

Words from the sound up: On the origin of sound-imitative words

Axel Ekström^{1,2}, Sonja Dahlgren³, Bonnie McLean⁴

¹ Centre for Cultural Evolution, Department of Psychology, Stockholm University, Sweden

² Speech, Music & Hearing, KTH Royal Institute of Technology, Sweden

³ Department of Languages, Helsinki University, Finland

⁴ Department of Linguistics and Philology, Uppsala University, Sweden

axel.ekstrom@su.se, sonja.dahlgren@helsinki.fi, bonnie.mclean@lingfil.uu.se

Abstract

We present a framework for the study of onomatopoeia, from their origins as spontaneous imitations of auditory events to their lexicalization as phonological words in specific languages.

Introduction

We present an interdisciplinary framework for tracing the origin and conventionalization of sound-imitative words. We examine the causal pathway from raw environmental acoustic signals to stable lexical items as shaped by perception, phonology, and morphology. We identify three stages in the emergence of onomatopoeia: (1) the perceptual encoding of acoustic signals by human listeners, (2) the spontaneous vocal imitation of these signals by speakers, (3) the lexicalization of imitations into phonological words. The framework integrates cognitive neuroscience, phonetics, psycholinguistics, and comparative linguistics, and offers an account of how perceptual biases and production constraints interact to produce iconic forms capable of entering lexical and grammatical systems.

Phase 1: Perception

First, the *perception* phase investigates how listeners represent non-speech as somehow speech-like. We examine whether onomatopoeic words elicit similar neural responses, as the sounds they imitate. Behavioral experiments test the relative influence of different acoustic cues – formant structure, fundamental frequency sweeps, and harmonic-to-noise ratios – in affecting auditory “speech-like” percepts; they also test whether formant-based imitations (e.g., mammal cries) are perceived more robustly across linguistic backgrounds, compared to imitating pitch (e.g., birdsong). In complementary experiments, we investigate whether naïve listeners match animal sounds more accurately and rapidly to iconic labels than to non-iconic descriptors, even if the sounds represent unfamiliar species.

Phase 2: Imitation

The *imitation* phase explores how speakers transform perceived sounds into speech. Recordings of spontaneous imitations of familiar and unfamiliar natural and anthropogenic sounds from speakers of typologically diverse languages – including Swedish, Finnish, Japanese, Limburgish, and Mandarin Chinese – are analyzed. The “filtering effects” of each language’s phonological inventory on imitation fidelity are examined for potential systematic divergences. Parallel studies with bilingual children examine code-switching and hybrid forms, probing the role of early imitation in scaffolding the transition from sound to lexeme.

Phase 3: Representation

The final *representation* phase focuses on the conventionalization process by harmonizing and expanding existing cross-linguistic resources, enabling large-scale cross-linguistic comparisons. As an example, the novel Dahlgren & Kittilä (in prep.) database of animal onomatopoeia includes categories of typological forms. For example, for dog barks, there are the “Hau” and “approximate” types, distinguishing barks from large and small dogs. In the project, this database is used alongside other corpora and etymological dictionaries of imitative words (e.g., Abbott 2004) as a basis for comparative analyses; to investigate how acoustic properties interact with potential phonological constraints to produce patterns of consistency and variation across languages, while also identifying priorities for future data collection and standardization.

Moving forward

In preliminary work (Ekström & Dahlgren, 2026), we have concentrated on common animal sound onomatopoeia, because they permit precise alignment between an external acoustic target and cross-linguistic word forms. Such comparisons have revealed both intriguing convergence and language-specific divergence driven by perceptual, articulatory, and structural constraints. We have demonstrated the framework’s utility by contrasting the cross-linguistic consistency of words for cat meows – driven by clear harmonic structure and formant transitions – with the greater variability of words for dog barks, where noisier signals interact more strongly with language-specific phonology. Through this integrated approach, the project refines theoretical accounts of the origin of sound-imitative words and establishes methodological standards and infrastructure for cumulative research, advancing cognitively grounded inquiry into the structure of words “from the sound up.”

Acknowledgements

AE was funded by the Swedish Research Council (2025-00209)

References

- Abbott, D. 2004. *Animal Sounds*. [Online] Available at: <http://www.eleceng.adelaide.edu.au/personal/dabbott/animal.html>
- Dahlgren, S. & Kittilä, S. In preparation. *Database of animal onomatopoeia*. [Online]
- Ekström, A., & Dahlgren, S. (2026). Sound symbolism a priori: Why are animal onomatopoeia (in)consistent? In I. H. Iida, I. Joo, & K. Akita (Eds.), *Joint conference of The 4th Iconicity Seminar & The 15th International Symposium on Iconicity in Language and Literature: Book of Abstracts*, pp. 30-32.

