204

Page 4

Chapter 2

Grouping in the stressing of words, in metrical verse, and in music

Nigel Fabb and Morris Halle

In this chapter we argue that computations of the same kind determine the well-formedness of certain structures in language, metrical verse, and music. In all three domains, elements are organized into groups (pairs and triplets) which are themselves organized into groups (and so on). When stress is placed in words, syllables are grouped; syllables are also grouped when the metricality of a line is determined; and in the metrical organization of music, timing slots are grouped. Grouping is accomplished by a set of iterative rules, which generate a bracketed grid from the initial material. Specific aspects of the word, metrical line, or piece of music are controlled with reference to particular features of the metrical grid. In the first three parts of the paper we introduce the iterative rules, and show how they explain the distribution of stress in a word, and how in metrical verse they both control the length of the line as well as its rhythm. In the fourth part of the paper, we discuss the grids assigned by Lerdahl and Jackendoff to music, and show that these grids can be generated by the same iterative rules as are used in language and poetic metre.

On the stress of words and on prosody in general

It is well known that our perception of physical stimuli goes well beyond the simple recording of the signals that impinge on our sense organs. For example, the perceptual effects which are often termed 'visual illusions' imply that our visual sense is somehow deceived into an incorrect judgement about the 'true' nature of the stimuli. In fact, when the matter is looked into with some care it appears that very few of our perceptions are direct records of the signals impinging on our peripheral sense organs; almost everything we perceive is a complicated construct quite remote from the physical—visual or acoustic—signals that impinge on our peripheral sense organs.

Perhaps the most striking example of a physical signal that evokes elaborate mental construction is our perception of speech. When we hear an utterance in a language we know, we normally perceive not only an acoustic signal, but also—and more importantly—a meaningful message that informs us about an aspect of this world or instructs us to perform an action of some kind. Although at this time we have only limited understanding of how a meaning is extracted from a noise, important advances have been made in the understanding of how certain aspects of language are processed. There is,

for example, general agreement that although the speech signal is quasi-continuous, humans perceive these acoustic signals as sequences of discrete sounds (phonemes, segments). The phonemes, moreover, are grouped simultaneously into two parallel kinds of units: morphemes and syllables.

Both morphemes and syllables are sequences of phonemes, but of a totally different kind. Viewed in terms of the morphemes that compose it, the English noun *con-fid-ent-ial-ity* is composed of five morphemes separated here by hyphens. The morphemes, moreover, are combined into a series of nested constituents as shown in (1).

```
(1)   con - fid - ent - ial - ity
       |     |     |     |     |
      pref-stem-suff-suff-suff
        \    /    /    /    /
          V      /    /    /
           \    /    /    /
             A      /    /
              \    /    /
                A      /
                 \    /
                   N
```

The tree in (1), however, represents only one aspect of the structure of the word. Simultaneously with being perceived as the nested object in (1), an utterance of the word *confidentiality* is also perceived as a sequence of the seven syllables shown in (2).

```
(2)   con.fi.den.ti.a.li.ty.
```

While the principles of grouping morphemes as in (1) are the domain of syntax and morphology, the very different principles of grouping syllables as seen in (2) are the domain of prosody.

When English speakers pronounce a sequence of syllables like that in (3), they do not pronounce all syllables in the same way, but differentiate them with respect to prominence or stress, as shown in (3), where the asterisk columns reflect different degrees of prominence. (The columns could be written either above or below the word; this makes no difference.)

```
(3)   con.fi.den.ti.a.li.ty
       *   *  *   *  * *  *
       *      *      *
       *             *
                     *
```

In the normal pronunciation of the word *confidentiality*, the odd-numbered syllables are more prominent than the even-numbered (except for the odd-numbered syllable at the end of the word), and the antepenultimate (the third from the end) syllable has greater stress than the rest.

Having established that differences in prominence are an important phonetic feature of English words, we are immediately faced with the necessity of providing a plausible mechanism that might account for the noted differences in prominence. It was suggested by Mark Liberman (1975) that the differences in prominence are the result of grouping the syllables. Liberman's suggestion was followed by a series of investigations into the nature of stress systems (e.g. Liberman and Prince, 1977; Hayes, 1981; Prince, 1983; Halle & Vergnaud, 1987; Hayes, 1995). These studies culminated in Idsardi (1992), where a powerful new mechanism for expressing prosodic facts of all kinds (not only the word stress) was first described. It is this mechanism that we now explain.

In the Idsardi account, the syllables of an utterance are projected as a sequence of asterisks, where each asterisk stands for a single syllable.[1] The theory also disposes of two junctures, represented here by ordinary parentheses, '(', the left parenthesis, and ')', the right parenthesis. The junctures are inserted among the asterisks, and this has the effect of grouping the syllables in accordance with principle (4).

(4) A left parenthesis groups the syllables on its right; a right parenthesis groups the syllables on its left. A syllable that is neither to the right of a left parenthesis, nor to the left of a right parenthesis is ungrouped.

Thus, in (5a) all syllables (asterisks) constitute one group, but in (5b) all asterisks are ungrouped.

```
(5)   (a)  (* * * * * * *
      (b)   * * * * * * *(
```

In both examples in (5) only a single juncture (left parenthesis) was inserted at one of the edges of the sequence. More than one juncture can be inserted and junctures can be inserted elsewhere than at the edges: junctures may freely be inserted anywhere into the asterisk sequence, thus turning a sequence of syllables into a sequence of groups of syllables. These groups have traditionally been called feet.

The groups are not just sub-sequences of syllables; our theory of groups says that they also possess the additional intrinsic property that one of the two syllables that terminates (i.e. begins or ends) the sub-sequence is the head of the group, and is projected to form a new sequence of asterisks. In addition to stipulating the nature of the parenthesis that is inserted into the asterisk sequence, it is necessary to stipulate whether the left-most or right-most asterisk in the group is the head. In (6) we illustrate the only two possible head locations in the group (5a). In (6a) the head is the left-most asterisk in the group, in (6b), the head is the right-most asterisk.

[1] In some rare stress systems, more than one asterisk is projected from certain of the syllables in the sequence. No such cases will be considered below.


```
(6)   (a)  (* * * * * * *
             *

      (b)  (* * * * * * *
                         *
```

Although we have shown some effects of parenthesis insertion, we have as yet not provided any information on the mechanism that carries out the insertion. In (5) and (6), the parenthesis is inserted by an ordinary rule in the phonology, which recognizes the left edge of a sequence. Perhaps the most important innovation of Idsardi's study was the discovery that parentheses are inserted not only in the same environments as those visible to ordinary phonological rules, but that parenthesis insertion is often due to a special iterative rule that starts at one edge of the sequence and ends at its opposite edge. Such iterative rules, which are not otherwise available in the phonology, insert parentheses at fixed intervals, and the intervals are of two kinds: binary or ternary. Iterative parenthesis-insertion can thus generate feet of only two lengths. We recall that head location in the groups also is subject to a binary choice: the left-most or the right-most asterisk in the group is the head. As a consequence the rule of iterative grouping of syllables can generate feet of only four basic kinds, the two binary feet called trochee (left-headed) and iamb (right-headed), and the two ternary feet, dactyl (left-headed) and anapaest (right-headed).

An account for the prominence distributions in (3) might begin with an iterative rule inserting right parentheses at binary intervals starting at the left edge. The effects of this rule are shown in (7), where we have labelled the two sequences of asterisks gridline 0 and gridline 1 respectively. This demonstrates the significant difference between the two types of parenthesis.

```
(7)   con.fi.den.ti.a.li.ty
      )*  *) *  *)*  *) *    gridline 0
       *     *    *          gridline 1
```

We note that if we had inserted left, instead of right parentheses from left to right in (7), we would have generated the four feet in (8), and implied, contrary to fact, that the word *confidentiality* has four, rather than only three prominences.

```
(8)   con.fi.den.ti.a.li.ty
      (*  * (*  *(* * (*    gridline 0
       *     *    *    *    gridline 1
```

The rest of the prominence pattern of (3) can be generated by positing for gridline 1 an iterative rule inserting left parentheses at binary intervals starting at the left edge and generating left-headed groups.[2] This is shown in (9).

[2] By analogy with the term feet on gridline 0, we might use the traditional terms metra for the groups on gridline 1, and cola for those on gridline 2.

```
(9)   con.fi.den.ti.a.li.ty
      )*   *) *   *)*  *) *     gridline 0
      (*      *   (*            gridline 1
       *           *            gridline 2
```

As shown in (10), the derivation of the prominence pattern of *confidentiality* is completed by the generation of a right-headed group on gridline 2 with the help of the (non-iterative) insertion of a right parenthesis at the right edge of the asterisk sequence.[3]

```
(10)  con.fi.den.ti.a.li.ty
      )*   *) *   *)*  *) *     gridline 0
      (*      *   (*            gridline 1
       *           *)           gridline 2
                   *            gridline 3
```

We summarize in (11) the formal rules, both iterative and other, that have been posited in the discussion of (3) above in order to assign the asterisk array in (10) to the word *confidentiality*.

```
(11)
  (a)  Project each syllable as an asterisk forming gridline 0.
  (b)  Gridline 0.  Starting at the L edge, insert a R parenthesis, form
  binary groups, heads L.
  (c)  Gridline 1.  Starting at the L edge, insert a L parenthesis, form
  binary groups, heads L.
  (d)  Gridline 2.  At the R edge, insert a R parenthesis, head R.
```

We use the term 'metrical grid' to name asterisk arrays like that in (10). It is to be noted that the grid of itself does not assign stress to the syllables of the word. To this end we require the rule (12).

```
(12)  Assign different degrees of prominence (of high pitch) to the
  syllables in direct proportion to the height of their asterisk column
  in the metrical grid.
```

It is the goal of the linguist to discover which combination of grid-building rules, and which stress-assigning rules, are required to explain the pattern of stresses which we find in the words of a specific language.

The need for rule (12) in addition to the rules of grid construction (11) is due to the fact that languages differ not only with respect to their rules of grid construction, but also

[3] The three rules stated above are simplified versions of the actual rules required for the assignment of stress to English words. For a discussion of the actual rules, see Halle (1998).

with respect to such rules of stress assignment as (12), which can vary from language to language.

An instructive example of the latter fact is provided by the stress system of the American Indian language Creek (our data is drawn from Haas (1977)). In a subset of the words of this language, only one syllable carries stress; this is the rightmost even-numbered syllable. In (13) we show two words which illustrate this pattern; we write numbers under the syllables to make it evident that it is the rightmost even-numbered syllable in each word which carries stress.

```
(13)
  (a)  imahicíta  'one to look after' (4th syllable stressed)
       1 2 3 4 5
  (b)  isimahicitá  'one to sight at one' (6th syllable stressed)
       1 2 3 4 5 6
```

The rules which generate grids for these words are stated in (14).

```
(14)
  (a)  Project each syllable as an asterisk forming gridline 0.
  (b)  Gridline 0.  Starting at the L edge, insert a R parenthesis, form
  binary groups, heads R.
  (c)  Gridline 1.  At the R edge, insert a R parenthesis, head R.
```

These rules generate the grids in (15) from the unstressed sequence of syllables.

```
(15)
  (a)
      imahicita
     )* *)* *)*        gridline 0
        *   *)         gridline 1
            *          gridline 2
  (b)
      isimahicita
     )* *)* *)* *)     gridline 0
        *   *   *)     gridline 1
                *      gridline 2
```

The main difference between stress in Creek and English is not due to different metrical grids, but rather to different rules assigning stress. In English stress is assigned by rule (12) to syllables that project to a gridline above gridline 0. In a long English word, like *confidentiality* above, there may be many stressed syllables with different degrees of stress. By contrast, in this subset of words in Creek no matter how long the word, only one syllable is stressed, namely, the right-most even-numbered syllable. But in order to locate this syllable, the metrical grids in (15) must be constructed. It is via these grids that stress is assigned by rule (16).

```
(16) Assign main stress (extra high pitch) to the syllable projecting to
  gridline 2.
```

The grid assigned to the word in Creek by the rules in (14) is periodic: gridline 0 projections of syllables are grouped into pairs (from left to right), and the rightmost syllable in each pair projects to gridline 1. This does not directly represent the stress pattern, or rhythm, of the word, which is not periodic. Instead, this grouping is required in order to locate the single syllable which projects to gridline 2 and to which stress is assigned. The grid in Creek is more complex (abstract) than the stress pattern of the word: in Creek the grid does not represent the rhythm of the word in any direct way, but only via a condition. In this, the grids for word stress are similar to grids for lines in metrical verse, where the grid is again more complex and more structured than the rhythmic pattern it controls.

Metrical verse

The cardinal difference distinguishing a text in verse from a text in prose is that verse is composed of lines. In this section we review the major formal aspects of metrical verse, which is the most common kind of spoken verse, and we leave aside poetry based on syntactic parallelism, like that of the Old Testament, or free verse (*vers libre*), which has enjoyed great popularity in Western Europe and America since the end of the nineteenth century. In all types of poetry the text is invariably split into lines: absent the line, the text is prose, no matter how 'poetic'.[4]

Lines in metrical verse are subject to two kinds of restriction: they are restricted in length, and they exhibit different kinds of restrictions on the placement of marked syllables—e.g. syllables bearing the word stress, or syllables with particular rhymes. Thus, lines in the English trochaic tetrameter, like those in (17), are eight or seven syllables in length, with a strong tendency of placing syllables with word stress in odd-numbered positions (counting from left to right).

```
(17)
  Slaves to London I'll deceive you;
  For the country now I leave you.
  Who can bear and not be mad,
  Wine so dear and yet so bad?
```

(Pierre Antoine Motteux, 'A Song', 1696.)

The most important result of our book on metrical verse (Fabb and Halle, 2008) was to show that these restrictions, as well as all restrictions encountered in other types of metrical verse, require the assignment of metrical grids to the syllable sequences that make up the line. The grids are of the same kind as those that were encountered in stress placement in the preceding section. In the case of the lines in (17) we explain the connection between the restrictions on line length and the restrictions on the

[4] For further discussion see Fabb (2002, chapter 5), and Fabb and Halle (2008, chapter 1).

placement of stressed syllables with the help of rules which generate a grid from the line, rules (18) and (20) below.

Rule (18) projects an asterisk from each syllable in the line. The result is shown in (19).

```
(18) Project each syllable as an asterisk forming gridline 0.
```

```
(19)
 Slaves to London I'll deceive you;
  *     *  *  *  *     * *     *     gridline 0

 Wine so dear and yet so bad?
  *    *  *   *    *   *  *          gridline 0
```

Next, the set of iterative rules (20), which are specific to the trochaic tetrameter, is applied to generate the grids in (21).

```
(20) Rules for English trochaic tetrameter
  (a) Gridline 0: starting just at the L edge, insert a L parenthesis,
  form binary groups, heads L.
     i. Incomplete groups are admitted.
  (b) Gridline 1: starting just at the L edge, insert a L parenthesis,
  form binary groups, heads L.
  (c) Gridline 2: starting just at the L edge, insert a L parenthesis,
  form binary groups, heads L.
```

```
(21)
  Slaves to London I'll deceive you;
   (*     * (*  * (*     *(*     *(    gridline 0
   (*        *    (*       *(          gridline 1
   (*              *(                  gridline 2
    *                                  gridline 3

  Wine so dear and yet so bad?
  (*    * (*    *  (*    * (*          gridline 0
  (*       *       (*      *(          gridline 1
  (*                *(                 gridline 2
   *                                   gridline 3
```

What necessarily limits the length of the line is that there are iterative rules on each gridline, that each of the three iterative rules in (20) inserts parentheses at binary intervals, and that the gridline which is the last to be generated (bottom-most in our grids) must have exactly one asterisk. We state the latter requirement formally in (22).

```
(22) The last-to-be-generated gridline in a well-formed metrical grid
  must contain a single asterisk.
```

In addition to line length, metrical verse in English places specific restrictions on the location of stressed syllables in the line, and it is these restrictions that provide the line with its characteristic rhythm. In a metre in which every odd-numbered syllable was stressed and every even-numbered metre unstressed, we would have the condition in (23).

(23) A stressed syllable must project to gridline 1, and a gridline 1 asterisk must project from a stressed syllable.

But this is not the right generalization, and hence not the right condition for English verse. Though the grid is fully periodic, the rhythm is not fully periodic, so the pattern of stressed and unstressed syllables cannot be fully controlled by a condition on the grid. It is true that it is possible to pronounce the line with a fully regular rhythm, for example stressing *I'll* in (21) because it is odd-numbered, but this is a performance practice rather than a fact about how the metre controls the composition of the line. It is also possible to perform the line with a more regular rhythm (in which *I'll* is not stressed), and this is just as metrical.

A better condition, which accounts more generally for the distribution of stressed syllables in English metrical verse requires a definition of the relevant type of syllable in (24) and a condition formulated as in (25).

(24) The syllable bearing the word stress in a polysyllabic word is a maximum, if it is preceded and followed in the same line by a syllable with less stress.

(25) Maxima must project to gridline 1.

Though stressed monosyllables tend to project to gridline 1, they do not have to; only the stressed syllable in a polysyllabic word must project to gridline 1, and this only when it is a maximum or 'peak' relative to surrounding syllables. There are two maxima in the first scanned line in (21) (the stressed syllables in *London* and *deceive*), and both project to gridline 1. Since there are no maxima in the second scanned line, the rule does not control the distribution of stressed and unstressed syllables at all in this line. Note that there is not a complementary requirement that gridline 1 asterisks must project from stressed syllables or from maxima. The line might in principle have no stressed syllables at all, and though this would be linguistically odd the line would not be unmetrical (so long as it had the right number of syllables).

It is important to note that conditions such as (25) which control for the presence of stressed syllables in the metrical line are not part of the grid-building process. The grid has already been generated from the line by the iterative rules (which know nothing of differences between syllables) at the point when the condition is checked. The job of the metrical analyst is to discover which combination of iterative and other rules, and which conditions on the resulting grid, best explain the characteristics of the metrical line.

It should now be clear how iterative rules such as those in (20) explain both why the metrical line has a specific length and how the syllables are patterned. We further illustrate our approach with the lines in iambic pentameter shown in (26).

(26)

Hungarians! Save the world! Renew the stories
Of men who against hope repelled the chain,
And make the world's dead spirit leap again!
On land renew that Greek exploit, whose glories
Hallow the Salamanian promontories,
And the Armada flung to the fierce main.

(Matthew Arnold, 'Sonnet to the Hungarian Nation', lines 9–14, 1849.)

The rules which generate a grid for iambic pentameter are given in (27).

(27) Rules for iambic pentameter

(a) Gridline 0: starting (just at the edge OR one asterisk in) at the R edge, insert a R parenthesis, form binary groups, heads R.

(b) Gridline 1: starting just at the R edge, insert a R parenthesis, form ternary groups, heads R.

 i. The last (leftmost) group must be incomplete -- binary.

(c) Gridline 2: starting just at the L edge, insert a L parenthesis, form binary groups, heads L.

In (28) we show the effect of the rules (just rule (27a) the second to sixth lines). We have indicated which syllables we think are likely to be stressed in a performance of the line, and we have underlined the maxima. Stressed syllables in general are not subject to a constraint on their location; only the subset of stressed syllables which we have identified as maxima must project to gridline 1. Because it is a grammatical word, 'against' does not contain a word stress, and hence not a maximum. /

(28)

```
    '        '        '        '        '
 Hungarians! Save the world! Renew the stories
 )*  *) *    *)   *   *)   * *)  *   *) *      gridline 0
     *       *)       *      *       *)        gridline 1
            (*                       *(        gridline 2
             *

     '        '      '       '          '
 Of men who against hope repelled the chain,
 )*  *)  *  *) *   *)   * *)    *   *)         gridline 0
     *      *      *      *         *          gridline 1
```

```
        ‘       ‘      ‘     ‘     ‘     ‘
 And make the world’s dead spirit leap again!
)*    *)    *  *)      *     *)*   *)  * *)     gridline 0
      *        *             *     *     *      gridline 1

       ‘     ‘        ‘       ‘           ‘
 On land renew that Greek exploit, whose glories⁵
)*   *)   * *)   *    *)  *   *)    *    *)*    gridline 0
     *     *          *       *          *      gridline 1

‘            ‘      ‘
Hallow the Salamanian promontories,
)*  *)   *  *)* *) *   *)* *)*                  gridline 0
    *       *  *       *  *                     gridline 1

                 ‘    ‘         ‘     ‘
 And the Armada flung to the fierce main.
)*    *) * *)*   *)   *   *) *     *)           gridline 0
      *  *   *        *        *                gridline 1
```

We note that parentheses on gridline 0 are inserted from right to left. There are two kinds of evidence for this. We have seen that the last group generated by an iterative rule can be incomplete, and in a trochaic metre this is the rightmost group; in an iambic metre the leftmost group can sometimes be incomplete (indicating that groups are formed from right to left). Furthermore, the rightmost syllable in the line can be 'extrametrical' (as seen in the first, fourth, and fifth lines of (28)). This arises because the iterative rule (27a) may skip the first asterisk encountered.

This is one way in which the line can have an extra syllable. But extra syllables can also appear within the line, which cannot be accommodated by skipping the first asterisk of gridline 0 in applying rule (27a). This is illustrated by some 11-syllable lines in iambic pentameter by John Donne in (29), where we indicate the syllables by writing the letter s below the line.

```
(29)
 And gluttonous death, will instantly unjoynt
 s    s s s    s     s   s  s  s s s
 My body, and soule, and I shall sleepe a space,
 s  s s  s    s     s   s  s    s     s  s
 But my ever-waking part shall see that face,
 s   s s s  s s    s    s     s   s    s
```

(John Donne, 'Holy Sonnet 6' lines 5–7.)

In some metrical theories extra syllables within the line are accommodated by treating a pair of syllables as a single metrical element. In contrast, we never allow two syllables to project as a single asterisk. In our theory, syllables are projected onto

[5] The noun *exploit* is given final stress in *Walkers Pronouncing Dictionary* of 1852.

gridline 0 by rule (18), but poets may (occasionally) violate rule (18) and not project a given syllable, so where two syllables appear to project as one asterisk in fact one syllable projects and the other does not.[6] In the first and second lines the third syllable does not project as an asterisk, and in the third line the fourth does not project; this is shown in (30), where the syllables which do not project are marked by Δ (for the convenience of the reader: this symbol has no formal status). Gridline 0 in each case consists of ten asterisks, even though the line has 11 syllables.

```
(30)
  And gluttonous death, will instantly unjoynt
 )*     *) Δ  *    *)    *   *)  *   *)*  *)      gridline 0
 My body, and soule, and I shall sleepe a space,
 )*  *)Δ  *    *)    *   *)  *    *)   *   *)     gridline 0
 But my ever-waking part shall see that face,
 )*  Δ *)*   *)*   *)    *    *)  *   *)          gridline 0
```

Projection of an asterisk on gridline 0 is governed by rule (18); if this rule does not apply then a syllable will not project as an asterisk, and we have seen that this is a possibility. However, no rule permits an asterisk to appear on gridline 0 if it is not projected from a syllable; that is, an asterisk cannot project from a pause or rest, but only from an actual syllable. (As we will see, music differs in this regard.)

The iterative rules in words and in metrical lines

There is a deep similarity between the computations of stress in a word, which we discussed in the first part, and the computations that test the metricality of a line, which we discussed in the second part of this chapter. The metrical grid is generated in both domains by an iterative (repeating) rule which inserts parentheses into a sequence of asterisks; the rule also projects a new line of asterisks, and the set of asterisks comprises the grid. Each iterative rule involves a choice from the following five parameters, of which all but the first involve a choice between two alternatives.

(31)

(i) Insertion starts either just at the edge of the gridline, or one asterisk in, or two asterisks in.

(ii) Edge of the sequence (Left (L) / Right (R)) at which insertion begins.

(iii) Nature of parenthesis inserted (L/R).

(iv) Interval between consecutive insertions (2/3 asterisks).

(v) Location of head in each group (L/R).

[6] There is no general principle in English poetry determining which syllables cannot project, but there are some tendencies (e.g. syllable ending in a vowel preceding a syllable beginning with a vowel, as in (29, 30)). By contrast, the non-projection of syllables in lines of Romance poetry is subject to strict rules. (For discussion, see Fabb and Halle (2008)).

However, there are two important differences between the assignment of stress in words and the ascertaining of metricality of a line. First, the metrical rules do not alter the phonetic form of the line in any way; instead, the rules generate a grid from the phonetic form, and test whether certain phonetic characteristics of syllables match the position of the syllable in the grid (e.g. that it projects to an asterisk on a specific gridline). In contrast, by assigning stress to particular syllables in the line, the stress rules alter the phonetic form of the line after generating a grid from the line.

The second important difference in the use of iterative rules in metrical verse and in words is that the rules function to limit line length in metrical verse. This is because in metrical verse every gridline is subject to iterative rules (this is not true in words, as we saw above), and by virtue of condition (22), the final gridline must contain one asterisk. The conjunction of these two facts means that a well-formed grid can be generated only from a line with a specific number of gridline 0 asterisks. In comparison, the rules (14) for Creek words can generate a well-formed grid from a word of any length (because gridline 1 asterisks are grouped by a non-iterative rule). We noted earlier that metrical verse is divided into lines, and that lines are not linguistic entities. Now we see another way in which a line is not a linguistic entity: it is limited in length. Linguistic entities are not controlled in their length by linguistic rules; for example, the length of words or sentences is not so controlled.

In this and the previous sections the same kinds of iterative rules were used to generate grids both for metrical lines of verse, and for words (in the assignment of stress). Though the two domains are distinct, similar computations are employed in both domains. In the next section, we see that metrical structure in music is also generated by the same kind of rule.

Metre in music

In their pioneering book *A Generative Theory of Tonal Music*, Lerdahl and Jackendoff note that music also has metre and describe the metrical structure as involving those aspects of music that relate to the 'regular alternation of strong and weak beats at a number of hierarchical levels' (1983, p. 8). We show below that the beat pattern of music is the result of grouping the timing slots. In this we differ from Lerdahl and Jackendoff (1983), for whom 'beats do not possess inherent grouping' (p. 28).

The metre of music is traditionally represented by an array of dots which in its essentials is identical with the metrical grids that were studied above. For Lerdahl and Jackendoff, the grid is not generated from the piece of music, but instead comes into existence independently, and represents the rhythmic organization of the music. The asterisks (dots) on the lowest gridline represent beats (defined by them as points without duration, which are equidistant in time); beats which project to a higher gridline are experienced as stronger than beats which project to a lower gridline. The rule system itself places no limit on the number of gridlines, but, as Lerdahl and Jackendoff note, there is a practical limit based on human auditory perception of rhythm.

We accept the general principle that metrical structure in music is represented as a grid. However, this is a grid not subject to condition (22), which requires that the last generated gridline contain one asterisk: this condition helps control the length of the line in poetic metre. This means that the last-to-be-generated gridline may contain

any number of asterisks, depending on how long the musical piece lasts. We suggest further that gridline 0 is projected not from the timeless beats, but from the time intervals between beats, with a defined, minimal time interval projecting as one asterisk. Both pitches and silences are assigned to time intervals. As a consequence, in music, a gridline 0 asterisk can project from a rest or silence, which is not possible in the metrical structure of verse.

However, we propose a major change in the explanation of why some grids are possible and others are not. Since for Lerdahl and Jackendoff any two-dimensional array of dots (asterisks) constitutes a potential grid, their major problem is to rule out the many arrays that are not possible as representations of musical form. For example, the grids in (32) are possible musical grids, but the grids in (33) are not.

```
(32)
 (a) * * * * * * * * *  0   (b) * * * * * * * * *  0
     *   *   *   *   *  1           *     *     *  1
         *           *  2           *           *  2

(33)
 (a) * * * * * * * * *  0   (b) * * * * * * * *  0
     *       *   *   *  1           *   *     *  1
         *           *  2           *         *  2
 (c) * * * * * * * * *  0   (d) * * * * * * *    0
       *       *     *  1         *   * *   *    1
       *             *  2         *     *        2
```

To rule out all the impossible grids, Jackendoff and Lerdahl (1983) formulate the metrical well-formedness rules (MWFRs) in (34).

(34)

MWFR 1 Every attack point must be associated with a beat at the smallest metrical level present at that point in the piece.

MWFR 2 Every beat at a given level must also be a beat at all smaller levels present at that point in the piece.

MWFR 3 At each metrical level, strong beats are spaced either two or three beats apart.

MWFR 4 The tactus and immediately larger metrical levels must consist of beats equally spaced throughout the piece. At the subtactus metrical levels, weak beats must be equally spaced between the surrounding strong beats.

(Jackendoff and Lerdahl, 1983, p. 97.)

Each note has an 'attack point', and a beat on the smallest metrical level corresponds to a gridline 0 asterisk in our grid. We have said that each timing slot projects as a gridline 0 asterisk; we can restate MWFR1 as this projection rule if we require that a note fully occupy a set of one or more timing slots. This is little more than a notational variant, but a more substantial difference can be seen when we consider the other three metrical well-formedness rules. In our account, the iterative rules take the place of the metrical well-formedness rules. None of the grids in (33) can be generated by iterative rules alone.

As we show below, all of the MFWRs are redundant if metrical grids are constructed with the help of the iterative rules of grid construction in the manner discussed in the preceding section. In particular, we assume that each note or pause has a duration that is a multiple of some minimal duration (most commonly, a 1/64 note) and we represent the minimal timing slot in a piece of music with a gridline 0 asterisk. Notes and pauses of duration longer than the minimum are assigned several consecutive timing slots. This is our equivalent of MFWR1.

MWFR2 filters out a grid such as (33a) where there is an asterisk on gridline 2 which does not have a vertically corresponding asterisk on gridline 1. In our account, such a grid cannot be generated because asterisks can appear on gridline n+1 only by being projected from gridline n. We do not need a filter such as MWFR2 to rule them out since our formalism generates only continuous asterisk columns.

MWFR3 filters out a grid such as (33c) where a pair of asterisks on gridline 1 is four (gridline 0) beats apart, or (33d) where a pair of asterisks on gridline 1 is one beat apart. Neither possibility can be generated by our rules, which generate asterisks on gridline n+1 by projecting one asterisk per group from binary groups on gridline n or from ternary groups on gridline n; but neither four-asterisk nor one-asterisk groups are possible on gridline n. Our theory does not need to invoke MWFR3 to exclude (33c).

MWFR4 filters out a grid such as (33b) where gridline 1 asterisks project from asterisks which on gridline 0 are sometimes two and sometimes three asterisks apart. We achieve this result by projecting gridline 1 asterisks as heads of groups of asterisks on gridline 0; the gridline 0 groups must be either consistently binary or consistently ternary, because parentheses are inserted by the iterative rule either at binary or at ternary intervals (but never a mixture).

Thus according to our account, all and only the well-formed grids can be generated by the iterative parenthesis-insertion rules which are based on setting the parameters in (31) and which are a part of universal grammar. In contrast the MFWRs do not have universal status, but instead function as ad hoc conditions which rule out some logically possible arrays of asterisks and parentheses. Our proposal that binary and ternary grouping is shared with the groupings in the assignment of stress to words and in the characterization of metres in metrical verse suggests a closer relationship between language and music than is found in Lerdahl and Jackendoff's account (1983, p. 85). (This gap between music and language is extended further in Lerdahl (2001, p. 285) who argues that this metrical grouping is related to grouping principles found in other aspects of musical form such 'pitch space'.)

We conclude this comparison by noting that Jackendoff and Lerdahl use the term 'grouping' to mean something different than we do. Their 'grouping structure' is the organization of the music into motives, themes, and so on (1983, p. 13); and they distinguish it from metrical structure. They consider grouping to be accessible to consciousness, and note that the elementary units of music (the notes and pauses) are perceived as being grouped together. It is largely for this reason that they insist that the metrical structure does not involve grouping: they say that there is no perception of a strong beat as grouped with the preceding weak beat but not grouped with the following (or vice versa). We agree with this claim about perception, while nevertheless

holding that metrical elements are grouped, where grouping is implemented by the insertion of parentheses as sketched above. A metrical grid is a structure which explains certain aspects of the organization of syllables in the words of a language, or in the lines of metrical verse, and also of timing slots in music. There is no conscious access to the rules which generate the grid or to the grid itself, and so no consciousness of the grouping of elements. The parentheses and the grouping of elements are part of the algorithm which generates the grid, but the grouping is not directly perceived in any of the three modalities: music, language, or poetry.

In (36) below, we show two lines of a song by George Peale, set to music by John Dowling. Above the line we indicate the notes,[7] and the musical grid; the musical grid is generated by the rules in (35). Gridline 0 asterisks project from timing slots corresponding to half-notes. The musical grid should be thought of as continuous over the two verselines. The last gridline to be generated, gridline 2, is ungrouped and contains any number of asterisks (it is not subject to condition (22)). (Below the lines we write the grid for the metrical structure of the poetry, which we return to shortly.)

(35)

(a) Gridline 0: starting just at the L edge, insert a L parenthesis, form ternary groups, heads L.

(b) Gridline 1: starting just at the L edge, insert a L parenthesis, form binary groups, heads L.

(36)

```
  *                   *            *              2
 (*          *       (*      *    (*              1
 (*   *  *  (*  *  * (* *  * (*  * *(*  *   *(    0
  𝅗𝅥    𝅗𝅥   𝅗𝅥   𝅝    𝅗𝅥    𝅝    𝅗𝅥    𝅗𝅥   𝅝   𝅝      ▬
 His golden locks time hath to silver turnde,
 )*   *) *   *)     *    *)   *  *) *   *)      0
      *      *)          *       *      *)      1
            (*                          *(      2
             *                                  3

              *                         *           2
 *           (*           *            (*           1
 (*  *    *  (*    *   *  (*     * *   (*   *   *(  2
  𝅗𝅥   𝅗𝅥    𝅗𝅥    𝅗𝅥     𝅗𝅥    𝅗𝅥    𝅗𝅥      𝅗𝅥 𝅗𝅥   𝅗𝅥    𝅗𝅥    ▬
 O time too swift, O swiftness never ceasing,
 )*  *)    *    *)   *   *)  *     *)*  *) *     0
     *          *)       *         *    *)       1
               (*                       *(       2
                *                                3
```

(George Peele (words), John Dowland (music), 'His golden locks', 1587.)

7 We assume a simplified representation of the musical rhythm, which we have adapted from the transcription in Greenberg, Auden, & Kallman (1956, p. 95). The time signature is 3/2; barlines in this edition precede the notes which we project to gridline 1. The word 'O' in the second line projects as a quarter-note preceded by a quarter-note rest, not shown here.

The relation between the musical grid and the rhythm of the music is captured formally by the rule (37).

(37) A note projecting to gridline n+1 is assigned greater prominence than the nearest note projecting to gridline n.

Below each of the lines of the poem we show the metrical grid that characterizes their metre, where each gridline 0 asterisk projects from a syllable; the lines are in iambic pentameter, and the grid is generated by the rules in (27).

The musical grid above the line and the poetic metrical grid below the line are both generated by rules based on the parameters in (31). This identity reflects a deep connection between music and language, but in almost every other respect, the musical and poetic representations are different. Though the rules are of the same kind, the actual rules used to generate the respective musical and metrical grids are different. The syllables which project (via their notes, and timing slots) to higher-numbered gridlines in the musical grid are not always the syllables which project to higher-numbered gridlines in the metrical grid of the poem. (The two lines of the song differ in that the first line has a closer correlation between text and tune than the second.) Note in particular that the words *golden* and *swiftness* contain maxima which must project to gridline 1 in the poetic metrical grid, but do not project as prominent in the musical grid.

For poetic metre, the grid functions to fix the length of the line, and this is achieved by a combination of requiring iterative rules on every gridline and the condition (22) requiring the last-to-be-generated gridline (the bottom-most in our grids) to have one asterisk. The first of these requirements is relaxed in the assignment of stress to words, which can therefore have groups of any length on certain gridlines and so no overall control on size. The second of these requirements is relaxed in music, where the last-to-be-generated gridline can contain any number of asterisks; unlike metrical poetry, musical pieces are not divided into metrical sections (lines) of pre-set lengths.

Conclusion

Above we examined three very different sets of data—the stress of words, metre in poetry, and metre in music—and we showed that these different bodies of fact are accounted for by grouping elements in a linear sequence into the hierarchical pattern of the metrical grid. In all three cases the elements were grouped into either pairs or triplets, marking one of the elements as head. The heads constitute a new (and sparser) sequence of elements, which in turn is subject to the further grouping. In the assignment of stress to words and in the construction of poetic metre the elements subject to grouping are syllables, whereas in music, it is timing slots that are grouped.

Metrical structure in all three cases is not a characteristic of the acoustic signal, but part of a mental representation of words or lines of verse, or sequences of music. Since language, poetry, and music are all products of the human mind it is perhaps not surprising that we find the same mechanism operating in these three domains. What is new here is the detail, namely that the entities subject to grouping are syllables in language, but timing slots in music; that grouping is implemented everywhere by junctures of the same kind (which we have represented by ordinary parentheses), and that grouping is recursive; i.e. that the heads of groups are projected on a separate gridline and are themselves subject to grouping.

Underlying the various groupings there must be specific neurobiological mechanisms, the nature of which is a mystery at this time. In view of the rapid advances that biology has achieved during the last fifty years, it can be hoped that this mystery will be solved in the not too distant future.

References

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

Fabb, N., & Halle, M. (2008). *Meter in poetry. A new theory.* Cambridge: Cambridge University Press.

Greenberg, N., Auden, W. H., & Kallman, C. (1956). *An Elizabethan song book. Lute songs madrigals and rounds.* Garden City, N.Y.: Doubleday Anchor Books.

Halle, M. (1998). The stress of English words. *Linguistic Inquiry, 29*, 539–68.

Halle, M., & Vergnaud, J-R. (1987). *An essay on stress.* Cambridge, MA: MIT Press.

Haas, M. R. (1977). Tonal accent in Creek. In L. Hyman (Ed.), *Studies in stress and accent* (Southern California Occasional Papers in Linguistics vol. 4, pp. 195–208). Los Angeles, CA: University of Southern California.

Hayes, B. (1981). *A metrical theory of stress rules* (PhD dissertation). MIT, Cambridge, MA.

Hayes, B. (1995). *Metrical stress theory.* Chicago, IL: University of Chicago Press.

Idsardi, W. (1992). *The computation of prosody* (PhD Dissertation). MIT, Cambridge, MA.

Lerdahl, F. (2001). *Tonal pitch space.* New York: Oxford University Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge MA: MIT Press.

Liberman, M. (1975). The intonational system of English (PhD dissertation). MIT Department of Linguistics, Cambridge, MA.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249–336.

Prince, A. (1983). Relating to the grid. *Linguistic Inquiry, 14*, 19–100.

In P. Rebuschat, M. Rohrmeier, J.A. Hawkins and I. Cross Language and Music as Cognitive Systems Oxford University Press 2012 pp. 4-21.