Typological tendencies in verse and their cognitive grounding

De Castro Arrazola, V.

Author: De Castro Arrazola, V.
Title: Typological tendencies in verse and their cognitive grounding
Issue Date: 2018-05-03

5 The detection of deviants in pseudo-verse lines

5.1 Introduction

Verse encompasses a variety of forms, such as song, poetry, chant or nursery rhymes. All of these contain words, but they also include a feature which is absent from everyday speech, viz. in verse, words are set to templates. These constrain the verbal material one wants to use in several ways. Often, for instance, the length of lines in poems or songs is limited by a fixed number of syllables or beats. In many languages, the relative prominence of the syllables is also constrained, so that e.g. when creating an English sonnet in iambic pentameter, lines are usually opened with an unstressed syllable, followed by a stressed one.

Notwithstanding these structural constraints, verse is not completely rigid, as poets and singers often deviate from the templates, introducing unexpected elements which can be exploited for aesthetic purposes by generating interest or surprise in the listener (Huron 2006). Studies of verse corpora show that, still, it is possible to generalise where deviations tend to occur: they are most frequent at the beginning of lines, and their incidence (progressively) decreases (Fabb 2002: 173–177). This phenomenon is referred to by terms such as *final strictness* or *initial looseness*, bringing out the fact that the asymmetry can stem from exceptional events at either edge of the line.

Despite the lack of a systematic typological survey, robust final strictness phenomena are reported for languages from unrelated families, such as Sanskrit (Arnold 1905), Finnish (Kiparsky 1968), Berber (Dell & Elmedlaoui 2008), and Greek (Allen 1973; Golston & Riad 2000). Chapter 4 provides further details on final strictness, as well as an overview of other phenomena usually considered examples of final strictness, such as rhyme or melodic cadence. Hence, there is some evidence that final strictness is not a property linked to a limited set of related languages which accidentally developed the tendency. Instead, the range of independent observations of final strictness, and the lack of a robust set of languages showing the opposite pattern (i.e. initial strictness) asks for a common explanation. There is the possibility, for instance, that the effect is driven by some aspect of cognition shared across populations. In the present chapter we explore one hypothesis within this context, namely, that the decrease in the frequency of deviations is due to an increase of attention along the line, which is disrupted between lines.

Previous studies show that if the occurrence of an event can be predicted, its processing is facilitated (Jongsma, Desain & Honing 2004; Niemi & Näätänen 1981). The internal regularities characteristic of verse lines allow for prediction building to take place. Nevertheless, this process may be disrupted by line boundaries, which are a defining feature of verse (Fabb 2015). Other constituent levels such as the stanza or the hemistich (see Chapter 2) may also show comparable disruptions, but we focus on the line because it is the constituent for which (1) final strictness is most often described, (2) universality has been argued.

In the present chapter we use sequences of drum strokes as a model of verse lines. Subjects are asked to detect deviations from a pattern under three different experimental conditions. We manipulate the relative timing of the strokes in order to test the extent to which the regularity of the stimuli and the pause between the sequences is driving the variation in reaction times. Overall, the results show that it takes longer to detect deviations closer to the beginning of the line, mirroring the data from verse corpora.

5.2 Method

5.2.1 Participants

A total of 45 subjects took part in the experiment (mean age = 23.1 year; 26 males, 19 females; all native Dutch speakers). Each participant was randomly assigned to one of the three conditions, reaching a total of 15 subjects per condition. The recruitment was done at Leiden University and Radboud University (The Netherlands). All participants signed an informed consent before performing the task (in accordance with Leiden University's LUCL procedure).

5.2.2 Procedure

The general procedure of the experiment was the same for all three conditions, i.e. subjects performed an auditory odd-ball experiment. Each participant listened to a total of 576 drum strokes; these could be of two kinds: (1) a probe stroke (n

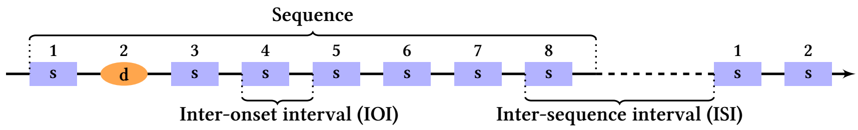

Figure 5.1: Temporal presentation of the stimuli, with time running from left to right following the arrow. The s symbol represents a standard drum stroke, and the d symbol represents a deviant stroke, i.e. the probe.

= 48), or a standard stroke (n = 528). Both sounds are publicly available studio recordings of a mridangam drum (Anantapadmanabhan, Bellur & Murthy 2013), comparable in frequency and intensity, but with differing timbre.[1]

Participants were instructed to press a key as soon as they detected a probe stroke. The general temporal configuration of the 576 strokes was similar across conditions: the strokes are played sequentially, with a short silent gap after every stroke, and a longer gap after every eighth stroke. Figure 5.1 depicts the temporal presentation of the stimuli, with time running from left to right following the arrow.

As seen in Figure 5.1, we refer to the group of eight strokes separated by a longer gap as a *sequence*. The duration from the beginning of a stroke to the beginning of the following stroke is called the *inter-onset interval* (IOI). The longer gap between sequences is called the *inter-sequence interval* (ISI). Each participant listened to a total of 72 sequences, two thirds of which (n = 48) contained a probe, and the remaining third (n = 24) served as fillers with no probe. None of the sequences contained more than one probe. The key measure taken during the experiment is the reaction time to detect the probe, i.e. the lapse of time between the onset of the probe and the subject pressing the key.

All three conditions contain the same number of sequences and probes, but they differ in their temporal presentation, as summarised in Table 5.1. The IOI is kept constant in conditions 1 and 2, i.e. strokes within sequences are isochronous. In condition 3, the IOI varies randomly and can take any value between 250 and 500 milliseconds. The difference between conditions 1 and 2 lies in the ISI, which is kept constant for condition 1, but varies in condition 2 between 1200 and 1800 milliseconds.

[1] Specifically, the standard sound is the *ta* stroke with identifier 224350, and the probe tone is the *num* stroke with identifier 224279.

Table 5.1: Summary of the parametres which define the three experimental conditions.

	Inter-onset interval (IOI)	Inter-sequence interval (ISI)
Condition 1	300 ms	1500 ms
Condition 2	300 ms	1200 $\sim$ 1800 ms
Condition 3	250 $\sim$ 500 ms	1500 ms

5.2.3 Statistical analyses

The main test we perform assesses whether probes occurring earlier within a sequence were detected more slowly than later probes. This is based on the observation that deviant syllables are more likely to occur earlier in verse lines, as formulated by the strict end hypothesis (Chapter 4). Thus, we build a mixed effects model with reaction time as the dependent variable, probe position as a fixed effect, and subject as a random effect. Subsequently, we run more complex models adding the experimental conditions as fixed effects, and controlling for potential confounds.

All the mixed models are implemented in `R` (R Core Team 2017) using the statistical package `lme4` (Bates et al. 2015). Significance of the predictors is calculated in two ways. First, we conduct maximum likelihood t-tests using Satterthwaite approximations to degrees of freedom, as implemented in the package `lmerTest` (Kuznetsova, Bruun Brockhoff & Haubo Bojesen Christensen 2016). Second, we build a null model, identical to the full model except that the variable of interest has been excluded. The fit of the model to the data is compared through a likelihood ratio test to determine whether the full model bears greater explanatory power, hence showing support for the predictor under consideration (Roberts, Winters & Chen 2015).

5.3 Results

Visual inspection of the reaction times to the probe plotted against the probe position within the line (Figure 5.2) reveals a strong negative correlation: probes occurring later in the line require less time to be detected.

However, there is an important confound to control for. Given the design of the experiment (maximally one probe per sequence), probes occurring earlier in the line can have a preceding probe closer by (i.e. if the previous sequence contains

Table 5.2: Summary of the fixed effects in the mixed model analysis.

Predictor	Estimate	Std. Error	df	t	Pr $(> t)$
(Intercept)	0.271	0.0925	51.85	2.93	0.00496
probe.dist	-0.0277	0.00534	1949.06	-5.18	2.39e-07
probe	-0.0346	0.0171	61.09	-2.02	0.0476
condition2	0.105	0.127	46.08	0.83	0.411
condition3	0.362	0.127	45.69	2.86	0.00637
probe:condition2	-0.0236	0.0228	48.29	-1.04	0.306
probe:condition3	-0.0791	0.0227	47.58	-3.48	0.00109

a late probe). Let *probe distance* be the number of strokes between a probe and its preceding probe. On the one hand, probes in position 1 have a mean probe distance of 4.5, while the mean probe distance in position 8 is 12.1 ($r = .64$). On the other hand, probe distance negatively correlates with reaction time: the longer the probe distance, the shorter it takes to react to a probe ($r = -.3$).

Our full model (see Equation 5.1) includes probe distance as a predictor of reaction time, plus an interaction between probe position and experimental condition. A random slope and intercept for the effect of probe position per subject is added. The results of the model are summarised in Table 5.2. Condition 1 is taken as a baseline with which the other two conditions are compared. It can be observed that probe position remains a robust predictor of reaction time after controlling for the confounding effect of probe distance. Besides, the effect of probe position is significantly greater under condition 3.

$$RT \sim \mathit{probe.dist} + \mathit{probe.pos} * \mathit{condition} + (1 + \mathit{probe.pos} | \mathit{subjectID}) \tag{5.1}$$

The statistical significance of the fixed effects on reaction time is confirmed by the comparison of the full model with equivalent null models where the predictor of interest has been removed. The addition of probe position improves the model ($\chi^2 = 38.5$, $p = 2.2e - 08$), the addition of the experimental condition does so too ($\chi^2 = 13.4$, $p = 0.0096$), and the model with an interaction between probe position and condition provides a better fit too ($\chi^2 = 11.3$, $p = 0.0035$).

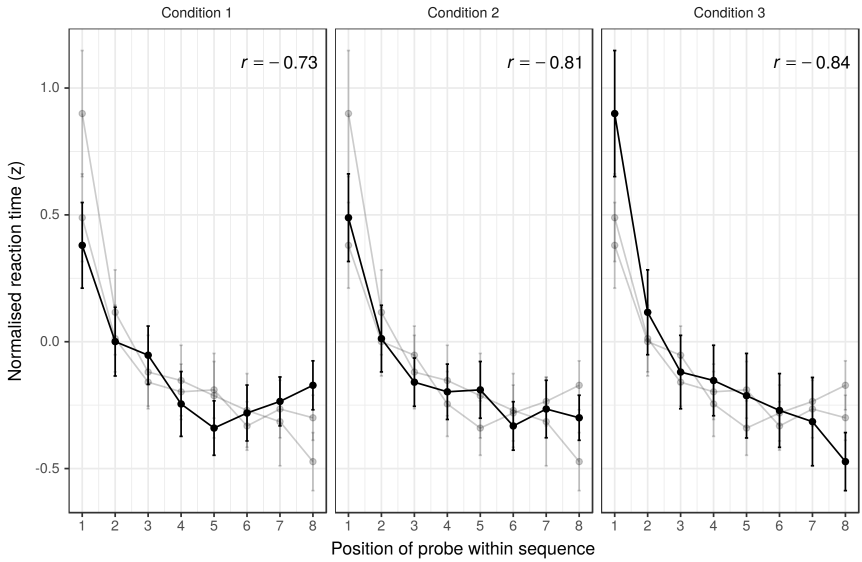

Figure 5.2: Mean reaction time to detect the probe in each of the eight possible positions within the sequence. Pearson's correlation coefficient added separately for each of the conditions. Error bars indicate the 95% confidence interval. In each panel, the condition of interest is highlighted, and the other two conditions are greyed out as a visual reference.

5.4 Discussion

Overall, our experimental results correlate with the strict end hypothesis posited for verse: (1) later in the verse line, deviations are less frequent, and (2) later in the experimental sequence, deviations are detected faster. Nevertheless, the three conditions under inspection do not pinpoint the defining rhythmic context under which the decreasing reaction time takes place.

Unlike similar odd-ball experiments (Schwartze, Farrugia & Kotz 2013; Bouwer & Honing 2015), our stimuli were organised into sequences separated by a longer silent gap. To be sure, the stimuli used in Bouwer & Honing (2015) have a recurring metrical structure which does evoke eight-beat sequences. Nevertheless, and despite the differences in design, we can hypothesise that the crucial difference which produces the decreasing reaction times lies in the sequence-dividing silent gaps.

That being said, we do find some noteworthy differences between the three experimental conditions. The effect of probe position on reaction time is smallest under condition 1, and it increments gradually under conditions 2 and 3.

The first condition is maximally regular, i.e. both the IOI and the ISI are kept constant throughout. This regularity entails that the timing of the events is also maximally predictable. In the second condition, the sounds within sequences are regularly spaced, but the onset of each sequence is unpredictable. Hence, by comparing conditions 1 and 2, we test whether the crucial factor producing a decrease in reaction time lies in the uncertainty of knowing *when* the first stroke of the sequence will be heard. An unpredictable beginning would produce a sequence-initial disadvantage, which would then disappear as further strokes are played with predictable timing. The results do confirm a slightly bigger initial-disadvantage under condition 2 compared to condition 1. Still, (1) with the current sample size, the difference fails to reach statistical significance, and, more importantly (2) the first condition still shows an initial disadvantage, even if the onsets of sequences are completely predictable.

The predictability-driven initial disadvantage (and final advantage) relies on the general readiness principle: if one can predict *when* an event will occur, the speed and accuracy with which we respond to the event is enhanced (Woodworth 1938; Niemi & Näätänen 1981). Despite the difference between the first two conditions, the longer gap which precedes sequences in both cases can be interpreted as a disruption of readiness.

Beyond readiness, finer-grained models of how attention is modulated as a function of predictability become relevant. According to the dynamic attention model (Large & Jones 1999), when we track an external regular rhythm such as a beat sequence, our attention is modulated at the same rate as the rhythm via entrainment. Empirical work (Jones et al. 2002; Fitzroy & Sanders 2015; Jongsma, Desain & Honing 2004) has shown that performance (a proxy for attention) peaks at the points where a beat is predicted, and decreases elsewhere. As the underlying mechanism, it is hypothesised that neural populations synchronise to the external rhythm by firing at the same rate.

The dynamic attention account can explain an increasing advantage later in the line, as the neural entrainment comes into place and attention tracks incoming strokes more precisely. However, this account relies on the regularity of the strokes for the increasing advantage to take place. An alternative account which does not rely on the isochrony of the input is the Bayesian predictive coding model (Vuust & Witek 2014). In this case, our prediction of events gets continuously updated as new stimuli are processed, regardless of isochrony. New strokes of the same kind reinforce our prediction, and performance is enhanced as a consequence.

The third experimental condition tackles this critical point: the IOI or within-sequence regularity. Our results show that, compared to the other two conditions, the initial disadvantage is further increased (Figure 5.2), and the overall effect of probe position on reaction time is significantly greater (Table 5.2). These results fit better the predictive coding rather than the dynamic attention model. Under the latter, the lack of isochrony of condition 3 would predict that the final advantage is diminished, but the opposite is true. The alternative explanatory mechanism is more general, since it relies on the building of predictions based on previous regularities, though not necessarily temporal. This has the potential of being applicable to a broader range of verse types, not restricted to prototypical metrical songs (where a beat can be felt), but including, for instance, non-isochronous poetry recitation.

Fluctuations of attention across verse lines offer a possible explanation of final strictness defined as a decrease in the frequency of deviations. Nevertheless, there exist other phenomena related to final strictness, such as rhyme, or the categorical control of word-length at the end of the line (Fabb 2002:174). An increasingly efficient use of attention is not well suited for these other kinds of final strictness, where the very end of the line is targeted. We should conclude, instead, that verse final strictness is driven by a variety of factors, including attention and the highlighting of constituent boundaries.

Further work is required in order to bridge two critical gaps. First, the low-level odd-ball task used here should be followed up with more ecological stimuli using e.g. verbal material (i.e. syllables), and rhythmic sequences. Many poetic metres, for instance, rely on the alternation of strong and weak positions; hence, deviations from the norm in that kind of context are more complex than in the present paradigm, where violations deviate from a single standard tone. Second, the gap between perception and production needs to be addressed, since the final strictness evidence which motivated the study relies on how poets and singers *produce* their lines of verse, not on how they *perceive* them. Unavoidably, the extent to which the attention mechanisms described here are applicable in a comparable way during the generation of lines (or other non-linguistic sound sequences) needs to be determined by production experiments.

5.5 Conclusion

Versification systems are cultural phenomena shaped by a complex interaction of factors. Typological tendencies such as final strictness can shed light on some of the underlying principles which both make possible *and* constrain the production of songs and poems. When the subjects in our experiment were asked to track the sequences of drum strokes and react to the deviant ones, their performance consistently dropped after the sequence-dividing gap. We propose that a similar drop of attention can play a role in the reduced faithfulness to templates found in songs and poems. Nonetheless, it should be kept in mind that, besides cognitive or anatomical constraints, verse is also shaped by aesthetic ideas which purposefully satisfy and violate our expectations.

References

Allen, W. 1973. *Accent and rhythm: prosodic features of Latin and Greek: a study in theory and reconstruction*. Cambridge University Press.

Anantapadmanabhan, A., A. Bellur & H. A. Murthy. 2013. Modal analysis and transcription of strokes of the mridangam using non-negative matrix factorization. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing: proceedings; 2013 May 26-31; Vancouver, USA. Piscataway, NJ: IEEE, 2013. p. 181-185*. Institute of Electrical & Electronics Engineers (IEEE).

Arnold, E.-V. 1905. *Vedic metre in its historical development*. Cambridge University Press.

Bates, D., M. Mächler, B. Bolker & S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Bouwer, F. L. & H. Honing. 2015. Temporal attending and prediction influence the perception of metrical rhythm: evidence from reaction times and ERPs. *Frontiers in psychology* 6.

Dell, F & M Elmedlaoui. 2008. *Poetic meter and musical form in Tashlhiyt Berber songs*. Köln: Rüdiger Köppe.

Fabb, N. 2002. *Language and literary structure: the linguistic analysis of form in verse and narrative*. Cambridge University Press.

Fabb, N. 2015. *What is poetry? Language and memory in the poems of the world*. Cambridge University Press.

Fitzroy, A. B. & L. D. Sanders. 2015. Musical meter modulates the allocation of attention across time. *Journal of cognitive neuroscience*.

Golston, C. & T. Riad. 2000. The phonology of Classical Greek meter. *Linguistics* 38(1). 99–167.

Huron, D. 2006. *Sweet anticipation: music and the psychology of expectation*. Massachusetts Institute of Technology.

Jones, M. R., H. Moynihan, N. MacKenzie & J. Puente. 2002. Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science* 13(4). 313–319.

Jongsma, M. L. A., P. Desain & H. Honing. 2004. Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biological Psychology* 66(2). 129–152.

Kiparsky, P. 1968. Metrics and morphophonemics in the *kalevala*. In C Gribble (ed.), *Studies presented to roman jakobson by his students*, 137–148. Cambridge (MA, USA): Slavica.

Kuznetsova, A., P. Bruun Brockhoff & R. Haubo Bojesen Christensen. 2016. *lmerTest: Tests in Linear Mixed Effects Models*. R package version 2.0-32.

Large, E.-W. & M.-R. Jones. 1999. The dynamics of attending: how people track time-varying events. *Psychological Review* 106(1). 119–159.

Niemi, P. & R. Näätänen. 1981. Foreperiod and simple reaction time. *Psychological Bulletin* 89(1). 133–162.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Roberts, S. G., J. Winters & K. Chen. 2015. Future tense and economic decisions: controlling for cultural evolution. *PloS one* 10(7). e0132145.

Schwartze, M., N. Farrugia & S. A. Kotz. 2013. Dissociation of formal and temporal predictability in early auditory evoked potentials. *Neuropsychologia* 51(2). 320–325.

Vuust, P. & M. A. Witek. 2014. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Frontiers in psychology* 5.

Woodworth, R. 1938. *Experimental psychology*. New York: Henry Holt.