Temporal co-ordination in conversation

Richard Ogden, University of York
Sarah Hawkins, University of Cambridge
Research context

• What are the commonalities and differences in the cognitive processes underpinning alignment in different modalities? How does speech relate to other physical (and visible) behaviours?

• How is social alignment achieved in conversation and music-making?

• We talk in time, and we talk in interaction.
This sub-project:

- How is rhythmic organisation used to handle speaker-to-speaker transitions?
- Focus: Question + Answer, a common adjacency pair, where speaker transition is strongly projected by the Q
- What is the temporal relation between Q and A?
Displaying alignment in answers to questions

A: are you enjoying your place now
B: yeah I am, it’s great

L: where is that
R: it’s near Girton

R: was that here as well
L: I walked in and saw the cameras
demographic questions; initial consent

Minutes

≥ 5 conversation
• How did you get here?
• What do you think of the room?

10 non-musical play
• card houses
• tallest tower: blocks
• market stall: playdough

10 musical play
• xylophone, kalimba
• drums, claves

10 conversation
• 9/11; Princess Diana’s death...
• important event you shared

detailed musical questions; final consent; £8
“Pilot”:
5 dyads,
• 3 musician pairs
• 2 non-musician
• various tasks

“Experiment I”

<table>
<thead>
<tr>
<th>8 dyads</th>
<th>M</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>musician</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>nonmusn</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Dyads:
• friends
• same-sex
• native speakers of English
• 18-31 years
• university educated
• both musicians, or both non-musicians
Recording

- 4 video cameras
- overhead omni mike
- stereo pair (music)
- 2 close-talking mikes
Labelling

• A single framework for labelling events, currently with:
  • 22 tiers for speech and movement
  • + 4 for music

• Gesture labelled without audio in Elan; speech labelled without video in Praat.

• Tracking beats in speech and music: pikes (Loehr 2007): f0 max/min + gesture peak currently, largely by hand; automatic (Cummins) tried but rejected—needs more exploration

Pikes: Loehr (2007) Gesture 7(2) 179–214
Cummins beats: http://cspeech.ucd.ie/~fred/beatExtraction.php
‘Pike’
(Loehr 2007, 2012)

- Peaks of physical activity tend to be temporally co-ordinated:
  - F0 prominence
  - blink
  - peak of gesture

- Loehr looks at individuals’ behaviour, not at the use of pikes to co-ordinate behaviour between individuals: can pikes be used as a method to co-ordinate interaction?
Hawkins, Cross & Ogden (2013)

- Looked at transitions between speech and music in musicians and non-musicians
- Showed that both groups use rhythmical pulses in speech and in music to coordinate interaction: successful interactions in both modalities show a tighter relation of pulses to pikes, and participants show stronger rhythmic entrainment.
Data analysed here

- A collection of 56 Question + Answer pairs
- Labelled using terms from Stivers et al. (PNAS, 2009), which draws on Raymond (ASR, 2003):
  - Q type: WHQ, YNQ
  - A follows Q?
  - visible component?
  - response preferred/dispreferred
Inter-turn temporal organisation
Couper-Kuhlen, *English Speech Rhythm*, 1993

- Presupposes rhythmicity in the Q: 1/5 Qs are arhythmic

- **Rhythmic**: first pike of A comes in on a beat established in the Q

- **Non-rhythmic**: first pike of A does not come in on a beat established in the Q

- **Timing**: early, on-beat, late
L's turn with Q

L: how did your meeting go;

pike π π π
interval 0.329 0.344

R's turn with A

R: pulse

Interval between adjacent pikes (sec)

pike π π π
interval 0.632 0.796 0.974 0.474

Presence of pike (words aligned left).

Pulse established by regular pikes

CLICK it (0.5) went all right (0.5)
Rhythmic vs. Arhythmic Questions

(n = 56)

- Rhythmic Questions: 79%
- Arhythmic Questions: 21%
Rhythmic vs. non-rhythmic entry into the turn space
(n = 44)

- 39% rhythmic (on beat)
- 25% rhythmic (late)
- 27% non-rhythmic (early)
- 5% rhythmic (early)
- 5% non-rhythmic (late)
Arhythmic Qs

• **Q too short to establish a pulse** (minimum of two stressed syllables needed):
  d’they `rEcognise you. ; <<all> what> `BY you?

• **Q produced too quiet or breathy** to measure anything

• **Expansion** by same speaker in same Turn Constructional Unit, no A

• **Perturbations** in production of Q (e.g. self-repairs)

• … or just not rhythmic!
Rhythmic, on-beat entry into the turn-space

L starts an interrogatively formatted Q. R can already prepare for a yes/no formatted answer.
Rhythmic, on-beat entry into the turn-space

Pike on ‘you’.
Rhythmic, on-beat entry into the turn-space

pike
interval
L
pike
interval
R
pulse

π π π

0.29

did you still go to

Next pike on ‘go’.
Interval of 290ms.
Rhythmic, on-beat entry into the turn-space

pike π π π
interval 0.29 0.36
L did you still go to school?
pike π π π
interval 0.36 0.33 0.31
R "Yeah." we still went to school-

Next pike on ‘school’. Interval of 360ms. (Weak) pulse established. R now knows what a fitted answer would be and has a time slot.
Rhythmic, on-beat entry into the turn-space

pike
interval
π π π

L
did you still go to school?

pike
interval
R

R produces a fitted response-token on beat with pulse in Q.

pulse

^ ^
Rhythmic, on-beat entry into the turn-space

Even though the pulse in L’s turn is not precise, 
R orients to it and uses it to time his A, which is syntactically and 
pragmatically fitted to the Q.

Temporal organisation is a shared resource 
(cf. Pickering & Garrod).
did you still go to school

it yeah we still went to school
Rhythmic, on-beat entry into the turn-space
Answers with rhythmic, on-beat entry

• Usually **fitted to Q**, e.g. ‘yes’/‘no’ to polar Q; fitted response to WH-word

• ... or an **account** for why no proper A (‘I can’t remember’, ‘I don’t know’)

• **A aligns with the Q**: it recognises that a Q has been asked, and a fitted A is provided in the appropriate temporal and sequential slot
Rhythmic, early entry into the turn-space

*Pulse maintained to 3rd position*

L starts a declaratively formatted turn
Rhythmic, early entry into the turn-space

**Pulse maintained to 3rd position**

... with a tag which makes relevant a response from R
(with preference for confirmation).
Rhythmic, early entry into the turn-space

*Pulse maintained to 3rd position*

pike
interval
L
it was a SUniday
wasn’t
it
pike
interval
R
pulse

R’s response comes immediately the Q is recognisable but a bit before it is due
Rhythmic, early entry into the turn-space

*Pulse maintained to 3rd position*

pike

interval

it was a Sunday wasn’t it

pike

interval

R

pulse

A in two parts: response token + S

A

in two parts: response token + S
Rhythmic, early entry into the turn-space

*Pulse maintained to 3rd position*

Overlapping, simultaneous pikes in third position; mutual orientation to pulse in the Q+A pair.

L and R display a shared understanding.
Rhythmic, late entry into the turn-space

pike \( \pi \) \( \pi \)
interval \( 0.33 \)
R
[how was your meeting.
L
[good
pulse
Rhythmic, late entry into the turn-space

pike π π π
interval 0.33
R [how was mEEting.
your
pike
interval
L [good
pulse ^
Rhythmic, late entry into the turn-space

pike π π
interval 0.33
R [hOw was mEEting.
your
pike
interval
L [good it was
pulse ∧ ∧
Rhythmic, late entry into the turn-space

pike π π
interval 0.33

R [How was your meeting.

pike π
interval 0.99
L [good

pulse ∧ ∧ ∧

Two beats of silence + pike of ‘good’ on third beat.
Rhythmic, late entry into the turn-space

pike $\pi \pi$

interval 0.33

R [how was meeting. your

pike $\pi \pi$

interval 0.99 0.33

L [good it was good; it long; it

was very

pulse ^ ^ ^
Rhythmic, late entry into the turn-space

pike \( \pi \) \( \pi \)
interval 0.33
R \[ \text{How was your mEEting.} \]
pike \( \pi \) \( \pi \) \( \pi \) \( \pi \)
interval 0.99 0.33 0.48 0.43
L \[ \text{good} \]
pulse \( \wedge \) \( \wedge \) \( \wedge \) \( \wedge \)

Pulse not maintained; complex answer.
Answers with rhythmic, late entry

- Usually fitted to Q (e.g. ‘yes’, ‘no’ to polar Q; fitted response to WH-word)
- but often presents a complication in the A, e.g. displaying some kind of effort (remembering a name or date)
- A aligns with the Q
Non-rhythmic, late entry into the turn-space

**But:** onset of A on second pulse after Q. **No ‘yes’**.

First pike of A 2.5 pulses after Q. **Resolves ambiguity in Q**.

pike \( \pi \) \( \pi \)
interval 0.361
\( R \) don’t/ YOU play hOckey
\( R \) di(d)n’t
pike
interval
L
pulse

\( \pi \) \( \pi \)
0.725 0.908 0.449

\( I \) do again nOW

\( \wedge \) \( \wedge \)
Non-rhythmic, late entry into the turn-space

First pike of A 2.74 pulses after Q

But: onset of A on second pulse after Q

pike \[\pi\] \[\pi\] \[\pi\] interval 0.268 0.298
R \( \text{what d'you think of the room} \)
pike interval
L
pulse \[\wedge\] \[\wedge\] \[\wedge\] \[\wedge\]

It's quite a nice room
Answers with non-rhythmic entry
Auer, Couper-Kuhlen & Müller 1999; Couper-Kuhlen 1993

• These answers tend to be more complex, e.g. correcting a presupposition of the Q

• Although the first pike of the answer is not rhythmically aligned, in answers that align with the Q, the production of speech often starts on beat

• A more marked format
Pulses and the pre-turn space

• Sometimes the A is prefaced with on-beat pre-turn material: in-breaths, clicks, *uhm*, etc. These are signs of ‘gearing up to speak’, so they project incipient speakership without yet taking a turn. (Cf. Ogden, *JIPA* 2013: clicks as metronomes)

• Such on-beat in-comings display an orientation to the relevance of an on-beat A, even if the first pike of the turn is not on beat.
Rhythmic, on-beat pre-turn material

Aligned but dispreferred response (‘no’)
Click provides a rhythmic and on-beat entry into the turn space

pike

π π π

interval

0.528 0.523

R has hE got to nIGHT nIGHT or

do a shift to-

pike

π π π

interval

0.551 1.033/0.482 0.271 0.582

L CLICK no hE’s on nIghts mOrrow from to-

pulse

∧ ∧ ∧
Rhythmic, on-beat pre-turn material

pike  π  π  π
interval  0.646  0.779

L  is it just college meet the rowing morning
the

pike  π  π  π
interval  0.855  0.859  0.495
R  UH (0.3) well
pulse  ∧  ∧  ∧

Complex response projected (‘well’) ‘Uh’ on beat; orients to relevance of a well-timed response
Rhythmic, on-beat pre-turn material

pike π π π
interval 0.329 0.344
L how did your mEEting go;
pike
interval 0.632 0.796 0.974 0.474
R CLICK It’ (0.5) wEnt all rIGHt (0.5)
pulse ^ ^ ^

Dispreferred response (gap; non-rhythmic entry) CLICK on beat; orients to relevance of a well-timed response
Non-rhythmic, late entry into the turn-space

• Pike of A is often not on beat with the pulse of the Q; but often, the onset of talk (rather than the pike) occurs on beat with the pulse.

• Suggests two sub-types of rhythmical alignment: pike-to-pike, pike-to-onset.

• Suggests measures need to include non-verbal (but vocal) material.
Further questions

- Best way to model pikes and rhythms?
- Most effective number of pikes to give a pulse?
- Can turns be designed so as to be long enough to produce a pulse?
- How are disfluencies and perturbations in the progressivity of social actions handled?
- Do we find cases where rhythm is repaired? (Would demonstrate orientation to wrong/right versions.)
Conclusions

• Alignment in interaction is handled through lexical and syntactic design (and parallelism); but also phonetic design

• Rhythmicity is a locally available resource and handles contingencies of interacting in time: it helps to reinforce intersubjectivity: participants’ displayed understandings of one another

• Our study contrasts with many studies of rhythm and timing in speech: it does not lead us to assume global properties language by language
Thanks to!

Hannah Leach

Ian Cross

Daniel Halford

Satinder Gill

David Greatrex

Rein Ove Sikveland

Newton Trust, U. Cambridge; BA/Leverhulme Foundation Small Grant