# Compression in killer whale pulsed calls

Javier Almunia[1], Jonas Philipp Luke[2], Fernando Luis Rosa González[2], and Ramon Ferrer-i-Cancho[*3]

[*]Corresponding author: rferrericancho@cs.upc.edu
[1]Loro Parque Foundation, Puerto de la Cruz, Spain
[2]Department of Industrial Engineering, Universidad de La Laguna, La Laguna, Spain
[3]Department of Computer Science, Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain

Decades of research in quantitative linguistics have unveiled that, in spite of the radical differences between languages spoken on Earth (Blasi, Henrich, Adamou, Kemmerer, & Majid, 2022), languages share general statistical patterns called linguistic laws (Zipf, 1949; Altmann, 1980; Bentz & Ferrer-i-Cancho, 2016). A robust instance is Zipf's law of abbreviation, namely, the tendency of more frequent words to be shorter. This law pervades languages independently of many relevant parameters: linguistic family, writing system or measurement unit (Bentz & Ferrer-i-Cancho, 2016; Petrini et al., 2023a). This and other linguistic laws have been found in a wide range of other species (Semple, Ferrer-i-Cancho, & Gustison, 2022). From a theoretical standpoint, these laws are seen as manifestations of principles of communication since Zipf's pioneering research. The law of abbreviation is a prediction of the principle of compression, namely pressure to reduce the magnitude (length or duration) of types (vocalizations or gestures) (Ferrer-i-Cancho et al., 2013; Ferrer-i-Cancho, Bentz, & Seguin, 2022).

Linguistic laws and their underlying principles remain underexplored in cetaceans, who communicate mainly through clicks, whistles and pulsed calls (Dudzinski & Hill, 2017). Evidence of Zipf's rank-frequency law has been reported for dolphin vocalizations (Markov & Ostrovskaya, 1990) and whistles (McCowan, Hanser, & Doyle, 1999). The law of abbreviation, Menzerath's law (longer linguistic constructs tend to be made of smaller parts) and Zipfian laws of word meaning (more frequent words tend to have more meanings) have been reported for dolphin whistles (Vradi, 2021; Ferrer-i-Cancho & McCowan, 2009). Here we expand this research program by adding killer whales, who regulate group movements and cohesion via acoustic communication and exhibit "dialects" that are culturally transmitted (Filatova et al., 2012).

As the view that languages are shaped by cost-cutting considerations is becoming popular (Gibson et al., 2019), recent research has quantified the actual degree of optimization of languages using two variables: the distance between syntactically related words (Ferrer-i-Cancho, Lusseau, & McCowan, 2022) and word lengths (Petrini et al., 2023b; Pimentel, Nikkarinen, Mahowald, Cotterell, &

Blasi, 2021; Ferrer-i-Cancho & Bentz, 2018).

Here we aim test for the presence of Zipf's law of abbreviation in killer whales and also to evaluate, for the 1st time in a non-human species, the degree of optimality of their vocalizations and its temporal evolution by means of a novel optimality score, $\Psi$, that measures the percentage of optimization of a system: $0\%$ in case of a system that maps type frequencies into type lengths arbitrarily; $100\%$ in case of an optimal coding system (Petrini et al., 2023b). To that aim, we use a dataset of spontaneous pulsed calls produced between 2007 and 2013 by six captive killer whales living in the Loro Parque facilities (Canary Islands, Spain).

We find a significant negative correlation between the frequency of a call type and its duration, in agreement with Zipf's law of abbreviation. To understand the strength of the finding, we also restrict the analysis to specific years. Then the correlation is only significant in 2013. However, three findings support some effect of compression on individual years: (a) the negative correlation that is predicted by optimal coding (Ferrer-i-Cancho et al., 2022) is found in all years except one (2010), (b) the mean duration of call types is below a novel random baseline (Petrini et al., 2023b) for all years and (c) crucially, the sum of the correlations that are obtained over all years is significantly low.

The $\Psi$ score indicates that pulsed calls are optimized to a $38\%$. This is a rather low degree of optimization compared to word durations in human languages: only Vietnamese, with $\Psi = 33\%$, exhibits a degree of optimality smaller than that of killer whales according to a recent study covering 46 languages (12 families, two constructed languages and one isolate; see results on Common Voice in Petrini et al. (2023b)). That indicates that the duration of killer whale pulsed calls has a degree of optimisation lower than most human languages.

Now we turn our attention to the evolution of call durations. In human languages, there is evidence that orthographic word lengths have been increasing over time (Chen, Liang, & Liu, 2015; Milička, 2018). In killer whales, we do not find any monotonic temporal trend, neither towards longer calls nor towards shorter calls over successive years or months. No monotonic trend is found for the optimality of their duration either. That suggests that, globally, vocalizers have neither increased nor decreased the duration of calls or its optimality in a way that changes persist over time.

To sum up, we conclude that coding efficiency is a property shared not only by humans and a long list of primates (see Safryghin et al. (2022)) but also cetaceans (dolphins and here killer whales). Our findings support the hypothesis that species with distant common ancestors may have converged to the law of abbreviation through the action of the principle of compression. Concerning compression in killer whale pulsed calls, we conclude that (a) the principle is acting with less intensity than in most human languages and (b) its intensity has neither decayed nor increased within the small group and small evolutionary scale (a period of 7 years) we have examined.

**Acknowledgements**

RFC is supported by a recognition 2021SGR-Cat (01266 LQMC) from AGAUR (Generalitat de Catalunya) and the grants AGRUPS-2022 and AGRUPS-2023 from Universitat Politècnica de Catalunya. The authors thank Isa Bregante, Jasmin Cirilo, Marco Fratini, Trine Hansen, Chiara Ivaldi, Charlotte Kirschner, William Kreikjes, Dorothee Kremers, Estela Lalueza, Yvonne Montenegro, Hector Morales, Marta Ojeda, Monica Vega and David Verchill for annotating parts of the dataset.

**References**

Altmann, G. (1980). Prolegomena to Menzerath's law. *Glottometrika*, *2*, 1-10.

Bentz, C., & Ferrer-i-Cancho, R. (2016). Zipf's law of abbreviation as a language universal. In C. Bentz, G. Jäger, & I. Yanovich (Eds.), *Proceedings of the Leiden Workshop on Capturing Phylogenetic Algorithms for Linguistics.* University of Tübingen.

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences*, *26*(12), 1153-1170.

Chen, H., Liang, J., & Liu, H. (2015). How does word length evolve in written Chinese? *PLOS ONE*, *10*(9), 1-12.

Dudzinski, K. M., & Hill, H. M. (2017). Cetacean communication. In J. Vonk & T. Shackelford (Eds.), *Encyclopedia of animal cognition and behavior* (pp. 1–12). Cham: Springer International Publishing.

Ferrer-i-Cancho, R., & Bentz, C. (2018). The evolution of optimized language in the light of standard information theory. In C. Cuskley, M. Flaherty, H. Little, L. McCrohon, A. Ravignani, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 12th international conference (evolangxii).* NCU Press.

Ferrer-i-Cancho, R., Bentz, C., & Seguin, C. (2022). Optimal coding and the origins of Zipfian laws. *Journal of Quantitative Linguistics*, *29*(2), 165-194.

Ferrer-i-Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M. J., & Semple, S. (2013). Compression as a universal principle of animal behavior. *Cognitive Science*, *37*(8), 1565-1578.

Ferrer-i-Cancho, R., Lusseau, D., & McCowan, B. (2022). Parallels of human language in the behavior of bottlenose dolphins. *Linguistic Frontiers*, *5*(1), 5-11.

Ferrer-i-Cancho, R., & McCowan, B. (2009). A law of word meaning in dolphin whistle types. *Entropy*, *11*(4), 688-701.

Filatova, O. A., Deecke, V. B., Ford, J. K., Matkin, C. O., Barrett-Lennard, L. G., Guzeev, M. A., Burdin, A. M., & Hoyt, E. (2012). Call diversity in the

North Pacific killer whale populations: implications for dialect evolution and population history. *Animal Behaviour*, *83*(3), 595-603.

Gibson, E., Futrell, R., Piantadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, *23*, 389-407.

Markov, V. I., & Ostrovskaya, V. M. (1990). Organization of communication system in Tursiops truncatus Montagu. In J. A. Thomas & R. A. Kastelein (Eds.), *Sensory abilities of cetaceans - laboratory and field evidence* (Vol. 196, p. 599-602). New York: Plenum Press.

McCowan, B., Hanser, S. F., & Doyle, L. R. (1999). Quantitative tools for comparing animal communication systems: information theory applied to bottlenose dolphin whistle repertoires. *Animal Behaviour*, *57*, 409-419.

Milička, J. (2018). Average word length from the diachronic perspective: The case of Arabic. *Linguistic Frontiers*, *1*(2), 81–89.

Petrini, S., Casas-i-Muñoz, A., Cluet-i-Martinell, J., Wang, M., Bentz, C., & Ferrer-i-Cancho, R. (2023a). Direct and indirect evidence of compression of word lengths. Zip's law of abbreviation revisited. *Glottometrics*, *54*, 58-87.

Petrini, S., Casas-i-Muñoz, A., Cluet-i-Martinell, J., Wang, M., Bentz, C., & Ferrer-i-Cancho, R. (2023b). The optimality of word lengths. Theoretical foundations and an empirical study.

Pimentel, T., Nikkarinen, I., Mahowald, K., Cotterell, R., & Blasi, D. (2021). How (non-)optimal is the lexicon? In *North American Chapter of the Association for Computational Linguistics* (pp. 4426–4438).

Safryghin, A., Cross, C., Fallon, B., Heesen, R., Ferrer-i-Cancho, R., & Hobaiter, C. (2022). Variable expression of linguistic laws in ape gesture: a case study from chimpanzee sexual solicitation. *Royal Society Open Science*, *9*, 9220849.

Semple, S., Ferrer-i-Cancho, R., & Gustison, M. (2022). Linguistic laws in biology. *Trends in Ecology and Evolution*, *37*(1), 53-66.

Vradi, A. (2021). *Dolphin communication: a quantitative linguistics approach.* Unpublished master's thesis, Barcelona School of Informatics, Barcelona.

Zipf, G. K. (1949). *Human behaviour and the principle of least effort.* Cambridge (MA), USA: Addison-Wesley.