

Understanding our perception of rhythm

July 9 2021



Credit: CC0 Public Domain

Scientists have long known that while listening to a sequence of sounds, people often perceive a rhythm, even when the sounds are identical and equally spaced. One regularity that was discovered over 100 years ago is the Iambic-Trochaic Law: when every other sound is loud, we tend to hear groups of two sounds with an initial beat. When every other sound

is long, we hear groups of two sounds with a final beat. But why does our rhythm perception work this way?

In a recent study in *Psychological Review*, McGill University Professor Michael Wagner shows that the rhythm we perceive is a result of the way listeners make two separate types of decisions, one about grouping (which syllables or tones group together) and the other about prominence (which syllables or tones seem foregrounded or backgrounded). These decisions about grouping and prominence mutually inform each other.

The findings may deepen our understanding of [speech](#) and [language](#) processing, with potential implications in a wide range of areas, including teaching, speech therapy, improving synthesized speech, and improving speech recognition systems.

What do scientists know about our perception of rhythm?

Sequences of tones and syllables are often perceived as rhythmically grouped. This is true even if all tones or syllables in a sequence are acoustically identical and equally spaced. In a sequence of otherwise equal sounds, listeners tend to hear a series of trochees (groups of two sounds with an initial beat) when every other sound is louder, and they tend to hear a series of iambs (groups of two sounds with a final beat) when every other sound is longer.

Since this generalization was first discovered by Thaddeus Bolton in 1894, it has been replicated in many studies, including those involving speech development in children. Today the consensus is out on whether Bolton's Iambic-Trochaic Law is a universal phenomenon, or whether it results from language experience. Although well-established for over a hundred years, the source of the phenomenon has remained unclear.

What did you discover?

We found that these rhythmic perceptions are not really about iambs or trochees. For a given stimulus, we make two separate decisions; grouping, or how we parse the signal into smaller chunks, and prominence, or which sounds are foregrounded or backgrounded. Together, these decisions result in our rhythmic intuitions. The two decisions are mutually informative, just like our visual system makes mutually informative decisions about the size and distance of an object. If we think of the object as close by, we infer that it's smaller than if we think of it as far away. This can lead to comical 'forced perspective effects,' as in this image of the Eiffel tower—we know that it is big and appears small because it's far away, but the girl apparently touching its peak makes it appear small and close by.

The results of the study suggest that it is these kinds of inferences that are the reason why, when listening to a series of syllables like ...bagabagaba..., we spontaneously perceive it as repetitions of either the word 'baga' or 'gaba.'" The words simply seem to pop out even though acoustically, it is just an unstructured sequence of sounds. In the case of tone sequences, where we can't recognize individual words, we simply perceive these effects as a regular iambic or trochaic rhythm.

You can try out the study and even participate yourself at the prosodylab's virtual field station.

What are the next steps?

If the effects observed in this study are universal and apply across languages, this would offer new insights into how newborns might begin to be able to parse the signal when they first get exposed to language, and it would also provide new opportunities for speech technology to

improve speech synthesis and speech recognition. However, earlier cross-linguistic work on the Iambic-Trochaic Law suggests that there is substantial variation between languages when it comes to [rhythm](#).

My team has recently started exploring how [different languages](#) really are once one teases apart the two dimensions of grouping and prominence, like what the present study did for English. Initial results show that once one disentangles the dimensions, there is substantial invariance across languages.

More information: Michael Wagner, Two-dimensional parsing of the acoustic stream explains the Iambic–Trochaic Law., *Psychological Review* (2021). [DOI: 10.1037/rev0000302](https://doi.org/10.1037/rev0000302)

Provided by McGill University

Citation: Understanding our perception of rhythm (2021, July 9) retrieved 11 May 2024 from <https://medicalxpress.com/news/2021-07-perception-rhythm.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--