rats were placed individually in the response boxes for each session. There were a total of forty training sessions, each lasting 30 minutes. In each session, thirty sequences (15 alternating sequences and 15 random sequences) were presented with an inter-stimulus interval of 60 secs. Order of presentation of sequences was randomized in each session. After the presentation of any alternating sequence, rats received reinforcement for their lever pressing responses. On the contrary, after the presentation of any random sequence, rats did not receive reinforcement. After the forty training sessions, there was a test session. It was identical to any training session, the only difference being that the eight test sequences were randomly interleaved and presented instead of eight training sequences. No reinforcement was delivered after any test sequence.

2.2. Results and Discussion

Rats learned to discriminate between alternating and random sequences, increasing the level of responses after alternating sequences and decreasing it after random sequences through training. This was confirmed with an ANOVA with data collapsed across participants, and session as a random factor ($F(1, 38)=41.65$, $p<0.005$; see Figure 1). Since the two sequences contained exactly the same elements, the only differentiating cue was the alternation of frequent and infrequent items. After training, we tested whether rats had used this alternation of frequent and infrequent items to assign either one or the other to initial positions in the sequence. We analyzed data from the test with subjects as a random factor, and found that rats pressed the lever more often ($t(11)=2.47$, $p<0.05$) after test sequences that put frequent items first ($M=137.1$, $SD=37.2$) than to sequences that put infrequent items first ($M=119.5$, $SD=40.5$, see Figure 2). This suggests that, facing a continuous stream of alternating elements contrasting in relative frequency, non-human animals have a tendency to put frequent items first, i.e. to group sequences as frequent-initial. In the absence of a native language and its specific word order, the frequent-initial and the infrequent-initial orders were equally possible parses of the sequences. This test, therefore, also served to explore whether a default grouping order exists in non-linguistic animals.

However, instead of a tendency to group frequency-alternating sequences as frequent-infrequent, results from Experiment 1 might rather reveal a tendency to only focus on frequent, and hence more familiar, items that appear at the beginning of test sequences. Thus, in Experiment 2 we wanted to explore further the tendency observed in animals to respond more to highly frequent items in a sequence.

3. Experiment 2. Focus on frequent items

This experiment was identical in all respects to the previous one, except for the test stimuli used. A new group of animals was trained to discriminate between alternating and random sequences (the same ones as in Experiment 1). The only difference with respect to Experiment 1 was that during test, instead of presenting AXBY and XAYB sequences, we pitted frequent-frequent items (as in AA, BB) against infrequent-infrequent items (as in XX, YY). If animals parse the alternating sequences into organized, frequency-alternating units during training, they should show no preference in this task. If they only use a simpler mechanism, favoring frequent familiar elements in the salient sequence-initial position, a preference for the frequent-frequent items might be seen.

3.1. Methods

3.1.1. Subjects

Subjects were a new group of 12 Long-Evans rats (six males) of 6 months of age. They were food-deprived until they reached 90% of their free-feeding weight. They had access to water ad libitum. Food was administered after each training session.

3.1.2. Stimuli

Training stimuli were the same as in Experiment 1. Test stimuli, on the contrary, consisted of four pairs of frequent-frequent (e.g. AA; BB) or infrequent-infrequent (e.g. XX; YY) syllables instead of the frequent-initial or infrequent-initial sequences used in Experiment 1. Stimuli were synthesized in the same manner as in the previous experiment.

3.1.3. Apparatus and procedure

The apparatus and the procedure were the same as in Experiment 1.

3.2. Results and discussion

As in the previous experiment, rats increased their lever pressing responses after alternating sequences and decreased them after random sequences during discrimination training ($F(1, 38)=73.04$, $p<0.005$). More interestingly, animals responded more ($t(11)=2.22$, $p<0.05$) to frequent-frequent test items ($M=112.2$, $SD=15.8$) than to infrequent-infrequent test items ($M=97.1$, $SD=24.4$, see Figure 2), suggesting that they are focusing on frequent items. These results suggest animals are not just memorizing specific pairs from the training set. Animals are still responding more to FF test items that never appeared during training. Also, given the high number of syllable combinations used in training, it is unlikely rats could memorize them.

Results from the training phase in experiments 1 and 2 suggest that animals are not only learning the differences in relative frequency between frequent and infrequent items. They are also learning to discriminate between alternating and random sequences. This suggests the animals are detecting the regular alternation of frequent and infrequent items in the alternating sequences as to differentiate them from sequences composed by exactly the same elements but organized randomly, This is a key finding that suggests that the basic computational abilities human infants might use to bootstrap word order from frequency information are shared with other animals. However, it is important to disentangle the effects of regularity (alternation of frequent and infrequent elements in the sequence) and sequential order (frequent elements placed in a fixed position, either the first or the last, in the sequence) in the present results. To tackle this issue, we ran a new experiment in which the random (and not the alternating) sequences were reinforced. If animals are focusing on relative frequency alone in the test phase, we should observe a similar pattern of results independently of sequential order, as differences in relative frequency were the same across sequences. On the contrary, if animals are focusing on sequential order, we should observe a contrasting pattern from that observed in the previous experiments.

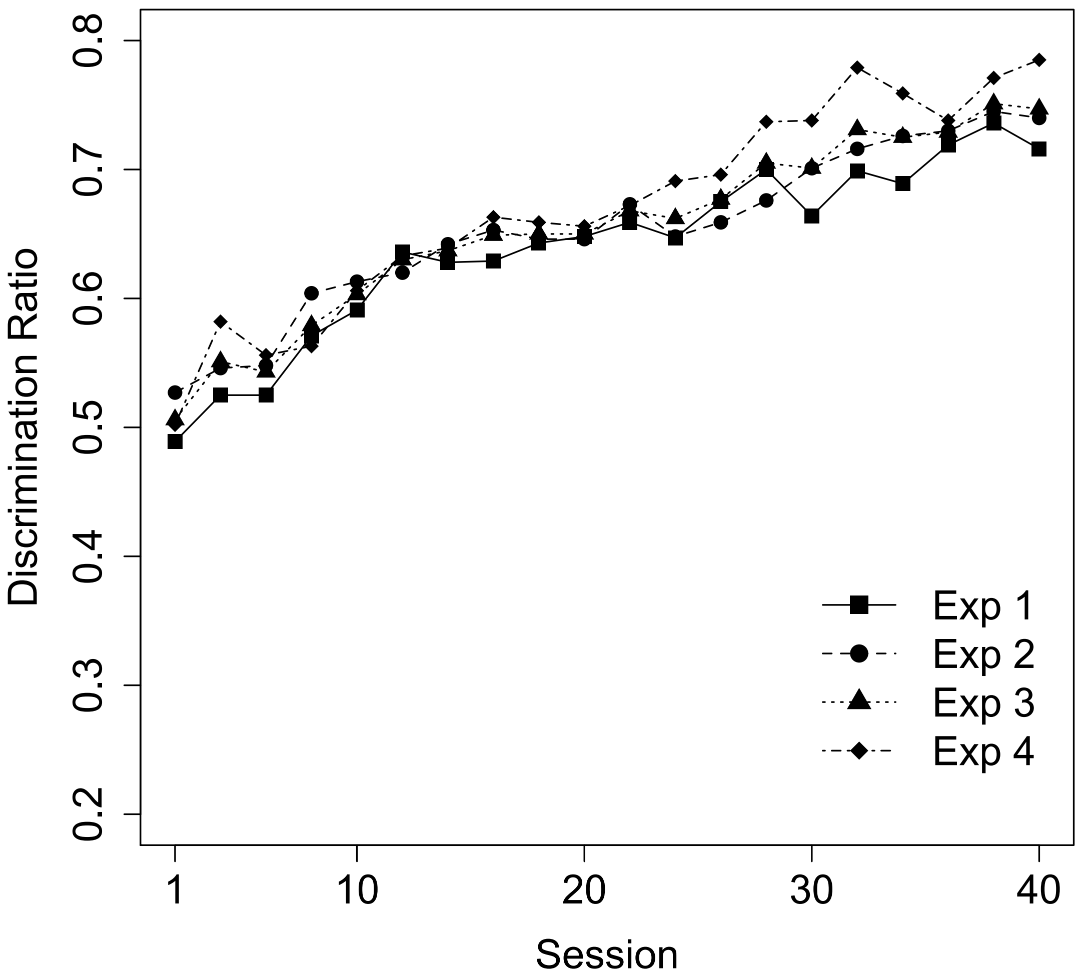

Figure 1. Discrimination ratio during training for experiments 1 to 4. The ratio is calculated by dividing the number of lever presses after reinforced stimuli by the total number of lever presses. The animals learned to discriminate among alternating and non-alternating sequences in all experiments.

4. Experiment 3. Frequency and word order

In the present experiment we explored preference for frequent items in the absence of regular alternations. Animals were trained to discriminate between alternating and random sequences. Contrary to Experiments 1 and 2, random sequences were reinforced while alternating sequences were not. In the test phase both test items alternating between frequent and infrequent items (AXBY; XAYB), and items composed of only frequent (e.g. AA; BB) or only infrequent (e.g. XX; YY) items were presented. The prediction is that we should observe the same pattern of results as in the previous experiments if the animals are mainly using relative frequency as a discrimination cue during test.

4.1. Methods

4.1.1. Subjects

Subjects were a new group of 12 Long-Evans rats (six males) of 6 months of age. They were food-deprived until they reached 90% of their free-feeding weight. They had access to water ad libitum. Food was administered after each training session.

4.1.2. Stimuli

Training stimuli were the same as in Experiments 1 and 2. Test stimuli included both types of test items as in the previous experiments. That is, there were four AXBY and four XAYB test sequences as in Experiment 1, and four pairs of frequent-frequent (e.g. AA; BB) or infrequent-infrequent (e.g. XX; YY) syllables as in Experiment 2.

4.1.3. Apparatus and procedure

The apparatus was the same as in Experiment 1 and 2. The procedure was very similar, and only differed from previous experiments in the fact that the random sequences (not the alternating sequences) were reinforced during training. That is, after the presentation of any random sequence, rats received food pellets for their lever pressing responses. On the contrary, after the presentation of any alternating sequence, rats did not receive pellets for their responses. During test, two types of test items were presented. The eight test items used in Experiment 1 (four AXBY and four XAYB), and the eight test items used in Experiment 2 (four frequent-frequent as in AA and BB, and four infrequent-infrequent as in XX and YY). During the test session all test items were presented randomly interleaved between training stimuli as in the previous experiments.

4.2. Results and discussion

Similarly to the previous experiments, rats learned to discriminate between random and alternating sequences during training ($F(1, 38)=54.33$, $p<0.005$), again suggesting that they were able to detect differences in word order between sequences as to tell them apart. During test, animals responded differently to the separate types of test items ($F(3, 44)=4.64$, $p<0.05$). They pressed the lever more often ($t(11)=3.86$, $p<0.05$) after AXBY ($M=133.8$, $SD=15.4$) than after XAYB ($M=113.3$, $SD=19.4$) sequences. They also responded more ($t(11)=2.24$, $p<0.05$) to FF ($M=120.4$, $SD=16.6$) than to II ($M=108.3$, $SD=19.3$) test items. That is, although random sequences were reinforced during training, animals still responded more to frequent items. This suggests that, even in the absence of alternations, animals are focusing on differences in relative frequency to discriminate among items presented during test. Together with the results observed in Experiments 1 and 2, the pattern observed in Experiment 3 provides strong evidence that relative frequency is being used as an important discrimination cue. That is, animals are focusing on the highly frequent items to respond during test.

In human infants, relative frequency among elements has been shown to act as a grouping cue to extract the order of words in their native language (frequent-infrequent in Italian infants, Gervain et al. 2008; Hochmann, 2013; and more crucially, infrequent-frequent in Japanese infants, Gervain et al. 2008). However, frequency is not the only information infants rely on to learn word order. Indeed, prosody is a strong cue humans use to learn key syntactic information such as relative word order (e.g. Christophe, Nespor, Guasti & van Ooyen, 2003, Bion, Benavides-Varela & Nespor, 2011). Further, frequency and phrase-level prosody have been shown to act in concert to cue word order in pre-lexical infants (Bernard and Gervain 2012, Gervain and Werker 2013). That is, infants appear to use several sources of information as grouping cues during speech processing. It thus may be the case that rats need more information in addition to frequency to display grouping biases over elements composing the sequences (e.g. de la Mora et al. 2013). We therefore wanted to explore if animals' focus on relative frequency would still be observed with the introduction of additional prosodic cues that might help them organizing the sequences. Thus, in our fourth experiment, we introduced a prosodic cue in the sequences presented to the animals.

5. Experiment 4. Is frequency overridden by prosody?

In the present experiment, we increased the pitch of infrequent syllables to 224 Hz, while keeping the pitch of frequent syllables at 200 Hz (similarly to Gervain and Werker, 2013). Rats are sensitive to these variations in pitch, and in fact can use them to group sequences of tones alternating in high pitch and low pitch following the principles described by the Iambic-Trochaic Law (thus putting items with higher pitch before items with lower pitch; see de la Mora et al. 2013). In all other respects the present experiment was identical to Experiment 1, including the presentation of exactly the same test items involving the alternation of frequent and infrequent syllables with constant 200Hz pitch (AXBY and XAYB). Given rats' previously shown preference for trochaic items when presented with pitch contrasts (de la Mora et al. 2013), this manipulation pitted frequency against prosody. A preference for frequent-initial items predicted an increased response to frequent-infrequent (AXBY) items, while a preference for high-pitch initial (trochaic) items predicted an increased response to infrequent-frequent (XAYB) items.

5.1. Methods

5.1.1. Subjects

Subjects were a new group of 12 Long-Evans rats (six males) of 4 months of age. They were food-deprived until they reached 90% of their free-feeding weight. They had access to water ad libitum. Food was administered after each training session.

5.1.2. Stimuli

In Experiment 3 we manipulated the pitch of infrequent syllables. We took the same training sequences as in the previous experiments, but this time synthesized infrequent syllables with higher pitch (224 Hz) than frequent syllables (200 Hz). Fundamental frequency changes were the same as those used by Gervain and Werker (2013), which allowed for more direct comparisons across species, but were less marked than those used by de la Mora et al. (2013). However, they are within the auditory limits that have been used with speech stimuli (e.g. Eriksson & Villa, 2006). Test stimuli were the same as in Experiment 1, thus pitting frequent-initial items (AXBY) against infrequent-initial items (XABY).

5.1.3. Apparatus and procedure

The apparatus and the procedure were the same as in Experiment 1 and 2.

5.2. Results and discussion

As in previous experiments, rats learned to respond more to alternating sequences than to random sequences during training ($F(1, 38)=128.29$, $p<0.005$). Also, during test, they pressed the lever more often ($t(11)=2.42$, $p<0.05$) to test sequences that began with a frequent item (AXBY; mean=142.5, SD=29.9) than to sequences that began with an infrequent item (XAYB; mean=115.2, SD=36.1, Figure 2). We can thus draw the conclusion that the animals' tendency to respond more to frequent items located in initial positions is very resilient. In fact, this focus on frequent items in initial positions is not overridden by the prosodic cue present in the sequences used in the present experiment. Animals continue to use the relative distribution of elements within a sequence as their primary discrimination cue. They do so, despite their documented preference for high-pitch initial items when no information about relative frequency is present (de la Mora et al. 2013).

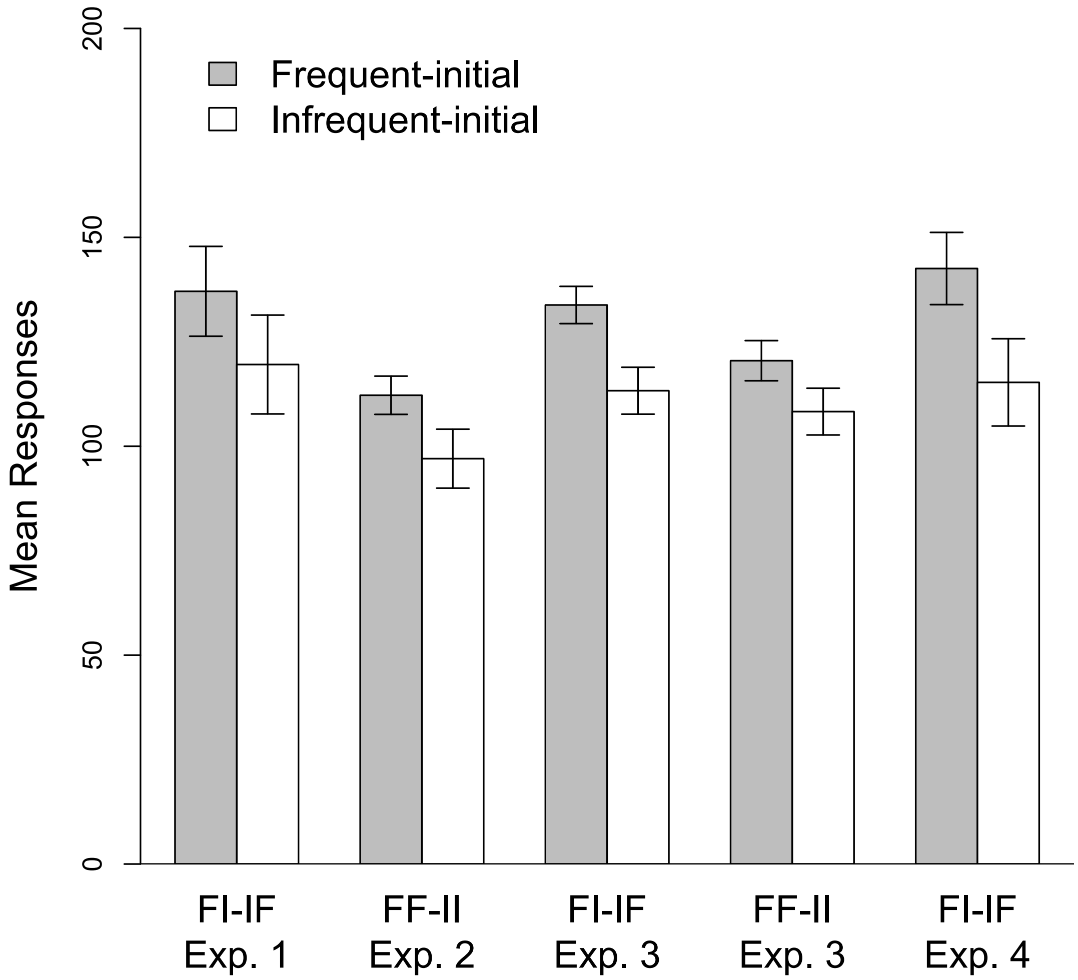

Figure 2. Mean number of responses (and standard error bars) for frequent-initial and infrequent-initial items during test. Animals consistently responded more after test sequences beginning with frequent items (Experiment 1), independently of whether the precise combinations appeared during training (Experiment 2). The same pattern was observed when random, instead of alternating sequences were reinforced (Experiment 3), and in the presence of conflicting prosodic cues (Experiment 4).

6. General Discussion

In four experiments we explored whether non-human animals can detect frequency difference across elements in acoustic sequences and use them to infer their relative order. This ability could form the basis for the capacity attested in human infants to track the relative frequency of elements in a sequence, that has been hypothesized to be one of the mechanisms that allows them to bootstrap word order. We observed that animals are able to discriminate between frequent and infrequent items in a sequence. They can also use differences in the sequential order of items in a stream to discriminate between the alternating and the random sequences. Our results provide evidence for a remarkable ability present in non-human animals to track relative frequency and regularities between elements in a sequence. Discrimination training data demonstrates that the basic mechanisms for detecting the regular alternation between frequent and infrequent elements are shared across species. Nevertheless, during test, animals appear to focus only on frequent items, but do not appear to use alternations across elements to form coherent groups. Furthermore, relative frequency does not appear to be overridden by prosodic information added to the sequences, as it appears to be the case for human infants.

From the perspective of animal behavior, these results are of theoretical importance in at least three ways. First, they show that even non-linguistic animals detect both the differences in relative frequency between elements in a speech sequence and the regularity according to which they are arranged. Second, data from the test demonstrates that the animals have a strong preference for frequent items. This even leads them to prefer sequences that have not appeared during training but contain such salient items (as in FF test items). And third, their focus on frequent items is so strong that it is not overridden by prosodic cues, which they are otherwise able to exploit.

More central for our study is the extent to which the current set of results might provide information regarding the development of grouping biases in human infants. Animals are computing the relative statistics between the elements in the sequences and differentiate between frequent and infrequent items. They are also keeping track of the regular alternation of elements in the sequence. The presence of these abilities in non-human animals suggests that the basic building blocks of the frequency-based bootstrapping capacity that young human infants might use to discover word order are shared across species. However, mounting evidence suggests that human infants also map frequency onto linguistic categories (functors and content words), so that infants with different linguistic backgrounds tend to assign frequent elements to different positions within the sequence following the order of their native language (e.g. Gervain et al. 2008; Gervain & Werker, 2013; Hochmann et al. 2010; Hochmann, 2013). That is, infants do not tend to consistently show a frequent-initial bias. Japanese monolinguals and OV-VO bilinguals readily show infrequent-initial preferences. Minimally, this suggests that if such a bias exists in humans, it is easily overridden by language experience. Finally, humans give more weight to prosodic than to frequency cues when processing language. This salient role of prosody is pivotal during language acquisition. Thus, whenever prosodic information is present, it helps human infants and adults to discover structure in speech (e.g. Gervain & Werker, 2013; Langus, Marchetto, Bion & Nespor, 2012; Peña, Bonatti, Nespor & Mehler, 2002; Shukla, Nespor & Mehler, 2007; Shukla, White & Aslin, 2011). In contrast, in the present experiments, we observed that rats kept their focus on the relative frequency of the elements, and we found no evidence of their switching behavior in the presence of prosodic cues.

The present set of results opens the possibility of further exploring the exact grouping cues human infants are using. For example, experiments with infants so far have tested the preference for FIFI vs IFIF sequences during test (e.g.

Gervain, Nespor et al. 2008; Gervain & Werker, 2013). It would be interesting to test their preference for sequences that have not been presented during familiarization such as FF or II (similar to the ones used in Experiment 2) to assess their reliance on frequent items. In a similar vein, infants could be familiarized with sequences composed of non-alternating frequent and infrequent items. This could allow us to determine whether the alternations play a pivotal role in the observed results. It could be the case that given non-alternating sequences, infants rely only on relative frequency without a strong focus on sequential order. A complementary issue is whether, with the appropriate amount and type of experience, non-human animals could learn to use differences in relative frequency for grouping elements in a sequence. That is a very interesting avenue of research. One could think of experiments with non-human animals in which experience could be a factor for the development of grouping biases based on relative frequency. Our results so far demonstrate that in the absence of such experience, there are no a priori grouping biases in animals.

Animals are avid for finding statistical regularities in the events surrounding them (Rescorla & Wagner, 1972), and it is well established that they can discriminate elements based on their relative frequency (e.g. Keen & Machado, 1999). Our results go further, by also demonstrating that they can use the order of frequent and infrequent elements in a sequence for discrimination. Humans have been proposed to use this ability to bootstrap word order in their native language, and our results suggest that its bedrock is already present in non-human animals. However, ample research with human adults and infants suggests that we map the statistical information observed in the world onto abstract linguistic categories (e.g. phonemes, syllables, syntactic categories, etc.). In the absence of such mapping, the computation of statistical regularities over acoustic stimuli does not appear to lead to the acquisition of abstract structures.

References.

Bernard, C., & Gervain, J. (2012). Prosodic cues to word order: What level of representation? *Frontiers in Psychology, 3,* 451.

Bion, R., Benavides-Varela, S., & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech, 54*, 123-140.

Christophe, A., Nespor, M., Guasti, M.T., & van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science, 6,* 211-220.

de la Mora, D., Nespor, M., and Toro, J.M. (2013) Do humans and non-human animals share the grouping principles of the Iambic-Trochaic Law? *Attention, Perception and Psychophysics, 75*, 92-100.

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*. ICSLP, Philadelphia.

Endress, A., & Bonatti, L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition, 105,* 247-299.

Eriksson, J. & Villa, A. (2006). Learning of auditory equivalence classes for vowels by rats. *Behavioural Processes, 73,* 348-359.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences, USA, 105*, 14222–14227.

Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology, 57,* 56-74.

Gervain, J. & Werker, J. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications, 4,* 1490.

Hay, J. & Saffran, J. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy, 17,* 610-641.

Hochmann, J.R. (2013). Word frequency, function words and the second gavagai problem. *Cognition, 128,* 13-25.

Hochmann, J.R., Endress, A., & Mehler, J. (2010). Word frequency as a cue for identifying function words in infancy. *Cognition, 115,* 444-457.

Iversen, J., Patel, A., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America, 124,* 262-274.

Keen, R., & Machado, A. (1999). How pigeons discriminate the relative frequency of events. *Journal of Experimental Analysis of Behavior, 72,* 151-175.

Kucera, H., & Francis, N. (1967). *A computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kuhl P., & Miller, J. (1975). Speech perception by the chinchilla: Voiced voiceless distinction in alveolar plosive consonants. *Science, 190*, 69-72.

Langus, A., Marchetto, E., Bion, R., & Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *Journal of Memory and Language, 66,* 285-306.

Marchetto, E., & Bonatti, L. (2013). Words and possible words in early language acquisition. *Cognitive Psychology, 67,* 130-150.

Marcus, G., Vijayan, S., Bandi Rao, S., & Vishton, P. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77-80.

Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio, 2*, 1-29.

Peña, M., Bonatti, L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*, 604-607.

Peña, M., Bion, R., & Nespor, M. (2011). How modality specific is the Iambic-Trochaic Law? Evidence from vision. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37*, 1199-1208.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Black, H. & Prokasy, W. (Eds.) *Classical conditioning II: Current research and theory*. (Appleton, New York), pp. 64-99.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology, 54*, 1-32.

Shukla, M., White, K., & Aslin, R. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-month-old infants. *Proceedings of the National Academy of Sciences, USA, 108,* 6038–6043.

Yoshida, K., Iversen, J., Patel, A., Mazuka, R., Nito, H., Gervain, J., & Werker, J. (2010). The development of perceptual grouping biases in infancy: A Japanese–English cross-linguistic study. *Cognition, 115*, 356–361.

Acknowledgements

The present research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 269502 (PASCAL) to MN; ERC grant agreement n° 312519 to JMT, and from the Agence Nationale de la Recherche (Programme JCJC 21373), the Emergence(s) program of the City of Paris as well as from the French Investissements d’Avenir - Labex EFL Program (ANR-10-LABX-0083) to JG.