5th European Masters in Language and Speech Summer School
Institute for Communications Research & Phonetics,
University of Bonn

07/12/2004 – 07/16/2004

Tutorial 5: Analyzing Speech Rhythm

Ingmar Steiner
ist@ikp.uni-bonn.de

Overview

The purpose of this tutorial is twofold:

1) Learn about and discuss recent studies of speech and language rhythm
2) Learn how to script Praat and use it to analyze speech data and manipulate it

We will proceed along the following course:

Monday session:

1. Background on rhythm research
2. Introduction to the BonnTempo corpus
3. First steps in Praat scripting
4. More background
5. Simple analysis of isochrony in speech

Tuesday session:

1. Recent approaches to acoustical correlates of rhythm
2. More analysis along those lines
3. Preparatory annotation of speech data for Session 3

Wednesday session:

1. "Isochronization" of speech data in Praat
2. End discussion

Session 1

Opening questions:

What is rhythm?
What is speech rhythm?
What units of speech are involved?
What characterizes these units?
Why does speech have rhythmic properties?

Literature for discussion:

3.6.2 Simple Rhythm Units (Stress-Timed and Syllable-Timed)

English sentences are spoken with recurrent bursts of speed, with long or short pauses or with intonation breaks between. A sentence or part of a sentence spoken with a single rush of syllables uninterrupted by a pause is a RHYTHM UNIT. […]

A rhythm unit which contains one, and only one, primary contour is a SIMPLE RHYTHM UNIT. […]

The timing of rhythm units produces a rhythmic succession which is an extremely important characteristic of English phonological structure. The units tend to follow one another in such a way that the lapse of time between the beginning of their prominent syllables is somewhat uniform. […]

(Controlled strictly and mechanically in poetry – and possibly partially so in some types of elegant prose, - the recurrent stress timing is perhaps even more important than the number of syllables in iambic or trochaic groups, or the like. Evidence of this fact is seen in the esthetic satisfaction obtained by English speakers from some lines of poetry […] [Cf. also footnote 50.]

The tendency toward uniform spacing of stresses in material which has uneven numbers of syllables within its rhythm groups can be achieved only by destroying any possibility of even time spacing of syllables. Since the rhythm units have different numbers of syllables, but a similar time value, the syllables of the longer ones are crushed together, and pronounced very rapidly, in order to get them pronounced at all, within that time limitation. This rhythmic crushing of syllables into short time limits is partly responsible for many abbreviations – in which syllables may be omitted entirely – and the obscuring of vowels; it implies, also, that English syllables are of different lengths, with their length of utterance controlled not only by the lexical phonetic characteristics of their sounds but also by the accident of the number of syllables in the particular rhythmic unit to which they happen to belong at that moment. […]

A single rhythm unit from such a sequence of units may be considered the regular or normal type. Because its length is largely dependent upon the presence of one strong stress, rather than upon the specific number of its syllables, it may conveniently be labelled a STRESS-TIMED rhythm unit (a phonemic type in contrast to syllable-timed units to be mentioned below, with both of them on a different level of contrast from the simple versus complex rhythm types).

Many non-English languages (Spanish, for instance) tend to use a rhythm which is more closely related to the syllable than the regular stress-timed type of English; in this case, it is the syllables, instead of the stresses, which tend to come at more-or-less evenly recurrent intervals – so that, as a result, phrases with extra syllables take proportionately more time, and syllables or vowels are less likely to be shortened and modified.

English also has a rhythmic type which depends to a considerable extent upon the number of its syllables, rather than the presence of a strong stress, for some of its characteristics of timing; in English, however, the type is used only rarely. In these particular rhythm units each unstressed syllable is likely to be sharp cut, with a measured beat on each one; this recurrent syllable prominence, even though the

stressed syllables may be extra strong and extra long, gives a "pattering" effect. The type may be called a SYLLABLE-TIMED rhythm unit (in a phonemic contrast to the stress-timed type).

If the unstressed syllables are each made quite abrupt, the unit becomes somewhat STACCATO. If the unstressed syllables are more or less equally timed, and somewhat prominent, but glided or smoothed together, the general impression is that of a SPOKEN CHANT. […]

For Latin-Americans, the shift from their tendency toward a syllable-timing rhythm to the English normal stress-timing rhythm is highly necessary, but comprises one of their greatest problems. In learning to speak English they must abandon their sharp-cut syllable-by-syllable pronunciation and jam together – or lengthen where necessary – English vowels and consonants so as to obtain rhythm units of the stress-timing type.

Words in a very close grammatical association are likely to belong to the same rhythm unit. […]

Words that have no innate lexical stress tend to join that rhythm groups preceding or following them with which they are grammatically most closely related […]

In traditional English orthography, a punctuation mark usually, but not always, represents (1) a pause, and, therefore, (2) the end of a rhythm unit; in addition, it sometimes gives (3) a partial indication of the attitude of the speaker – a fact which, in turn, conditions the stress placement, or degree of stress, or intonation placement, or specific intonation, or even the quality of the voice, or some combination of these. Punctuation marks are often supplemented by italics, or capital letters, and so on, to make the stress and intonation type and placement more specific.

50 […]

The subject of rhythm has a very large literature, which I shall not review here. One of the most interesting of the instrumental studies is that of A. Classe, The Rhythm of English Prose (Oxford, 1939). Classe measures the quantity of syllables of different phonetic types, in different phonetic places in relation to stress groups and grammatical structure. Although he has (3) "purposely avoided the question of intonation," nevertheless his studies are applicable to it, since, in part, he is testing some of the rhythm theories of Daniel Jones. He concludes (132) that "An English sentence is normally composed of a number of more or less isochronous groups which include a varying number of syllables"; (101) "…stress-groups containing various numbers of syllables tend, everything else being equal, to assume approximately the same duration, the length of syllables must vary…" On the other hand, he shows many measurements which are not completely equal, and for which subrules can often be given. In groups of more that three syllables, these subrules oftentimes do not hold, since all the syllables tend to be of the same speed; (104) "The effect is unquestionably due to the increased speed of articulation for the longer groups, resulting from the tendency to make groups isochronous."

[…]

Pike, Kenneth L. (1945) *Intonation of American English*. Ann Arbor: University of Michigan Press, 34-36.

4. Rhythm

Although hesitations and other pauses tend at times to disguise the fact, all human speech possesses rhythm. This emerges clearly during those moments when speech is fluent and uninterrupted. Rhythm, in speech as in other human activities, arises out of the periodic recurrence of some sort of movement, producing an expectation that the regularity of succession will continue. The movements concerned in the rhythm of

speech are those of the syllable- and stress-producing processes, which together make up the pulmonic air-stream mechanism (p. 36). Speech rhythm is essentially a muscular rhythm, and the muscles are the breathing muscles.

There are important implications for perception here. Speech-rhythm is experienced as a rhythm of movement. Obviously it is directly so experienced by the speaker, but what of the hearer? We can say that he, too, is vicariously experiencing a rhythm of movement – he, in a sense, is speaker also. As was pointed out above (p. 23), recognition of 'the identity of speaker and hearer' is essential to an understanding of many aspects of speech perception. We talk, for convenience, about 'hearing' rhythm, but in fact we feel it, entering empathetically into the movements of the speaker, to which the sounds we hear are clues. But in order to have this immediate and intuitive apprehension of speech rhythm it is necessary, of course, that speaker and hearer should have the same mother-tongue – otherwise 'phonetic empathy' will not work: the sounds will not be recognized as accurate clues to the movements that produce them.

It is the way in which the chest-pulses and the stress-pulses recur, their mode of succession and co-ordination, that determines the rhythm of a language. There are two basically different ways in which the chest-pulses and the stress-pulses can be combined, and these give rise to two main kinds of speech-rhythm. As far as is known, every language in the world is spoken with one kind of rhythm or with the other. In the one kind, known as syllable-timed rhythm, the periodic recurrence of movement is supplied by the syllable-producing: the chest-pulses, and hence the syllables, recur at equal intervals of time – they are isochronous. French, Telugu, Yoruba illustrate this mode of co-ordinating the two pulse systems: they are syllable-timed languages. In the other kind, known as a stress-timed rhythm, the periodic recurrence of movement is supplied by the stress-producing process: the stress-pulses, and hence the stressed syllables, are isochronous. English, Russian, Arabic illustrate this other mode: they are stress-timed languages.[7]

When one of the two series of pulses is in isochronous succession, the other will not be. Thus in a syllable-timed rhythm, the stress-pulses are unevenly spaced, and in a stress-timed rhythm the chest-pulses are unevenly spaced. […]

The rhythm of language is the foundation of verse, in most languages. Thus French verse is based on syllable-timed rhythm, and English verse on stress-timed rhythm.[9] Unless this close connexion between ordinary speech rhythm and verse rhythm is recognized, prosodic theory is likely to be unfruitful – as a great deal of it has been in the past. Since verse has its basis in natural speech rhythms, it is not necessary to learn to listen differently from the way one listens to conversation in order to appreciate verse – provided it is composed in one's mother tongue: the rhythm is intuitively experienced by 'phonetic empathy'. (If verse is composed in a language which is not the mother tongue but has been acquired, and which has a different rhythmic basis, then learning to listen differently is unlikely to help to yield the direct intuitive appreciation one has for verse in the mother tongue. Thus English speakers find it difficult to feel French verse to be verse at all.) It is, however, necessary to learn to listen differently in order to be able to analyse speech rhythm, whether of one's mother tongue or another language, and to describe it in general phonetic terms. Few succeed in doing this without training.

7 The existence of the two basically different kinds of speech rhythm was pointed out by Arthur Lloyd James in Speech Signals in Telephony (1940), p. 25: he called them 'machine-gun rhythm' and 'morse-code rhythm'. The more apt terms 'syllable-timed' and 'stress-timed' were coined by K.L. Pike and first used by him in The Intonation of American English, p. 35. Many writers since the eighteenth century have pointed out that in English stressed syllables tend to be isochronous.

9 See David Abercrombie, 'A phonetician's view of verse structure' (1964), for an account of the basis of English verse.

Abercrombie, David (1967) *Elements of General Phonetics*. Edinburgh: Edinburgh University Press, 96-98.

Discussion.

More questions:

How can we verify or refute the claims laid out here?
More generally: how can we analyze speech rhythm?

Empirical examination of speech rhythm would require a speech corpus.
Which points need to be considered to use such a corpus for speech rhythm research?

The BonnTempo Corpus

Read speech at various speech rates. Male and female speakers, German, English, French and Italian native speakers, as well as non-native material. Semantically equivalent text for all languages (translation of a passage from the novel *Selbs Betrug* by Bernhard Schlink (Diogenes Verlag, 1994, S. 242).

German:

> Am nächsten Tag fuhr ich nach Husum. Es ist eine Fahrt ans Ende der Welt. Hinter Gießen werden die Berge und Wälder eintönig, hinter Kassel die Städte ärmlich und bei Salzgitter wird das Land flach und öde. Wenn bei uns Dissidenten verbannt würden, würden sie ans Steinhuder Meer verbannt.

English:

> The next day I went to Falmouth. It is a voyage to the end of the world. After Lincoln the hills and woods become monotonous, after Bristol the towns get boring and near Saintsbury the countryside becomes flat and desolate. If dissidents were banned in our country, they would be banned to the Portishead Bay.

French:

> Le jour suivant, je me suis rendue à Albi. C'est un voyage au bout du monde. Après Lisieux, les montagnes et la forêt deviennent monotones, après Châtel, les villes désolées, et près de Chartreville, la campagne devient plate et déserte. Si chez nous les dissidents étaient exilés, ils seraient alors exilés à Clermont-Ferrand.

Italian:

> Il giorno dopo andai a Bologna. È un viaggio fino alla fine del mondo, dopo Rovereto i colli e i boschi diventano monotoni, dopo Verona le città diventano misere, e presso Revere il paessaggio diventa pianeggiante e deserto. Se da noi un dissidente venisse esiliato, verrebbe esiliato a Ostilglia.

Examples of corpus data.

First steps in Praat:

Praat consists of the **Object Window** as well as the **Picture Window** (for export into text processing applications or as PostScript files). The object window contains the **List of Objects** and to the right, the **Dynamic Menu**.

Each Object in the list has a **Type**, a name, and a session internal **ID** (not visible). Depending on its Type, the Dynamic Menu will change to allow a multitude of actions, including querying (which prints information to the **Info Window**), manipulation, viewing (which opens an **Editor Window**), conversion (which creates a new Object), and painting to the Picture Window.

Objects can be selected with the mouse, and more than one object can be selected by dragging the mouse or holding down the Shift or Ctrl key.

Important: Praat Objects are not files!

They reside in memory only, and are lost when Praat exits. Therefore, it is possible to write all Objects to files. Since all such files contain an identifier in their header, Praat will recognize them as the correct Object Type when it reads them again, regardless of file name extension, but it is good practice to name the files appropriately. Files are read into Objects with the **Read from File…** command.

Initially, the most useful Object Type is **Sound**. To analyze a Sound, use the **Edit** command. This opens a **Sound Editor** window. This consists of a waveform view (oscillogram) and a **spectrogram**. The spectrogram is displayed (or not) according to certain settings, which can be accessed through commands such as **Show analyses…**, **Spectrogram settings…**, etc. Other available analysis options include **pitch**, **formants**, **pulses** and **intensity**.

Sounds can be annotated through **TextGrid** Objects. These contain any number of tiers, either **TextTier** or **IntervalTier** Objects. A TextGrid can be selected along with a Sound to annotate it. In essence, these tiers associate points in time either with points or interval boundaries, as well as with strings of text.

Tiers can be processed into **TableOfReal** Objects, which contain labeled rows and columns, and numeric cells. These can be statistically analyzed either with built-in commands (such as **Get column stdev…**) or exported into spreadsheet applications.

It is possible to use **IPA symbols** for phonetic transcription. For this, Praat contains an implementation of the SIL Encore Fonts, and it is possible to display character

combinations as IPA symbols not only in the Picture Window but also directly in the TextGrid, during annotation.

Praat also includes a host of other interesting sections, including **Multidimensional Scaling**, **Articulatory Synthesis**, **Neural Networks**, and **Optimality Theory**.

To become really productive with Praat, one must take advantage of **Scripting**. Praat is fully scriptable, so that any number of commands can be executed as a batch command for any number of files or Objects, with additional script internal functions.

All commands in Praat, whether they appear in Dynamic or Fixed Menus, in Editors or in the Object or Picture Window, can be used in Praat Scripts, one command per line (spaces and all). Commands that end in **…** require a number of **arguments** that are specified in the same line, separated by spaces. Also, all Praat Commands begin with a capital letter.

Script internal functions, however, begin with a lower-case letter and are used for object selection, string handling, output to the Info Window and most importantly, **variable names**.

Scripts can be written in the **Script Editor** (using the **New Praat script** command) or loaded there from text files. They are executed (in full or in part) with the **Run** command. One useful tool is the **command history**, which enters Praat Commands into the Script Editor exactly as they were executed through the Praat interface.

Our first Praat Script is:

```
printline Hello world!
```

The function `printline` takes the rest of the line and prints it to the Info Window, followed by a line break.

It would be just as easy to create a **string variable** and have `printline` output that to the Info Window. All variables in Praat Scripts are either string variables or numeric variables. String variable are identified by a trailing `$` at the end of their name. Whenever variable names are used as arguments to script functions or Commands, they are enclosed in apostrophes.

```
hello$ = "Hello world!"
printline 'hello$'
```

There are a few **predefined variables**, such as `newline$` and `tab$`, which can be used to control tabs and line breaks:

```
hello$ = "Hello world!"
print 'hello$''newline$'
```

Loops and **conditions** are also important scripting language elements:

```
for myInteger from 1 to 5
 print 'myInteger''tab$'
endfor
```

```
if myBoolean$ = "true"
 printline Condition met!
endif
```

Variables are vital to Praat Scripts. For instance, the output that **Query** commands send to the Info Window can be redirected to string variables. For example, if a TextGrid is selected (these examples assume that the TextGrid contains only IntervalTiers), the following will, of course, output the number of intervals to the Info Window:

```
Get number of intervals... 1
```

The following alternative uses a variable and prints that:

```
numberOfIntervals = Get number of intervals... 1
echo 'numberOfIntervals'
```

(`echo` will print to the Info Window, but erases its contents first, something that can also be done with `clearinfo`.)

This approach also allows us to get the label of the last interval in a TextGrid, by using the variable as an argument to another Command:

```
ni = Get number of intervals... 1
lastLabel$ = Get label of interval... 1 'ni'
```

In fact, we can easily write a loop that will print every interval's label to the Info Window:

```
ni = Get number of intervals... 1
for currentInterval to ni
 Get label of interval... 1 'currentInterval'
endfor
```

One thing "missing" from the Praat scripting language are **arrays**. However, this limitation can be circumvented by using variables within other variables' names. This way, for example, we can read all the labels in a TextGrid into one "array" of variables, which can be accessed later in the script:

```
ni = Get number of intervals... 1
for ci to ni
 label'ci'$ = Get label of interval... 1 'ci'
endfor
for ci to ni
 currentLabel$ = label'ci'$
 printline 'currentLabel$'
endfor
```

This example illustrates one caveat of using arrays in this way: apostrophized variable names in Command arguments cannot be nested; the following will *not* work as intended:

```
printline 'label'ci'$'
```

Multidimensional array are also possible:

```
nt = Get number of tiers
for ct to nt
 ni = Get number of intervals... 'ct'
 for ci to ni
  tier'ct'label'ci'$ = Get label of interval... 'ct' 'ci'
 endfor
endfor
```

This will place the label of the fifth interval in the second tier of a TextGrid in the variable named `tier2label5$`. It should be kept in mind, however, that since these are not real arrays, there is no way to return their size, sort their contents, or otherwise address them as a whole. This can only be done through loops.

Just like all Praat Commands that require arguments, the same can be done for scripts, through **Forms**. A script form presents a window to the user, which can contain values and/or options that can be modified, as well as helpful text. Forms let the user define variables which are then available throughout the script. An example:

```
form Please enter a short text
 comment This text will be displayed in the Info Window.
 text MyTextString This is the initial text.
endform
echo 'myTextString$'
```

Various types of input are available for different variable types.

It is also possible to add Scripts to the Fixed or Dynamic Menu as **Buttons** by using the appropriate Command from the Script Editor. This allows instant access to often-used Scripts without having to load and run them through the Script Editor.

Under Windows, there is an alternative to running Praat Scripts under the graphical Praat interface. The console application **Praatcon** is intended for the quick and efficient use of Praat Scripts from the command line. Here, arguments to the Scripts are given via the command line, and all output to the Info Window is printed to the standard output. Praatcon is in many ways much faster than its graphical counterpart, so for analysis of large corpora or CPU intensive procedures, it pays off to design a Script to be run in the console by Praatcon.

Scripts can also be used to write to and read from text files that are not Praat Object files. Similarly, Scripts can send **System Commands** to execute simple commands or run Praat external programs. This is done with the function `system`. By combining this feature with the Praatcon application, it is actually possible to run a Script with Praatcon from within another Script run under the graphical Praat, which can be more time-efficient in some cases:

```
system praatcon myActualScript.praat
```

Other tricks with Praat Scripts include the use of **Procedures** and **including other Scripts** as "modules", as well as combining these:

myScript.praat has the following contents:

```
clearinfo
include myModule.praat
form This will multiply two integers
 comment Please enter a whole number (default is 0):
 integer MyInteger 0
 comment Please enter another whole number:
 integer MyOtherInteger 0
endform
call multiply myInteger myOtherInteger
print 'myInteger' times 'myOtherInteger' is: 'product'.
```

myModule.praat has the following contents:

```
printline myModule.praat has been included! =)
procedure multiply firstNumber secondNumber
 product = firstNumber * secondNumber
endproc
```

The **Picture Window**, neglected up to now, is a powerful output tool for Scripts and their results. All output was up to now directed to the info window, which can be rather inappropriate for tables, etc. By drawing to the Picture Window, we can let Scripts create plots, tables, or other figures, using various fonts (including IPA symbols), sizes, lines, and colors. These can in turn be exported in PostScript or to meta files (EMF) that can be imported into text processing applications for publication. This export capability makes the Picture Windows useful even when using the command line based Praatcon.

For detailed reference on scripting and all other Praat topics, see the Praat **Manual**, which is presented as fully searchable hypertext documentation, as well as online at

```
http://www.fon.hum.uva.nl/praat/manual/Intro.html
```

More literature for discussion:

```
http://www.personal.rdg.ac.uk/~llsroach/phon2/frp.pdf
```

Roach, Peter (1982) "On the distinction between 'stress-timed' and 'syllable-timed' languages." In: David Crystal (ed.), Linguistic Controversies. London: Edward Arnold, 73-79.

Exercise:

Compute the standard deviation of syllable duration for the syllable-timed languages in Praat (normal speech rate only for now). Use either the formula

$$StDev = \sqrt{\frac{n\sum x^2 - \left(\sum x\right)^2}{n(n-1)}}$$

or Praat's built in Commands for statistical analysis.

Next, compute the standard deviation of Inter-Stress Interval duration for the stress-timed languages. This requires prior annotation of syllable stress.

What can we observe regarding isochrony?

End Discussion.

Session 2

Literature:

ABSTRACT

Rhythm, or the grouping of elements into larger units, is a property of all languages. The particular rhythm of a language is the result of the interaction of a number of components, such as the relative length, pitch, and segmental quality of accented and unaccented syllables, and phonological components, such as syllable structure and the function of accent. A system of rating whereby these components are broken down into features which can be assigned a plus or minus value allows us to compare the rhythm of languages or language varieties. Languages which have "strong stress" or which have been labeled "stress-timed" are seen to share certain features. Rhythm is a total effect involving phonetic and phonological as well as segmental and prosodic phenomena.

INTRODUCTION

Is it possible to develop a phonetic concept of rhythm that can be applied to all languages, in the same way that we use the system of cardinal vowels or the IPA chart of consonants? The distinction between stress-timed and syllable-timed languages [Pike 1945, Abercrombie 1967] is just such an effort at a general phonetic definition. In this theory, stress-timed languages show a tendency for stresses to recur at regular time intervals, and in syllable-timed languages, syllables are said to recur at regular intervals; all languages are believed to have one or the other rhythmic basis. Although many linguists have adopted the distinction, some have criticized the theory for being too simplistic (after all, it only divides all the languages in the world in half) and for grouping together languages which are felt to have noticeably different rhythms, such as English and Arabic […] or Spanish and French. In addition, many native speakers of "syllable-timed" languages have objected to the designation, as if it somehow meant that their language had no rhythm. Indeed, Crystal and Quirk [1970] refer to the lack of regular stress-timed pulses as "arhythmic." Linguists have had difficulty applying the concept to languages. Attempts to do so by instrumental analysis have been futile. Numerous experiments have shown that a language can not be assigned to one or the other category on the basis of instrumental measurements of interstress intervals or syllabic durations [Roach 1982, …].

Should we then give up the only phonetic theory of rhythm that we have, or perhaps turn to a purely phonological approach? Phonologists in the tradition of Trubetzkoy have treated rhythm in terms of the function and location of accent in the word. Metrical phonologists […] have assumed that all languages have an underlying strong-weak distinction and show a tendency towards alternation which can be shown in a grid or tree structure of the word. Although this approach brings out the importance of *grouping* of elements into larger units, which is considered essential in all psychological definitions of rhythm, it tends to make all languages look alike, at least on paper, and makes no attempt to specify further how these patterns are realized in spoken language in continuous speech. […]

It seems that an adequate description of rhythm in a language or across languages requires both phonetic and phonological information […]. We can define rhythm as the grouping of elements into larger units; the units need to have some similarity and be marked off from each other in some way in order to be perceived as groups […]. In language, most would agree that the elements that are grouped are syllables, and that, in some languages at least, stresses (or accents) serve to set off groups. Neither "syllable" nor "stress" have general phonetic definitions, which from the start makes a purely phonetic definition of language rhythm impossible. All instrumental studies as well as all phonological studies have had to decide in advance where the stresses (if any) fall and what a syllable is in the language under investigation in order to proceed. Although rules for syllable division and inventories of syllable types have been worked out for many languages on the basis of phonological criteria, stress is more problematic, and definitions of it have varied widely. In this paper, I shall use the term "accent" as it has been defined by Trubetzkoy [1935] as the phonological feature which when realized promotes the perception of one particular syllable (or mora) in relation to others. Accent can then serve as a basis of rhythmic grouping. The term "stress" will be reserved for the phonetic realization of certain kinds of linguistic accent. Rhythm is a total effect (also probably a grouped series of motor commands in production) that involves the interaction of a number of components, of which the following appear to be the most important for the purposes of comparing languages. It is most evident in continuous speech through the repetition of rhythmic groups at a natural speed for the speaker. Obviously, some speakers and some styles exhibit better rhythm than others and seem to be more representative of a particular speech community. […] In each style, a plus, zero, or minus is assigned to a language depending on the extent to which it exhibits the feature in question.

COMPONENTS OF LANGUAGE RHYTHM

1. Length

Duration

+ Accented syllables, and especially accented vowels, are regularly longer than unaccented syllables (by 1.5 or more). {e.g. English, Serbo-Croatian}

0 Accented syllables are slightly longer than unaccented syllables. {e.g. Spanish, Greek}

− Accent does not affect the length of syllables, or the language has no accent. {e.g. Japanese, Yoruba}

Syllable Structure

+ The language has a variety of syllable types (both heavy and light syllables with many different possible syllable structures), and heavy syllables tend to be accented, whereas light syllables tend to be unaccented. {English, Arabic}

− There is a very limited number of syllable types (predominantly CV or CVC), and accent and syllable weight are independent. There may be active proc-

esses such as final cluster simplification, epenthesis, or liaison to break up or prevent the formation of unusually heavy syllables. {Spanish, French}

Quantity

+ Quantity distinctions, if present in the language, are only permitted in accented syllables; in unaccented syllables they are neutralized (only short). {some Arabic dialects}

0 All quantity distinctions occur in accented syllables, but only a small subset can occur in unaccented syllables. {Estonian}

− Quantity distinctions are permitted in both accented and unaccented syllables. Restrictions on quantity are not conditioned by accent. {Hungarian, Finnish}

[…]

3. Quality

Vowels

+ The maximal vowel system exists in accented syllables; vowels in unaccented syllables tend to be reduced or centralized (especially open vowels). {English, Swedish}

0 The unaccented vowel system is smaller than that of accented vowels, but unaccented vowels are not necessarily centralized. There may be processes of devoicing or raising which occur only to unaccented vowels. {Russian, Portuguese}

− There is the same vowel system and similar articulation in all syllables. If elision or devoicing processes exist, they affect accented and unaccented vowels equally and are determined by phonetic environment rather than accent. {Spanish, Japanese}

[…]

CONCLUSION

By applying these categories to various languages, one should be able to come up with a comparative rhythm "score." The more pluses a language has, the more likely we are to say that the language has "strong stress" ("dynamic" or "expiratory" accent) and is "stress-timed." The differences between accented and unaccented syllables are maximized, and accent would clearly be the principle for grouping. We would expect that naïve native speakers – as well as trained non-native speakers – could fairly consistently identify accented syllables in continuous speech. In a language with many minuses in these categories, we would have to look elsewhere for the principle of grouping: what is it that permeates the entire linguistic system, binds units together and helps listeners segment the flow of speech into meaningful chunks? It could be patterns of tone, of syllable or vowel length, or even the repetition of certain segmental or grammatical features. Although the language may have some kind of accent, naïve native speakers would have difficulty identifying the place of accent consistently in continuous speech, and linguists would have difficulty finding its acoustic correlates, even in words said in isolation. This does not necessarily mean that this kind of language is somehow defective or arhythmic because it is lacking a feature that certain prestige languages have. All languages have rhythm, but more independent research needs to be done to discover exactly what the rhythmic principles are in languages which do not show a tendency towards "stress-timing."

[…]

Dauer, Rebecca M. (1987) “Phonetic and phonological components of language rhythm.” Proceedings of the 11th International Congress of Phonetic Sciences. Tallinn, Vol. 5, 447-450.

Discussion.

More literature:

```
http://www.ehess.fr/centres/lscp/persons/ramus/Cognition99.pdf
```

Ramus, Franck, Marina Nespor & Jacques Mehler (1999) “Correlates of linguistic rhythm in the speech signal.” Cognition 73, 265-292.

Discussion.

Exercise: Compute *%V* and *ΔC* in Praat for all TextGrids in our data. If possible, use the Picture Window to plot the results.

Question: Does this really capture the essence of speech rhythm? If not, what are we measuring?

Discussion.

Question: What would speech sound like if the isochrony hypothesis were true?

For this, we need to annotate syllables by prominence in stress-timed data.

Preliminary exercise: Mark stressed syllables in stress-timed data with a trailing “*”.

Session 3

Exercise: Create a Praat script that does the following:

- computes the average unit duration for each recording, ignoring pauses
- creates a TextGrid which contains these isochronous intervals
- modifies the duration of each unit to match that average (use PSOLA)
- writes the “isochronized” Sound to a WAV file
- use script modules, so that several groups can work in parallel

Final discussion.