

Speech rhythm as naturally-occurring and culturally-transmitted behavioral patterns

Mikhail Ordin

BCBL – Basque Centre on Cognition, Brain and Language and Ikerbasque – Basque Foundation for Science, Sebastian, Spain

Address for correspondence: Basque Centre on Cognition, Brain and Language, Paseo Mikeletegi 69, San Sebastian, 20009 Spain. m.ordin@bcbl.eu

Short title: *Rhythm: speech processing, acquisition, evolution*

Abstract:

Rhythm is fundamental to every motor activity. Neural and physiological mechanisms that underlie rhythmic cognition in general and rhythmic pattern generation in particular are evolutionary ancient. As speech production is a kind of motor activity, investigating speech rhythm can provide insight into how general motor patterns have been adapted for more specific use in articulation and speech production. Studies on speech rhythm may further provide insight into the development of speech capacity in humans. As speech capacity is putatively a pre-requisite for developing a language faculty, studies on speech rhythm may cast some light on the mystery of language evolution in the human genus. Hereby, we propose an approach to exploring speech rhythm as a window on speech emergence in ontogenesis and phylogenesis, as well as on diachronic linguistic changes.

Keywords: *speech evolution, speech acquisition, speech processing, speech rhythm, language change*

Speech production is a kind of motor activity, and rhythmic patterns are characteristic of movements produced by biological systems. The mechanisms for controlling motor rhythm have been inherited by mammals from their distant evolutionary ancestors and are shared by all vertebrates¹⁻⁵. Although every human individual possesses the same neural, anatomical, and physiological mechanisms for controlling speech rhythm, rhythmic patterns differ across languages⁶. These language-specific patterns of motor activity are based on the same physiological mechanisms, yet they are shaped by the peculiarities of the ambient language and culturally transmitted within human populations sharing a given language. Consequently, studying speech rhythm opens a window onto nature–nurture interactions in language use and acquisition.

As speech production is a motor activity based on neural, anatomical, and physiological mechanisms, rhythmic patterns are naturally occurring behavioral patterns. At the same time, rhythmic patterns are language specific, i.e., they differ between languages, and thus are culturally transmitted patterns of behavior. The best methodological approach for investigating such naturally-occurring yet culturally-transmitted behaviors is Tinbergen's framework.⁷⁻⁹ Tinbergen suggested that study of behavioral patterns be conducted in four different directions:

- 1) Development of a certain pattern of behavior in ontogenesis (i.e., considering how rhythmic patterns are acquired by individuals);
- 2) Development of a certain pattern of behavior in phylogenesis (i.e., considering how language-specific rhythmic patterns emerge in populations sharing the same language(s), and which components of rhythmic cognition humans share with non-human animals);
- 3) Functional load (i.e., considering how rhythmic patterns facilitate communication and how better capacity for rhythmic cognition might enhance the fitness of individual organisms);
- 4) Emergence and mechanisms (i.e., considering the underlying mechanisms that lead to the emergence of speech rhythm and allow for culture-specific rhythmic modulations within biological constraints as well as constraints on learnability and processability).

A recent special issue of *Ann. N.Y. Acad. Sci.*¹ addresses each of these four major lines of research.

Development in ontogenesis

Addressing this broad research issue requires an exploration of the development of speech rhythm during the course of first and second language acquisition in humans and development of rhythm in communicative signals emitted by non-human animals. Earlier studies showed that speech rhythm develops from more regular (isochronous, with quasi-equal syllable and vowel durations) towards more irregular rhythmic patterns, and that this trajectory is typical of both first and second language acquisition¹⁰⁻¹³. The development of speech rhythm in language acquisition is governed not only by language-specific, and thus culturally-transmitted factors (i.e., dominant rhythmic patterns in the native language of an adult learning a second language), but also by universal factors (i.e., by mechanical constraints on speech production and the properties of the mammalian auditory system). Thus, adult native speakers of languages that exhibit irregular rhythmic patterns cannot simply transfer these rhythmic skills from their native to a second language, but have to develop the required degree of irregularity anew, similar to the way that children develop rhythmic skill in their first language^{10,11}.

To the best of my knowledge, there have been no systematic studies of the developmental aspects of rhythm acquisition across different species to date. We do not know whether the emergent rhythmic patterns in vocal signals are refined in accordance with the norms of specific species or even specific populations within various species, and, if so, whether there are similarities in developmental trajectories across species.

Another understudied aspect of the ontogenesis of rhythmic cognition is rhythm (re)acquisition when specific neural circuits engaged in language processing and speech production or related underlying physiological mechanisms are no longer available, as may be the case for patients recovering from stroke or people with cochlear implants. The paper by Jiam and Limb¹⁴ focuses on rhythm processing by cochlear implant users.

Development in phylogenesis

¹ *Ann. N.Y. Acad. Sci.* XXXX: XX- XX (2019).

The development of rhythmic patterns over the course of historical language changes has also received insufficient attention. We know that certain phonological, phonotactic, and syntactic parameters have been shown to correlate either with irregular or regular distribution of vowel onsets^{6,15,16}. The existence of complex syllables with consonantal clusters, for example, increases irregularity in the distribution of vowel onsets. In many languages, syllable complexity increases over the course of diachronic change—as happened, for instance, with the loss of vowels in Germanic¹⁷ and Romance¹⁸ languages. The appearance of geminates (double consonants) in Australian languages also resulted in higher syllable complexity and thus more irregular rhythms¹⁹. Changes in stress patterns in typologically diverse languages is easily modelled as a shift away from isochronous syllable durations²⁰. Thus, diachronic language changes in some respects recapitulate the rhythmic changes observed during first and second language acquisition in ontogenesis. However, it remains a question for further exploration whether changes in grammar drive phonological changes, thus leading to different rhythmic patterns in spoken speech, or whether phonetic changes modulate diachronic changes in syntactic and morphological structures.

Phylogenetic changes in rhythm can be both culturally transmitted and biologically determined. Culturally transmitted language changes have been explored within the iterated language learning paradigm^{21–23}. Iterated learning means acquiring a behavioral pattern by observing a different individual performing that behavior. The output of “one generation of learners” is used as the model input for the “following generation”, and such iterations continue for multiple generations (a) revealing the emergence of a structured behavior from a chaotic mix of incoherent behavioral patterns, (b) leading to ritualization of behavior in later generations, and (c) capturing the dynamics of cultural cross-generational transmission. Importantly, in relation to rhythmic cognition, iterated learning paradigms show the clearly biological basis of cultural transmission, which makes it difficult to determine which rhythmic aspects are naturally determined and which are culturally transmitted²⁴.

Biologically determined aspects of rhythmic cognition are best investigated in cross-species studies. Importantly and ideally, all aspects of rhythmic cognition should be studied, including periodic motor pattern generation, pulse extraction, and, finally, beat entrainment and meter induction²⁵. Different aspects of rhythmic cognition may be developed similarly or to different extents across species (while periodic motor pattern generation is well established across species, beat entrainment seems to be restricted to species capable of vocal learning, and meter induction has so far been confirmed only for primates, even those who are not vocal learners). These components of rhythmic cognition have different functional roles, and probably also have different biological bases and evolutionary histories²⁵. In the recent *Ann. N.Y. Acad. Sci.* issue, several papers focus on the developmental and evolutionary aspects of speech rhythm, including reviews by Ravignani et al.²⁶, Filippi et al.^{xxx}, and Fitch²⁷ whose discussions masterfully encompasses both ontogenetic and phylogenetic views on the emergence of rhythmic patterns in speech. As speech capacity—the capacity for vocal communication—is putatively a pre-requisite for developing a language faculty²⁸, studies on speech rhythm may cast some light on the mystery of language evolution in the human genus.

Functional load

Complex behaviors cannot evolve without being functional, and intermediate stages in the evolution of each component of rhythmic cognition should exhibit adaptive value and gradual development in all genera (see also Lieberman²⁹). At the same time, strict parallel or consecutive development of components is not a prerequisite for the emergence and evolution of a specific component of rhythmic cognition. Thus, for a comprehensive understanding of the role of speech rhythm in language evolution, acquisition, and change, it is useful to address the question of the ultimate mechanisms of rhythmic behavior: how rhythmic cognition improves the fitness of individuals (and, maybe, languages as well).

Speech rhythm has often been studied as a cue for controlling the flow of conversation^{30–32}. Synchronization of the rhythmic patterns of conversation partners has frequently been mentioned as a reliable correlate of communication success: the better the degree of synchronization of speech rhythm around turn transitions, the higher the communication success.^{30,32–34} Rhythmic continuity across utterances found in typologically distinct languages is scarcely, if at all, modulated by cultural differences, suggesting strong universal influences on how rhythm is used to coordinate interlocutors' utterances in live interactions³⁵. The evolutionary roots of this faculty are explored in the cross-species study by Ravignani *et al.*²⁶

Acoustic and temporal characteristics inherent to different rhythms, especially the degree of temporal isochrony between salient acoustic events in vocal signals, have been associated with various affective states, affiliative and pro-social behaviors, and cooperative urges^{36,37}. Corncrakes (a bird in the rail family), for example, switch from regular rhythm with isochronous intervocalization intervals to irregular rhythm when signaling aggression³⁸. Monkeys, including gelades, baboons, macaques, and marmosets, produce specific rhythms with lip-smacking frequency at around 4–5Hz to signal affiliation^{39–41}. In dolphins, specific motor rhythm patterns and synchronization of motor behavior between interacting individuals signals cooperation⁴². Polyanskaya *et al.*⁴³ take this further to explore whether humans make pragmatic inferences regarding the degree of cooperation between interacting agents when listening to utterances with different rhythmic patterns.

Another possible function of speech rhythm is expressing the identity of the speaker. Speaker's identity is always echoed in speech, and prosody is one of the most important carriers of a speaker's identity⁴⁴. Recently, Ordin and Mennen⁴⁵ found within-speaker between-language differences in the pitch range of Welsh-English bilinguals. They suggested that speakers who want to emphasize their "Welshness" varied their pitch range between languages to a greater extent than those speakers who cared less about expressing their identity as Welsh speakers. Presumably, prosody encoded by rhythmic properties can also be used to emphasize identity, which has strong evolutionary implications for the role of speech rhythm in language change^{46,47}. This possibility needs to be empirically explored.

Finally, rhythm is crucial to speech processing, e.g. for segmenting a continuous stream of syllables into its discrete constituents (phrases and words)^{48,49}, for adapting to non-native accented speech and evaluating the degree of proficiency of foreign language learners,^{50,51} and for the purposes of language acquisition^{52,53}. This important function of speech rhythm is further highlighted by Goswami⁵⁴.

Mechanisms underlying rhythm cognition

The fourth research direction for exploring how naturally occurring behavioral patterns are further modified by the cultural environment relates to proximate mechanisms: the biological and computational bases of rhythm processing, including the neural underpinning of rhythmic cognition.

A major biological mechanism involved in speech rhythm is the ability to couple neural oscillations to external acoustic rhythms. Speech comprehension is also correlated with the success of this acoustic signal-cortex frequency coupling (i.e., the similarity between the frequencies of the temporal envelopes of the speech signal and the listener's cortical activity), and with phase locking between these two temporal envelopes⁵⁵. The remarkable correspondences detected between the amplitude modulation spectrum of the acoustic waveform of spoken utterances and the EEG response in different frequency bands (the so-called beta band, 12–30Hz, and theta band, 4–8Hz) reveal the entrainment of neural to acoustic rhythms. Information in these frequency ranges of the EEG waveform is sufficient for discriminating the specific words and sentences that were perceived by the listener when the EEG was recorded^{56,57}. Several years ago, Suppes, Han, Epelboim, & Lu⁵⁸ recruited participants to listen to a large set of sentences while EEG was recorded. Averaged EEG waveforms from 10 listeners were produced for each separate sentence and then compared sentence-wise to the individual EEG waveforms of the other 10 listeners. Pair-wise matching between averaged and individual EEG waveforms was used to identify (by means of the best match) which sentence each individual participant was listening to when a certain EEG response was obtained. The recognition rate was above 80%, and in some cases above 90%. This means that brain oscillations coupled with the acoustic signal at certain frequencies, and this coupling was similar across all listeners.

Proximate mechanisms for rhythmic expression and perception closely interact with ultimate mechanisms, and the emergence and evolution of proximate mechanisms is dependent on these ultimate mechanisms. The proximate mechanism of speech rhythm processing – entrainment of neural oscillations to external rhythms – has been found in many animals, not restricted to mammalian species^{59,60}. The mystery of these ubiquitous cross-species similarities in neural architecture and specifically the neural mechanisms honed for processing rhythmic acoustic input was tackled by Richard Turner in 2010⁶¹. He explored the statistical properties of natural sounds (such as running water, wind, fire and rain) and showed that the auditory system is optimized for processing natural sounds in the time-frequency domain. The “auditory texture” of artificial sounds (e.g., produced by mechanical means and not occurring in the natural environment) is distinctly different from that of natural sounds, in terms of summary statistics. Importantly, the summary statistics of natural sounds, obtained using amplitude demodulation, time-frequency decomposition and sub-band demodulation, are

similar to those found in natural speech. Thus, the neural mechanisms that initially evolved for processing a range of natural sounds also determined the parameters of speech.

Speech was shaped by the neural architecture in order to be processable by pre-existing neural mechanisms. Syllable durations in naturalistic speech in human languages universally fall within the limit of 100–240 msec., which is also within the theta range of brain oscillations⁶². Manipulations that leave syllable durations within the 100–240 msec. limits may affect comprehensibility to a certain extent but do not make speech totally unintelligible. However, if syllable durations are artificially manipulated to fall outside these limits, speech becomes unintelligible to human listeners. If fast (compressed) speech is divided into 40 msec blocks and each speech interval is followed by a 160-msec silent pause (providing 200 msec timeframes for each syllable) intelligibility is restored. This may be due to the fact that each speech interval falls within 1 period of theta oscillation at 5 Hz⁶³ suggesting that speech is intelligible as long as the speech rhythm and brain rhythms in the theta frequency range are coupled. Acoustic-cortical coupling is a prerequisite for speech comprehension, and acoustic-neuronal phase locking is set during speech processing by the brain⁵⁵.

Based on the results outlined above, it is possible to hypothesize that brain oscillations are probably related to auditory processing either because they provide a temporal window for speech segmentation (e.g., segmenting the continuous speech stream into syllables, phrases and larger constituents), and/or because they constitute neural representations of linguistic structural units, e.g., syllables. This coupling of syllabic and brain rhythms can further be modelled as a system of coupled oscillators, with changes in one oscillator influencing the other. Brain rhythms impose certain constraints on what rhythmic patterns in the speech signal are processable, and which are not. Rhythmic patterns in the ambient language, in turn, influence these neural oscillations within physiologically possible constraints. This model can be described in terms of dynamical system theory and allow brain rhythm to be both a neural representation of a linguistic unit and a temporal window for segmentation of the incoming speech stream through mutual influence.

Considerable work is still required for a complete description of the rhythmic interactions between cortical (and subcortical) regions underlying sensory-motor synchronization, beat induction and meter perception. Some of the studies in [Ann NY Acad Sci XXXX \(2019\)](#) tackle the problem of the proximate mechanisms of rhythmic cognition using different methodological approaches. Tichko and Large⁶⁴ use a modeling approach to explore how neural oscillations emerge and develop in ontogenesis in order to fine-tune to culture-specific musical rhythms. Lizarazu, Lallier, and Molinaro⁶⁵ report an experimental study. They focused on the interaction of slow, delta (<4Hz) and theta rhythms (4–8Hz), as well as fast, gamma (30–40Hz) rhythm emerging in the left and right hemispheres during speech perception to understand the nested coupling of neural oscillations at different frequencies, and the entrainment of the whole system to the speech signal. Lieberman²⁹ provides an extensive literature review to give an overview of the genetic underpinning of neural substrates and networks involved in rhythm processing. He advocates the claim of evolutionary continuity in proximate mechanisms of rhythmic cognition.

The original papers in [Ann NY Acad Sci XXXX \(2019\)](#) will find the readership among an interdisciplinary audience whose research spans the fields of ecology, evolution and ethology, linguistics, comparative psychology, human evolutionary sciences, and neuroscience, which should open the door to the exciting field of rhythmic cognition for scientists who might not even have realized before that their studies are relevant to a wider community of ardent professionals working in other, sometimes seemingly distant, disciplines.

Acknowledgments

This research is supported by the Basque Government through the BERC 2018 -2021 program and by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation SEV-2015-0490 and through project RTI 2018-098317-B-I00 funded by AEI/FEDER.

References

1. MacNeilage, P. F. Evolution of the mechanism of language output: Comparative neurobiology of vocal and manual communication. in *Approaches to the evolution of language: Social and cognitive bases*. (ed. Hurford, J.R., Studdert-Kennedy, M.G., Knight, C.) 222–241 (Cambridge University Press, 1997).
2. MacNeilage, P. F. The frame/content theory of evolution of speech production. *Behav. Brain Sci.* **21**, 499–511 (1998).
3. MacNeilage, P. F. *The origin of speech*. (New York, 2010).
4. Cohen, A. H. Evolution of the vertebrate central pattern generator for locomotion. in *Neural control of rhythmic movement in vertebrates*. (ed. Cohen, A.H., Rossignol, S., Grillner, S.) 129–166 (Wiley, 1988).
5. Rossignol, S., Lund, J. & Drew, T. The role of sensory inputs in regulating patterns of rhythmical movements in higher vertebrates. A comparison between locomotion, respiration and mastication. in *Neural control of rhythmic movement in vertebrates*. (ed. Cohen, A.H., Rossignol, S., Grillner, S.) 201–283 (Wiley, 1988).
6. Ramus, F., Nespor, M. & Mehler, J. Correlates of linguistic rhythm in the speech signal. *Cognition* **73**, 265–92 (1999).
7. Tinbergen, N. On aims and methods of Ethology. *Z. Tierpsychol.* **20**, 410–433 (2010).
8. Bateson, P. & Laland, K. N. Tinbergen's four questions: an appreciation and an update. *Trends Ecol. Evol.* **28**, 712–718 (2013).
9. Hogan, J. A. A framework for the study of behavior. *Behav. Processes* **117**, 105–113 (2015).

10. Ordin, M. & Polyanskaya, L. System. *System* **42**, 244–257 (2014).
11. Ordin, M. & Polyanskaya, L. Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *J. Acoust. Soc. Am.* **138**, 533–44 (2015).
12. Payne, E., Post, B., Astruc, L., Prieto, P. & Vanrell, M. del M. Measuring Child Rhythm. *Lang. Speech* **55**, 203–229 (2012).
13. Allen, G. D., Hawkins, S. The development of phonological rhythm. in *Syllables and Segments*. (ed. Bell, A., Hooper, J.) 173–185 (North-Holland, 1978).
14. Jiam, N. T. & Limb, C. J. Rhythm processing in cochlear implant-mediated music perception. *Ann. N. Y. Acad. Sci.* nyas.14130 (2019). doi:10.1111/nyas.14130
15. Dauer, R. Stress-timing and syllable-timing reanalyzed. *J. Phon.* **11**, 51–62 (1983).
16. Schiering, R. The phonological basis of linguistic rhythm: cross-linguistic data and diachronic interpretation. *Lang. Typology Universals* **60**, 337–359 (2007).
17. Venneman, T. Systems and changes in Early Germanic phonology: A search for hidden identities. in *Germania: Comparative studies in the Old Germanic languages and literatures* (ed. Calder, D., Christy, T.) 45–65 (Woodbridge, 1988).
18. Horne, M. The clitic group as a prosodic category in Old French. *Lund Work. Pap. Linguist.* **35**, 99–111 (1989).
19. Blevins, J. *Evolutionary Phonology*. (Cambridge University Press, 2004). doi:10.1017/CBO9780511486357
20. Bybee, J. L., Chakraborti, P., Jung, D. & Scheibman, J. Prosody and Segmental Effect Some Paths of Evolution for Word Stress. *Stud. Lang.* **22**, 267–314 (1998).
21. Smith, K., Kirby, S. & Brighton, H. Iterated Learning: A Framework for the Emergence of Language. *Artif. Life* **9**, 371–386 (2003).
22. Smith, Kenny, Brighton, Henry, Kirby, S. Complex systems in language evolution: the cultural emergence of compositional structure. *Adv. Complex Syst.* **06**, 537–558 (2003).
23. Brighton, H., Smith, K. & Kirby, S. Language as an evolutionary system. *Phys. Life Rev.* **2**, 177–226 (2005).
24. Ravnani, A., Delgado, T. & Kirby, S. Musical evolution in the lab exhibits rhythmic universals. *Nat. Hum. Behav.* **1**, 0007 (2017).
25. Tecumseh Fitch, W. The biology and evolution of rhythm: Unravelling a paradox. in

- Language and Music as Cognitive Systems* 73–95 (Oxford University Press, 2012).
doi:10.1093/acprof:oso/9780199553426.003.0009
26. Ravignani, A. *et al.* Rhythm in speech and animal vocalizations: a cross-species perspective. *Ann. N. Y. Acad. Sci.* (2019). doi:10.1111/nyas.14166
 27. Fitch, W. T. Sequence and hierarchy in vocal rhythms and phonology. *Ann. N. Y. Acad. Sci.* nyas.14215 (2019). doi:10.1111/nyas.14215
 28. Fitch, W. T. Unity and diversity in human language. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **366**, 376–88 (2011).
 29. Lieberman, P. The antiquity and evolution of the neural bases of rhythmic activity. *Ann. N. Y. Acad. Sci.* nyas.14199 (2019). doi:10.1111/nyas.14199
 30. Cowley, S. J. Conversational functions of rhythmical patterning: A behavioural perspective. *Lang. Commun.* **14**, 353–376 (1994).
 31. Auer, P., Couper-Kuhlen, E. & Müller, F. *Language in time : the rhythm and tempo of spoken interaction.* (Oxford University Press, 1999).
 32. Szczepek Reed, B. Speech rhythm across turn transitions in cross-cultural talk-in-interaction. *J. Pragmat.* **42**, 1037–1059 (2010).
 33. Couper-Kuhlen, E. & Auer, P. On the contextualizing function of speech rhythm in conversation. in 1 (1991). doi:10.1075/pbns.6.2.02cou
 34. Couper-Kuhlen, E. *English Speech Rhythm.* **25**, (John Benjamins Publishing Company, 1993).
 35. Stivers, T. *et al.* Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci.* **106**, 10587–10592 (2009).
 36. Lang, M. *et al.* Lost in the Rhythm: Effects of Rhythm on Subsequent Interpersonal Coordination. *Cogn. Sci.* **40**, 1797–1815 (2016).
 37. Lang, M., Bahna, V., Shaver, J. H., Reddish, P. & Xygalatas, D. Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biol. Psychol.* **127**, 191–197 (2017).
 38. Ręk, P. & Osiejuk, T. S. Sophistication and simplicity: conventional communication in a rudimentary system. *Behav. Ecol.* **21**, 1203–1210 (2010).
 39. Ghazanfar, A. A. & Takahashi, D. Y. Facial expressions and the evolution of the speech rhythm. *J. Cogn. Neurosci.* **26**, 1196–207 (2014).
 40. Ghazanfar, A. A. & Takahashi, D. Y. The evolution of speech: vision, rhythm, cooperation. *Trends Cogn. Sci.* **18**, 543–53 (2014).

41. Bergman, T. J. Speech-like vocalized lip-smacking in geladas. *Curr. Biol.* **23**, R268–R269 (2013).
42. Connor, R. C., Smolker, R. & Bejder, L. Synchrony, social behaviour and alliance affiliation in Indian Ocean bottlenose dolphins, *Tursiops aduncus*. *Anim. Behav.* **72**, 1371–1378 (2006).
43. Polyanskaya, L., Samuel, A. G. & Ordin, M. Regularity in speech rhythm as a social coalition signal. *Ann. N. Y. Acad. Sci.* nyas.14193 (2019). doi:10.1111/nyas.14193
44. de Jong, G., McDougall, K., Nolan, F. Sound change and speaker identity: an acoustic study. in *Speaker Classification*. (ed. Müller, Ch., Schötz, S.) 130–141 (Springer-Verlag, 2007).
45. Ordin, M. & Mennen, I. Cross-Linguistic Differences in Bilinguals' Fundamental Frequency Ranges. *J. Speech, Lang. Hear. Res.* **60**, 1493–1506 (2017).
46. Nettle, D. Using Social Impact Theory to simulate language change. *Lingua* **108**, 95–117 (1999).
47. Nettle, D. & Dunbar, R. I. M. Social Markers and the Evolution of Reciprocal Exchange. *Curr. Anthropol.* **38**, 93–99 (1997).
48. Cutler, A. Segmentation problems, rhythmic solutions. *Lingua* **92**, 81–104 (1994).
49. Ding, N. *et al.* Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* **81**, 181–187 (2017).
50. Polyanskaya, L., Ordin, M. & Busa, M. G. Relative Salience of Speech Rhythm and Speech Rate on Perceived Foreign Accent in a Second Language. *Lang. Speech* **60**, 333–355 (2017).
51. Polyanskaya, L. & Ordin, M. The effect of speech rhythm and speaking rate on assessment of pronunciation in a second language. *Appl. Psycholinguist.* **40**, 795–819 (2019).
52. Langus, A., Mehler, J. & Nespors, M. Rhythm in language acquisition. *Neurosci. Biobehav. Rev.* **81**, 158–166 (2017).
53. Nazzi, T. & Ramus, F. Perception and acquisition of linguistic rhythm by infants. *Speech Commun.* **41**, 233–243 (2003).
54. Goswami, U. Speech rhythm and language acquisition: an amplitude modulation phase hierarchy perspective. *Ann. N. Y. Acad. Sci.* (2019). doi:10.1111/nyas.14137
55. Ahissar, E. *et al.* Speech comprehension is correlated with temporal response

- patterns recorded from auditory cortex. *Proc. Natl. Acad. Sci.* **98**, 13367–13372 (2001).
56. Suppes, P., Lu, Z. L. & Han, B. Brain wave recognition of words. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 14965–9 (1997).
 57. Luo, H. & Poeppel, D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* **54**, 1001 (2007).
 58. Suppes, P., Han, B., Epelboim, J. & Lu, Z. L. Invariance between subjects of brain wave representations of language. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 12953–8 (1999).
 59. Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I. & Schroeder, C. E. Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science (80-.)*. **320**, 110–113 (2008).
 60. Wilson, M. & Cook, P. F. Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychon. Bull. Rev.* **23**, 1647–1659 (2016).
 61. Richard E. Turner. *Statistical Models for Natural Sounds*. (University College London, Gatsby Computational Neuroscience Unit, 2010).
 62. Buzsáki, G. *Rhythms of the Brain*. (Oxford University Press, 2006).
doi:10.1093/acprof:oso/9780195301069.001.0001
 63. Ghitza, O. & Greenberg, S. On the Possible Role of Brain Rhythms in Speech Perception: Intelligibility of Time-Compressed Speech with Periodic and Aperiodic Insertions of Silence. *Phonetica* **66**, 113–126 (2009).
 64. Tichko, P. & Large, E. W. Modeling infants' perceptual narrowing to musical rhythms: neural oscillation and Hebbian plasticity. *Ann. N. Y. Acad. Sci.* nyas.14050 (2019).
doi:10.1111/nyas.14050
 65. Lizarazu, M., Lallier, M. & Molinaro, N. Phase–amplitude coupling between theta and gamma oscillations adapts to speech rate. *Ann. N. Y. Acad. Sci.* nyas.14099 (2019).
doi:10.1111/nyas.14099