with the collaboration of Wolfgang Kesselheim & Stavros Skopeteas 770–783. Berlin: Mouton de Gruyter.


**Reviewers’ address:**

Anna Grashchenkova & Pavel Grashchenkov
Russia, Moscow, 109044
1aja Dubrovskaja str., 8/12, 85
izmaja@mail.ru, gra-paul@mail.ru

---


Reviewed by Klaus J. Kohler (Institut für Phonetik und digitale Sprachverarbeitung, Christian-Albrechts-Universität Kiel)

**Scope and overview**

This monograph is a slightly revised and shortened version of the author’s PhD thesis, which was written at the University of Munich under the supervision of Theo Vennemann. It consists of five chapters. A two-page Introduction gives an overview of the contents and sets the theme: the critical assessment of relevant publications on rhythmicity in languages and the subsequent development of a new linguistic typology of rhythmic contour formation.

Chapter 1 traces the history of what the author calls “the phonology of rhythm”, thus making it clear right from the start that his approach to speech rhythm is phonological and that phonetic manifestation is subserviant to underlying abstract phonological coding requirements in a specific language. Two types of rhythm studies are distinguished:

1. The investigations that have their roots in phonetics regard rhythmicity as an isochronous recurrence of prosodic units, either feet or syllables, and
classify the languages of the world as being either stress-timed or syllable-timed.

2. On the other hand, Metrical Phonology relates rhythmicity to the regular sequencing of prominent and less prominent sound portions, with languages organising this alternation in various different ways.

In Chapter 2, a distinction is first introduced between rhythmicity and rhythmic contour, the former being defined as the euphonic quality of phonological gestalt which manifests itself to varying degrees in utterances, the latter as the phonological gestalt features that are relevant for the assessment of rhythmicity (p. 83). The aspects of rhythmic contours comprise distinctive duration and prominence patterns in the phonological structures of particular languages, and their different weighting in the individual phonologies determines different rhythmic contours and through them different rhythmicities between languages. Then research into syllable isochrony and into the accentual basis of rhythm is critically evaluated, summarised in the conclusion that the theory of rhythmicity in the isochrony tradition went to far, whereas Metrical Phonology was too narrow as a universal phonological framework for rhythm.

Chapter 3 elaborates a new typology of rhythmicity in languages, starting from the premiss that phonological coding requirements of duration and accent take precedence over eurhythmic preferences. The question therefore is how rhythmic contours can arise given the language-specific distinctivity constraints. Since the two rhythmically relevant prosodic dimensions, time and prominence, can contribute to this coding power, but not simultaneously, the author sets up four possible types of rhythmic contour formation:

- A preponderant use of the time dimension in word phonology results in the mora-based rhythm, as in Japanese, West-Greenlandic or Finnish, languages with very different morphosyntactic structures.
- The use of the time dimension for syntactic coding leads to a phrase-based rhythm, as in French, and less clearly so in Korean.
- In a prominence-based rhythm, the temporal structure of utterances is guided by extra-phonologically determined word and sentence accents, with different syntactic, information-structural, and morpholexical coding potentials in, e.g., English and German versus Russian.
- Languages that completely, or at least largely, lack the distinctive phonological use of both time and prominence constitute the alternating rhythm type, such a Italian and Spanish.
In Chapter 4, the author adduces empirical justification from experimental phonetics, speech technology, psycholinguistics, and comparative metrical poetics for the priority of prosodic contrastiveness and for the proposed inventory of rhythmic types. A scant two pages constitute Chapter 5 titled “Summary and Outlook”, ending with the negative statement that, to this day, we lack a theory of rhythmic language change.

**The author’s phonological approach**

Dufter deserves praise for a critically annotated, comprehensive survey of research into speech rhythm in languages, for introducing some order into the overwhelming jungle of publications from a bewildering variety of perspectives, and for attempting a new classification of rhythmical contour types. It is equally noteworthy that he makes a strong point for the separation of accentuation in word and sentence phonology, criticising Metrical Phonology not only for using English as the yardstick of any discussion on rhythm and for dragging representatives of other languages, such as François Dell, into their conceptual framework, but more particularly for treating lexically and morphologically conditioned stress as well as sentence accent and rhythmic structure indiscriminately in the same theoretical paradigm.

However, the argumentation is too narrowly phonological and does not give full credit to the signalling power of phonetic properties in speech production and perception as a guiding function in speech communication: the temporal recurrence of similar signal stretches helps listeners to decode speech more efficiently because incoming signals can be projected onto an ongoing pattern. Rhythm is thus essentially a dynamic bodily action that exhibits varying degrees of regularity in time and intensity for a listener, and it goes beyond static distinctive phonological objects of sound duration and accentuation. Rhythm is phonetic performance operating on prosodic structures set by the language. The prosodic structures may be highlighted or overruled in performance; there are good rhythmical speakers, and there are bad ones; on the political scene, former German Chancellor Gerhard Schröder is an example of the first type, former GDR head of state Erich Honecker, a clutterer, and Bavarian Minister President Edmund Stoiber, a very disfluent speaker, are examples of the second. The more regular speakers are in their action dynamics the more easily will they be understood. How important this performance aspect is can easily be judged by the intelligibility or rather the usual lack of it in monotonous, get-it-over-with-fast announcements at airports and on planes.
This principal attribute of rhythm in speech communication is not focused on in this monograph, the discussion of rhythm thus lacks an essential component. The reference to experimental phonetics in Chapter 4 is anecdotal rather than an integral part of a theory of rhythm.

The author also misses an important point in the scientific evaluation of the isochrony discussion. The search for phonetic isochrony is an example of behaviourist surface analysis which tries to soak up structural principles from empirical measurement. At the time the hypothesis of isochronous feet or syllables was put forward it should have been realised a priori that it would have to be rejected. It was known then that duration varies through stress, syllable structure, phrasing and other sentence and performance variables, as well as from repetition to repetition. Thus syllables in a so-called syllable-timed language like French cannot be isochronous in the strict sense of the term, and the compression in a so-called stress-timed language like German cannot be complete to yield isochrony from accent to accent. Furthermore, no threshold was ever proposed to separate isochronous from non-isochronous stretches of speech. So time and effort were not well spent on measuring the durations of syllables or feet, because they could never support the claim, but could not falsify it either in the absence of an independently motivated variability margin. The attempts were simply bad science and should have been given a more cursory and at the same time more illuminating discussion in this book.

An alternative phonetic approach

In Dufter’s static linguistic structural framework, rhythm is a deep-seated characteristic of a language type that permeates all language performance. What we still need is a flexible dynamic framework that accounts for rhythm’s variability in the same language from spontaneous speech to text reading, to the reading of constructed isolated sentences and finally to verse recitation. Rhythm is a feature of speech communication and needs to be investigated as such, taking into account its guiding function and the varying constraints imposed on it by different communicative situations and tasks.

There are three timing strands in speech production: subglottal (breathing), glottal (phonation, voice pitch), and supraglottal (vocal tract articulation). Each one has its own cyclicity, and they are synchronized in structured patterns incorporating features that are universal in human language, characteristic of language types, and specific to individual languages, and which are disrupted in speech and language pathologies, such as stuttering and cluttering. Subglottal
and glottal timing patterns are first and foremost long-term in aerodynamic cycles of vocal effort and breath group organization, and in rising–falling or falling–rising pitch cycles, respectively. Supraglottal timing, on the other hand, is more short-term in opening–closing and closing–opening vocal tract dynamics, with a concomitant waxing and waning of radiated acoustic energy, resulting in syllables and, in the last resort, in sound segments. This difference between the three strands gave rise to the classificatory dichotomy suprasegmental vs. segmental. However, the typical temporal courses in these strands may be reversed. Subglottal and glottal timing are short-term in force accents (Kohler 2003a, 2006) or in phonatory sound distinctions (voiced, voiceless, breathy, laryngealized), riding on the long-term-settings. And on the other hand, vocal tract gestures are integrated into long-term settings, providing the basis of articulation, over-all speech rate, and rhythmic groupings of syllables. Different coordinations of these strands in different languages determine different rhythmic patterns (Kohler 2003b).

**Dufter's prominence-based rhythm**

In Germanic languages, more particularly in West-Germanic ones, but also in Russian, pitch cyclicity creates a sequence of prominences by pitch changes from high to low or vice versa (Isačenko and Schädlich 1970). In these languages, words also have the phonological attribute of a stressed syllable, which is a place holder for pitch prominence to be hooked onto at the phrase level if the word is to be accented, and which can be captured in a citation form utterance. Such stressed syllables have a greater range of complexity in their vocal tract dynamics, i.e. in their segmental make-up, than unstressed ones, where we find fewer vowel contrasts, frequent reduced, central vowels, schwa elision, and less complex consonant clusters. Stressed syllables are thus likely to have greater supraglottal articulatory weight. When they are inserted into a chain of pitch prominences, which is synchronized with the stressed syllables, this weight is heightened and the perceptual prominences are increased. In addition, prominent syllables are slowed down, non-prominent syllables are speeded up and simplified in movement trajectories (resulting in reduction and elision of segments). This produces a recurrence of prominent–non-prominent alternations (feet), which is additionally regularized by prominent syllables being shorter in plurisyllabic as against monosyllabic feet. This does not create isochrony, but it creates recurrent regularity. The more clearly the accent patterns are marked by pitch peaks and valleys, duration, and concomitant acoustic energy differentials between prominent and non-prominent sections of feet, the stronger the rhythm becomes for a hearer and the more powerful is its guide function.
From the point of view of prominence-based rhythm it is thus irrelevant whether a language has free lexical stress, like Russian, or rule-governed stress assignment, like English and German, or fixed stress, like Hungarian; it is the prominence regularity, spanning lexical items, that counts. The contrastivity of lexical stress is usually quite limited anyway, and extremely so in German. On the contrary, if the accent pattern is associated wrongly with stress places, even if there is no lexical differentiation in the language by stress in a particular segmental chain, the false prominence pattern interferes more with intelligibility in such a language than any other segmental or prosodic feature. This is the reason why French speakers are so difficult to understand when they introduce French accentuation and rhythm into speaking a prominence-based language. See further below.

In verse, this rhythmicity is stylized to even greater regularity by various devices, such as the selection, by the composer, of words that provide a more uniform number and weight of syllables across feet, or the realization of more narrowly timed, more finely stuctured pitch cycles by the performer. Finally, the prominent syllables may be additionally marked by alliteration to increase the impression of regularity. This was, of course, the poetic device of Germanic verse, but it is also a common feature of modern phraseology, such as “mit Kind und Kegel”, which is more regular and therefore more rhythmic than the reverse word order because it has two syllables per foot, and this regularity is underscored by alliteration. Similarly, we get “mit Pfeil und Bogen” in German, but “with bow and arrow” in English, which both achieve the same rhythmic pattern by the opposite semantic sequence.

So prominence cyclicity is the recurrent regularity in these languages, which may for this reason be called stress-timed. The regularity is most stringent in verse recitation; in spontaneous speech, at the opposite end of the rhythmicity scale, regularity may be constantly disturbed, without being absent altogether. Clear speakers realise it better than slurring and monotonous speakers. But even so, the corollaries of prominence cyclicity — compression, reduction, elision in non-prominent syllables, vowel centralization, preferred word sequences — are always operative. However, even quite strict regularity does not mean isochrony.

Dufter’s alternating rhythm

In Spanish and Italian, there are also recurring pitch prominence cycles hooked onto lexical stress positions, but non-prominent syllables in these cycles are not reduced in spectrum and duration to the same extent as in the Germanic
languages. Moreover, syllables in these languages are of very restricted structural
types, with CV sequences dominating, and complex clusters being absent, espe-
cially finally. Spanish has no distinctive vowel quantity, in Italian long vowel +
consonant and short vowel + geminate tend to occupy the same time slot. There
is therefore a very strong regularity of syllable cycles within the prominence cy-
cles. So the auditory impression of recurring events is far less accent, but more
syllable oriented. On the basis of this rhythmic organization it is only natural that
verse metre should be based on fixed numbers of syllables, rather than accents,
per line which can again be regularised more strongly by selection in verse com-
position and recitation. And lines would be most efficiently marked at their ends
by rhyme of final syllables, rather than by alliteration of their accented syllables.

Dufter’s phrase-based rhythm: The case of French

There is agreement that French has no lexical stress, although Dufter maintains
it does on the final syllable (p. 93). He misinterprets the stronger prominence
of the final syllable in a phonetic phrase, which may of course also be a word
in citation form, with stress as a phonological place holder in lexical items.
The word in French is not a phonetically marked unit, neither by culminative
stress, nor by phonetic boundaries, as is demonstrated by the abundance of
word play of the type

(1) “De quelle couleur est toujours un coffre-fort quand on le vide? — Il est tout
vert. (Il est ouvert.)”

(2) “Un abbé et un athée tombent dans un bourbier. Il en sort deux provinces
de la Grèce. — La Thessalie et la Béotie. (L’athée sali, et l’abbé aussi.)”

French does not highlight words by pitch accents either, but marks the end of
phrases by a phrase accent, based on pitch and lengthening, in close association
with syntactic parsing and more generally with sense groups. Coustenoble and
Armstrong (1934) give the following assessment of accentuation in French:
“That, in connected unemphatic speech stress does not belong to the word, as in
English, but to the word group: the word does not bear it unless it occurs finally
in that group.” (p. 4)

Vaissière (2006) refers to French as une langue à frontières (p.100) with
different degrees of boundary strength (p. 97) marking the end of sentences
and mots prosodiques within them (p. 109ff). The boundary of the sentence-in-
ternal prosodic word (group) is signalled by lengthening and rising (continu-
arion) pitch of the final syllable, the greater the extent of both, the stronger the
perceived internal boundary. This use of pitch to mark prosodic phrases that coincide with semantic and syntactic groupings in French thus differs from the use of pitch in accent units that may cut across sense groupings in, e.g., English and German. The consequence of this phrasing role of pitch, heightened by final-syllable lengthening, is that all the syllables in a prosodic phrase before the last one have a similar degree of prominence that differs from the much more prominent final syllable. Furthermore, inside such prosodic phrases (mots prosodiques) the syllable type (C)V is extremely frequent, because of the disappearance of lexical word boundaries. A typical sentence like

(3) “Marie vient à Paris demain.”

has no other syllable structure (vient à may be either CVCV at a stylistically higher level or CVV more colloquially). In

(4) “Cet homme | est énormément bête.”

and

(5) “Cet homme est énorme, | et m’embête.”,

Vaissière (2006, p. 112) describes the same segmental sequence as organised into two different prosodic phrases (which I have marked by |). In both sentences, both phrases end in closed syllables, whereas the phrase-internal syllables are open, with the exception of “-nor-” in the first sentence. Closed syllables in the already prominent phrase-final position increase this prominence and thus contrast it more strongly with the more even flow of prominence preceding it, which has a high frequency of occurrence of (C)V.

The presence or absence of e caduc is also a means to regulating the complexity of syllables inside prosodic phrases and at the end. “une semaine (vs. la s(e)maine), ours[e] blanc, portefeuille, quatre-vingt, quart(e) chevaux (vs. deux ch(e)vaux)” are examples of avoiding heavy consonant clusters inside prosodic phrases by creating additional CV[=ә] syllables. If there is another simple syllable before the phrase-final one, [ә] can be absent, as in “port(e)manteau” (Léon 1966; Malécot 1976, Kohler 2002). Similarly, Passy and Rambeau (1918) transcribe

(6) “Les choses en restent là pour l’instant.”
[le ʃɔs ә rεstә la pur lεstә.] (p. 16, ә in italics = optional)

(7) “Le bonhomme, qui n’avait pas grande littérature, ne comprends pas et reste là, tout confus.”
[lә bɔnm, ki nә vε prә prә la, tu kәfy.] (p.18).
Bold type indicates prominent syllables, single blanks separate lexical items, larger blank spaces prosodic phrases. The strongest boundaries are at the end of each sentence and after the 5th prosodic phrase in (7). Next in strength are the boundaries after the 1st, 3rd, and 4th prosodic phrases in (7) (also marked by final consonant lengthening), which may be further graded 1st, 3rd > 4th; the 2nd boundary is the weakest. In the case of _reste là_ preceding such a strong boundary, [ә] is obligatory before the immediately following final syllable. In the first sentence, however, the boundary between _en restent là_ and _pour l’instant_ may be weaker, due to the greater semantic unity, and [ә] may be absent or present depending on how strong the speaker wants to mark the boundary.

All these sentence-phonetic processes produce a staccato effect in the perception of spoken French (Coustenable and Armstrong 1934, p. 3), or what Lloyd James (1940, p.25) called _machine-gun rhythm_, as against _morse-code rhythm_ in English. The prosodically marked phrases recur less regularly than the articulatory cycles composing each one of them, thus French exhibits a rhythmic structure that is the opposite of Dufter’s prominence-based type, found in, e.g., the Germanic languages. In this sense, French is a _syllable-timed_ language, with a regular recurrence of vocal tract gestures inside prosodically marked phrases. The essential characteristic of these syllabic events is their regularity within a margin of variability, not their isochrony.

However, in prosodic phrases, specific words may receive an emphatic reinforcement — _accent d’insistance_ (Grammont 1934, pp. 1), which is not controlled primarily by pitch, as in the Germanic languages, but by force on the intensified syllable. It increases the acoustic energy, pitch, and vowel duration, but, most importantly and different from the phrase-final rhythmic prominence, it interferes with the supraglottal cyclicity through lengthening of the initial consonant of the first syllable in the word that has a consonantal onset. If the word starts with a vowel the consonantal onset may be the consonant of the preceding word, or it may be created as a glottal stop. Thus the last word of _c’est impossible_ and _c’est épouvantable_ may either have a lengthened [p] initially in the second syllable, or begin with a lengthened [t] (_c’est im-/é-) or an initial [ʔ]. Example (7) shows this emphatic accent in [tu] of the last prosodic phrase.

The extension of prosodic phrases depends on the degree of delicacy in the marking of sense units. For example, (4) and (6) may also constitute a single prosodic phrase, each integrating two more CV syllables. Or contrariwise, long prosodic phrases may be more finely graded to map greater delicacy of syntactic and semantic structuring. By doing this, prosodic phrases recur more regularly with more regular internal syllable cyclicity, which contributes greatly to
effective, high-quality prose reading and verse recitation. The same considerations apply to verse metre as for Italian and Spanish.

Dufter’s mora-based rhythm

Finally, a brief word on Dufter’s mora-based rhythm. The basic cyclicity for Japanese rhythm is the vocal tract alternation of closing and opening in sequences of CV. Either the close beginning or the open end of the movement may be lengthened, so we get long consonants and long vowels, and the close beginning may be absent, i.e. we get V as well. The lengthening of vowels and consonants has been interpreted as doubling the basic CV cycle. The consequence of this is the postulate of the morae CV, V, C(CV) and (V)V. The structural interpretation of this mora concept (Warner and Arai 2001 a,b) refers to recurring simple (C)V structures, to the lack of an accent effect on duration and to the irregular placing of pitch accents in syllable strings. All these factors guarantee a high regularity in the sequencing of morae by themselves, without the need for an independent temporal mora control. But even this very simple cyclicity does not create isochrony, as has been clearly shown by Sagisaka and his research group, summarized in Sagisaka (2003). Rather, the lack of isochrony is expected variance that does not interfere with the perceived regularity. Under this perspective Japanese turns out to be one of the most basic syllable-timed languages.

Conclusion and Outlook

Rhythm is related to events that unfold in time, the recurrence of similar structured events at regular intervals. Regularity and temporal recurrence are essential, not isochrony of static units (mora, syllable, foot). Dufter has presented a phonological approach with reference to units the linguist sets up in a framework of categorical distinctions. However, the resulting phonological typology lacks the perspective of speech communication. This is the point where the phonetician has to continue. On the basis of the phonetic theoretical framework which I have outlined, we can now carry out scientifically sound empirical experiments. What we should start with is to collect data of reading a standard text from trained and untrained speakers, providing a spread along the scale from good to bad rhythmicity. These data should then be evaluated by listeners on, say, a 5-point rhythmicity scale from very good to very bad. The clustering of the data resulting from this perceptual evaluation are to be analysed with regard to the differences in syllabic timing and pitch patterns.
From these differences hypotheses can be drawn as to the margins of regularity that have to be observed if speech is to be rhythmical. These hypotheses are the basis for systematic parameter manipulation in speech synthesis and can be tested by perceptual assessment of the degradation or improvement of examples classified as good or bad rhythmic speech, respectively. All these empirical steps need to be carried out for a variety of languages from all four typological categories. Speech synthesis has already been applied in an engineering environment to rhythm studies in Japanese with such a goal in mind (Kato 1999).

Dufter has set a milestone in the phonological discussion. We can now proceed to do empirical research within the wider framework of speech communication. A vast and long overdue undertaking!

References


**Reviewer’s address:**

Prof. Dr. Klaus Kohler
Holm 4
D-24113 Molfsee, Germany

---


Reviewed by Martin J. Endley (Hanyang University, Seoul)

This volume contains a selection of papers delivered at a conference on modality in contemporary English held at the University of Verona in September 2001. One particularly pleasing feature of the book is that it contains contributions from scholars who subscribe to different theoretical positions and who adapt different methodologies. As a result, it provides a series of valuable snapshots into the types of research currently being done in this particular subfield of linguistics. Indeed, having worked one’s way through the book, the overall impression one is left with is that the study of modality within contemporary English continues to be a rich linguistic seam, one marked by considerable diversity of outlook and methodological approaches.

The book is arranged into four broad sections, each corresponding to a major theme. The four sections are: the semantics and pragmatics of core modal verbs, the status of emerging modal items, stylistic variation and change, sociolinguistic variation and syntactic models. A further noteworthy feature of this book, and to my mind a highly welcome one, is that discussions are not restricted to the familiar closed set of core modal auxiliary verbs (i.e. *can, could, may, might, must, shall, should, will*, and *would*) which share the so-called NICE properties (see Huddleston and Pullam, 2002, 92–3): instead several contributions deal with other forms and expressions which involve modality to various degrees; indeed, in some papers these more marginal realizations of modality are the main focus. The overall effect of this volume then is to present modality as very much an “open” field, with speakers using numerous linguistic items to realize what has to be understood as a crucial communicative need.