
A database as a method of raising typological questions about poetic form

Stefano Versace & Nigel Fabb

Proceedings of Conference on
Language Documentation & Linguistic Theory 3

Edited by Peter K. Austin, Oliver Bond, Lutz Marten &
David Nathan

19-20 November 2011 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project
Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London WC1H 0XG
United Kingdom

Department of Linguistics:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4679
linguistics@soas.ac.uk
<http://www.soas.ac.uk/academics/departments/linguistics>

Hans Rausing Endangered Languages Project:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4349
elap@soas.ac.uk
<http://www.hrelp.org>

© 2011 Stefano Versace & Nigel Fabb

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

ISBN: 978-0-7286-0398-1

This publication can be cited as:

Stefano Versace & Nigel Fabb. 2011. A database as a method of raising typological questions about poetic form. In Peter K. Austin, Oliver Bond, Lutz Marten & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 3*, 289-296. London: SOAS.

or:

Stefano Versace & Nigel Fabb. 2011. A database as a method of raising typological questions about poetic form. In Peter K. Austin, Oliver Bond, Lutz Marten & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 3*. London: SOAS. www.hrelp.org/eprints/ldlt3_30.pdf

A database as a method of raising typological questions about poetic form

STEFANO VERSACE & NIGEL FABB

University of Strathclyde

1. WHY THERE ARE TYPOLOGICAL ISSUES RELATING TO LANGUAGE AND POETIC FORM?

All languages seem to be used to produce verbal art in the form of spoken, sung or written texts, of which the more linguistically interesting are the kinds of text called ‘verse’. A text is called verse when it is divided into sections – called lines – which have no independent and consistent linguistic status; lines are verbal sequences which are not the same as sentences, or phrases, or intonational units. Particularly in verse, the language of verbal art is subject to additional regulation which is not characteristic of language in ordinary use; such regulation might include rhyme, meter and parallelism, among others. The consequences of these regulations are that language in verbal art might be organized into sections of a certain length as in the twelve-syllable lines of French alexandrine verse, or where a certain syllable must echo the sound of another syllable as in rhyme, or where adjacent parts of a text must repeat a similar syntactic and semantic structure, as in the example cited in (1), which is taken from a Tswana poem praising the Kgatla chief Lentswe.

- (1) (a) *Mme ere ella tshadi yôôMosenyi*
and when it weep women of.family.mosenyi
‘and when it weeps, the women of the Mosenyi family/household’
- (b) *ere ella thepa yakwagaSuwê*
when it weep maiden at.home.of.Suwe
‘when it weeps, the maiden at the home of Suwe’
(Schapera 1966: 13)

Example (1) shows adjacent lines forming a couplet in which there are similar syntactic structures in the two lines, with partially different but semantically related words. This is parallelism, one of the kinds of regulation of verbal art which is widespread in literatures which are culturally unrelated, as can be seen by the fact that this kind of parallelism is also found, for example, in Khalka Mongolian long epics *Baatarlag Tuul* (cf. Pegg 2001: 51, Poppe 1958), and some Old Hittite poetic texts (cf. Melchert 2006).

The extra regulation of language in verbal art might be a development of regulations applying more generally in language – for example, it has been argued that rhyme is a development of reduplication. Alternatively, it might be some other organizing principle which is not usually found in language, e.g. when a line is required to have twelve syllables; no ordinary linguistic process counts up to twelve, so even though the elements which are being counted are linguistic constituents – syllables – the procedure for counting those linguistic elements is itself non-

linguistic, at least in part. The question of whether some regulation of verbal art is a development of regulation of ordinary language is the topic of Fabb (2010).

Documentation and archiving of languages systematically recognizes many aspects of the ordinary linguistic structure of languages. However, linguists rarely pay much attention to the extra kinds of regulation which are found in verbal art. This derives in part from the lack of any general typological investigation of these regulations of verbal art. The purpose of this paper is to put forward some questions that need to be asked regarding verbal art when documenting a language. These are not just questions which can be asked in new investigations, but also questions which can be asked about already existing archives. Hymes (1965) makes this point in his discussion of what can be reconstructed from transcriptions of Haida and Kwakiutl poems made by Boas and others in the late nineteenth and early twentieth century. There are many archives of this kind in which verbal art is preserved in a form understood by its original transcribers, but where new questions can be asked to uncover different kinds of regulation which were previously unsuspected.

In 1966 Jakobson said that ‘our information about the distribution of parallelism in the folklore of the world and its character in various languages is still sparse and fragmentary’ (1987: 150); his comment remains true almost fifty years later, and can be extended to most kinds of regulation of verbal art. There are no general typologies relating to poetic forms, and the need for such an approach to verbal art is still as strong as it was for Burling (1966: 1435): ‘If we knew more about enough languages, it should be possible to describe the general pan-human features of verse, the features within which our human nature limits us’. Although work in this direction has already been done, the main focus has been to either develop a historical-typological perspective (cf. Gasparov 1996), or to corroborate very specific frameworks (cf. Hanson & Kiparsky 1996, Golston & Riad 1997, Fabb & Halle 2008).

Hence one goal (which we take as our goal) is to develop a general typology in the form of a publicly accessible searchable database, which raises questions that can be addressed to verbal art in any language, and which allows the answers to be gathered in such a way as to allow for cross-linguistic comparison. A second goal is to make the database a theory-forming tool. The way the data is formatted must, on the one hand, be as open as possible to different theories. On the other hand, data of different kinds and in different languages must be subjected to the same set of questions, so that we understand every verbal art in relation to every other, and so that we do not interpret less familiar arts in terms of more familiar ones, as has often been the case in comparative poetics. A third goal is to consolidate theory-related practices of analysis of verbal art by providing a questionnaire for field workers and typologists, perhaps along the lines of the *Lingua Descriptive Studies Questionnaire*, developed by Bernard Comrie and Norval Smith in 1977.

We illustrate these points by discussing the design of a completed prototype database called *VerseTyp*, which we constructed as an output of a British Academy funded project. We outline its results as well as its potentials for future bigger projects on the typology of poetic forms in the languages of the world. The general aims of this typology are: 1) to create a tool for typological comparison between poetic forms in different languages; 2) to create a procedure for describing different

realizations of the same form in different languages/traditions; 3) to develop fine-grained descriptions for every poetic form in a given language; 4) to allow the user to test whether poetic forms are related to the host linguistic structure, and how; 5) to develop hypotheses on how different poetic forms (e.g. meter and parallelism) may interact with each other (cf. Fabb 2009).

2. VERSETYP: A TYPOLOGY FOR VERSE

VerseTyp is a digital database that has collected information about twenty-five poetic forms in seven languages. We picked the languages for variety, and to create problems for ourselves, which we would then be forced to solve in this pilot. They are: Bora drummed language (which we chose as an example of a non-spoken form of language), Hittite (chosen because the surviving evidence is so fragmentary), English and Italian (chosen because they present the opposite problem of being so extensively manifested and having had so much work done on them), Maltese (chosen because it has a relatively young poetic tradition, partly derived from Italian, partly from Arabic), Mongolian (chosen because it is a living oral tradition, with parallelism as a key structural feature), and Somali (chosen as a living oral tradition with highly complex meters which are sometimes set to music). VerseTyp manifested a systematic way of asking questions of different data, relating both to older documentation and to more recent data gathered by linguists.

In doing this kind of work, a first question we must ask – though it might be answered only after answering other questions – is whether the text is divided into sections, perhaps several levels of section. The most obvious example of a section is a line, but it is not always obvious that a text is in lines, and this may need to be discovered: a verse text might survive only in prose form, and as Hymes showed some kinds of oral narrative appear to be verse but lack many of the obvious diagnostic features of verse. Often a text will be divided into two layers of sections, so that *Beowulf* – a text which incidentally survives as written prose – is both in lines and in verses, where two verses make up a line. Texts might also be in lines which are in couplets, another kind of dual structure.

The issue of division into sections is important for other questions about the text, because other kinds of regulation such as meter, or rhyme, or parallelism are defined relative to the division into lines and other types of section: meter, for example, presupposes lineation, because meter controls the number of syllables in a line. Here is another interesting typological issue: a verbal art text may be subjected to several kinds of regulation at the same time, but do these co-occur or do they interact? Is parallelism or rhyme ever sensitive to meter? Is some aspect of the linguistic form of a text dependent on the music of the text as it is sung? More broadly, can we associate aspects of poetic form with something outside language altogether, such as cultural meaning? It is common for specific poetic forms to be associated with specific moods, or specific ways of speaking.

Here are some of the questions we asked in VerseTyp about each poetic form:

- What is the form called? (Sample answers: an Italian sonnet, English iambic pentameter, etc.)
- Is it metrical, parallelistic, both or other?
- Is the text of a certain size?
- Is it related to other types of text either in this language or in some other language? (Poetic forms are often borrowed).
- What are the sections, at different levels, into which the text is organized?
- What kinds of elements are counted and what are the principles of counting?
- What asymmetries can be discovered in sections and are the beginning or end distinctly different? (There are sometimes scalar asymmetries, such as having increasingly long words later in a line, and sometimes rules are applied more or less strictly at the beginning or at the end of a section).
- Where is a section-internal boundary required (caesura) or forbidden (bridge)?
- How are units (e.g. syllables) divided into marked and unmarked and how are the marked and unmarked units distributed? (i.e., the distribution of markedness is related in English and Latin iambic verse, but the manifestation of markedness is quite different, being stress in the former and weight in the latter).
- Is there rhyme, alliteration, or some other kind of sound patterning and how are these related to the sectioning of the text?
- Is there parallelism and of what kind?
- How does the sectioning of the text relate to the syntactic structure of the text: can phrases or words be split across line boundaries?
- Are oral formulae used?
- Is the text set to music?
- Are there specific cultural associations of the text?

The above is a subset of the various questions which might be asked. It should be said that the openness of some of them reflects the fact that no-one yet has an extensive understanding of the typological possibilities and, in particular, the various kinds of regularity which might exist. Furthermore, they reflect how one kind of regularity might depend on another for any specific kind of text.

What might a general typology for a very large number of poetic (including song) traditions from around the world tell linguists? We might be able to separate cultural influence from universal drivers: we know that literatures influence each other, but we assume that there are also universal factors which might be cognitively or otherwise functionally driven. We could ask in what ways forms can combine (e.g., kinds of rhyme with kinds of metre), if any combinations of formal rules are never found, and when forms do co-occur, whether they interact with one another or are best understood as entirely independent constraints on the text. With a sufficiently wide range of data we would be able to test hypotheses about how forms differ (e.g., the possibility that alliteration is always local, whereas rhyme can be non-local). We could also correlate poetic forms with linguistic forms, particularly in the phonology,

to test the common assumption that the specific forms in a given poetry arise from phonological and other forms specific to its language (i.e. a poetry is matched to, or determined by the language).

With a general typology we could also test and refine our hypotheses on the nature of poetic universals, whether they exist and how they might look. We could ask if they can be related to linguistic universals in any meaningful way, and if they can be treated as implicational universals. For example, is there any language-specific reason why most Somali meters are modelled on a quantitative metrical system, comparable to Greek and Latin metrical poetry (cf. Banti & Giannattasio 1998)? Similarly, is there any interaction between SOV word order and the fact that *Baatarlag Tuul* and other poetic forms in Khalka Mongolian constantly display line-initial rhyme? What linguistic feature makes possible the development of a poetic form in a given language? Last but not least, it is by allowing comparison among unrelated traditions that we could overcome a still widespread Western bias in the analysis of poetic forms (cf. Banti & Giannattasio 2004: 292). In addition, we would be able to acquire a deeper knowledge about structural facts about how poetic forms can be manifested. In order to ask such questions and achieve our goal in a meaningful and non-sporadic way, we need to construct as wide as possible a typology, making use of both existing and new documentation.

3. AN EXAMPLE: TSWANA PRAISE-POEMS

We take Tswana *mabôkô* (praise-poems) as an example of how linguistic and poetic theory and language and verbal art documentation can feed into each other. There are a lot of data available on verbal arts in Sub-Saharan languages (e.g., Greenberg 1960 to Schapera 1966, and, for Northern Sotho, Joubert 2004). However, previous analyses of Bantu verbal art never discovered, to put it in Greenberg's words, any 'prosodic system', i.e. a systematic metricality. This, we argue, is because previous analyses did not rely on questions formulated specifically for the investigation of verbal arts, nor did such analyses rely on a sufficiently formalized metrical theory.

Praise-poems are a native poetic form of Tswana, which are also found in other languages of the area. An analysis that shows they are metrical would have significant importance for a typology of metrical form, as it would imply at least two related generalizations: 1) metrical form is built on innate cognitive principles; 2) metrical forms may arise in a culture by endogenous development, and not only by cultural spread.

We analysed the praise-poems collected by Schapera (1966) who acknowledges certain facts about lines in Tswana praise-poems, for example, that the number of words per line seems to be constrained in some way. He thus hypothesizes that some kind of control on the form of the line may be active in praise-poems, but concludes that he has 'no direct evidence in support of this'. To find the evidence we need to ask questions of a kind not envisioned in the early 1960s.

In fact these poems are metrical. Their metricality can be demonstrated by assuming that the counting of syllables is the primary feature of most metrical traditions and that counting can be achieved by a specialized computation, as claimed by Fabb & Halle (2008). Among the tools available in Fabb & Halle's framework, the concept of a 'loose meter' can be applied to such lines. A loose meter is one whose control over syllable counting is not absolute, but is instead mediated by assigning some elements of the phonology of the line a special status; a loose meter runs a rule-based computation over such pre-determined elements. The computation generates a grid from the line and the grid is then subject to constraints; in this way, the meter is regulated. This is shown in the notation below each line, which represents the first level of the grid (GL 0) constructed for each line of the poems edited by Schapera. This gridline is produced by the interaction between a rule that assigns a square left-bracket to the penultimate syllable in each word, and by another rule that constructs a left-headed binary group by inserting right brackets from the left edge of the line. When a group is interrupted by a square bracket, the grouping restarts from the next available asterisk, e.g. in (2a-b):

- (2) (a) *Mme ere ella tshadi yôôMosenyi* GL
) * [* *) [* *) [* *) **)* [* *) 0→
- (b) *ere ella thepa yakwagaSuwê* GL
 [* *) [* *) [* *) * *)* [* *) 0→

The number of words and syllables differ from (2a) to (2b), but the groups constructed by the automatic procedure are five in both cases. The number of groups is constant throughout a number of other poems, and we suggest that the metrical rules explained above can be taken to represent the meter of the praise-poems. This is also true in the case of the lines in (3a-b), where we notice another interesting fact about the praise-poems:

- (3) (a) *e.lê.tse ## tshi.mong ##*
 * [* *) [* *)
 ya.ga.Mma.mo.tlha.sê.di. GL
 * *) * *) * [* *) 0→
- (b) *mo.bo.lê.meng ## ja.ga.Mma.mo.tlha.sê.di.* GL
) * *) [* *) * *) * *) * [* *) 0→
 (Schapera 1966: 14)

Lines (3a-b) only have 3 and 2 words respectively, but still have the same number of groups as (2a-b). We can express this fact by generalizing that, in a line, 'the fewer words, the longer': in other words, the formal correlation is not between, e.g., lines and number of words, or number of stress, but between number of words and length of the words themselves within a line. This scansion can be consistently applied to all the poems edited by Schapera.

Should this scansion hold for other languages of the area, we would be able to say that this form is a Bantu poetic form in its own right. It should be noted that the

scansion we propose disregards important aspects of the phonology of Tswana, such as lexical tone. Only one phonological feature is taken into account by the metrical computation: the phenomenon of penultimate lengthening which is characteristic of Tswana and other Bantu languages (cf., e.g., Hyman 2009). Analyses of further traditions are thus required in order to state how widespread this way of controlling line length is and if this underspecification of the phonology is a phenomenon typical of the area, of the language family, or of Tswana only.

As shown in section 1, Tswana praise-poems are also regulated by parallelism. The co-occurrence and the interaction of different kinds of regulation in the same tradition remains one of the most interesting questions in poetic typology; this can be investigated by the procedure we have outlined in section 2. We have shown how a theory has enhanced the results of the database procedure: the metrical theory of Fabb & Halle (2008) guided the discovery of metricality in the set of poems documented by Schapera (1966). In addition, existing documentation has provided complementary support for a theory in metrical studies.

4. CONCLUSIONS

As the database includes any kind of language-based regularity in poetry it enables a more holistic account of the connection between poetic form and linguistic form, as well as of the relations between poetic forms than has previously been available.

The structure of the database also implies a re-thinking of linguistic approaches to verbal art. This is because: 1) the categories represented in the database need to be thoroughly thought out in order for all poetic forms to be formatted as discrete entries and inserted into the database itself; 2) we need to abstract theoretical questions regarding, e.g., poetic metrical structure in order to discover meters where they were not expected to exist, or in order to analyze metrical traditions that differ from the better-known ones. Furthermore, applying a standardized set of questions to forms of different kinds may trigger new discoveries in (and of) structures previously thought to be well understood. The construction of a database, as an empirical methodology in the investigation of poetics in a wide variety of languages, thus raises new theoretical and typological questions about verbal arts in the languages of the world, as well as about the human use of language in the making of poetic forms.

REFERENCES

- Banti, Giorgio & Francesco Giannattasio. 1996. Music and Metre in Somali Poetry. In R.J. Hayward & I.M. Lewis, *Voice and Power - The culture of language in North-East Africa - Essays in honour of B. W. Andrzejewski*, SOAS, ALC Supplement 3: 83-127.
- Banti Giorgio & Francesco Giannattasio. 2005. Poetry. In Alessandro Duranti (ed.), *A Companion to Linguistic Anthropology*, 290-320. Oxford: Blackwell.

- Burling, Robbins. 1966. The Metrics of Children's Verse: A Cross-Linguistic Study. *American Anthropologist* 68, 1418-1441.
- Fabb, Nigel. 1999. Verse Constituency and the Locality of Alliteration, *Lingua* 108, 223-245.
- Fabb, Nigel. 2009. Formal interactions in poetic meter. In Tonya Kim Dewey and Frog (eds.), *Versatility in Versification: Multidisciplinary Approaches to Metrics*, 147-165. New York: Peter Lang.
- Fabb, Nigel. 2010. Is literary language a development of ordinary language? *Lingua* 120/5, 1219-1232.
- Fabb, Nigel, & Morris Halle. 2008. *Meter in Poetry: a new theory*. Cambridge: Cambridge University Press.
- Gasparov, M. L. 1996. *A History of European Versification*. Oxford: Clarendon Press.
- Golston, Chris & Tomas Riad. 1997. The phonology of classical Arabic meters. *Linguistics* 38 (1), 111-132.
- Greenberg, Joseph. 1960. A Survey of African Prosodic Systems. In S. Diamond (ed.), *Culture in History: Essays in Honour of Paul Radin*, 925-950. New York: Columbia UP.
- Hanson, Kristin & Paul Kiparsky. 1996. A parametric theory of poetic meter. *Language* 72, 287-335.
- Hyman, Larry. 2009. Penultimate Lengthening in Bantu: Analysis and Spread. In *UC Berkeley Phonology Lab Annual Report*, 195-209.
- Hymes, Dell. 1965. Some North Pacific Coast poems a problem in anthropological philology. *American Anthropologist*, New Series, 67, 316-341.
- Jakobson, Roman. 1987. Grammatical parallelism and its Russian facet. In K. Pomorska and S. Rudy (eds.) *Roman Jakobson. Language in literature*, 145-179. Cambridge, MA: Harvard UP.
- Joubert, Annekie. 2004. *The power of performance: Linking past and present in Hananwa and Lobedu oral literature*. Berlin: Mouton de Gruyter.
- Melchert, Craig, H. 2006. New Light on Hittite Verse and Meter? In *Proceedings of the 18th Annual UCLA Indo-European Conference*, Journal of Indo-European Studies Monograph Series, 53. Washington, DC: Institute for the Study of Man, 117-128.
- Pegg, Carole. 2001. *Mongolian Music, Dance & Oral Narrative: performing diverse identities*. Washington, D.C.: University of Washington Press.
- Poppe, Nikolaus. 1958. Der Parallelismus in der epischen Dichtung der Mongolen. *Ural-Altischen Jahrbücher*, 30, 195-228.