RHYTHM AS A RESOURCE TO GENERATE PHONETIC AND PHONOLOGICAL COHERENCE IN LISTS

Alexandra Imrie

Abstract

This paper uses natural conversational data to examine the use of rhythm in conversation as a resource to generate phonetic and phonological coherence in English lists. This paper does not continue the search for all-pervasive rhythm in language, but considers one particular activity in conversation, that of list making. Data are presented that show that rhythm can indeed be produced and orientated to within the activity of list-making. Rhythm production and orientation is seen to be important to interaction.

1. Introduction

In conversation, speakers must manage turn-taking; who can speak, when they can speak, and how long the turns are. One recurrent problem for speakers is the production of turns or turn constructional units (TCUs) which belong to a larger structure or sequence (for example, extended turns and increments to previous turns). Such turns or TCUs must first be secured, and then they must be interpreted as belonging together by other participants. That is, coherence is a general problem in conversation.

Coherence is used in the following to mean the matching of prosodic, lexical and/or syntactic features to produce TCUs which are orientated to as being part of a larger structure. Phonetic coherence, then, means that the various prosodic (phonetic) features across TCUs display similarities and are orientated to as being part of a larger sequence. Phonological coherence relates to phonological categories; the structures or representations of which these (and other) features are exponents.¹

This paper looks at the conversational activity of list production. It examines the phonetic and phonological properties of lists used to generate coherence, and also their effect on the demarcation of the list from the surrounding talk, and on turn-holding and turn-taking. The focus is on how rhythm can be used to achieve these things, as this is an area which has been underrepresented in studies on talk-in-interaction.

Previous work on lists is briefly introduced before turning to work done on rhythm in speech, music and cognition. Using the insights this provides, the analysis addresses the use and effects of rhythm in naturally occurring lists. It is hoped that the results of the analysis provide evidence that rhythm is indeed a device for coherence in lists, even if there remains a lot to be done to establish the details of how it works.

2. Literature Review

2.1. Previous work on lists

Studies on lists have revealed that they are part of a larger three-part structure. Sánchez-Ayala (2003: 331) and Selting (2007: 488) both refer to this structure, in which the list is the middle part. Selting (2007: 491), for example, proposes a “context-free structure” (492) for her observations from German data, which includes a pre-detailing component, the list, and the...
post-detailing component. Sánchez-Ayala (2003: 331) calls these three parts the onset, the body and the coda. The problems for speakers in list construction are linked to this three-part structure. A speaker must project a list (Jefferson (1990: 77)) to secure a longer turn to achieve this. Within the list, the speaker must generate coherence between the list items (Selting (2007: 487)) and then signal the end of the list (Lerner (1994: 23)) to mark the end of the list activity. Because lists themselves usually contain three items (Jefferson (1990: 63)), ending is often programmatically signalled by the production of the third item. However, this does cause a further problem for the speaker when the list will be longer: keeping the turn to continue the list.

Within the list proper, various recurring features have been found that speakers regularly produce and orient to. Lists are lexically and semantically coherent. Selting (2007: 491) mentions that list items are collocationally linked, and Sánchez-Ayala (2003: 325) notes the recurrence of conjunctions in lists. Jefferson (1990: 67) and others after her mention that the third list item is often a generalised completer, and remarks on the use of alliteration within the list (68). Syntactically, lists also show parallelisms (Sánchez-Ayala (2003: 325)).

In terms of prosody, there have been some attempts to classify the behaviour of lists. Sánchez-Ayala’s 2003 paper mentions ultimate stress lengthening and prosodic parallelism (325) as features of lists, and suggests that the beginning of each new unit is marked by a pause, a pitch reset and anacrusis (329).

However, it is the intonation contour of list items as a device for producing coherence which has attracted the most attention in the literature. Selting (2007: 523) argues that the matching of intonation contours within the list “enhances the interpretation of the sequence of items as a cohesive structure and practice”. One of her examples of a collaborative list deals deals specifically with the similarity of intonation contours to show how listeners orient to prosody in collaborative lists (515).

Rhythm has not gone completely unmentioned in the literature on lists. Selting (2007: 487), for example, makes reference to timing and rhythm in her paper, but does not go into detail. A more detailed perspective is provided by Auer, Couper-Kuhlen, and Mueller (1999). As part of a wider ranging study on rhythm in conversation, they present some examples of English lists (41) and conclude that isochronous beats are favoured in conversational lists in English (58). Their work on Italian lists is more in-depth, showing that Italian lists are often tripartite (a three part list corresponding to three beats (192), whereby the last item may span more than one beat (193)). They conclude for Italian that rhythm is a property of “parallel structures” (195) and that rhythmic structures help to organise “formulaic pairs” such as lists (200). Such findings can be compared with the findings from the English lists in the analysis to follow.

The conclusion from these studies on lists is that there is a great deal of coherence at multiple levels within a list. What stands out, however, is that rhythm has not received as much attention as other aspects. The following sections introduce findings from studies on the neural and cognitive aspects of rhythm and work done on rhythm as a feature of language.

2.2. Previous work on rhythm as a capability

2.2.1. Defining rhythm

A definition of rhythm has not been agreed on, but Gibbon and Gut (2001) propose a general definition:

the recurrence of a perceivable temporal patterning of strongly marked (focal) values and weakly marked (non-focal) values of some parameter as constituents of a tendentially con-
stant temporal domain (environment) (Gibbon and Gut (2001: 91)).

In terms of language, Auer et al. (1999: 25) suggest that rhythm has two parameters: the recurrent patterning of events (beats), (referred to as prosodic prominences, based on stress and/or accentuation (37)), and the “regular succession of beats in time” (25). In their summary, they write that rhythm is:

a perceptual gestalt which imposes isochronous patterns on speech within and across turns at talk (Auer et al. (1999: 202)).

For this study, which refers both to music and to language, Auer et al. (1999: 25)’s definition of ‘regular beats in time’ will be used. The regularity of the distance between the beats is isochrony.

2.2.2. Rhythm and music

Before beginning the search for rhythm in language, it is worth looking for evidence that rhythm is a feature that humans can orient to. Studies have shown that this is indeed the case: beat perception and synchronisation (BPS) is a complex psychological process, which seems to be unique to humans (Patel, Iversen, Chen, and Repp (2005)). Humans quickly and subconsciously orient to a beat, and synchronise at different levels of periodicity (226). Beat perception remains even when faced with stress at non-beat locations and variability in the distance between beats (227). Listeners orient to an established beat by anticipating the next beat to come (Patel (2006: 100)). This means that we can expect speakers to orient to a rhythm, if there is one there. The next step is to examine whether there is evidence that language could have a beat.

To do this, we can look at the links between music and language; the two organised sound systems in human cultures (Patel (2005: 59)). Perceiving and proving rhythm in music is not a problem (Patel and Peretz (1997: 204)). We could therefore reasonably expect that language is also rhythmic if it can be shown that language and music are related and share properties. Various studies have done just this, showing that prosody and music share neural resources (Patel, Peretz, Tramo, and Labreque (1998: 123)), and that there is a link between syntax and music (Patel (2003a: 679)). Another interesting link between language and music has been described by Patel and Daniele (2003). Using a measure of durational contrast, the nPVI (normalised pairwise variability index (cf Grabe and Low (2002)), they showed that the timing of units (feet or syllables) in a language is reflected in the music of the composers of that language.

These studies conclude that music and language share many resources. More importantly for this paper, rhythm is a common feature of the two (Patel (2005), Patel (2003b)), involving similar cognitive abilities. However, studies which note the similarities in rhythm are careful to distinguish between grouping and metre.

Grouping is the property of dividing a rhythm into phrases, in which the units are not equally separated (Patel (2003b: 141)). The commonalities between music and speech rhythm are accepted to be mostly, if not exclusively in this domain (140). Patel (2006: 99) writes that phrase endings are similarly marked in both music and speech, with boundaries being signalled with lengthening, and remarks that the perception of grouping is neurally linked in the two domains.

Metre, however, is the beat or rhythm, a “periodic temporal-accentual scheme” (Patel et al. (1998: 125)), and it is the conclusive proof of this in language which has eluded researchers. Indeed, Patel (2006: 100) writes that, contrary to music, the stressed elements of speech do not create a rhythmic beat. This is reflected in the scant success of studies searching for isochrony (see following section).
It seems strange, however, that two systems which are linked at other levels should differ only in this. Stranger still, because language, like music, has a metrical hierarchy (Patel (2006: 100)), from which a beat can often be perceived, just not proved. The neural and cognitive links between music and language lead to the prediction that speech is rhythmic. As the next section shows, this has not been easy to demonstrate.

2.3. Previous work on rhythm in speech

Most of the work on rhythm in speech has focused on the search for isochrony as a property which is always present in language. This has largely been interpreted as the search for proof of a difference between stress-timed and syllable-timed languages, with studies in English looking for isochrony in time distance between stressed syllables (i.e. foot duration) (Auer et al. (1999: 12)). The evidence for isochrony from such experiments has been scarce (Ramus, Nespor, and Mehler (2006), Grabe and Low (2002), Gibbon and Gut (2001)), and has lead to claims that it is a perceptual phenomenon (Grabe and Low (2002: 516)), or that silent beats may also play a role (Cummins and Port (1989: 147)).

A move away from foot measurement brought more success, with Grabe and Low (2002) finding weak categoricity between syllable- and stress-timed languages based on the nPVI of vowel and inter-vowel durations. Ramus et al. (2006) also had some success measuring vocalic intervals and the standard deviation of consonantal intervals. These studies, however, are concerned with rhythm as a universal property of language – something that is continually present. There has been relatively little work done on rhythm as a tool in interaction.

One study which adopted this stance is Auer et al. (1999). Based on the assumption that speech is not “exhaustively rhythmic” (36), but that some parts are more rhythmic than others, they showed that a rhythmic beat (even including silent beats (48)) can be used by co-participants as a tool for turn-taking, preference organisation and closings in English and German (33). In particular, their work on Italian and (to a lesser degree) English lists showed that rhythm is often a property of list constructions, and may go hand-in-hand with assonance, alliteration (194), isometry (equal numbers of syllables) (40) and parallelism/repetition (27).

It is this route that the following analysis will take. In searching for rhythm in language, most studies have ignored the possibility researched by Auer et al. (1999), that the creation of and orientation to a regular pulse is a conversational resource. Based on the results of cognitive experiments proving that language and music are linked, and that humans are unique in possessing the ability to spontaneously orient to a beat, and encouraged by Auer et al. (1999)’s results, this analysis will look for rhythm in English speech. Rhythm will be considered as a resource for the interactional activities of list construction and the turn-taking in their vicinity. The study does not attempt to describe the wider integration of rhythm into the language’s phonetic/phonological system, nor how rhythm is used in other interactional activities.

3. Methodology

3.1. Terms used and assumptions made

In their study on timing, Auer et al. (1999: 63) refer to anacrusis - syllables occurring before the beat. Because of the more widespread use of the term to mean utterance-initial unstressed syllables (Cruttenden (1996: 21)), the more musical term “upbeat” will be used. The distinction between the two phenomena is an important one; although the syllables in an upbeat belong
syntactically and intonationally to the following talk, rhythmically, they belong to the preceding phrase (see figure 1 which uses a structure taken from Ogden, Hawkins, House, Huckvale, Local, Carter, Dankovicova, and Heid (2000) to show how the final word, a, belongs to the same rhythmic group as robin). This non-mapping of prosodic and syntactic structure is mentioned by Ladd (1996: 237), and is shown in the analysis to be relevant to turn taking, allowing speakers to project more talk before the beat falls.

Figure 1: Metrical structure frl-sum02-bluebird-01

The second point to note is an assumption that will be made throughout this analysis. Auer et al. (1999: 24) claim that rhythm can first be established after three beats. However, when they talk about silent beats, they state that a silent beat must come after two clear beats (48), confirming the intuitive assumption that, after two beats, a third can be meaningfully expected, equally placed from the second as was the second from the first. After this third beat, the rhythm is fully established.

3.2. Data and methods

The data under analysis are taken from a corpus of around two hundred and thirty lists collected by University of York students in summer 2002 and 2004. Of these lists, 117 (51%) were identified as being perceptually isochronous on listening by the author (a native speaker of English, and a trained musician). The perception of isochrony was based on:

1. The sense that the stressed/accented syllables in the list could form a beat, and
2. The ease of tapping along to this beat.

Twenty-six lists were chosen as a subset for analysis for the purposes of this study. These had the strongest perceived isochrony – the accented or stressed syllables were particularly easy to identify and tap along with in perceptually equidistant beats.

The methodology followed for the analysis mirrors that of Auer et al. (1999). The list was first transcribed in a word processing program. Prosodic prominences were auditorily identified based on stressed and accented syllables, and the resulting beat was established by tapping along
with the list on repeated hearings. The beat was marked on the transcription by underlining the syllable which corresponded to the beat. Any pauses which contained possible silent beats were also noted. The aim of working perceptually in the initial stages was to ensure that the later measurements reflected the auditory perception of the beat, since this is the only way participants have to identify and orient to rhythm.

Once the perceived beat had been noted, the list transcriptions were transferred to the TextGrid in Praat. Three tiers were used as standard: for the transcription, the beat position and the interval between the beats. More transcription tiers were added as necessary.

The intervals between the beats were then measured (in seconds, using Praat). This provided numerical values which could be compared, in terms of percentage deviation (see Auer et al. (1999: 53)). The smaller the difference between consecutive intervals, the higher was the measured isochrony. The analysis then concentrated on how this rhythm could be managing coherence within the list, turn-taking and demarcation of the list from the rest of the turn. These judgements were based on the participants’ orientation to the rhythm in the subsequent turns, and the conversation was treated as an emergent activity, i.e. participants do not have the view of the analyst, who can see what happens next.

3.3. Placing the beat and measuring isochrony

The decision about where to place the beat in the beat tier was not easy. As Cummins and Port (1989: 147) mention, beat placement is not straightforward. Some studies have found the P-Centre (ibid) and nPVI to be more reliable for measurements than feet. However, these measuring techniques were not used, based on Auer et al. (1999: 53)’s remarks that rhythm is stronger when beats co-occur with stress, and that vowel-onset to vowel-onset is more reliable than the P-Centre. The beat was therefore placed at the vowel onset (the third glottal pulse), and annotated with a “V” on the points tier in Praat.

What exactly counts as an isochronous interval is also difficult to determine. Auer et al. (1999: 14) quote Couper-Kuhlen (1993) as giving a 20% permitted deviation. She mentions that the perception of rhythm may be affected, and the permitted deviation increased, by factors such as speaker switches, and syntax and text structure (specifically lists). Auer et al. (1999: 53) conclude that perceptual isochrony does not have deviations of more than “roughly 35%” (my emphasis) between intervals, and that the border between perceptual isochrony and anisochrony is “somewhere between 31% and 47%” (54) (my emphasis). Not only are these three statements different, they are also not rigorously defined. Fortunately, most of the intervals were well below 35%, and often below 20%.

4. Analysis: features and uses of isochrony

4.1. Evidence of isochrony

The first finding of the analysis is that perceived isochrony in lists can be measured in the majority of cases. In twenty-five of the lists, the measured intervals between the beats reflected the perception of isochrony. The one example for which isochrony could not be shown is addressed in the discussion section below.

One example is meo-sum02-grieving-01, shown in example 1 and figure 2. As can be seen from figure 2, the perceived beat was marked, and the distances measured. The deviations between the intervals are: -4.3%, -6.8% and +4.8% respectively, which is well within the allowed
margins mentioned above. The special conventions for the transcriptions are mostly taken from Auer et al. (1999), and are as follows:

- * strong phonetic prominence
- ' weak phonetic prominence
- / ... / regular isochronous rhythm. The material between the slashes corresponds to one beat
- / ... an on-beat beginning, after which isochrony breaks down
- ^= silent beat
- {f} louder talk

1. A .hh uhm psychologists and psychiatrists describe you know you go through
2. / *this stage /
3. / *this stage and /
4. / *this stage /
5. / ^= .hh
6. B yeah /
7. A / *no two people experience it the same way

<table>
<thead>
<tr>
<th>stage</th>
<th>through</th>
<th>this</th>
<th>stage</th>
<th>this</th>
<th>stage and this</th>
<th>stage</th>
<th>.hh</th>
<th>no</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td>v</td>
<td>yeah</td>
</tr>
<tr>
<td>0.46</td>
<td>0.46</td>
<td>0.41</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Isochrony in meo-sum02-grieving-01

4.2. Evidence of active isochrony generation

The subsequent finding from the analysis is that rhythm is not just a coincidental product of speech, but seems to be actively created. The use of pauses between list items which are of different lengths is an indication the speakers are consciously generating a rhythm, which creates coherence within the list.

2. A .hh but when you think of some of the world leaders at the moment .hh
3. {f} / *Blair (0.3) /
4. {f} / *Bush (0.3) /
5. {f} / *Berlusconi /
6. / ^= (0.6) eh /
7. / ^= (0.7) /
8. B / *(inc)/
9. C / *Chi’rac /
Figure 3: Isochrony in sjb-sum02-chirac-03

Figure 3 shows the isochrony in the list beginning on line 2 of example 2. The silence after the first two list items (visible in the spectogram) makes them isochronous with each other, and with the third item, which is longer. By using pauses between the items, the speaker seems to be actively creating isochrony by “filling” the space up to the beginning of the next stressed syllable. An interesting point to note about this example is the alliteration of the list items.

Another example of using pauses to generate isochrony is shown in figure 4 and example 3.

(3) 1 A see how many of them are named=
    2 B =mm
    3 A / *blu:ebird /
    4 / *do:ve (0.3) /
    5 / *re:dwing /
    6 / *ro:bin /
    7 t- a whole selection named after birds
    8 B red-

Figure 4: Isochrony in frl-sum02-bluebird-01

Here, the first item (line 3) is long, with a noticeably long stressed vowel (marked with a colon in the transcription). The vowel in the second item is also long, but not long enough to make the production of the third item isochronous. By pausing before producing the third item,
the speaker keeps the isochrony in the list, which continues for the fourth item, uniting them into a rhythmic group.

4.3. Positioning of talk relative to the beat

Once a rhythm has been established, co-participants can position their talk and inbreaths in three places with respect to the beat: before the beat (as an upbeat), on the beat, or after the beat. The following data show that talk and inbreaths positioned before the beat are a device for turn-holding or claiming a turn, on-beat events show coherence with the preceding speech, and talk or inhalation which falls after the beat breaks the rhythm of the preceding talk.

One example in the data showed all three types of positioning. In figure 5, the list frl-sum02-bluebird-01 is reproduced.

![Figure 5: Isochrony in frl-sum02-bluebird-01](image)

This list, which is shown in example 4 consists of an introduction (line 1), a four-item list (lines 3-6), and a coda (line 7) (Sánchez-Ayala (2003)'s term).

(4) 1 A see how many of them are named=
   2 B =mm
   3 A /*blu:ebird /
   4 /*do:ve (0.3) /
   5 /*re:dwing /
   6 /*ro:bin /
   7 t- a whole selection named after birds
   8 B red-

There is quite strong isochrony between the stressed vowels in the list. The percentage deviations between the intervals are: +4.6%, -4.4% and -1.1% respectively. The resulting rhythm generates coherence, and is used by the coparticipants to position their speech.

The speaker’s first problem to overcome is introducing the list. He does this with a general introduction about the names of the cars. This is accepted with a minimal token of “mm” by
the second speaker. The next problem is to demarcate the list from the rest of the talk, by producing a first item which is hearable as the first item of a list. Because of the time it takes to establish a rhythm, it is most likely that this is achieved intonationally. As shown in figure 5, the pitch contour for the first item “bluebird” (and subsequent items) can be described as an “upward staircase with slightly falling final pitch” (Selting (2007: 507)), one of the frequently used intonation contours in Selting’s (ibid.) German lists (and a recurring contour in these English data). Using such typical list intonation can signal a list even in the first item, Selting argues (485).

As mentioned above, the speaker pauses between items to create a rhythm in the list. The resulting on-beat production of the third and fourth items reinforce the intonational coherence in the list.

At this point, the final problem of list construction is encountered: getting out of the list and into the coda or the following turn. To do this, the list producer must employ specific devices to hold the turn if he wants to carry on speaking. Again, rhythm seems to play a part here. After the production of the final list item, the move into the coda is on an upbeat, as seen from this simplified metrical structure. The stressed syllables are at the beginning of each bar:

| bluebird | dove | redwing | robin a |

An established beat therefore allows participants to carry on talking by producing speech which will not be pragmatically, syntactically or intonationally complete by the time the next beat is due. The second speaker orients to this by not coming in.

The talk after this beat no longer gives the perception of isochrony. As can be seen from the spectogram in figure 6, the stressed syllables are quieter than those in the list, which may be one reason for this perception. The intervals between the first stressed syllables after the list (marked with “x” in the figure) are actually isochronous; however, the isochronous intervals are very different to those of the list. The next two intervals are not isochronous. The change from the isochrony of the list serves doubly to strengthen the coherence of the list and to separate the following talk from that list. The second speaker comes in with his turn after the coda is pragmatically, syntactically and prosodically complete. As well as “gluing” the list items together then, the placement of talk relative to a beat can be used to manage turn-taking.

The list meo-sum02-grieving-01 also displays onbeat and upbeat talk, as shown in figure 7 and example 5.
Psychologists and psychiatrists describe you go through this stage / this stage and / this stage / ^ .hh

Yeah /

A / *no two people experience it the same way

The third item in this list is rhythmically secured by the upbeat positioning of “and”, which belongs rhythmically to the second item, but pragmatically to the third:

| this stage | this stage and | this stage | .hh | no |

This is interesting about this example is that the inbreath at the end of the list is onbeat (with a silent beat). In other examples, a lack of upbeat talk often means that the turn is surrendered, especially at the end of an otherwise complete list. The second speaker does indeed come in, with a minimal acceptance token, but only once the inbreath has been produced, and not timed with the beat. After this minimal token, the first speaker carries on into the coda, which starts on beat. In this example, then, it could be argued that as well as creating coherence by setting the list apart from the surrounding talk, the isochrony was used as a strategy of surrendering the turn by not producing upbeat speech or inhalation. When the second speaker responded only minimally, the first speaker carried on with the turn, starting on beat to link his speech back to the previous talk.

Perhaps the strongest evidence of upbeat talk as a continuation resource is found in cjb-sum02-crown-01 (figure 8 and example 6).

I’m not interested in either of the personalities of the Queen or Prince / Charles .hh or / {f} / *any of them /

The problem for the speaker of this list is that his use of “either” projects two items (Jefferson (1990: 75)). Using the rhythm established by the second beat, he positions an inbreath and says “or” before the third beat falls:

Queen or Prince | Charles .hh or | any of them .hh |
In this way, he uses rhythm to secure the third list item despite his projection of two items. In doing so, he creates coherence within the list by producing an onbeat third item. The problem of continuing speaking for the coda is similarly resolved, except that he inhales as an upbeat to a silent beat.

4.4. Continuation of rhythmic talk outside of the list

In a few examples, rhythmic talk continued outside of the list, showing how rhythm can also be used to link talk within a larger activity, such as question-answer sequences, or two lists which are part of one larger turn.

In the list mfs-sum04-knife-03 (figure 9 and example 7), upbeat talk is used to secure the turn for the second and third items.

(7) 1 A is it still the same deal that you can choose your basic knife
    2 and then you can have little things added like you might want a b-.hh
    3 / *bottle opener or /
    4 / *scissors or /
    5 / *whatever=
    6 B =I be /
    7 / *lieve [so yeah]
    8 C [I think so]
turn before the anticipated beat which was set up in the list. His stressed syllable in “believe” is on-beat with the isochrony set up in the list. In this way, his turn generates coherence by relating the answer to the question not only pragmatically but also rhythmically.

A similar example is klm-sum04-class-1 (example 8 and figure 10). Again, the list is a part of a question. As in example 7, the speaker asking the question does not produce a coda. The difference here is that there is no upbeat answer from speaker B. Instead, there is a silent beat. After this, speaker B does produce an answer, whose stressed syllable is on beat with the rhythm which was set up in the list and held through the silent beat. Her laughter which follows also starts on beat (annotated with ‘-’). The delay in answering the question could be because the answer given was not one of the choices offered in the list. Despite this delay, speaker B’s answer is rhythmically matched to the question, linking the two turns together.

(8) 1 A so it’s going to cost us even more money
2 B so what- how would you rate the service then
3 / *first /
4 / *second or /
5 / *third /
6 / *class /
7 / * (0.4) /
8 A / *fourth /
9 / *(laughs) /
10 B fourth

Figure 10: Isochrony in klm-sum04-class-1

The use of rhythm to unite the talk of two participants is also used in the list jlw-sum04-caterer-05 (figure 11 and example 9). This example also involves a question-list which does not have a coda.
In this example, speaker A is asking a question about the catering company that speaker B works for. The question involves a list, which is rhythmic. In line 5, both speakers orient to this rhythm by projecting their continuing or incoming talk as an upbeat. Speaker A projects the continuation of the list with the upbeat “your”, making the list syntactically incomplete by the time the next beat falls. Speaker B’s inbreath projects incoming talk to coincide with the end of the second list item. She actually takes her answer turn in line 8, which begins in overlap with the final list item.

The turn-taking in this example is particularly interesting. On the one hand, we know that lists usually contain three items (Jefferson (1990)). Based on this, speaker B’s inbreath in line 6 projects talk at a place where she does not yet have the turn. On the other hand, speaker A has already produced the right answer. Also, his list contained ‘big’ and ‘small’, which are binary terms and therefore could exhaust all of the possibilities for asking about a company. Based on this, it could indeed be relevant for speaker B to take her turn at this place. However, speaker A projects that he will produce a third item, using the upbeat “your”. Speaker B waits until speaker A is at a TRP in this third item before coming in with her turn. She produces it as a two-part list in which one of the items was the item in overlap. This coincides with Lerner (1994: 21)’s observation that lists can be used to acknowledge receipt of information. In a similar fashion to the examples above, her answer is also (fairly) isochronically constructed (figure 12).

It would seem that as well as acknowledging receipt of the token in overlap, she also constructs her list to be prosodically similar (in terms of rhythm) to speaker A’s. As well as showing how two participants can orient to a beat, this provides another example of how an established
beat can be used to create coherence within individual lists and in the larger activity of question-answer.

Figure 13 and example 10 show an instance of two successive lists. One of the problems here is producing the second list straight after the first one, and it is here that rhythm plays a role. With the beat well established by the third item of the first list, the upbeat “you” projects more talk. The continuing isochrony generates coherence within and between the lists, and also provides a continuing beat which the speaker uses to position to more upbeats: the beginning of the final list item, and an inbreath to introduce the coda. Thus, the upbeat placement of syntactically incomplete talk projects more to come to hold the turn, and links the two lists to each other.

(10) 1 A council have said
2 / *no /  
3 / *no /  
4 / *no you /  
5 / *can’t have /  
6 / *this you /  
7 / *can’t have /  
8 / *that .hh /  
9 and it’s left the parents

This example and the other data in this section show that rhythm can also be used and oriented to outside of a list. This would suggest that rhythm may be part of the larger action of “cohesion”, of which list-making is a sub-type. Or perhaps it is a “contextualization cue”.  

Rhythm as a Resource to Generate Phonetic and Phonological Coherence in Lists

This example and the other data in this section show that rhythm can also be used and oriented to outside of a list. This would suggest that rhythm may be part of the larger action of “cohesion”, of which list-making is a sub-type. Or perhaps it is a “contextualization cue”.

Figure 12: Isochrony in jlw-sum04-caterer-05

Figure 13: Isochrony in lda-sum02-nonono-03
as Auer et al. (1999: 202) suggest, and serves to delimit activities and places in sequences and turns.

4.5. Use of silent beats

The data show that a rhythm does not necessarily need stressed syllables to continue. Speakers can both create and orient to up to two silent beats, which continue the coherence of the talk even through periods of silence.

An example which shows the use of a silent beat is jlm-sum02-prof-01 (figure 14 and example 11).

(11) 1 A I see it in
     2 / *policemen I see it in /
     3 / *teachers I see it in /
     4 / ^ .hhh in /
     5 / *health professionals

![Figure 14: Isochrony in jlm-sum02-prof-01](image)

After the second item, the repeated production of “I see it” is before the beat and followed by a long inbreath, during which a silent beat falls. At this point, the list is not complete because the final item which has been projected by the upbeat “I see it” has not been produced. When the item does come, the stressed syllable is double the distance from the last stressed syllable, suggesting that speakers can produce isochronous talk even based on silent beats.
Proof of orientation to silent beats can be seen in the list klm-sum04-class-1 (reproduced as figure 15 and example 12). The turn by speaker B contains a three part list which is isochronous (more so after the second beat) and is a question. Because the question makes an answer relevant, speaker B stops talking. No answer comes initially, but when the answer does come, it is timed as double the length of the previous isochrony. In this way it is marked as belonging to the preceding talk, and orients to the previous rhythm. This gives some proof that a silent beat can not only be set up by a participant, but also oriented to by another.

(12) 1 A so it’s going to cost us even more money
      2 B so what- how would you rate the service then
      3 / *first /
      4 / *second or /
      5 / *third /
      6 / *class /
      7 / ^ (0.4) /
      8 A / *fourth /
      9 / *(laughs) /
      10 B fourth

Figure 15: Isochrony in klm-sum04-class-1

This type of orientation to silent beats is also shown in the list sjb-sum02-chirac-03 (reproduced as figure 16 and example 13). After completing an isochronous list, the first speaker does not produce a coda, and no other speaker takes the turn. As an upbeat to a second silent beat (the first being at the end of “Berlusconi”), the first speaker produces a minimal token to invite further talk. The second speaker comes in on beat, orienting to the isochrony set up in the list and preserved in the silent beats. During his talk, the third speaker makes an addition to the list, also orienting to the rhythm. In this way, the rhythm of both heard beats and silent beats sets the list apart as a rhythmic entity from the preceding speech, and joins the following talk to the list.

(13) 1 A .hh but when you think of some of the world leaders at the moment .hh
      2 / *Blair (0.3) /
      3 / *Bush (0.3) /
      4 / *Berlus’coni /
      5 /^ (0.6) eh /
      6 /^ (0.7) /
      7 B / *(inc)/
      8 C / *Chirac /
4.6. Different levels of rhythm

There is also evidence of different levels of rhythm, with two stressed beats in place of one. Because there is only one example of this (figure 17 and example 14), it is unclear what this achieves.

(14) 1 A after the municipal corporations act of eighteen thirty seven
2 mu
3 /*nicipal /
4 /*housing mu/
5 /*nicipal /
6 /*hospitals mu/
7 /*nicipal /
8 /*gas mu/
9 /*nicipal elec/
10 /*tricity mu/
11 /*nicipal /
12 /*tran /
13 /*sport .hh mu
14 /*nicipal mu/
15 /*seums mu/
16 /*nicipal /
17 /*art /
18 /*galleries .hh mu /
19 /*nicipal
20 .hh erm airop:- fields in the end where I learnt to fly during the war

This very long list involves a lot of repetition, and shows isochrony, up until lines 17-18 (“art galleries”). Both initial syllables are stressed, and perceived as beats. The perceived isochrony is measured as being 48% shorter than the preceding interval. However, it is conceivable that this is an example of a sub-hierarchy within the metric structure, similar to having two 1/8 beats in the place of a 1/4 beat, or periodicity at different levels (Patel et al. (2005: 226)), since the time distance between these two beats, when added together (0.54) is only 14% more than the previous interval.
4.7. Summary

These data have shown that isochrony can be established and oriented to by participants in a conversation, even when the beat is not overtly realised. It seems to aid the projection of more talk by placing talk or inbreaths as an upbeat, and creates coherence within the list and the larger activity. This coherence is not only based on the establishment of the beat in itself, but is also due to the fact that the rhythmic unit often encompasses speech from two or more syntactic phrases, thus joining them together.

5. Results and observations

One of the main conclusions that can be drawn from the analysis above is that perceived isochrony in lists can be measured instrumentally. Furthermore, participants can, and do, construct and orient to isochronous lists in naturally occurring data. Even the presence of silent beats can be oriented to. These are significant findings, since it has previously been claimed (Patel (2006: 100)) that there is no evidence for a beat in speech. At least on a preliminary level, this claim can be refuted; not only is there a beat, but it also serves a purpose.

The rhythmic pulse set up can be used to promote coherency in a list by creating a unit of rhythmic talk which stands out from other talk in the turn. Within the list, syntactic phrases can be joined together because of their placement relative to the beat. The analysis has also shown that participants use rhythm as a device for turn-taking management within lists: talk can be positioned so as to hold a turn or to surrender a turn.

An interesting feature of the lists which were perceptually isochronous is the length of the list items. The general trend was that lists with shorter items are more likely to be perceptually isochronous.

5.1. Comparisons with Auer et al. (1999)

The predictions drawn from Auer et al. (1999)’s work on Italian lists were generally upheld in these data. Rhythmic talk does indeed seem to be feature of parallel structures (195). They
remark that prosody is one of the “part-and-parcel features of lists, organizing their construction and being organized by it” (192). The data from this analysis support this.

Auer et al. (1999: 194)’s conclusions about rhythm include observations about features which co-occur with rhythm. They include assonances, alliterations, vowel lengthening and list structures (grammatical parallelisms). There is evidence in the examples that supports their observations. One list mentioned in the analysis above that showed alliteration was sjb-sum02-chirac-03 (Blair, Bush, Berlusconi). Other examples in the data also displayed alliteration, for example vfh-sum04-spelling-5, shown in example 15.

(15) 1 A information that you’re putting forward to .hh let’s say a
  2 / *client or a con /
  3 / *temporary or a /
  4 / *colleague .hh /
  5 and it’s very important that the impression you receive

Vowel lengthening was remarked as being a property of frl-sum02-bluebird-01 in example 3 above. Parallel structures with syntactic recycling were also present, for example in jlm-sum02-prof-01, reproduced here as example 16:

(16) 1 A I see it in
  2 / *policemen I see it in /
  3 / *teachers I see it in /
  4 / ^ .hhh in /
  5 / *health professionals

Auer et al. (1999: 192) also mention rhythmic tripartition as a property of some of their lists – that there are three beats to match the three items, making the constructional and rhythmic lengths equal. Again, the majority of the lists in the data fit in with this pattern. There were also examples (not mentioned in the analysis) where the last item is more weighted, and receives two beats instead of one. This corresponds with another of Auer et al. (1999: 193)’s observations.

In general then, the English lists in these data show very similar properties with the Italian lists discussed by Auer et al. (1999).

5.2. Lists in questions

A rather new aspect that these data have brought to light is the use of lists in question sequences. Three list examples discussed in this paper were questions. All three had structural similarities: a lack of coda, and the continuation of rhythmic talk in the answer turn. In two of the examples (mfs-sum04-knife-03 in example 7, and jlw-sum04-caterer-05 in example 9), the answer turn began as an upbeat. The third example (klm-sum04-class-1 in example 8) was slightly different in this respect, as the answer turn was produced on-beat, after a pause containing a silent beat. However, the answer in this example did not answer the question by using one of the options offered. It is possible that the lateness of the answer compared to the other examples may be related to this. Auer et al. (1999) devote a section to rhythm and preference, which could be used as the basis for further work in this area to determine the interactional work done by talking on-beat with other speakers’ rhythms or with self-created rhythms. More study on lists in questions will shed light on whether the structure and properties observed here are recurring, and if there are any other features specific to questions which contain lists.
6. Discussion

The results of the analysis can be interpreted as evidence for a phonological category of rhythm. The exponents of this category may include the prominence of stressed syllables, lengthening of the nuclear syllable, tempo (Sánchez-Ayala (2003: 335)), isometry (Auer et al. (1999: 40)) and the reduction of unstressed syllables, though the present analysis has not focused on these aspects. As mentioned above, there are features which co-occur with these exponents in list structures in particular, including alliteration. Syntactic matching also often co-occurs with rhythmicity in lists, as phrases that are syntactically similar are easier to match in rhythm, as Cummins and Port (1989: 145) mention, quoting Classé (1939).

6.1. Problems with the analysis

Despite these encouraging initial results, there are still large problems with the study of rhythm. Some of the examples from the data were considerably more complicated. In particular, there was one list (shown in example 17 and figure 18) which gave a strong perception of isochrony, but this isochrony could not be measured. The respective deviations were: +66%, -10%, -48%, +52%, -5%, +64% and -12.5%.

(17)  
1 A well you certainly look happy this morning cuz you’re stood together  
2 you’re just  
3 / *shoulder to /  
4 / *shoulder you- you’re /  
5 / *holding each others’ /  
6 / *hands it’s/  
7 / *absolutely de /  
8 / *lightful to /  
9 / *see and you’re /  
10 / *actually a /  
11 / *real inspi/  
12 / *ration /  
13 .hh but now Alan I believe

Figure 18: Isochrony in krm-sum04-holdinghands-5

Even if the repair of “you’re” in line 4 were taken into account as a temporary instability or ‘wobble’ (Auer et al. (1999: 49)), the list shows large deviations in time between the prominences. Perhaps the part of the difficulty with this list lies in the fact that the syntax between the four list items does not match well. The list items are also longer than the examples where
isochrony could be shown. There is clearly still work to be done on lists such as this, where isochrony is perceptible, but, as yet, not measurable.

That being said, even measurable isochrony is still an issue in this area. As mentioned above, there are differing views on what constitutes isochrony, ranging between 20% and 35% allowed deviation. Most of the data do fall into this margin, but not all of them. Despite this, isochrony is hearable. Perceptual experiments would shed some light on the exact margins of isochrony, on the amount of discrepancy allowed for wobbling, and on the amount of time really necessary to establish a beat. Linked to these questions is the issue of subjectivity. To what degree is isochrony subjective? Auer et al. (1999: 45) mention that not every prominence is necessarily a perceptual beat. Would another analyst or phonetician (without musical training) have a different perception of if and where a rhythmic pulse falls?

6.2. Rhythm outside of lists

A few examples in the data have shown that perceptual (and measurable) isochrony is not necessarily limited to the list proper. A beat may continue across a speaker switch (especially in question-answer sequences, or collaborative lists).

Extract ell-sum04-fruits-07 (figure 19) is a list with only two items, which has both a rhythmic introduction and coda. This example supports the observation made previously that rhythm is not only used to construct lists, but is used as some other type of action, of which lists are a sub-type. Whether rhythm is a general tool for cohesion or contextualisation in English could be researched in further analyses.

More practical questions about the use of rhythm also remain unanswered from the analysis, such as whether speakers orient to the actual spoken beat, or to the “underlying beat” as Laver (1994: 525) suggests. Rhythm has been shown to create coherence and manage turn-taking, but there is still a large amount of work to be done to address the issues above.
7. Conclusion

Despite the problems highlighted above, the analysis has shown that rhythm is used and oriented to in naturally occurring lists. The exponents of rhythm have yet to be fully analysed and understood, and their interaction with other prosodic features, such as intonation, must be examined. It is very unlikely that rhythm acts alone in the cases discussed above.

The analysis has shown that by concentrating on one activity within the language, instead of the whole language, and by using the principles of conversation analysis, the prediction that rhythm could be used as a device in conversation has been upheld. Further work should attempt to create a more detailed analysis of this, and should also search for other activities in talk-in-interaction where rhythm is a device available to participants.
Acknowledgements

My thanks go to Richard Ogden for his course on Prosody in English, and his extremely helpful comments on this paper. I would also like to thank the reviewers for their comments and suggestions.

Notes

1. The terminology used here is from Firthian Prosodic Analysis. A good overview of FPA can be found in Anderson (1985). A more detailed explanation of terms, including of phonetic features and phonological categories can be found in Ogden (1995), along with a Firthian Prosodic Analysis of Finnish.

References


