Do humans and nonhuman animals share the grouping principles of the Iambic - Trochaic Law?

Daniela M. de la Mora[a] Marina Nespor[b] & Juan M. Toro[a]

a. Center for Brain and Cognition. Universitat Pompeu Fabra

b. ERC - PASCAL at SISSA/ISAS

Word count: 5733

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

The Iambic-Trochaic Law describes humans' tendency to form trochaic groups over sequences varying in pitch or intensity (i.e., the loudest or highest sound marks group beginnings), and iambic groups over sequences varying in duration (i.e., the longest sound marks group endings). The extent to which these perceptual biases are shared by humans and nonhuman animals is yet unclear. In Experiment 1, we trained rats to discriminate pitch-alternating sequences of tones from sequences randomly varying in pitch. In Experiment 2, rats were trained to discriminate duration-alternating sequences of tones from sequences randomly varying in duration. We found that nonhuman animals group as trochees sequences based on pitch variations, but they do not group as iambs sequences varying in duration. Importantly, humans grouped the same stimuli following the principles of the Iambic-Trochaic Law (Experiment 3). These results suggest an early emergence of the trochaic rhythmic grouping bias based on pitch, possibly relying on perceptual abilities shared by humans and other mammals as well, whereas the iambic rhythmic grouping bias based on duration might depend on language experience.

Introduction

Tick tock, goes the clock/And what now shall we play?/Tick tock, goes the clock/Now summer's gone away? As this song illustrates[1], humans tend to perceive the isochronous ticks of a clock as a sequence of two paired sounds, an example of what is known as perceptual grouping (Bolton, 1894). Furthermore, variations in intensity within a sequence of tones lead to the perception of initial-prominence groups (i.e., the loudest sound marks the beginning of the group), whereas differences in duration lead to the perception of final-prominence groups (i.e., the longest sound marks the ending of the group; Woodrow, 1909). These principles of perceptual grouping depending on intensity and duration variations have been described as the Iambic - Trochaic Law (ITL; Hayes, 1995), where iambs correspond to groups with final-prominence (weak-strong) and trochees correspond to groups with initial-prominence (strong-weak).

Research has suggested the ITL may play an important role during language processing, supporting speech segmentation based on prosody (Trehub & Trainor, 1993; Hayes, 1995; Hay & Diehl, 2007). More importantly, recent evidence suggests there is a strong correlation between prosody and syntax at different hierarchical levels (Nespor, et al., 2008), and that infants might use this information to bootstrap into some aspects of the grammatical structure of their native language (e.g., Christophe, Gout, Peperkamp, & Morgan, 2003; Gout, Christophe, & Morgan, 2004; Jusczyk, Cutler, & Redanz, 1993; Jusczyk, et al., 1992). More specifically, in Nespor, et al. (2008) it is shown that at the phrasal level, prominence in trochaic grouping is signaled not only by increased intensity, but also by increased pitch. Since the different realizations of prominence reflect word order - i.e., whether heads precede or follow their complements - it is proposed that the specific type of prominence an infant is exposed to might be exploited to acquire the basic word order of its native language. Thus,

general perceptual biases described by the ITL might serve as the stepping-stones for the acquisition of some basic aspects of syntactic structure.

The relevance of the ITL to language processing raises the question of the extent to which these perceptual grouping biases might depend on language experience. Recent research across languages supports the hypothesis that such grouping principles are present in human adults regardless of the stress pattern of their native language (Hay & Diehl, 2007; Bion, Benavides-Varela, & Nespor, 2011). In their study, Hay and Diehl (2007) presented sequences of tones and sequences of the syllable /ga/ alternating either in duration or intensity to English and French speakers. They instructed participants to group sequences into a two-beat rhythmic pattern and to indicate whether the rhythm consisted of a strong sound followed by a weak sound or a weak sound followed by a strong sound. Researchers found that both English and French speakers perceived sequences varying in duration as having iambic rhythm (i.e., weak-strong) whereas they perceived sequences alternating in intensity as having trochaic rhythm (i.e., strong-weak). Hence, results suggested grouping principles of the ITL are not modulated by the participants' native language. In a similar vein, Bion, et al. (2011) asked Italian speakers to listen to a sequence of syllables alternating either in pitch or in duration. They were then presented with two pairs of syllables with constant pitch and duration – one respecting and one violating iambic-trochaic grouping during familiarization. Participants were asked to judge which of the two pairs of syllables were adjacent during the familiarization phase. Participants familiarized with pitch-varying sequences remembered better the pairs that had initial prominence during familiarization. Participants familiarized with duration-varying sequences remembered better the pairs that had final prominence. In contrast, Iversen, Patel and Ohgushi (2008) tested whether the phrasal prominence of one's native language could influence perceptual grouping.

They thus choose to test speakers of English – a head-initial language - and speakers of Japanese – a head-final language. They familiarized English and Japanese adult speakers with a sequence of tones alternating in either duration or intensity, and found that both groups segmented intensity-varying sequences as trochees. However, only English speakers, but not Japanese speakers, segmented duration-varying sequences as iambs. The authors suggested this pattern reflected an influence of the linguistic environment on individuals' perceptual grouping biases. That is, the results mirrored the difference between the acoustic correlates of phrasal prominence signaling word order in the participants' native languages.

Still, it could be the case that the grouping principles described by the ITL are present early in development, but are modulated once infants interact with their linguistic environment. In fact, research on infants' perceptual grouping biases has suggested developmental differences between the two principles of the ITL. In a recent study, Yoshida, et al. (2010) familiarized 5 and 7 month-old English- and Japanese-learning infants to a stream of tones alternating in duration. During testing, they measured infants' preference for either iambic or trochaic groups. Results showed that only 7 month-old English infants segmented the sequence as iambs. In contrast, 5 month-old English infants and 5 and 7 month-old Japanese infants showed no preference for either trochaic or iambic sequences, suggesting exposure to a given linguistic environment might be necessary for the iambic grouping bias to appear. Parallel findings were reported by Bion, et al. (2011) with 7 month-old Italian-learning infants. The authors familiarized infants with a stream of syllables alternating in either duration or pitch. Whereas infants familiarized with a stream alternating in pitch showed a preference for trochaic pairs of syllables, infants familiarized with a stream alternating in duration did not show a clear preference for either iambic or trochaic

pairs. Together, these studies suggest a late emergence in development of the iambic grouping bias based on duration cues, pointing to the idea that it might depend on language experience. On the contrary, they suggest that the trochaic grouping bias based on intensity (Hay & Saffran, 2012) and pitch (Bion, et al., 2011) might appear early in development, and hence not be dependent on experience with a given linguistic environment (but see Höhle, Bijeljac-Babic, Herod, Weissenborn, & Nazzi, 2009).

A complementary aspect of the ITL is its presence across perceptual modalities. Iambic and trochaic grouping biases – first observed for music perception (Bolton, 1894) - apply to both linguistic and non-linguistic tone sequences (Hay & Diehl, 2007; Hay & Saffran, 2012) and are even present in the visual domain (Peña, Bion, & Nespor, 2011). This opens the possibility that the ITL reflects a general perceptual ability that is not necessarily related to language, but that can still be modulated given certain linguistic exposure. One way to address this issue is through a comparative approach. To the extent that the grouping principles described by the ITL are general and have not evolved for linguistic processing, they might also be present in other species. Even more, any differential effect that linguistic experience might have on iambic or trochaic grouping biases might be reflected in experiments using animals that, putatively, have no such experience. Research on comparative cognition has shown that humans and other species share some perceptual abilities we use for language processing (Yip, 2006). For example, previous studies found that cotton-top tamarin monkeys (Ramus, Hauser, Miller, Morris, & Mehler, 2000) and rats (Toro, Trobalon, & Sebastián-Gallés, 2003) can discriminate between two languages using their prosodic cues. It is thus possible that human and nonhuman animals share the grouping biases through which they extract prosodic information. The existence of these perceptual principles in a nonhuman animal would point towards the possibility that infants might use general

grouping principles, not evolved for language processing, to bootstrap some basic linguistic components.

In the present study, we wanted to investigate whether the principles of the ITL are uniquely human or might also be present across species. More specifically, we tested this possibility in a nonhuman animal that does not use complex vocalizations as a mean of inter-specific communication, such as the rat (*Rattus norvegicus*). We thus ran two experiments. In Experiment 1 we explored whether nonhuman animals can group as trochees sequences varying in pitch. In Experiment 2 we approached the complementary question of whether they can group sequences varying in duration as iambs.

Importantly, this research also allowed us to explore the extent to which the perceptual grouping biases described by the ITL reflect an influence of humans' linguistic environment or, on the contrary, they are independent of experience with language. Our hypothesis was that, if the perceptual grouping biases observed in human adults and infants are the result of language experience, we would not find a preference for either iambs or trochees in the experiments with animals. On the contrary, if grouping biases based on pitch and duration are differentially sensitive to experience with language (duration being more sensitive to experience than pitch; see Bion, et al., 2011; Hay & Saffran, 2012; Yoshida, et al., 2010), we might observe parallels across species for one of the cues, and not for the other.

Experiment 1: Grouping of sequences alternating in pitch

In Experiment 1 we explored whether we can observe in a non-human animal a bias to group as trochees sequences alternating in pitch. Studies with human adults and infants have reliably observed such bias for both intensity and pitch variations (e.g.

Bion, et al., 2011; Hay & Diehl, 2007; Hay & Saffran, 2012; Iversen, et al., 2008). In our study, however, we focused on pitch for several reasons. First, the aim of our paper is to investigate whether the ITL, hypothesized to be involved in syntactic bootstrapping (Nespor, et al., 2008), is a grouping mechanism shared by non-human animals. For language, it has been proposed that at the phrasal level, while duration marks iambic grouping, pitch is a much more important correlate of trochaic grouping than is intensity. Intensity alone, in fact, cannot mark prominence, and it always works together with other prosodic features, while duration and pitch can mark prominence on their own (Turk & Sawusch, 1996). In addition, intensity differences between stressed and unstressed vowels are very small, about 3-4 dB (Ortega-Llebaria & Prieto, 2011), while the minimum perceptual threshold for differences in intensity varies between 1 and 2 dB. Thus the increase in intensity caused by stress is perceptually very small. In addition, infants are not very sensitive to differences in intensity (Saffran, Werker, & Werner, 2006), and thus they are less likely to exploit intensity for the perception of phrasal prominence. Since the ultimate goal of our study is to investigate whether a mechanism exploited by infants to acquire language is shared by a non-human mammal, we have not included intensity in our study.

Methods

Subjects

Subjects were 6 Long-Evans rats (four males) of 4 months of age. They were food-deprived until they reached 80% of their free-feeding weight. They had access to water ad libitum. Food was administered after each training session.

Stimuli

Stimuli were sixteen Pitch Sequences (PS) and 16 Pitch Random Sequences (PRS). PS were composed by the concatenation of sixteen 200 ms pure tones each alternating in pitch. Importantly, sequences always included the alternation of a low (420 Hz) and a higher tone (525, 630, 735 or 840 Hz, all of which are within the range of hearing frequencies of rats; e.g., Heffner, Heffner, Contos, & Ott, 1994). For example, the sequence of tones in a PS would be (in Hz) 420-525-420-840-420-630-420-735-420-840-420-630-420-735-420-525. Half of the PS started with a low tone, and half with a high tone. The same tones used in the PS were combined at random to form the PRS (e.g., 420-420-525-630-420-420-420-735-840-420-420-420-735-630-840-525), so no systematic alternation of low and higher tones was present. A 200 ms inter-stimulus-interval (ISI) separated all tones. Every sequence lasted 6.2 sec and was faded 1 sec at its onset and offset. The tones were synthesized with Amadeus II software at a sampling rate of 44.4 KHz, and a sampling size of 16Bit.

Apparatus

Rats were placed in Letica L830-C Skinner boxes (Panlab S. L., Barcelona, Spain) while a laptop computer using a custom made program presented stimuli, recorded the lever-press responses and provided reinforcement. A Pioneer Stereo Amplifier A-445 and two E. V. (s-40) speakers, located besides the boxes, were used to present the stimuli.

Procedure

Rats were trained to press a lever until they reached a stable response rate at a variable ratio of 10 (+/- 5) (VR-10 schedule; that is, the lever-pressing response rate at which food was delivered varied between 5 and 15 times from trial to trial). During this

time, no stimuli were presented. Training to discriminate across stimuli started once rats reached a stable rate of responses. Discrimination training consisted of 30 sessions, 1 session per day. The logic behind this training procedure is that it leads rats to discriminate alternating from random sequences, and to associate the former with food delivery. Response rates during training and test can be used as a measure of sequence differentiation and grouping. For example, previous experiments have shown changes in response rates to sentences varying in rhythmic class when rats learned to discriminate among them (Toro, et al., 2003). Complementarily, rats tend to press more often a lever after test items that have been grouped through their high statistical coherence in a continuous speech stream than after test items with low statistical coherence (Toro & Trobalón, 2005). Thus, during each training session rats were placed individually in a Skinner box while 32 sequences (16 PS and 16 PRS) were presented with an inter-sequence interval of 60 sec. Sequence presentation was balanced within each session and across sessions, so all sequences were presented the same number of times across training. Every time a PS was presented, food was delivered at a variable ratio of 7 (+/- 3) (VR-7 schedule; that is, the lever-pressing response rate at which food was delivered varied between 4 and 10 times from trial to trial). Food delivery continued during 60 sec after PS presentation. On the contrary, after the presentation of each PRS no food was delivered, no matter how often the rat pressed the lever. Rats' lever-pressing responses were registered simultaneously with the presentation of the stimulus and during 60 sec of inter-sequence interval.

After 30 training sessions a test session was run. Instead of sequences, only pairs of tones were presented. There were four low-high pairs (420-525, 420-630, 420-735, 420-840 Hz), and four high-low pairs (525-420, 630-420, 735-420, 840-420 Hz). Pair presentation was randomized with the only restriction that no more than two pairs of the

same type were presented in a row. Each pair was presented only once, so there were a total of eight test trials. As in the training phase, there were 60 sec between the presentations of each pair. Lever-pressing responses were registered simultaneously with the presentation of a pair and the 60 sec following presentation. Food was delivered after both high-low and low-high pairs in order to avoid any confound of stimuli discrimination with reinforcement schedule. Hence, any difference observed in lever-pressing responses would be due to a difference in the way rats segmented the stream during training. That is, if rats pressed the lever more often for high-low test pairs than for low-high ones, this would suggest that rats associated these pairs more strongly with the PS sequences, and would imply that they grouped the sequences as trochees (high-low groups). If rats grouped sequences as iambs, they should press the lever more often for low-high pairs. If they show no preference, this would mean that they did not segment the PS sequences in either way: neither as trochees, nor as iambs.

Results and Discussion

During training, rats' increasingly responded to PS. To explore how lever-pressing responses changed across sessions, we ran a repeated-measures ANOVA over the average of lever-pressing responses to PS and PRS, with *session* (1 to 30) and *stimuli* (PS and PRS) as within-subjects factors. This analysis showed a non-significant difference between sessions ($F_{(29, 145)} = 1.380$, $p = 0.111$), but a significant difference between stimuli ($F_{(1,5)} = 33.959$, $p < 0.005$) and a significant interaction between both main factors ($F_{(29, 145)} = 13.174$, $p < 0.001$). To account for differences in overall levels of responding, mean lever presses were converted to a percentage of responses to PS and PRS. A repeated-measures ANOVA over the percentage of lever-pressing responses to the reinforced stimuli (PS), with *session* (1 to 30) as the within-subjects

factor, yielded a significant difference between sessions ($F_{(29, 145)} = 11.738$, $p < 0.001$; see Figure 1) due to the increment of the percentage of responses throughout the training phase, from session 1 (M=45.40%) to session 30 (M=66.82%). Importantly, during the test phase, out of the total number of responses to test trials, the percentage of responses to trochaic (i.e., high-low) over iambic (i.e., low-high) pairs was significantly above what is expected by chance (M=53.20%, SD=2.39; $t(5) = 3.275$, $p < 0.05$, $d = 1.893$; with chance being an equal percentage of responses to trochaic and iambic trials; see Figure 2), suggesting rats grouped the PS into trochees.

Together, these results suggest that rats learned to discriminate sequences alternating in pitch (PS) from random sequences (PRS), as they responded differently to PS and PRS during training. More relevant to the present study, results from the test phase suggest that they grouped the PS into trochees and not into iambs. This is reflected in a higher percentage of responses to high-low pairs that exceeds what would be expected if rats were responding at chance after test pairs. This points towards the idea that, as human adults and infants, rats show a trochaic bias for grouping sequences alternating in pitch. Moreover, it provides support to the hypothesis that the trochaic bias observed in humans might be a universal feature that might appear independently of language experience.

Experiment 2: Grouping of sequences alternating in duration

In Experiment 2 we turned to investigate the complementary question of whether the first principle of the ITL, that is, the iambic grouping of sequences varying in duration, is present in nonhuman animals. So far, research with human infants suggests this principle might heavily depend on language experience (Bion, et al., 2011;

Hay & Saffran, 2012; Yoshida, et al., 2010). If so, we should not observe this bias to group as iambs sequences varying in duration in other species.

Methods

Subjects

Subjects were 7 new Long-Evans rats (five males) of 4 months of age that had not participated in Experiment 1. They were food-deprived until they reached 80% of their free-feeding weight. They had access to water ad libitum. Food was administered after each training session.

Stimuli

Stimuli were sixteen Duration Sequences (DS) and 16 Duration Random Sequences (DRS). The structure of these sequences was the same as structure of sequences in Experiment 1. DS were composed by the concatenation of 16 pure tones with a fundamental frequency of 440 Hz each alternating in duration. Importantly, sequences always included the alternation of a short (200 ms) and a longer tone (350, 400, 450 or 500 ms, which are all tone durations and intervals that rats easily perceive; see for example Kelly, Cooke, Gilbride, Mitchell, & Zhang, 2006; Roger, Hasbroucq, Rabat, Vidal, & Burle, 2009). For example, the sequence of tones in a DS would be (in ms) 200-350-200-500-200-400-200-450-200-500-200-400-200-450-200-350. Half of the DS started with a short tone, and half with a long tone. The same tones used in the DS were combined at random to form the DRS (e.g., 450-500-200-350-200-200-200-200-450-400-400-200-500-350-200-200), so no systematic alternation of short and longer tones was present. A 200 ms ISI separated all tones. Every sequence lasted 8 sec

and was faded 1 sec at its onset and offset. The tones were synthesized with Amadeus II software at a sampling rate of 44.4 KHz, and a sampling size of 16Bit.

Apparatus and Procedure

The apparatus and the procedure were the same as in Experiment 1 except that in this case the test items were four short-long pairs (200-350, 200-400, 200-450, 200-500 ms) and four long-short pairs (350-200, 400-200, 450-200, 500-200 ms).

Results and Discussion

During training, rats' responses to DS and DRS did not significantly vary. A repeated-measures ANOVA over the average of lever-pressing responses to DS and DRS, with *session* (1 to 30) and *stimuli* (DS and DRS) as within-subjects factors, showed a non-significant difference between sessions ($F(29, 174) = 1.4126$, $p = 0.086$) and stimuli ($F(1, 6) = 1.003$, $p = 0.335$), but a significant interaction between them ($F(29, 174) = 5.762$, $p < 0.001$). As in Experiment 1, mean lever-pressing responses were converted to percentage of responses. A repeated-measures ANOVA over this percentage of lever-pressing responses to the reinforced stimuli (DS), with *session* (1 to 30) as the within-subjects factor, yielded a significant difference between sessions ($F(29, 174) = 6.508$, $p < 0.001$; see Figure 1). This difference is explained by the increased in lever-pressing responses through out the training, from session 1 (M=39.37%) to session 30 (M=55.79%).

More importantly, during the test phase, a t-test analysis showed that the percentage of responses to iambic pairs (i.e., short-long) was not significantly above chance (M=49.99%, SD = 7.67; $t(6) = -0.002$, $p = 0.998$, $d = -0.001$; see Figure 2). Together, these results suggest that, during training, rats did not discriminate between

DS and DRS, nor did they group the DS into iambs as reflected by chance performance during test. Moreover, it could mean that the iambic grouping principle observed in human adults and infants is not a universal bias, but a language experience dependent trait.

A comparison of the percentage of responses to the reinforced stimuli (PS in Experiment 1, and DS in Experiment 2) during the training phase, with *session* (1 to 30) as within-subjects factor, and *experiment* (1 and 2) as between-subjects factor, yielded a significant difference between sessions ($F(29, 299) = 17.895$, $p < 0.001$), and experiments ($F(1, 11) = 11.583$, $p < 0.01$), as well as a significant interaction between them ($F(29, 319) = 3.067$, $p < 0.001$). These results suggest that the difference in rats' performance during both experiments was due to a differential processing of the stimuli independently of the training procedure. That is, rats easily extracted information patterns over pitch-varying sequences but not over duration-varying sequences. This difference is further reflected by above-chance response rates during test for trochaic pairs based on pitch variations (Experiment 1), but not for either iambic or trochaic pairs based on duration variations (Experiment 2).

A remaining question regarding the results of Experiment 2 is if they could be explained by the rats' lack of sensitivity to the acoustic changes we implemented in the stimuli. However, according to previous studies, rats can discriminate between sounds with even shorter durations (e.g., 50 ms) and smaller time intervals than those present in our stimuli (Kelly, et al., 2006; Roger, et al., 2009). For example, Roger, et al. (2009) reported rats' mismatch negativity signatures in response to deviant stimuli with an interval difference of 50 ms with respect to the standard tone. In our study, the shortest duration of a tone was of 200 ms and the smallest interval difference between two tones was of 150 ms. Hence, our results of Experiment 2 can neither be interpreted as rats'

inability to process the duration of the tones used, nor to distinguish their differences in duration. Likewise, it is unlikely that greater interval differences between longer tones would yield a different result since our stimuli fit within the discrimination threshold observed in Roger, et al. (2009).

Nevertheless, to directly test the possibility that longer tones could trigger iambic grouping in rats, we ran a control condition with 9 new rats. Stimuli and procedure were exactly the same as those of Experiment 2, except that the shortest tone had a duration of 500 ms whereas the longer tones lasted 800, 1100, 1400 or 1700 ms (more than twice the duration of tones used in Experiment 2). The results from this control experiment closely replicated the results of Experiment 2. Throughout the training phase rats increased their lever-pressing responses, but during the test phase they did not press the lever more often for iambic test pairs (short-long) than for trochaic ones (long-short) (M=50.59%, SD = 3.96; t(8) = 0.453, p = 0.663, d = 0.210), suggesting they did not tend to group the alternating sequences as either iambs or trochees. Moreover, a comparison between the test phase of Experiment 2 and the control experiment yielded a non significant difference between them (t(14) = -0.206, p = 0.840, d = 0.098), suggesting that the rats where equally unable to group either as iambs or as trochees sequences of longer tones varying un duration. Thus, the results from Experiment 2, and from this control experiment with longer durations, point in the same direction. They suggest that, although rats increased their responses to DS, they were unable to correctly group the tones forming the reinforced sequences presented during the training phase (e.g., short-long groups or long-short groups) in order to discriminate them from the non-reinforced sequences. A final concern is whether the stimuli used in the present study are actually grouped by humans following the principles of the ITL. To test this, we run a third experiment with human adults.

Experiment 3: Grouping of alternating sequences by human participants

In the previous experiments we observed that rats tend to group as trochees sequences alternating in pitch (Experiment 1), but do not tend to group as iambs sequences alternating in duration (Experiment 2). We proposed that this lack of iambic grouping observed in animals might indicate that some experience (for example with language) might be necessary for an iambic grouping bias to emerge. However, it could also be the case that the specific sequences of tones varying in duration we used in Experiment 2 are not well suited to trigger iambic grouping even in humans. In fact, so far, experimental evidence concerning the ITL using tones in human adults (Hay & Diehl, 2007, Iversen, et al., 2008) and infants (Yoshida, et al., 2010, Hay & Saffran, 2012) have used sequences in which the same pair of tones alternated along the sequence, whereas in our stimuli the pair of tones varied within the sequence. Therefore, our aim in Experiment 3 was to test if the alternating sequences presented to the rats in the previous experiments would elicit in humans the grouping biases predicted by the ITL.

Methods

Participants

Twenty undergrad students from the Universitat Pompeu Fabra took part in this experiment. They were all native speakers of Spanish, and received monetary compensation for their participation.

Stimuli

Stimuli were the same alternating sequences used in Experiments 1 (PS) and 2 (DS).

Procedure

We presented participants with the alternating sequences used in Experiment 1 and 2. The order of presentation of sequences varying in pitch and sequences varying in duration was balanced (with no more than 2 sequences of the same type presented in concatenation). After each sequence, participants were presented with two test pairs (high-low and low-high for the sequences alternating in pitch; long-short and short-long for the sequences alternating in duration; these were the same test pairs used in Experiment 1 and 2). Participants were asked to indicate which pair better corresponded with the sequence they previously heard. There was a pause of 500 ms between test pairs. Participants had no time limit to answer. All participants were tested in a silent room, wearing headphones. The experiment was presented on a Macintosh OS X based laptop using the experimental software PsyScope X B57.

Results and Discussion

After listening to sequences alternating in pitch, participants significantly preferred trochaic (high-low) pairs (M=59.17%, SD = 15.03; $t(19) = 2.727$, $p < 0.05$, $d = 0.863$). After listening to sequences alternating in duration, participants significantly preferred iambic (short-long) pairs (M=67.08%, SD = 19.40; $t(19) = 3.938$, $p < 0.005$, $d = 1.245$). These results indicate that participants grouped as trochees the sequences alternating in pitch and as iambs the sequences alternating in duration. Thus, all stimuli used in the present study are grouped by human adults following the principles of the ITL. Interestingly, if we compare the test results across humans and animals, we find

that both segment pitch alternating sequences in a similar manner (t(24) = 0.96, p = 0.347, d = 0.555), but they perform significantly different for sequences alternating in duration (t(25) = 2.25, p < 0.05, d = 1.159). This suggests there is a trochaic rhythmic grouping bias based on pitch independent of language experience. It also provides support to the suggestion that such experience could be necessary in order to group sequences alternating in duration (Bion, et al., 2011; Iversen, et al., 2008; Yoshida, et al., 2010).

General Discussion

The presence of the perceptual grouping biases described by the ITL in a nonhuman animal was probed by testing rats' discrimination and segmentation of sequences alternating in pitch (Experiment 1) and sequences alternating in duration (Experiment 2). The ITL states that sequences varying in duration are segmented as iambic groups (i.e., weak-strong), whereas sequences varying in pitch or intensity are segmented as trochaic groups (i.e., strong-weak). Results showed that rats present a trochaic bias for the stream alternating in pitch, but they showed no grouping preference for the stream varying in duration. When we tested human participants with the same stimuli as animals (Experiment 3), we found they grouped both streams following the principles described by the ITL. Regarding the two aims of the present work, these findings allow for two conclusions. First, they show that some perceptual grouping principles that humans use during language processing might be shared across species. Second, they suggest that the two grouping principles described by the ITL are differentially affected by experience.

Our results coincide with previous findings from infant and adult studies that suggest perceptual grouping biases based on duration (Yoshida, et al., 2010), pitch and

duration (Bion, et al., 2011), as well as on intensity and duration (Iversen, et al., 2008; Hay & Saffran, 2012), are differently modulated by experience. They suggest that the trochaic grouping bias based on pitch might be a widely general perceptual principle mostly independent of language experience, while the iambic grouping bias based on duration might be modulated by the linguistic environment and thus might appear in later stages of development. Results such as the ones presented here - suggesting that human and nonhuman animals share the trochaic grouping bias based on pitch - point in this direction and strengthen the idea that the trochaic bias emerges independently of linguistic experience. On the contrary, the fact that we did not observe any evidence of an iambic grouping bias based on duration in a nonhuman animal fits well with the suggestion that this principle might be more dependent on experience with speech stimuli.

In addition, the present results point against the proposal that a trochaic bias is universal and should appear for all sequences varying either in pitch or duration (Allen & Hawkins, 1978). Rats' non preference for either iambic or trochaic pairs during the test phase of Experiment 2 and the control experiment, together with previous research with human adults (Hay & Diehl, 2007; Iversen, et al., 2008), and infants (Bion, et al., 2011; Yoshida, et al., 2010), suggests that the trochaic rhythmic grouping bias is only present in both humans and nonhuman animals under pitch or intensity variations, but not under variations in duration.

Could it be the case that duration random sequences are harder to discriminate from alternating duration sequences than their equivalent in the pitch condition? Research with human adults suggests that irregular temporal patterns might disrupt performance over regular patterns within a session (e.g., Jones & Yee, 1997). Thus, random sequences might be disrupting processing of alternating sequences in our

duration condition. Nevertheless, there was a relatively long ISI (60 sec) between any RDS and any DS in our experiment. This might have mitigated such disrupting effects. Also, we are not aware of any literature suggesting that sequences as the ones used in the present study could disrupt discrimination in animals. We are also not aware of literature testing whether random changes in duration (Experiment 2) could have a greater impact on alternating sequences than random changes in pitch (Experiment 1). However, to compare across experiments we are assuming that changes in pitch in both the alternating and the random sequences are equivalent for the animals to changes in duration. As we have described above, changes in the tones used in the present study are well within the processing range of rats in both dimensions (frequency for pitch and time for duration). This is a good indicator that animals might be processing in a similar way changes across these two dimensions. Thus the differences in our results are not due to changes in one dimension being more easily processed than changes in the other dimension. However, more research would be needed to empirically establish the extent of this equivalence and whether sequences randomly varying in duration (DRS) might have more disrupting effects over more regular sequences (DS) than sequences randomly varying in pitch (PRS) over sequences with regular pitch changes (PS). The results of the present experiments suggest that rats easily learn to discriminate alternating from random sequences in the pitch condition, and that such discrimination leads to a trochaic grouping bias during test. On the contrary, under equivalent conditions, animals did not learn to discriminate alternating from random sequences in the duration condition, and no grouping bias was observed during test.

The fact that both humans and nonhuman animals share the trochaic perceptual grouping bias based on pitch suggests that this might be based on a general perceptual mechanism, neither exclusive to humans nor specific to language, and likely

independent of experience. In addition, our findings might reflect the absence of a universal grouping bias based on duration. As an alternative, we suggest that perceptual grouping based on duration might require previous experience that would direct perception towards the relevant acoustic cues within the input. However, though differentially sensitive to experience, once they are active, both grouping biases may help to bootstrap word order information based on cues of prominence present in speech (Bion, et al., 2011; Nespor, et al., 2008). Finally, the present findings add evidence to research on comparative cognition suggesting that some important aspects of language might be processed by basic perceptual abilities present in both humans and other species. Furthermore, they point towards the idea that these abilities have not evolved for linguistic purposes but that are, nevertheless, used by humans when analyzing the speech input.

Acknowledgements

Footnotes

1. Song by Mark Gatiss, from the popular television series "Doctor Who".

References.

Allen, G. & Hawkins, S. (1978). The development of phonological rhythm. In A. Bell
& J. Hooper (Eds.), *Syllables and segments.* Amsterdam: North-Holland.

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.

Bolton, T. (1894). Rhythm. *The American Journal of Psychology, 6,* 145-238.

Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in
the continuous speech stream: The role of prosody. *Journal of Phonetics, 31,* 585-
598.

Gout, A., Christophe, A., & Morgan, J. (2004). Phonological phrase boundaries
constrain lexical access: II. Infant data. *Journal of Memory and Language, 51,*
547-567.

Hay, J. & Diehl, R. (2007). Perception of rhythmic grouping: Testing the
Iambic/Trochaic law. *Perception & Psychophysics, 69,* 113-122.

Hay, J. & Saffran, J. (2012). Rhythmic grouping biases constrain infant statistical
learning. *Infancy, in press.*

Hayes, B. (1995). *Metrical stress theory: Principles and case studies.* Chicago: The
University of Chicago Press.

Heffner, H., Heffner, R., Contos, C., & Ott, T. (1994). Audiogram of the hooded
Norway rat. *Hearing Research, 73,* 244-247.

Höhle, B., Bijeljac-Babic, R., Herod, B., Weissenborn, J., & Nazzi, T. (2009).
Language specific prosodic preferences during the first half year of life: Evidence
from German and French infants. *Infants Behavior and Development, 32,* 262-
274.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America, 124,* 2263-2271.

Jones, M.R, & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 693-709.

Jusczyk, P., Hirsh-Pasek, K., Kemler Nelson, D., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology, 24,* 252-293.

Jusczyk, P., Cutler, A., & Redanz, L. (1993). Infants' preference for the predominant stress pattern of English words. *Child Development, 64,* 675-687.

Kelly, J. B., Cooke, J. E., Gilbride, P. C., Mitchell, C., & Zhang, H. (2006). Behavioral limits of auditory temporal resolution in the rat: amplitude modulation and duration discrimination. *Journal of Comparative Psychology, 120,* 98-105.

Nespor, M., Shuckla, 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.

Nespor, M., & Vogel, I. (2008). *Prosodic Phonology.* Berlin: Mouton de Gruyter. 1st edition 1986. Dordrecht. Foris.

Ortega-Llebaria, M., & Prieto, P. (2011). Acoustic correlates of stress in central Catalan and castilian Spanish. *Language & Speech, 54*, 73-97.

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.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language

discrimination by human newborns and by cotton-top tamarind monkeys. *Science,*

288, 349-351.

Roger, C., Hasbroucq, T., Rabat, A., Vidal, F., & Burle, B. (2009). Neurophysics of

temporal discrimination in the rat: A mismatch negativity study.

Psychophysiology, 46, 1028-1032.

Saffran, J., Werker, J., & Werner, L. (2006). The infant's auditory world: Hearing,

speech, and the beginnings of language. In: Siegler R, Kuhn D, editors. *Handbook*

of Child Development. Wiley; New York. pp. 58–108.

Toro, J. M., & Trobalón, J. B. (2005) Statistical computations over a speech stream in a

rodent. *Perception & Psychophysics 67,* 867-875.

Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in

language discrimination tasks by rats. *Animal Cognition, 6,* 131-136.

Trehub, S., & Trainor, L. J. (1993). Listening strategies in infancy: The roots of music

and language development. In S. McAdams & E. Bigand (Eds.), *Thinking in*

sound: Cognitive perspectives on human audition (pp. 278-327). Amsterdam:

Elsevier.

Turk, A., & Sawusch, J. (1996) The processing of duration and intensity cues to

prominence. *Journal of the Acoustical Society of America, 99*, 3782-3790.

Woodrow, H. (1909). A quantitive study of rhythm: The effect of variations in intensity,

rate and duration. *Archives of Psychology, 14,* 1-66.

Yip, M. (2006). The search for phonolgy in other species. *Trends in Cognitive Sciences,*

10, 442-446.

Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J., &

Werker, J. F. (2010). The development of perceptual grouping biases in infancy:

A Japanese-English cross-linguistic study. *Cognition, 115,* 356-361.

Figure legends.

Figure 1.

Mean percentage (and standard error bars) of rats' responses during 30 training sessions to sequences varying in pitch (Experiment 1; black triangles) and sequences varying in duration (Experiment 2; white circles). A performance of 50% indicates rats responded equally to alternating sequences and random sequences. Animals did not show any evidence of discriminating sequences varying in duration, while quickly learned to discriminate sequences varying in pitch.

Figure 2.

Mean percentage (and standard error bars) of rats' responses to target pairs (high-low for pitch; short-long for duration) during test. A performance of 50% indicates rats responded equally to trochaic and iambic pairs. Animals in Experiment 1 tended to respond more to pairs with initial prominence (high-low). Animals in Experiment 2 did not show any tendency to respond more to pairs with either initial (long-short) or final (short-long) prominence.
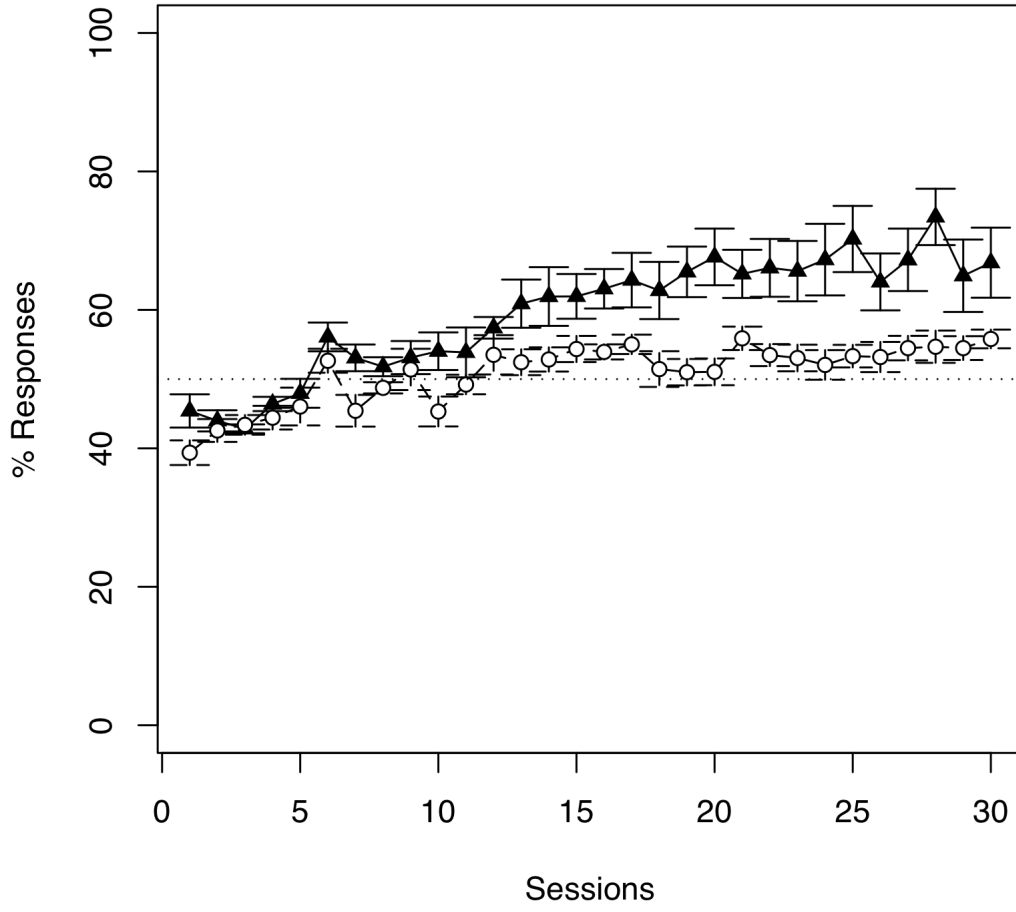
Figure 1.

Figure 2.