



Universiteit
Leiden
The Netherlands

Typological tendencies in verse and their cognitive grounding

De Castro Arrazola, V.

Citation

De Castro Arrazola, V. (2018, May 3). *Typological tendencies in verse and their cognitive grounding*. LOT dissertation series. Retrieved from <https://hdl.handle.net/1887/61826>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/61826>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/61826> holds various files of this Leiden University dissertation.

Author: De Castro Arrazola, V.

Title: Typological tendencies in verse and their cognitive grounding

Issue Date: 2018-05-03

1 Introduction

This dissertation is about verse, some of its recurrent features, and cognitive aspects which can explain their prevalence. On the one hand, I present computational tools to describe systematically regularities in verse corpora; on the other hand, I conduct experiments to show how aspects of human cognition can shape the verse traditions of the world.

1.1 On verse

By *verse* I refer to a range of verbal phenomena, most typically songs and poems, but also nursery rhymes, religious chants or demonstration slogans. Compared to everyday speech, all these forms show additional layers of structure, like a regular alternation of accented syllables, a fixed melody, or a systematic number of syllables per utterance. This kind of verse-specific structures can be conceived in terms of a template to which words are set; examples of templates are the tunes of strophic songs, a poetic form like a limerick, or four-beat structures typical of slogans.

Besides introducing rhythmic or melodic regularities to the discourse, verse templates introduce their own constituent structure, such as stanzas, couplets, lines or feet. Crucially, verse constituents cannot be defined just by relying on syntactic, phonological or pragmatic criteria. Often, the presence of one of these extra-speech constituency levels, namely, the line, is considered a necessary and sufficient condition to define verse: “one can accept a minimal definition of poetry as discourse organized in lines” (Hymes 1977:454); “verse is text which is divided into lines” (Fabb 2014:29).

The distinction between everyday conversation and verse is also described in terms of a continuum, where aspects of language such as prominence or pitch are subject to increasing restrictions or regularities (List 1963). This view accommodates emic taxonomies where, beyond forms we would name *song* or *speech*, there exist intermediate forms such as chanting and recitation (e.g. Seeger 2004:45 for such an analysis among the Suyá). Even in the lack of a clear-cut definition, and

1 Introduction

without a comprehensive survey of every human society, there is a consensus that verse is a universal phenomenon (Brown & Jordania 2011).

Notwithstanding its unifying features, the fact that verse is both widespread and diverse has led to it being studied from an array of angles within fields such as metrics, linguistics, musicology, literary studies or psychology. This dissertation takes an explicitly inclusive approach by drawing from these different types of sources.

1.2 Explaining verse

As with other cultural universals such as language (Christiansen & Kirby 2003) and cooking (Wrangham et al. 1999), verse traditions include, on the one hand, an extreme diversity typical of culturally evolved practices, and, on the other hand, a shared set of principles, which, minimally, serve us to identify the phenomenon as universal. The alleged universality of verse poses questions related to its origins in phylogeny, and its function or evolutionary advantage within the biological and cultural context. These issues in relation to music more generally have been of interest to science for a long time (Darwin 1871), and are receiving renewed attention, mainly due to the development of music cognition research and to comparative studies of music-like phenomena across cultures and species (Wallin, Merker & Brown 2001; Honing et al. 2015; Fitch 2015; Hoeschele et al. 2015; Richter & Ostovar 2016; Patel & Demorest 2013).

The current thesis takes for granted the universality of verse, and tackles the issue of why, among its conspicuous diversity, certain features of verse remain fairly constant across traditions. This belongs to the more general question of causality in verse: why does a given instance of verse show a feature *x*? Leading ethnomusicologist Bruno Nettl, in his overview of the discipline, states that the central question of ethnomusicology concerns “what it is that determines the nature of the musical style (...) of a human society or ethnic group” (Nettl 2010:340).

Researchers have tackled the problem from a number of perspectives, often dividing the potential determinants according to the nature–nurture dichotomy: “the function of tones in relation to each other cannot be explained adequately as part of a closed system without reference to the structures of the sociocultural system of which the musical system is a part, and to the biological system to which all music makers belong” (Blacking 1974:30). Underlying, there is the idea that songs (or other cultural practices) are constrained by (1) highly variable, culture-specific or even idiosyncratic principles or values, and (2) fixed, innate

properties which keep constant across populations.

Even if popular, the concept of *innateness* has proven highly problematic given the complex interactions between genetic endowment and environment during the development of individuals (Mameli & Bateson 2006; 2011). Hence, bearing in mind the readiness of humans to construct and modify the environment they live in (i.e. niche construction, Odling-Smee, Laland & Feldman 2003), the distinction between fixed and variable determinants is better regarded as a fuzzy one. In this respect, we can regard the complete set of variables potentially explaining the verse features of a tradition as the *verse niche* surrounding the poets and songwriters (comparable to the linguistic niche paradigm proposed by Lupyan & Dale 2015). The verse niche extends from the microscopic anatomical details of an individual larynx or auditory cortex, to aesthetic and economic values of the closest community or broader society, and to wide-ranging properties of the ecosystem, such as the climate. The dissertation only tackles cognitive determinants of verse, but I will first discuss other types of determinants in order to establish the appropriate context.

As an illustration of the potential effect of the broader environment, consider the observation by Blacking (1965) about the relation between terrain and musical tempo in two Bantu languages: “the tempo of Venda music is related to the steady walking pace of the people, thrust upon them by the mountainous environment”, whereas the neighbouring Tsonga people “live in the flatlands, walk faster, and have music of more rapid tempo” (quoted in Nettl 2010:345). According to Blacking (1967:25), properties of the terrain make certain types of dances and their accompanying music more suitable than others. This link between the walking affordance promoted by the surrounding and the preferred tempo has been generalised as a bias active in all human music:

Since adults take about 10,000 steps per day (Tudor-Locke & Myers 2001) the presentation of the sound/sensory stimuli to the fetus is ubiquitous under normal circumstances. Music is often played at tempos similar to walking (Changizi 2011) and it is reasonable to propose that one of the bases for this connection results from the concurrent stimuli of the tactile and sonic sensations of maternal footfalls informing the developing brain of the fetus. (Teie 2016:3)

A note of caution applies to this and the following potential determinants of musical or linguistic traits: given their correlational nature (e.g. walking tempo correlates with musical tempo), the hypotheses call for thorough data on phylogenetically diverse populations, and/or convergent evidence from e.g. experiments.

1 Introduction

Otherwise, spurious correlations with little or no explanatory value may accumulate given the increasing availability of relatively large but unbalanced digital datasets (Roberts & Winters 2013).

Climate constitutes another popular source of determinants from the so-called-natural environment. Higher humidity and temperature have been related to a higher incidence of a number of interconnected phonological features, such as vocalic sounds (Everett 2017), sonority (Fought et al. 2004), CV syllables (Munroe & Silander 1999), lower phonological complexity (Maddieson et al. 2011), and tonal languages (Everett, Blasi & Roberts 2015). As in the case of the terrain biasing tempo via walking speed, climate is argued to have an indirect effect on the linguistic system. For instance, higher temperatures may increase the physical distance between speakers, and higher humidity may facilitate certain aspects of phonation; these, in turn, can promote some phonological features and not others.

The effect of the social niche on a verse system can be considered more direct, although proving its robustness is comparably challenging. For instance, even though statistical correlations have been found between sexuality-related values and phonetic/phonological aspects of language and music, the causal mechanism is not well understood. Societies with higher levels of extra-marital sex exhibit languages with greater phonological sonority (Ember & Ember 2007), a correlation comparable to the one between warm climate and sonority mentioned before. Similarly, the singing style of societies imposing severe sexual control on female pre-marital sex is associated with narrower and more nasal phonation (Lomax 1976:23). Nonetheless, the explanatory value of these observations remains limited, as reflected by the author's conjectural tone: "it is as if one function of song was to voice the level of sexual tension in culture" (Lomax 1976:262).

Cultural features tend to be inherited in bundles, which means that, if an early population (e.g. Proto-Bantu) shows a correlation between two features, it is likely that its descendant populations (in the case of Bantu, hundreds) will also show the correlation; not because of a causal link between the two, but as a result of inheritance. In order to increase the explanatory value of correlational studies, it becomes critical to tackle this issue (i.e. Galton's problem, Simonton 1975). For instance, Lupyán & Dale (2010:1) show that "languages spoken by large groups have simpler inflectional morphology than languages spoken by smaller groups as measured on a variety of factors such as case systems and complexity of conjugations." The correlation is robust within a number of independent linguistic families, controlling thus for the inheritance problem. Besides, they defend an explicit causal mechanism, namely that languages spoken by larger groups tend

to have a higher proportion of speakers who learnt the language as adults, and that complex morphology has been shown to be easily acquirable by children, but challenging for adult learners.

Among the properties of the social niche most directly relatable (regarding causality) to features of verse, language plays a central role. More specifically, it is common among metricists to consider that “the possible versification systems for a language” can be derived “from its phonology” (Hanson & Kiparsky 1996:288). Following this line of thought, Golston & Riad (1997; 2000) analyse the metrics of Arabic and Greek as substantially shaped by their phonological properties. Indeed, prosodic features such as stress are frequently mentioned as determining the shape of songs and poems: “the Czechs and Hungarians have languages in which utterances begin with stressed syllables, and their folk songs typically begin with stressed beats” (Nettl 2010:346). The effect has been shown even within the work of composers of instrumental music, where (1) there is no overt use of words, and (2) there is a particular emphasis on creativity, with its tendency to counteract structures inherited from tradition (Patel & Daniele 2003).

Finally, one can consider that the smallest niche with a potential to influence properties of verse is contained within each person’s body, e.g. at the anatomical and cognitive levels. Phoneticians have produced a vast literature on how our anatomical features influence the sounds of speech, and, more relevant here, how these influences can explain typological tendencies in the world’s phonological systems (for an overview, see Ohala 2010, Gordon 2016, and references therein). Linguists tend to accept that the average anatomical features of our species as a whole can account for many phonological generalisations in typology; nevertheless, they are often sceptical about anatomic variation as a source of phonological differences between populations:

One idea is that languages change because their speakers have particular physical characteristics: short tongues, or thick lips, or gappy teeth, or something else in this vein. This is of course nonsense: the size of your tongue has no more to do with the way you speak than has the size of your feet. But that doesn’t stop people making this kind of argument. (Trask 2010:19)

The preceding argument is made within the context of diachronic accounts of language change. Notwithstanding an understandable suspicion for simplistic anatomical explanations which can be incorporated too easily into racist discourses, large-enough datasets may provide evidence for subtle anatomical effects over long periods of time:

1 Introduction

very small, quantitative differences between populations in the anatomy of the vocal tract (here, the alveolar ridge) can influence the articulatory and acoustic properties of a class of sounds (here, post-dental/alveolar clicks) facilitating their inclusion in, and their further elaboration as part of, the phonological inventory, resulting in cross-linguistic variation. (Dediu, Janssen & Moisik 2017:16)

The interactive nature of most forms of speech (and verse to a lesser degree) leads to the conceptualisation of individuals' biases as a trade-off between two competing forces: "minimization of articulatory effort and maximization of perceptual distinctness" (Gordon 2016:17). Most forms of verse also exhibit this double dimension of production and perception, although the pressure for efficient, pragmatic transmission of the signal is not always as fundamental as in everyday conversation. To be sure, engaging in verse can be notoriously challenging in many traditions, e.g. by explicitly making the phonological parsing obscure or imposing complex formal restrictions on the choice of words (for an overview, see "The difficulty of poetry", Fabb 2015:63–71).

The fact that verse is made cognitively demanding for aesthetic or social reasons diminishes the effect of effort-minimising constraints (e.g. Zipf 1949), but all these explicit difficulties need to be framed within our cognitive capabilities nonetheless. Hence, the prevalence of particular verse features can be explained also by the cognitive niche, as it has been proposed for other aspects of culture more generally (Sperber & Hirschfeld 2004; Hauser 2009). Within the broad spectrum of cognition, it is particularly relevant to focus on traits which appear reliably in human development (Mameli & Bateson 2006). The framework of Core Knowledge Systems (Spelke & Kinzler 2007) describe some such traits, which arguably guide our basic representations of e.g. objects, number or space, and which "may provide some of the foundations for uniquely human cognitive achievements, including the acquisition of language and other symbol systems" (Spelke & Kinzler 2007:92). In Chapter 2, for instance, I argue that limits on our numeric cognition (Feigenson, Dehaene & Spelke 2004; Mandler & Shebo 1982) constrain the size of verse constituents.

All the different factors mentioned above interact in a variety of ways, as evidenced generally by the fact that humans adapt to, *and* modify, their environment (Laland, Odling-Smee & Myles 2010). In order to address these factors and their interactions systematically, we can explain their causal effect at different temporal levels. Vygotsky (1962) and, more recently, Smith, Brighton & Kirby (2003) propose a tripartite division between effects occurring in the process of (1) biological

evolution, (2) cultural evolution, and (3) ontogeny. Even smaller temporal levels can be distinguished, such as effects active at the level of milliseconds during on-line production and processing of words (Rączaszek-Leonardi 2010). As Enfield (2014) argues, these and similar proposals provide “conceptually distinct but interconnected causal frames”, which can be useful to understand a phenomenon such as verse. This dissertation focuses on a temporally low-level frame, that is, on causes related to cognitive processing.

1.3 Outline of the dissertation

The dissertation is divided into three parts, each containing a corpus-based chapter followed by an experimental chapter. Each of the parts tackles a different aspect of verse deemed to be typologically widespread. The corpus-based chapter of each part focuses on how we can derive structural tendencies by analysing verse productions which have already been documented (in collections of poems, songs, or descriptions of metrical systems). The experimental chapters try to replicate these observations using controlled verse-like stimuli. In this way, I elucidate the extent to which aspects of general cognition can explain the patterns described in the corpora.

The first part deals with the cornerstone of verse: the structure of the templates. I review (in Chapter 2) the strategies used in different languages to demarcate constituents such as stanzas and lines. Based on these chunking strategies, I examine the numeric demands imposed by templates, reflected in the number of objects contained within each constituent. I conclude that, when the number of identical adjacent objects within a constituent exceeds four, an additional subgrouping is created. This limit may be related to subitizing and working memory constraints. Further (Chapter 3), I demonstrate that this kind of chunking can emerge spontaneously when subjects are asked to reproduce long, random sequences of syllables. Using an iterated learning paradigm, we observe that the small chunking tendencies showed by each person get amplified in the process of transmission, thus offering a model for the cultural emergence of verse-like structures.

The second part tackles the issue of how templates are realised into actual poems or songs; more precisely, I investigate patterns of deviations from a template. First, I offer statistical support for a widely held proposal that verse lines deviate more often at the beginnings rather than towards the end (Chapter 4). Second, I show experimentally that this asymmetry also holds for the perception of verse-like sound sequences: subjects take longer to detect deviant tones towards the

1 Introduction

beginning of the sequences (Chapter 5). These results can be explained by theories of dynamic attention and prediction, according to which later events benefit of the regularity of the preceding ones to build more accurate representations, and hence make it less likely for poets to produce deviations, and for subjects to miss them.

The third and final part also deals with the realisation of templates, but in the specific case of words being set to musical tunes (i.e. textsetting). A number of previous studies have described how phonological features of a language such as stress or tone are aligned with musical features such as metrical prominence or melodic contours. I propose a computational method to address this alignment issue in a systematic way, as a means to ease, standardise and encourage comparisons across genres and languages (Chapter 6). The chapter takes a large corpus of Dutch folk songs as a case study, showing that opposite stress and prominence contours are strongly avoided, particularly within content words and when no melismas are involved. An important limitation of most approaches to textsetting is that they rely on introspection or song corpora. I address this issue by developing a simple perception experiment (Chapter 7) which supplies reliable and detailed textsetting judgements from a given language community, i.e. Dutch, providing converging evidence for the corpus-based hypotheses from the preceding chapter.

References

- Blacking, J. 1965. The role of music in the culture of the Venda of the Northern Transvaal. *Studies in Ethnomusicology* 2. 20–53.
- Blacking, J. 1967. *Venda children's songs: a study in ethnomusicological analysis*. University of Chicago Press.
- Blacking, J. 1974. *How musical is man?* Seattle: University of Washington Press.
- Brown, S. & J. Jordania. 2011. Universals in the world's musics. *Psychology of Music* 41(2). 229–248.
- Changizi, M. 2011. *Harnessed: how language and music mimicked nature and transformed ape to man*. BenBella Books.
- Christiansen, M. H. & S. Kirby (eds.). 2003. *Language evolution*. Oxford University Press.
- Darwin, C. 1871. *The descent of man, and selection in relation to sex*. London: John Murray.
- Dediu, D., R. Janssen & S. R. Moisik. 2017. Language is not isolated from its wider environment: vocal tract influences on the evolution of speech and language. *Language & Communication* 54. 9–20.
- Ember, C. R. & M. Ember. 2007. Climate, econiche, and sexuality: influences on sonority in language. *American Anthropologist* 109(1). 180–185.
- Enfield, N. J. 2014. *Natural causes of language: frames, biases, and cultural transmission*. Language Science Press.
- Everett, C. 2017. Languages in drier climates use fewer vowels. *Frontiers in Psychology* 8.
- Everett, C., D. E. Blasi & S. G. Roberts. 2015. Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences* 112(5). 1322–1327.
- Fabb, N. 2014. The verse-line as a whole unit in working memory, ease of processing, and the aesthetic effects of form. *Royal Institute of Philosophy Supplement* 75. 29–50.
- Fabb, N. 2015. *What is poetry? Language and memory in the poems of the world*. Cambridge University Press.
- Feigenson, L., S. Dehaene & E. Spelke. 2004. Core systems of number. *Trends in cognitive sciences* 8(7). 307–314.
- Fitch, W. T. 2015. Four principles of bio-musicology. *Phil. Trans. R. Soc. B* 370(1664). 20140091.

1 Introduction

- Fought, J. G., R. L. Munroe, C. R. Fought & E. M. Good. 2004. Sonority and climate in a world sample of languages: findings and prospects. *Cross-Cultural Research* 38(1). 27–51.
- Golston, C. & T. Riad. 1997. The phonology of Classical Arabic meter. *Linguistics* 35(1). 111–132.
- Golston, C. & T. Riad. 2000. The phonology of Classical Greek meter. *Linguistics* 38(1). 99–167.
- Gordon, M. K. 2016. *Phonological typology*. Oxford University Press.
- Hanson, K & P Kiparsky. 1996. A parametric theory of poetic meter. *Language* 72(2). 287–335.
- Hauser, M. D. 2009. The possibility of impossible cultures. *Nature* 460(7252). 190–196.
- Hoeschele, M., H. Merchant, Y. Kikuchi, Y. Hattori & C. ten Cate. 2015. Searching for the origins of musicality across species. *Phil. Trans. R. Soc. B* 370(1664). 20140094.
- Honing, H., C. ten Cate, I. Peretz & S. E. Trehub. 2015. Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370(1664). 20140088–20140088.
- Hymes, D. 1977. Discovering oral performance and measured verse in American Indian narrative. *New Literary History* 8(3). 431–457.
- Laland, K. N., J. Odling-Smee & S. Myles. 2010. How culture shaped the human genome: bringing genetics and the human sciences together. *Nature reviews. Genetics* 11(2). 137–148.
- List, G. 1963. The boundaries of speech and song. *Ethnomusicology* 7(1). 1–16.
- Lomax, A. 1976. *Cantometrics: an approach to the anthropology of music: audio-cassettes and a handbook*. Berkeley: University of California Extension Media Center.
- Lupyan, G. & R. Dale. 2010. Language structure is partly determined by social structure. *PloS one* 5(1). e8559.
- Lupyan, G. & R. Dale. 2015. The role of adaptation in understanding linguistic diversity. In R. De Busser & R. J. LaPolla (eds.), *Language structure and environment: social, cultural, and natural factors*, vol. 6, 289–316. John Benjamins Publishing Company.
- Maddieson, I., T. Bhattacharya, D. E. Smith & W. Croft. 2011. Geographical distribution of phonological complexity. *Linguistic Typology* 15(2). 267–279.
- Mameli, M. & P. Bateson. 2006. Innateness and the sciences. *Biology and Philosophy* 21(2). 155–188.

- Mameli, M. & P. Bateson. 2011. An evaluation of the concept of innateness. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 366(1563). 436–443.
- Mandler, G. & B. J. Shebo. 1982. Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General* 111(1). 1.
- Munroe, R. L. & M. Silander. 1999. Climate and the consonant-vowel (cv) syllable: a replication within language families. *Cross-Cultural Research* 33(1). 43–62.
- Nettl, B. 2010. *The study of ethnomusicology: thirty-one issues and concepts*. University of Illinois Press.
- Odling-Smee, J. F., K. N. Laland & M. W. Feldman. 2003. *Niche construction. The neglected process in evolution*. Princeton University Press.
- Ohala, J. J. 2010. The relation between phonetics and phonology. In W. J. Hardcastle, J. Laver & F. E. Gibbon (eds.), *The handbook of phonetic sciences, second edition*, 653–677. Wiley-Blackwell.
- Patel, A.-D. & J.-R. Daniele. 2003. An empirical comparison of rhythm in language and music. *Cognition* 87. B35–B45.
- Patel, A. D. & S. M. Demorest. 2013. Comparative music cognition: cross-species and cross-cultural studies. In D. Deutsch (ed.), *The psychology of music*, 647–681. Elsevier.
- Rączaszek-Leonardi, J. 2010. Multiple time-scales of language dynamics: an example from psycholinguistics. *Ecological Psychology* 22(4). 269–285.
- Richter, J. & R. Ostovar. 2016. “It don’t mean a thing if it ain’t got that swing” – an alternative concept for understanding the evolution of dance and music in human beings. *Frontiers in human neuroscience* 10.
- Roberts, S & J Winters. 2013. Linguistic diversity and traffic accidents: lessons from statistical studies of cultural traits. *PLoS ONE* 8(8). 1–13.
- Seeger, A. 2004. *Why Suyá sing. A musical anthropology of an Amazonian people*. Cambridge University Press.
- Simonton, D. K. 1975. Galton’s problem, autocorrelation, and diffusion coefficients. *Behavior Science Research* 10(4). 239–248.
- Smith, K., H. Brighton & S. Kirby. 2003. Complex systems in language evolution: the cultural emergence of compositional structure. *Advances in Complex Systems* 6(04). 537–558.
- Spelke, E.-S. & K.-D. Kinzler. 2007. Core knowledge. *Developmental Science* 10(1). 89–96.
- Sperber, D. & L. A. Hirschfeld. 2004. The cognitive foundations of cultural stability and diversity. *Trends in cognitive sciences* 8(1). 40–46.

1 Introduction

- Teie, D. 2016. A comparative analysis of the universal elements of music and the fetal environment. *Frontiers in Psychology* 7. 1–8.
- Trask, R. L. 2010. *Why do languages change?* Cambridge University Press.
- Tudor-Locke, C. E. & A. M. Myers. 2001. Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. *Research quarterly for exercise and sport* 72(1). 1–12.
- Vygotsky, L. S. 1962. *Thought and language*. Cambridge, MA: MIT press.
- Wallin, N.-L., B Merker & S Brown (eds.). 2001. *The origins of music*. Massachusetts Institute of Technology.
- Wrangham, R., J. H. Jones, G. Laden, D. Pilbeam & N. Conklin-Brittain. 1999. The raw and the stolen: cooking and the ecology of human origins. *Current Anthropology* 40(5). 567–594.
- Zipf, G. K. 1949. *Human behavior and the principle of least effort: an introduction to human ecology*. Cambridge, MA: Addison-Wesley Publishing.