

# Do humans and nonhuman animals share the grouping principles of the iambic–trochaic law?

Daniela M. de la Mora · Marina Nesp̄or · Juan M. Toro

Published online: 7 September 2012  
© Psychonomic Society, Inc. 2012

**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 sounds mark group beginnings), and iambic groups over sequences varying in duration (i.e., the longest sounds mark 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 sequences based on pitch variations as trochees, but they do not group sequences varying in duration as iambs. Importantly, humans grouped the same stimuli following the principles of the iambic–trochaic law (Exp. 3). These results suggest the early emergence of the trochaic rhythmic grouping bias based on pitch, possibly relying on perceptual abilities shared by humans and other mammals, whereas the iambic rhythmic grouping bias based on duration might depend on language experience.

**Keywords** Iambic–trochaic law · Comparative cognition · Perceptual bias · Speech perception

*Tick tock, goes the clock/And what now shall we play?/Tick tock, goes the clock/Now summer’s gone away?* As this song

D. M. de la Mora · J. M. Toro (✉)  
Center for Brain and Cognition, Universitat Pompeu Fabra,  
C. Roc Boronat, 138,  
CP. 08018, Barcelona, Spain  
e-mail: juanmanuel.toro@upf.edu

M. Nesp̄or  
European Research Council, PASCAL Project,  
International School for Advanced Studies (SISSA),  
Trieste, Italy

illustrates,<sup>1</sup> 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), according to which iambs correspond to groups with final prominence (weak–strong) and trochees correspond to groups with initial prominence (strong–weak).

Research has suggested that the ITL may play an important role during language processing, supporting speech segmentation based on prosody (Hay & Diehl, 2007; Hayes, 1995; Trehub & Trainor, 1993). More importantly, recent evidence has suggested a strong correlation between prosody and syntax at different hierarchical levels (Nesp̄or 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, Nesp̄or et al. showed 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—that is, whether heads precede or follow their complements—it is proposed that the specific type of prominence that an infant is exposed to might be exploited to acquire the basic word order of its native language. Thus, the general perceptual biases described by the ITL might serve as the stepping

<sup>1</sup> Song by Mark Gatiss, from the popular television series *Doctor Who*.

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 has supported the hypothesis that such grouping principles are present in human adults, regardless of the stress pattern of their native language (Bion, Benavides-Varela, & Nespors, 2011; Hay & Diehl, 2007). In their study, Hay and Diehl presented sequences of tones and sequences of the syllable /ga/ that alternated in either duration or intensity to both English and French speakers. The researchers 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. The results revealed 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, these results suggested that the grouping principles of the ITL are not modulated by the participants' native languages. In a similar vein, Bion et al. asked Italian speakers to listen to a sequence of syllables alternating either in pitch or in duration. The participants were then presented with two pairs of syllables with constant pitch and duration—one respecting and one violating the iambic–trochaic grouping during familiarization—and were asked to judge which of the two pairs of syllables had been adjacent during the familiarization phase. The participants familiarized with pitch-varying sequences remembered better the pairs that had had initial prominence during familiarization. Likewise, the participants familiarized with duration-varying sequences remembered better the pairs that had 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 that 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, Yoshida et al. measured infants' preference for either iambic or trochaic groups. The results showed that only 7-month-old English infants segmented the sequences as iambs. In contrast, 5-month-old English infants and both 5- and 7-month-old Japanese infants showed no preference for either trochaic or iambic sequences, suggesting that 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 this bias might depend on language experience. In contrast, the studies suggest that the trochaic grouping bias based on either intensity (Hay & Saffran, *in press*) or 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 nonlinguistic tone sequences (Hay & Diehl, 2007; Hay & Saffran, *in press*) and are even present in the visual domain (Peña, Bion, & Nespors, 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 that we use for language processing (Yip, 2006). For example, previous studies have found that cotton-top tamarin monkeys (Ramus, Hauser, Miller, Morris, & Mehler, 2000) and rats (Toro, Trobalon, & Sebastián-Gallés, 2003) can discriminate between two languages on the basis of 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 toward 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 interspecific 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 that vary in pitch. In Experiment 2, we approached the complementary question of whether they can group as iambs sequences that vary in duration.

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, 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 the grouping biases based on pitch and duration are differentially sensitive to experience with language (with duration being more sensitive than pitch to such experience; see Bion et al., 2011; Hay & Saffran, *in press*; 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 could observe in a nonhuman animal a bias to group as trochees sequences alternating in pitch. Studies with human adults and infants have reliably observed such a bias for both intensity and pitch variations (e.g., Bion et al., 2011; Hay & Diehl, 2007; Hay & Saffran, *in press*; Iversen et al., 2008). In our study, however, we focused on pitch for several reasons. First, the aim of our study was to investigate whether the ITL, hypothesized to be involved in syntactic bootstrapping (Nespor et al., 2008), is a grouping mechanism shared by nonhuman 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 are less likely to exploit intensity for the perception of phrasal prominence. Since the ultimate goal of our study was to investigate whether a mechanism exploited by infants to acquire language could be shared by a nonhuman mammal, we did not include intensity in our study.

### Method

**Subjects** The subjects were six Long-Evans rats (four male, two female) of 4 months of age. They were food deprived until they reached 80 % of their free-feeding weight but had access to water *ad libitum*. Food was administered after each training session.

**Stimuli** The stimuli were 16 pitch sequences (PSs) and 16 pitch random sequences (PRSs). The PSs were composed of concatenations of sixteen 200-ms pure tones, each alternating in pitch. Importantly, the sequences always included the alternation of a low tone (420 Hz) and a higher tone (525, 630, 735, or 840 Hz, all of which are within the range of hearing frequencies of rats; see, e.g., Heffner, Heffner, Contos, & Ott, 1994). For example, the sequence of tones in a PS could be (in hertz) 420–525–420–840–420–630–420–735–420–840–420–630–420–735–420–525. Half of the PSs started with a low tone, and half with a high tone. The same tones used in the PSs were combined at random to form the PRSs (e.g., 420–420–525–630–420–420–420–735–840–420–420–420–735–630–840–525) so that no systematic alternations of low and higher tones were present. A 200-ms interstimulus interval (ISI) separated all tones. Every sequence lasted 6.2 s and was faded 1 s 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 16 bits.

**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 leverpress responses, and provided reinforcement. A Pioneer Stereo Amplifier A-445 and two E.V. (s-40) speakers, located beside the boxes, were used to present the stimuli.

**Procedure** The rats were trained to press a lever until they reached a stable response rate at a variable ratio of 10 ( $\pm 5$ ) (VR-10 schedule; i.e., the leverpressing response rate at

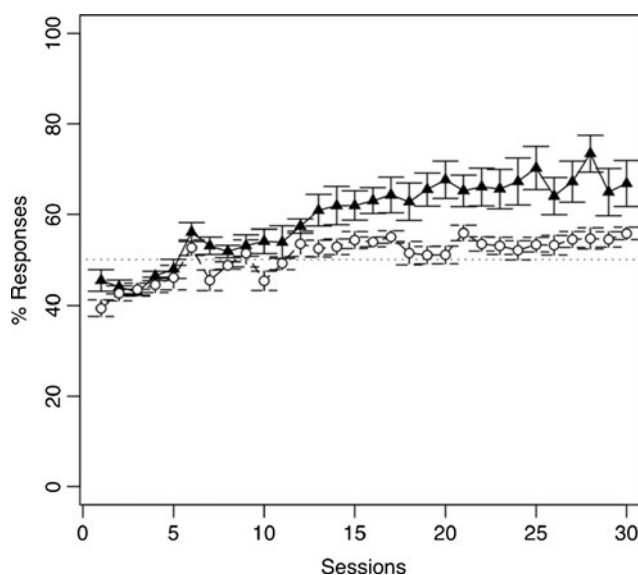
which food was delivered varied between 5 and 15 times from trial to trial). During this time, no stimuli were presented. Training to discriminate across stimuli started once the rats had reached a stable rate of responses. The discrimination training consisted of 30 sessions, one session per day. The logic behind this training procedure was that it would lead rats to discriminate alternating from random sequences and to associate the former with food delivery. The 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 a lever more often after test items that have been grouped on the basis of 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 intersequence interval of 60 s. Sequence presentation was balanced within each session and across sessions, so that 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 ( $\pm 3$ ) (VR-7 schedule; that is, the leverpressing response rate at which food was delivered varied between 4 and 10 times from trial to trial). Food delivery continued for 60 s after a 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' leverpressing responses were registered simultaneously with the presentation of the stimulus and for the 60-s intersequence interval.

After 30 training sessions, a test session was run. Instead of sequences, only pairs of tones were presented: 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 presentations were randomized, with the only restriction being 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, 60 s elapsed between the presentations of consecutive pairs. Leverpressing responses were registered simultaneously with the presentation of a pair and for the 60 s following presentation. Food was delivered after both high–low and low–high pairs in order to avoid any confound of stimulus discrimination with reinforcement schedule. Hence, any difference observed in leverpressing responses would be due to a difference in the way that 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 PSs, and likewise 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 showed no preference, this would mean that they did not segment the PSs in either way, as trochees or as iambs.

## Results and discussion

During training, rats increasingly responded to PSs. To explore how leverpressing responses changed across sessions, we ran a repeated measures ANOVA over the average of leverpressing responses to PSs and PRSs, with Session (1–30) and Stimuli (PS and PRS) as within-subjects factors. This analysis showed a nonsignificant difference between sessions [ $F(29, 145) = 1.380, p = .111$ ], but a significant difference between stimuli [ $F(1, 5) = 33.959, p < .005$ ] and a significant interaction between the two factors [ $F(29, 145) = 13.174, p < .001$ ]. To account for differences in overall levels of responding, mean leverpresses were converted to percentages of responses to PSs and PRSs. A repeated measures ANOVA over the percentages of leverpressing responses to the reinforced stimuli (PSs), with Session (1–30) as the within-subjects factor, yielded a significant difference between sessions [ $F(29, 145) = 11.738, p < .001$ ; see Fig. 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



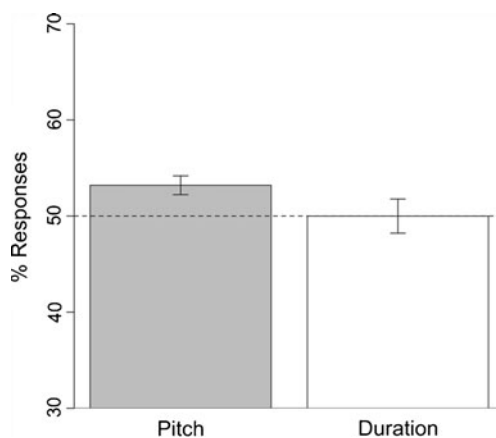
**Fig. 1** Mean percentages (and standard error bars) of rats' responses during 30 training sessions to sequences varying in pitch (Exp. 1; black triangles) and sequences varying in duration (Exp. 2; white circles). A performance of 50 % indicates that rats responded equally to alternating sequences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

test trials, the percentage of responses to trochaic (i.e., high–low) over iambic (i.e., low–high) pairs was significantly above what would be expected by chance [ $M = 53.20\%$ ,  $SD = 2.39$ ,  $t(5) = 3.275$ ,  $p < .05$ ,  $d = 1.893$ ; with chance being equal percentages of responses to trochaic and iambic trials; see Fig. 2], suggesting that rats grouped the PSs into trochees.

Together, these results suggest that rats learned to discriminate sequences alternating in pitch (PSs) from random sequences (PRSs), as they responded differently to PSs and PRSs during training. More relevant to the present study, results from the test phase suggested that they grouped the PSs into trochees and not into iambs. This is reflected in the higher percentage of responses to high–low pairs that exceeded what would be expected if rats were responding at chance after test pairs. This points toward the idea that, like 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 has suggested that this principle might heavily depend on language experience (Bion et al., 2011; Hay & Saffran, *in press*; Yoshida et al., 2010). If so, we should



**Fig. 2** Mean percentages (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 that rats responded equally to trochaic and iambic pairs. The animals in Experiment 1 tended to respond more to pairs with initial prominence (high–low), whereas the animals in Experiment 2 did not show any tendency to respond more to pairs with either initial (long–short) or final (short–long) prominence

not observe this bias to group as iambs sequences varying in duration in other species.

## Method

**Subjects** The subjects were seven new Long-Evans rats (five male, two female) 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 but had access to water ad libitum. Food was administered after each training session.

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

**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 DSs and DRSs did not vary significantly. A repeated measures ANOVA over the averages of leverpressing responses to DSs and DRSs, with Session (1–30) and Stimuli (DS and DRS) as within-subjects factors, showed a nonsignificant difference between sessions [ $F(29, 174) = 1.4126$ ,  $p = .086$ ] and stimuli [ $F(1, 6) = 1.003$ ,  $p = .335$ ], but a significant interaction between them [ $F(29, 174) = 5.762$ ,  $p < .001$ ]. As in Experiment 1, mean leverpressing responses were converted

to percentages of responses. A repeated measures ANOVA over these percentages of leverpressing responses to the reinforced stimuli (DSs), with Session (1–30) as the within-subjects factor, yielded a significant difference between sessions [ $F(29, 174) = 6.508, p < .001$ ; see Fig. 1]. This difference is explained by an increase in leverpressing responses throughout 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 [ $M = 49.99\%$ ,  $SD = 7.67$ ;  $t(6) = -0.002, p = .998, d = -0.001$ ; see Fig. 2]. Together, these results suggest that, during training, rats did not discriminate between DSs and DRs, nor did they group the DSs into iambs, as reflected by chance performance during the test. Moreover, it could mean that the iambic grouping principle observed in human adults and infants is not a universal bias, but rather a language-experience-dependent trait.

A comparison of the percentages of responses to the reinforced stimuli (PSs in Exp. 1 and DSs in Exp. 2) during the training phase, with Session (1–30) as a within-subjects factor and Experiment (1 and 2) as a between-subjects factor, yielded significant differences between sessions [ $F(29, 299) = 17.895, p < .001$ ] and experiments [ $F(1, 11) = 11.583, p < .01$ ], as well as a significant interaction between the factors [ $F(29, 319) = 3.067, p < .001$ ]. These results suggest that the differences in rats' performance during both experiments were due to a differential processing of the stimuli, independently of the training procedure. That is, rats easily extracted information patterns over pitch-varying but not over duration-varying sequences. This difference was further reflected by above-chance response rates during test for trochaic pairs based on pitch variations (Exp. 1), but not for either iambic or trochaic pairs based on duration variations (Exp. 2).

A remaining question regarding the results of Experiment 2 was whether they could be explained by the rats' lack of sensitivity to the acoustic changes that 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. 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 be interpreted neither as rats' inability to process the durations of the tones used nor as an inability 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 by Roger et al.

Nevertheless, to directly test the possibility that longer tones could trigger iambic grouping in rats, we ran a control condition with nine new rats. The 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, 1,100, 1,400 or 1,700 ms (more than twice the duration of the tones used in Exp. 2). The results from this control experiment closely replicated the results of Experiment 2. Throughout the training phase, rats increased their leverpressing responses, but during the test phase, they did not press the lever more often for iambic (short–long) than for trochaic (long–short) test pairs [ $M = 50.59\%$ ,  $SD = 3.96$ ;  $t(8) = 0.453, p = .663, d = 0.210$ ], suggesting that they did not tend to group the alternating sequences as either iambs or trochees. Moreover, a comparison between the test phases of Experiment 2 and the control experiment yielded a nonsignificant difference between them [ $t(14) = -0.206, p = .840, d = 0.098$ ], suggesting that the rats were equally unable to group as either iambs or trochees sequences of longer tones varying in 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 DSs, 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 nonreinforced sequences. A final concern was whether the stimuli used in the present study would actually be grouped by humans following the principles of the ITL. To test this, we ran 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 (Exp. 1), but do not tend to group as iambs sequences alternating in duration (Exp. 2). We proposed that this lack of iambic grouping observed in animals might indicate that some experience (e.g., with language) might be necessary for an iambic grouping bias to emerge. However, it could also be that the specific sequences of tones varying in duration that 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 (Hay & Saffran, *in press*; Yoshida et al., 2010) has been based on 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 whether the alternating sequences presented to the rats in the previous experiments would elicit in humans the grouping biases predicted by the ITL.

## Method

**Participants** A group of 20 undergraduate students from Universitat Pompeu Fabra took part in this experiment. They were all native speakers of Spanish and received monetary compensation for their participation.

**Stimuli** The stimuli were the same alternating sequences used in Experiments 1 (PSs) and 2 (DSs).

**Procedure** We presented participants with the alternating sequences used in Experiments 1 and 2. The order of presentation of sequences varying in pitch and sequences varying in duration was balanced (with no more than two 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 Exps. 1 and 2). A pause of 500 ms separated the test pairs. Participants were asked to indicate which pair better corresponded with the sequence that they had previously heard, and they had no time limit to answer. All of the participants were tested in a silent room, wearing headphones. The experiment was presented on a Macintosh OS X–based laptop using the PsyScope X B57 experimental software.

## Results and discussion

After listening to sequences alternating in pitch, participants significantly preferred trochaic (high–low) pairs [ $M = 59.17\%$ ,  $SD = 15.03$ ;  $t(19) = 2.727$ ,  $p < .05$ ,  $d = 0.863$ ]. After listening to sequences alternating in duration, participants significantly preferred iambic (short–long) pairs [ $M = 67.08\%$ ,  $SD = 19.40$ ;  $t(19) = 3.938$ ,  $p < .005$ ,  $d = 1.245$ ]. 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 were 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 groups segmented pitch-alternating sequences in a similar manner [ $t(24) = 0.96$ ,  $p = .347$ ,  $d = 0.555$ ], but that they performed significantly differently for sequences alternating in duration [ $t(25) = 2.25$ ,  $p < .05$ ,  $d = 1.159$ ]. This suggests that 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 (Exp. 1) and sequences alternating in duration (Exp. 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). The 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 the animals (Exp. 3), we found that 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 have suggested that perceptual grouping biases based on duration (Yoshida et al., 2010), pitch and duration (Bion et al., 2011), and intensity and duration (Hay & Saffran, *in press*; Iversen et al., 2008) 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 argue against the proposal that a trochaic bias is universal and should appear for all sequences varying in either pitch or duration (Allen & Hawkins, 1978). Rats' nonpreference 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 that duration random sequences are harder to discriminate from alternating duration sequences than are their equivalents in the pitch condition? Research with human adults has suggested 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 s) between any DRSs and any DSs in our experiment, which might have mitigated such disrupting effects. Also, we are not aware of any literature suggesting that such 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 (Exp. 2) could have a greater impact on alternating sequences than do random changes in pitch (Exp. 1). However, by comparing across experiments, we have assumed that changes in pitch in both the alternating and random sequences are equivalent for animals to changes in duration. As we have described above, the 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 (DRSs) might have more disrupting effects over more regular sequences (DSs) than sequences randomly varying in pitch (PRSs) over sequences with regular pitch changes (PSs). 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, that is 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 toward the relevant acoustic cues

within the input. However, though the biases are differentially sensitive to experience, once they are active, both may help to bootstrap word order information on the basis of 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 toward the idea that these abilities have not evolved for linguistic purposes but are, nevertheless, used by humans when analyzing speech input.

**Author note** This research was supported by grants Consolider Ingenio CSD2007-00012 and PSI2010-20029, as well as by ERC grant agreement n.312519 to J.M.T. and by (FP7/2007-2013)/ERC grant agreement n.269502 (PASCAL) to M.N. We thank Tere Rodrigo and the staff from the Laboratorio de Psicología Animal of the Universitat de Barcelona for their help with the experiments, and three anonymous reviewers for their fruitful comments.

## References

- Allen, G., & Hawkins, S. (1978). The development of phonological rhythm. In A. Bell & J. Hooper (Eds.), *Syllables and segments*. Amsterdam, The Netherlands: 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. F., & Diehl, R. (2007). Perception of rhythmic grouping: Testing the Iambic/Trochaic law. *Perception & Psychophysics*, *69*, 113–122.
- Hay, J. F., & Saffran, J. R. (in press). Rhythmic grouping biases constrain infant statistical learning. *Infancy*. doi:10.1111/j.1532-7078.2011.00110.x
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago, IL: 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. *Infant Behavior & Development*, *32*, 262–274.
- Iverson, 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., Cutler, A., & Redanz, L. (1993). Infants' preference for the predominant stress pattern of English words. *Child Development*, *64*, 675–687.



- 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.
- 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., 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.
- Ortega-Llebaria, M., & Prieto, P. (2011). Acoustic correlates of stress in central Catalan and Castilian Spanish. *Language and 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 tamarin 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 R. Siegler & D. Kuhn (Eds.), *Handbook of child development* (pp. 58–108). New York, NY: Wiley.
- Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, *67*, 867–875. doi:10.3758/BF03193539
- Toro, J. M., Trobalón, 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, The Netherlands: 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 quantitative study of rhythm: The effect of variations in intensity, rate and duration. *Archives of Psychology*, *14*, 1–66.
- Yip, M. (2006). The search for phonology 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.