Oral Tradition, 31/2 (2017): 355-372

Poetic Parallelism and Working Memory

Nigel Fabb

Parallelism Depends on Memory

The type of parallelism considered here is a relation between sections of text such that each resembles the other in linguistic form, or in lexical meaning, or in form and meaning.[1] The focus is specifically on cases where parallelism is found frequently or systematically, usually as a generic characteristic, or as a part of its poetic tradition. I suggest that this type of parallelism depends on both parallel items being held simultaneously in working memory. For example, Phillips (1981:114) describes a Minangkabau (West Sumatra) oral narrative called *sijobang*, in which "one notices how frequently there occur pairs of lines in which the sense of the first line is repeated in fresh words in the second." Here are two such couplets (Phillips 1981:114-15):

bukan mbo ka salah tanyo,	I shall not make offensive enquiries,
olun badan ka salah sudi,	nor shall I ask offensive questions,
Santan pikie dalam-dalam	Think carefully, Santan
cubolah inok pamonuengkan	ponder and consider well

In the first couplet the second line has a similar meaning and a similar linguistic form to the first, and some words are repeated. In the other couplet the second line has a similar meaning, but not a similar form. Both kinds of parallelism are common in oral verbal art, and are sometimes found in written verbal art. I assume that where parallelism is systematically used in a text, then it is psychologically real: that is, that the author is guided in the composition of the text by the goal of forming parallel sections, and that the hearer is guided in his or her reception and evaluation of the text by attributing parallelism to it.

If we assume that parallelism is psychologically real, such that parallelism is assigned to a text by some psychological process, then memory is crucial, because while composing or listening to the second line, the first line must be remembered: it is the relationship between the first and second line that constitutes the text as parallelistic. Two very different kinds of memory

[1] I would like to thank Frog, Julia Kind, Stefano Versace, and the three anonymous reviewers for ideas, advice, and comments. This article explores the special effects of poetic form as part of a three-year Major Research Fellowship funded by the Leverhulme Trust, titled "Epiphanies in Literature: A Psychological and Literary Linguistic Account."

may be relevant: long-term memory and working memory. Long-term memory can contain an unlimited amount of information. Working memory is limited, so as to contain only a very small amount of information. In this essay I consider the possibility that systematic parallelism is established over a section of text that constitutes a small amount of information. The couplets quoted above are short enough to fit as wholes into working memory. Many parallelistic traditions have texts characterized by parallelism over short couplets of this kind (for example, Rotenese and Rindi [Indonesian], K'iche' [Mayan], Chinese, Finnish, Ugaritic, and so on, all discussed in Fabb 2015), and I suggest that this is not an accident. Instead, I propose that systematic parallelism is optimally constructed in working memory, because there are various functional advantages—epistemic and aesthetic—that arise when parallelism is established in working memory. Nevertheless, parallelism does not need to be established over a text held in working memory; in some cases parallelism holds over texts that are too large to fit and so parallelism in these cases must exploit long-term memory for texts. Kinds of parallelism which exploit long-term memory are found intermittently, rather than holding systematically or consistently throughout a text.

Studies about memory in oral poetics, such as those by Rubin (1995) or Frog (2012), have mainly focused on long-term memory for texts, and have asked whether meter, oral formulae, rhyme, and other poetic forms play a role in long-term retention. Parallelism may also aid long-term retention. Mitchell (1988) argues that if Eastern Sumbanese parallel couplets are misremembered, the ritual may fail, and posits that parallelism aids retention; parallelism here aids long-term memory. This is compatible with my proposal that working memory plays a role in the moment-by-moment composition and reception of parallelistic texts.

There have been experimental studies of long-term memory for poetry; an early example is Henmon (1917), and more recent examples include Rubin (1995) and Tillmann and Dowling (2007), who argue that if a text is divided into lines, the textual material is better remembered. However, I know of no experimental work that looks at how poetry is held and aspects of its form are constructed in working memory. Most of what we know about working memory comes from tests that involve recall for lists of items, and it is unclear how one would begin to establish how poetry is held in working memory. As such, any account of the role of working memory in the processing of poetry must be speculative.

A type of parallelism not discussed here is parallelism in ordinary speech, or in various types of interactions when one speaker produces a text parallel to that of another speaker, as discussed by Tannen (1987:581) in her treatment of repetition in conversation. If these parallelisms are sustained systematically, then I would predict that parallel texts should be of limited size so as to fit into working memory.

Parallel Couplets May Be Held in the Episodic Buffer of Working Memory

Working memory is a limited-capacity system for processing small amounts of material, including the words we speak or hear at any given moment. Its immediacy and limited capacity makes it distinct from long-term memory, which is usually the sole focus of work on memory in oral verbal art. I use the Baddeley and Hitch (1974) model of working memory, as further

developed developed by Baddeley (2012). In the Baddeley-Hitch model, working memory is a multi-component system that contains three distinct short-term memory stores. Two of these stores take information from what is perceived, heard and seen, as well as drawing on long-term memory (including linguistic knowledge), in order to categorize this input. One input store is the visuospatial sketch-pad that holds visual information; I do not discuss it here. Another input store is the phonological loop, which holds auditory information including speech; it is subject to a limit on duration to about two seconds of speech. The visuospatial sketch-pad and the phonological loop both take sensory input and send output into the main memory store, which is called the episodic buffer. In the episodic buffer, information from the two input stores is combined with material drawn from (and sent back to) long-term memory. The capacity of the episodic buffer is based on information size, not duration: the information in the episodic buffer is limited to about four coherent chunks, where each chunk could for example contain several words. Baddeley (2012:15) describes the capacity of the episodic buffer this way: "Memory span for unrelated words is around 5 [words], increasing to 15 when the words make up a sentence." The episodic buffer is not time-limited. The various parts of working memory are controlled by another working memory component called the central executive, which focuses attention—not necessarily consciously—on material in working memory. Though it is possible to give a normative estimation of working memory capacity, as Baddeley does, and as I assume in this essay, working memory capacity in fact varies between individuals. Engle et al. (1999:313) note the correlations between working memory capacity and other cognitive abilities, and in particular that working-memory capacity correlates with a person's fluid intelligence ("the ability to solve novel problems and adapt to new situations"). There is, however, no reason to think that working-memory capacity varies systematically by culture. Instead, what can vary is the way in which poets draw on long-term memory to reconstruct "remembered" texts, including by using oral formulae or "memory houses" or other techniques (as discussed by Rubin 1995).

When we listen to spoken literature, the linguistic input is first passed through the phonological loop, where some of its linguistic form is established, and then the partially analyzed verbal material passes into the episodic buffer. The episodic buffer takes verbal material from the phonological loop as well as from long-term memory, and the material is chunked. There is no requirement that specific kinds of linguistic processing be undertaken over material held in the episodic buffer (Gathercole 2007:761). Like many kinds of complex information processing, the overall processing of a sentence can move material into and out of long-term memory, even if much of the processing takes place in the episodic buffer.

I have argued previously (Fabb 2015:140-70) and here that parallelism is one of a number of kinds of poetic form that are processed and established over material held in the episodic buffer. The episodic buffer sends material into long-term memory; Rubin (1995:72) argues that both the content and form of oral poetry can be stored as gist, that is to say, its substance or essence, in long-term memory. All parts of working memory are limited in capacity. The phonological loop is limited by the duration of the material held in it. About two seconds worth of speech can be held at a time; this need not be the last two seconds of what was heard, because it is possible to refresh the held speech by subvocal rehearsal, so long as it does not exceed two seconds in duration. In crucial contrast the episodic buffer—where I suggest poetic form is processed—is limited not by duration but by how material can be combined into chunks: about

four chunks can be held at one time. Several words may be combined into a single chunk, particularly if they form a coherent phrase. This allows a text of about fifteen English words to be held at one time in the episodic buffer, if the text is syntactically coherent enough to divide it into about four chunks. This rough measure of fifteen English words depends on the syntactic structure of the sequence and whether the words form idiomatic or fixed combinations. It also depends on language-specific factors. For example, if words in a language are more complex than English words, fewer may be held in working memory (as Cohen-Mimran et al. 2013:247 argue for Arabic). Note that this does not mean that the capacity of working memory varies by culture or language, only that different kinds of linguistic form place different demands on working memory capacity.

The phonological loop can hold speech of about two-second duration, which is a shorter stretch of verbal material than can be held in the episodic buffer. This is too short a duration to fit the parallel material as a whole. Consider for example Minangkabau *sijobang* oral poetry, as described by Phillips (1981:21, 41), in its recited and sung forms. When recited, each parallel line takes about two seconds, so the single line could fit into the phonological loop but not a parallel couplet. When sung, each parallel line takes three to four seconds, so the sung version of each parallel member is in itself too long to fit into the phonological loop. This is typical of spoken and sung poetry: poetic forms and sections are not limited by duration. In performed poetry, whether parallelistic or metrical, and in any language, there is no psychologically determined limit on the duration of a performed line or other poetic section. This is contrary to Turner and Pöppel's (1988) proposal that performed lines of metrical verse are in all languages constrained to last no longer than three seconds, in order to fit into their proposed three-second window of consciousness. They provide, however, no good evidence for their proposal, and I have previously (Fabb 2013) disproved it by showing that for a corpus of over one thousand lines of recorded performed English metrical poetry, about sixty percent of lines were longer than three seconds. Thus there is no evidence that the durations of performed sections of verse are generally controlled by any psychological factor, either for metrical or for parallelistic verse. Instead of duration, I suggest that the crucial constraint on poetic forms involves the amount and organization of information (words, for example), based on chunking in the episodic buffer.

Many of the current ideas about working memory come from the work of George Miller and his colleagues in the 1950s, including the term "working memory" itself and the notion of chunking. Miller (1956) has often been read, however, as suggesting that there is some privileged status for "the magic number seven" (more specifically, between five and nine units) in working memory. Some earlier claims (discussed in Fabb 2015:181-83) about working memory and poetry have referred to seven units, for example by claiming that approximately seven-syllable lines have some privileged status. No current theory of working memory, however, supports a seven-unit measure; Cowan (2001:88, 104) suggested that if there is a "magic number," it is likely to be four (that is, four chunks).

Short Parallel Sequences

In this section I discuss various kinds of systematic parallelism that hold over textual sequences that are short enough to fit easily as wholes into working memory. The simplest example is where parallelism involves two short adjacent sections forming a short couplet. This is a common pattern that can be sustained throughout a text; here are some examples from different traditions:

Pessüt penkit hierelömmä	Scrubbed benches we would spoil
Hüväñ tuvañ turmelomma.	The good house we would ruin.
	(Karelian, cited by Frog 2014:191)

A re: Selothe, utlwa, Morena o a bitsa	He said: Selothe, listen, the King is calling,
Utlwa, Lerothi o a memetsa.	Listen, Lerotholi is calling aloud.
	(Basotho, cited by Kunene 1971:90)

Dyoos kuuk'aʔn sin aanimaa alaq	God He has with him your souls
loq'chajin sin animaa alaq	He reverently guards your souls
	(Quiché Mayan, cited by Norman 1980:388)

An occasional variant is for parallelism to be extended to three or more members. Forth (1988:155) describes an invocation in Rindi (Eastern Sumba, Indonesia), mainly in couplets but also including this triplet:

Tomanggunya na tula pakajanga	I reach the notched support
Tomanggunya na rehi pakawuku	I reach the knotted time
Tomanggunyaka na kanduruku handàkangu	I have reached the first thunder

The triplet as a whole sequence may be too large to fit into the episodic buffer. In some triplets, however, it is equally possible to say that parallelism is established just between two adjacent members, and so is always in couplets. For the example above, lines one and two can be taken as a couplet, and lines two and three can be taken as a couplet. There is no evidence that the third line has a specific parallel relation to the first, independent of any relation it might have with the second. Kunene (1971:78) comes to the same conclusion in discussing parallel multi-line sequences in Basotho oral poetry. Though in principle parallelism could be processed within working memory in this way, the fact that parallelism seems never to be systematically in triplets suggests that triplet parallelism is still more costly for processing than ordinary couplet parallelism.

Sometimes in couplet parallelism, two parallel couplets intersect, in an ABA′B′ pattern (where the first couplet is repeated with minor changes) as in the following example from Rindi in Indonesia (Forth 1988:146, 160):

Luananyaka la uma mandamobu duna	He has gone to the house which does not rot,
Na papameranda la lima	The one who has made our hands equal,
Hi luananyaka la kaheli mandambata	He has gone to the house floor which does not break,
Na papahamanda la ngaru	The one who has made our mouths the same.

Whether we analyze patterns of this kind as involving intersecting parallelism or reanalyzed as a single couplet of two long lines AB and A′B′, in either case, parallelism requires keeping all four lines in play at the same time. This is too long a sequence to hold in working memory. However, these intersecting parallelisms are never systematic (and belong to a type discussed in the next section). I claim only that systematic parallelism must have its two parts small enough to fit into working memory.

In some parallelistic traditions, couplets deploy a fixed pair of words, one in each line. These conventionalized word pairs sometimes have a meaning in combination that they do not have separately, a meaning that may be fixed. Here is a pair (called a dyad by Fox [2014:114]) from Rotenese, in which the parallel combination of "a cock's tail feathers" and "a rooster's plume" together form the ritual expression for a male child (110):

De bongi-na popi-koak	She gives birth to a cock's tail feathers
Ma lae-na lano-manuk.	And she bears a rooster's plume.
	= she gives birth to a male child

The term *difrasismo* **(**a pair of words that takes on a metaphorical meaning) describes a similar metaphorical pairing in Nahuatl. In the following couplet, the combination of "water" and "mountain" together means "the city," so the couplet as a whole means "the lords of the city" (Bright 1990:440):

in a-hua-que	the lords of the water
in tepe-hua-que	the lords of the mountain
	= the lords of the city

It may be that word pairs of this type generally appear only when the words are relatively close, such that they can be held together in working memory. In the final section of this paper, I suggest that the metaphorical meanings of such word pairs are well suited to being established in working memory, which would favor their use in short parallel sequences. It is also possible that the paired words count as a single information unit for working memory purposes; thus this might be another formal practice that is adapted to fit the parallel material into the limited capacity of working memory.

Longer Parallel Sequences

Parallelism is a simple formal device that does not depend on any specific set of rules, where those rules might reflect some specific psychological process (in this it differs from

meter). Though I have suggested that there are advantages in processing parallelism in working memory, the non-specificity of parallelism as a kind of form means that it must also be possible to process parallelism over any kind of element, including textual sequences that are too large to fit into working memory. In this section I look at some examples of longer sequences.

Consider the following parallel couplet from an Ipili text (PNG [Papua New Guinea]) (Borchard and Gibbs 2011:181):

Lipi ongane kii pipi tupa yoko yata wato,	Cutting down *pipi*, putting them up here,
Lipi ongane kai maukale tupa yoko yata wato,	Cutting down *maukale*, putting them up here,

Given the long lines, this couplet looks at first as though it might exceed the limits of working memory. However, all but two of the words in the second line are repeated from the first. The repeated words of the second line may take up less capacity in the episodic buffer in comparison with the new words of the first line, and so the sequence may not be overall as large as at first seems. Another question to consider is whether parallel sequences are processed in a special way, which means that the couplet need not be held as a whole in order to establish parallelism. For example, where the sequence of words in the first line is parallel to the sequence in the second, words in the first line could be dropped from working memory once their match has been found in the second; this means that parallelism could be established for a couplet over a sequence that at any time is a continuous sequence of words just over one line in length. These special processing strategies might be learned by expert composers and listeners. Hu et al. (2014:1764) suggest that hearers can apply specific strategies to working memory that, though they do not extend total capacity, nevertheless manipulate what can be held within the fixed capacity.

Another larger-scale type of parallelism is what Poppe (1958:196) called strophic parallelism (*Strophenparallelismus*), as seen in this Mongolian excerpt (Pegg 2001:196):

Mösön degegür güyüdel tei,	Running along the ice
Mönggün coqur mori mini,	Oh, my silvery dappled horse,
Mönggü sirü-ber cimegsen,	Wearing silver and coral decorations
Keüken ür-e mini hümün-ü-düü	My daughter belongs to another
Sugul degegür güyüdel tei,	Running along a very narrow path
Suqai jegerde mori mini,	Oh, my tamarisk chestnut horses
Subud sirü-ber cimegsen,	Wearing pearl and coral decorations
Keüken ür-e mini hümün-ü-düü	My daughter belongs to another

Here a four-line sequence is parallel to another four-line sequence, each forming a strophe. These types of parallelism are likely to involve long-term memory, given that the size of the sequences involved appears to be too large to fit into the episodic buffer. However, it is possible that here too there are formulaic and other devices that reduce the amount of memory capacity required. For example, we might ask whether the repeated sections of these strophes constitute formulae that can be held in working memory with little demand on the available capacity, despite their length.

An extended example of strophic parallelism can be found in a Toraja (Indonesia) memorized oral ritual in the *ma'pakumpan* genre, whose central section is a sequence of thirty-three parallel six-line strophes. This part of the text is performed collectively, as noted by Zerner and Volkman (1988:289): "Almost everyone in the house wakes up and participates in this portion of the *Ma'pakumpang* by shouting suggestions to the *tominaa.*" This is the text that they sing together:

Ia kumpung lako Jawa	It bends like a tree-top toward Java,
Ia kakumbaya baya	It sways for a while,
Ia mentangkean moya	It becomes branches of maa' cloths,
Mendaunan sanda sanda	Becomes leaves of all kinds,
Angga dipokalalanna	Everything that is used,
Mintu'dikande kandena	All that is eaten.

Ia kumpung langan Seko	It bends like a tree-top up to Seko,
Ia kakumbaya baya	It sways for a while,
Ia mentangkean bassi	It becomes branches of iron,
Mendaunan sanda sanda	Becomes leaves of all kinds,
Mintu'dipokalalanna	Everything that is used,
Mintu'dikande kandena	All that is eaten.

Perhaps the repeated parts of the strophe are stored as a formula that takes up minimal space in the episodic buffer, and is reproduced repeatedly by changing a few words in specific slots.

A different kind of strophic parallelism comes from an Asmat song (PNG) in four-line strophes, where every strophe has the same meaning as the first strophe, but the words are different. Consider these three lines, each of which is the second line in its respective strophe, each line having a different form but (as shown in the English translation) the same meaning (Voorhoeve 1977:30):

áya na mewero-awocaia	hi! I am the red-parrot woman	strophe 1
aya na isama-awoca	hi! I am the red-parrot woman	strophe 2
aya na yewer-awoca	hi! I am the red-parrot woman	strophe 3

The word *mewero* in the first strophe means "red parrot." In the second strophe it is replaced by the word *isama* ("fire"), and in the third strophe by the word *yewar,* which names a different kind of parrot. But all three lines have the meaning of the first line. The word in the first strophe is called an *arcer* (a word from the everyday speech in its literal meaning) word by the Asmat. The words in the second and third strophes are *ta-poman* words, which are defined as follows: these are words which may have ordinary; however, they lose these ordinary meanings within the parallel structure and take on the meaning of the first word. This unusual type of strophic parallelism is likely to involve long-term memory rather than working memory. Each subsequent strophe is parallel to the first strophe, and is not an adjacent strophe-to-strophe parallelism of the kind seen in couplet parallelism. In principle, this type of long-distance relationship is something

we might expect only where long-term memory is involved. However, there is a reason to be cautious. As I note at the end of this paper, working memory plays a role in suppressing literal meanings in metaphors, and the literal meanings of *arcer* words must be suppressed; this might suggest that working memory is playing a specific role here.

It is worth noting that sequences of metrical patterns can be repeated, in ordinary verse, as for example when a four-stress line is followed by a three-stress line and the four-three pattern is repeated in the next two lines. A more complex example is seen in classical Greek odes, where the strophe consists of a sequence of metrically different lines, and exactly the same metrical sequence is then copied in the antistrophe (this is called *responsion*). These are not instances of parallelism, but of repetition as part of building a complex metrical pattern: parallelism as defined in this paper is a repetition of linguistic form or lexical meaning, not a repetition of poetic forms such as meter. Hence it places no specific demands on working memory.

Poetry, and Its Added Forms

I have previously proposed this defining difference between prose and poetry (Fabb 2015:9-10):

> A poem is a text made of language, divided into sections which are not determined by syntactic or prosodic structure.
>
> Prose is text made of language, which is divided into sections on the basis of syntactic or prosodic structure.

Poetry and prose may be spoken, sung, written, or signed. The sections of poetry are often called "lines," with other sizes of sections called couplets, strophes, half-lines, and so on: nothing substantial depends on the nomenclature. The evidence that poetry is divided into sections comes from the added forms that presuppose this sectioning. For example, a poetic meter measures the length of a section, usually by counting syllables and defining a rhythm over it, and so meter is a non-linguistic way of defining a part of the text as a section, and hence defines the text as poetry.

Note that this definition of poetry does permit every line to coincide with a sentence, or with some other specific syntactic constituent, as is true in some traditions. However, even if every line were a sentence of the language, it cannot be that every sentence of the language is a line—or everyone would be speaking poetry at all times. The same applies to the relation between lines and ordinary prosodic constituents such as intonation phrases: the line might always be an intonation phrase in some language's poetry, but in that language an intonation phrase is not always a line. This is why the sections are (as the definition above specifies) "not determined" even in a poetry where there is a requirement that every section (that is, line) is also a syntactic constituent.

The added forms of poetry are meter, rhyme, alliteration, and parallelism, which all depend in part on the division of the text into sections. As I have claimed about the added forms (Fabb 2015:177): "A poetic section on which systematic added forms depend must be able to fit

as a whole unit into the episodic buffer in working memory." Meters hold over lines whose lengths vary within strict limits, and these lines are short enough to fit into working memory. When rhyme is systematic, the rhyming word is always located relative to a relatively short section such as a line; there are, for example, no traditions in which only the last word in a stanza rhymes. The same is true of alliteration. Where there is systematic couplet parallelism the parallel members are short, such that even when they are combined the parallel couplet remains short.

Many kinds of parallelism resemble meter in extending from beginning to end of a section of text, but there are differences; differences that may relate to how meter and parallelism are processed in working memory. While meter tends to hold consistently for every line throughout a text, even systematic parallelism may be intermittent, such that some lines in the poem do not involve parallelism; in contrast, poems do not mix metrical with non-metrical lines. Another difference is that a meter extends across a certain size of unit throughout a text: for example, a meter generally governs the rhythmic structure of a line-sized section of text, such that the scope of the meter can be seen by characteristic rhythmic looseness at the beginning and strictness at the end of the section. If the meter sometimes governs line-sized sections, it will not sometimes shrink in scope so that sometimes it governs half-line sized sections or expand in scope so that it sometimes governs couplet sized sections. . But this variation in size of unit is exactly what we find in some parallelistic texts. Thus Forth (1988:147) notes that in the Rindi text he examines, most lines are in parallel pairs, but there are also a number of single lines in which there is instead a parallelism between the two parts of the line. Another difference between meter and parallelism involves cadence, an increased rhythmic regularity towards the end of the line that is found across many metrical traditions. There is no equivalent of cadence across traditions of parallelism. Consider Rindi (Forth 1988:151), where parallelism holds between a pair of lines, with a beginning, middle and end. The two ends can be identical, the two beginnings can be identical, the beginning and end can be identical, and the middle and end can be identical. These fixed patterns in various parts of the Rindi line do not resemble the cross-linguistic characteristics of metrical cadence. Another difference between meter and parallelism is that many meters are much more complex in their organization. For example, Homer's dactylic hexameter characterizes lines of thirteen to seventeen syllables, in alternating subsequences of either heavy-heavy or heavy-light-light, ending on a fixed cadence of heavy-light-light-heavy-heavy, and forcing a word boundary in one of three possible positions near the middle of the line. Similarly complex, the Serbo-Croatian decasyllable has ten-syllable lines, with a tendency to stress odd-numbered syllables (in a trochaic rhythm), and again with a controlled rhythmic ending: if the seventh and eighth syllables are heavy they may not carry stress, if the ninth syllable is light it may not carry stress (and any stressed syllable must here be a heavy syllable). The ninth and tenth syllables must be part of the same word; the third and fourth syllables must be part of the same word, and the fourth syllable must be word-final (Jakobson 1966:418). Meters require a complex system of rules, which Halle and I (2008:11) argue can be derived from a universal theory of meter.

I propose that parallelism is not subject to rules of this type of complexity, systematicity, or linguistic abstractness. Meter is a type of hierarchical grouping (Hayes 1989; Fabb and Halle 2008:11-20). It is processed by adapting general psychological mechanisms of hierarchical

grouping that are also in musical cognition, event segmentation, and so on (Cohen 2000:1). These psychological mechanisms both constrain the range of metrical patterns, but also enable them to be more complex. Because parallelism is not based on any specific psychological mechanism of this kind, it is both freer and less complex. Parallelism is just a matter of composing one section of text so that it resembles another. In some traditions, parallelism will be used in specific ways, constituting local rules within traditions, but these are arbitrarily chosen conventions. The one example I know of a generative rule system for parallelism analogous to that which is required for meter is that of Norman (1980:395-97), who formulates rules to generate parallel pairs in Quiché Mayan ritual language. However, unlike generative rules for meter that resemble generative rules for linguistic stress, Norman's generative rules for parallelism do not resemble any other kind of linguistic generative rule, such as those for meter, stress, phonology, or syntax. As such, Norman's system does not improve on a simpler—and hence preferable—approach to parallelism that just treats parallelism as repetition with variation, subject to tradition-specific rules requiring similarities or variations in specific parts of the line.

Why Might Parallelism Favor Working Memory?

Parallelism can in principle be established over texts larger than can be held in working memory, where the parallel material can be moved in and out of long-term memory as it is being processed. Large-scale or long-distance parallelisms are presumably managed in this way, along with devices such as formulae, allusions and intertextuality, main and subplot, and many types of parallelism of meaning. In this essay I have suggested that when parallelism of linguistic form or lexical meaning is systematic in a text, it is organized into couplets sufficiently short that their parallelism can be established in working memory. I now consider three advantages of limiting the size of the text in this way to enable it to fit into working memory, relating to contrastive valence, fluency of processing, and metaphor; arguing in each case that working memory plays a specific role in enabling emotional and epistemic effects.

I begin with Huron's (2006:21) account of expectation and the aesthetic effect of contrastive valence when listening to music (drawing also on Kind 2016:40). Contrastive valence arises when different kinds of processing operate at different rates to produce an effect where expectations are first disappointed and then satisfied as the listening proceeds. Huron focuses on contrastive valences where a negative affect is followed by a positive affect, as is characteristic in many musical forms. For example, an anticipated musical cadence may not be completed as expected and this may produce negative affect, which is reversed to positive affect when the delayed cadence arrives correctly; this contrast of negative to positive is a strong aesthetic experience. The listener hearing a parallel text learns to expect that each line is followed by another line that is "the same"; this is always a local expectation because it relates one line to the next and so does not need to draw on long-term memory. The expectation of sameness is always initially violated, because the second line is not exactly the same as the first, but then secondarily is met, because the line is underlyingly recognized as the same once the parallelism is established. Hence there is contrastive valence.

Contrastive valence may have enhanced effects when it arises in working memory, and this would favor the processing of parallelism in working memory, and hence favor parallelism over short couplets. A reason for thinking that working memory is involved is this: the psychological processes of attention are focused on the material in working memory (Engle et al. 1999:310); in the Baddeley-Hitch model, the central executive controls the memory stores by focusing attention on the material in working memory. Focusing attention on material in working memory may boost the arousal generated by material in working memory: that is, the material that we immediately process has greater potential to generate arousal. This would make evolutionary sense: one of the functions of arousal is to prepare an animal to freeze, fight, or flee in response to immediately present environmental changes (that is, processed in working memory), and Huron (2006:35) argues that these types of arousal may be co-opted in aesthetic experience. There thus may be an advantage in generating contrastive valence from material while it is held in working memory. This is one possible reason why the added forms are processed over material in working memory, which explains why the sections must be relatively small, and it also explains why parallelism characteristically holds over relatively short sections of adjacent text because the sections must be held together order to establish parallelism.

Now I consider parallelism and fluency relative to working memory. Reber et al. (2004:377) have argued that the regularities of poetic form enable a text to be processed more fluently, which, in turn produces pleasure in the hearer, along with a subjective effect that the processed material is more true (the "illusory truth effect") and more familiar. We might conclude that any form that increases fluency of processing may generate rewards, either emotional or epistemic. For example, Frog (2012:50) suggests that the choice of a specific poetic genre reduces the range of vocabulary that is likely to be used in the poem; we might suggest that this restriction of the selection space means that it should be easier for the hearer to identify the words of the text, making processing more fluent. Parallelism may increase fluency of processing in various ways. For example, the second parallel member will be easier to process than the first because its syntactic structure is already known (in most cases), and the words that vary may be part of fixed pairs such that once the first word is heard, the second member can be predicted. In ordinary language even an unanticipated syntactic parallelism has a priming effect, both in production by making a matched word or structure more easily available and also in comprehension (Sturt et al. 2010:347). In parallelistic traditions, parallelism will be expected and so should ease processing effort since there will be a greater degree of prediction of the next section once the first has been heard. It may be that the ease of processing is improved if all the material is kept in working memory where it is easier to access (because there is no need to search and retrieve from long-term memory). Hence fluency effects would be enhanced by keeping all the parallel material in working memory. This might be another motivation for organizing parallelism into short adjacent units, such that both parts of the parallel structure can be held in working memory at the same time, thus maximizing the effects of fluency of processing.

Finally, consider meaning and parallelism. Parallelism requires us to infer two kinds of relation between meanings. Each of the two parallel lines expresses a meaning; these meanings are different but, somehow, also the same. The hearer must infer the relation (the ground) between the two lines. Furthermore, the paired lines may together have a coherent meaning: the

two lines are the vehicle in a metaphor where their combined meaning is the tenor. What is interesting about both these cases is that the inference is required not to determine the ultimate meaning but instead to determine the ground that connects the two lines, or the pair and their meaning. Consider Central American *difrasismo*, where a pair of words takes on a metaphorical meaning. Norman (1980:392) says that in Nahuatl *b'iineem* and *chakaneem* separately mean "walking" and "crawling" but when combined in a parallel pair they mean "daily activities." Similarly, *eeqaʔn* and *pataal* separately mean "load" and "burden," and in a parallel pair they mean "family of groom." All the meanings—the vehicles and tenors of the metaphors—are fixed and known, but the ground that connects them is not specified and must be inferred, including the relation between the two terms and the relation between the pair and their metaphorical meaning. The point is made by Forth (1988:135) about Rindi parallelism, "it may not be the simpler reference of terms and phrases . . . in ritual language . . . which is screened off or disguised, so much as the precise sense in which terms are appropriate to their denotata." It is the relations between the parts of the metaphor that produce the peculiar effect of the parallelism. Working memory may play a particular role in coping with the multiple meanings, and connections between them, as required for the interpretation of parallelism. Pierce et al. (2010:403) summarize evidence that working memory is required in order to temporarily ignore literal meanings, which is essential for establishing the meanings "in between" the parts of parallelism. This would fit with the idea that parallelism is optimally processed within working memory, not only for formal but also for interpretive reasons.

Conclusion

In this essay I have noted that in many traditions parallelism holds between parallel members that are adjacent and short. When short enough, such a pair of parallel members can fit as a whole into the episodic buffer in working memory. Other types of parallelism, including strophic parallelism, may depend on long-term memory, because it is likely that they exceed the capacity of the episodic buffer. However, we do not know how much the redundancies and idiomatic forms in these larger structures reduce the actual capacity demand, and so even these larger structures might be able to fit, though their relative rarity suggests that they are non-optimal, perhaps because they put a strain on processing.

There may be advantages in holding all of the parts of a parallel structure in working memory. Parallelism may generate arousal by contrastive valence through the satisfaction and denying of expectations, and this arousal may be boosted by the attentional focus on material in working memory. Parallelism enables fluency of processing, which has effects on hedonic and epistemic aspects of aesthetic experience, and this fluency of processing may have increased effects if it is generated in working memory. Parallelism requires the production of non-literal meanings, for which working memory is optimal. In all of these ways parallelism that holds between short adjacent sections can exploit the characteristics of human working memory to achieve particular psychological effects, both aesthetic and epistemic.

University of Strathclyde

References

Baddeley 2012 — Alan D. Baddeley. "Working Memory: Theories, Models, and Controversies." *Annual Review of Psychology*, 63:1-29.

Baddeley and Hitch 1974 — Alan D. Baddeley and Graham J. Hitch. "Working Memory." In *The Psychology of Learning and Motivation: Advances in Research and Theory*. Ed. by G. A. Bower. New York: Academic Press. pp.47-89.

Bauman 1975 — Richard Bauman. "Verbal Art as Performance." *American Anthropologis*t, 77.2:290-311.

Borchard and Gibbs 2011 — Terrance Borchard and Philip Gibbs. "Parallelism and Poetics in Tindi Narratives Sung in the Ipili Language." In *Sung Tales from the Papua New Guinea Highlands: Studies in Form, Meaning, and Sociocultural Context*. Ed. by Alan Rumsey and Don Niles. Camberra: Australian National University Press. pp.165-95.

Bright 1990 — William Bright. "'With One Lip, with Two Lips': Parallelism in Nahuatl." *Language*, 66:437-52.

Cohen 2000 — Gillian Cohen. "Hierarchical Models in Cognition: Do they have Psychological Reality?" *European Journal of Cognitive Psychology*, 12:1-36.

Cohen-Mimran et al. 2013 — Ravit Cohen-Mimran, Jasmeen Adwan-Mansour, and Shimon Sapir. "The Effect of Morphological Complexity on Verbal Working Memory: Results from Arabic Speaking Children." *Journal of Psycholinguistic Research*, 42.3:239-53.

Cowan 2001 — Nelson Cowan. "The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity." *Behavioral and Brain Sciences*, 24:87-185.

Engle et al. 1999 — Randall W. Engle, Stephen W. Tuholski, James E. Laughlin, and Andrew R. Conway. "Working Memory, Short-term Memory, and General Fluid Intelligence: A Latent-variable Approach." *Journal of Experimental Psychology: General*, 128:309-31.

Fabb 2013 — Nigel Fabb. "There is no Psychological Limit on the Duration of Metrical Lines in Performance: Against Turner and Pöppel." *International Journal of Literary Linguistic*s, 2.1:1-29.

Fabb 2015 — _____. *What is Poetry? Language and Memory in the Poems of the World*. Cambridge: Cambridge University Press.

Fabb and Halle 2008 — Nigel Fabb and Morris Halle. *Meter in Poetry: A New Theory.* Cambridge: Cambridge University Press.

Forth 1988 — Gregory Forth. "Fashioned Speech, Full Communication: Aspects of Eastern Sumbanese Ritual Language." In *To Speak in Pairs. Essays on the Ritual Languages of Eastern Indonesia.* Ed. by James J. Fox. Cambridge: Cambridge University Press. pp.129-60.

Fox 2014 — James Fox. *Explorations in Semantic Parallelism.* Canberra: Australian National University Press.

Frog 2012 — Frog. "On the Case of Vambarljóð II: Register and Mode from Skaldic Verse to Sagnakvæði." *RMN Newsletter*, 5:49-61.

Frog 2014 — _____. "Parallelism, Mode, Medium and Orders of Representation." In *Verbal Art and Performance*. Pre-print Papers of the Seminar-workshop 26th-27th May 2014 Helsinki Finland. Ed. by Frog. *Folkloristiikan toimite*, 21. Helsinki: University of Helsinki, pp.185-207.

Gathercole 2007 — Susan E. Gathercole. "Working Memory and Language." In *The Oxford Handbook of Psycholinguistics*. Ed. by M. Gareth Gaskell. pp.757-70.

Hayes 1989 — Bruce Hayes. "The Prosodic Hierarchy in Meter." In *Rhythm and Meter*. Ed. by Paul Kiparsky and Gilbert Youmans. San Diego, CA: Academic Press. pp. 201-60.

Henmon 1917 — V. A. C. Henmon. "The Relation between Learning and Retention and Amount to be Learned." *Journal Of Experimental Psychology*, 2:476-84.

Hu et al. 2014 — Yanmei Hu, Graham J. Hitch, Alan D. Baddeley, Ming Zhang, Richard J. Allen. "Executive and Perceptual Attention Play Different Roles in Visual Working Memory: Evidence from Suffix and Strategy Effects." *Journal Of Experimental Psychology: Human Perception And Performance*, 40.4:1665-78.

Huron 2006 — David Huron. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.

Jakobson 1966 — Roman Jakobson. "Grammatical Parallelism and its Russian Facet," *Language*, 42.2:399-429.

Kind 2016 — Julia Kind. "Elizabeth Bowen's 'Narrative Language at White Heat': A Literary Linguistic Perspective." Unpubl. Ph.D. Dissertation, Johannes Gutenberg University of Mainz.

Kunene 1971 — Daniel P. Kunene. *Heroic Poetry of the Basotho*. Oxford: Clarendon Press.

Miller 1956 — George A. Miller. "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information." *Psychological Review*, 101: 343-52.

Mitchell 1988 — David Mitchell. "Method in the Metaphor: The Ritual Language of Wanukaka." In *To Speak in Pairs. Essays on the Ritual Languages of Eastern Indonesia*. Ed. by James J. Fox. Cambridge: Cambridge University Press. pp. 64-86.

Norman 1980 — William M. Norman. "Grammatical Parallelism in Quiché Ritual Language." Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society, Vol. 6. pp. 387-99.

Pegg 2001 — Carole Pegg. *Mongolian Music, Dance and Oral Narrative*. Seattle: University of Washington Press.

Phillips 1981 — Nigel Phillips. *Sijobang: Sung Narrative Poetry of West Sumatra*. Cambridge: Cambridge University Press.

Pierce et al. 2010 — Russell Pierce, Rick MacLaren, and Dan Chiappe. "The Role of Working Memory in the Metaphor Interference Effect." *Psychonomic Bulletin and Review*, 17:400-04.

Poppe 1958 — Nikolaus Poppe. "Der Parallelismus in der epischen Dichtung der Mongolen." *Ural-Altaische Jahrbücher*, 30:195-228.

Reber et al. 2004 — Rolf Reber, Norbert Schwarz and Piotr Winkielman. "Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience?" *Personality and Social Psychology Review*, 8:364-82.

Rubin 1995 — David C. Rubin. *Memory in Oral Traditions: The Cognitive Psychology of Epic, Ballads, and Counting-Out Rhymes*. New York: Oxford University Press.

Sturt et al. 2010 — Patrick Sturt, Frank Keller, and Amit Dubey. "Syntactic Priming in Comprehension: Parallelism Effects with and without Coordination." *Journal of Memory and Language*, 62:333-51.

Tannen 1987 — Deborah Tannen. "Repetition in Conversation: Toward a Poetics of Talk." *Language*, 63:574–605.

Tillmann and Dowling 2007 — Barbara Tillmann and W. Jay Dowling. "Memory Decreases for Prose, but Not for Poetry." *Memory & Cognition*, 35:628-39.

Turner and Pöppel 1988 — Frederick Turner and Ernst Pöppel. "Metered Poetry, the Brain, and Time." In *Beauty and the Brain: Biological Aspects of Aesthetics*. Ed. by Ingo Rentschler, Barbara Herzberger, and David Epstein. Basel: Birkhäuser Verlag. pp.71-90.

Voorhoeve 1977 — Clemens Lambertus Voorhoeve. "Ta-poman: Metaphorical Use of Words and Poetic Vocabulary in Asmat Song." *Pacific Linguistics, Series C*, 38:19-38.

Zerner and Volkman 1988 — Charles Zerner and Toby Alice Volkman. "The Tree of Desire: a Toraja Ritual Poem." In *To Speak in Pairs: Essays on the Ritual Languages of Eastern Indonesia*. Ed. by James J. Fox. Cambridge: Cambridge University Press. pp. 282-305.

This page is intentionally left blank.