Studies from the periphery of language

Why phonetics is more than larynx-to-lips

Richard Ogden
richard.ogden@york.ac.uk
University of York
Centre for Advanced Studies in Language and Communication
Department of Language & Linguistic Science
Conventional assumptions about speech

- Sounds have linguistic (especially *lexical*) value
- Sound types = consonant, vowel, suprasegmental (under-theorised on IPA chart)
- Speech is made in the range: larynx to lips/nostrils
Outline:

Conventional assumptions about speech can be challenged:

1. Issues of modality.

2. The prosody/segment division accounts poorly for many kinds of meaning.

3. Some sounds are ‘in’ speech but not ‘of’ it.
I. Language and modality
1.1. Speech and gesture aren’t as distinct as we thought
• There is a **continuum between co-speech gesture and sign language** (MacNeill 2005, Slobin 2008)

• Gesture may be deeply embedded in sign language (Liddell & Metzger 1998, Liddell 2003)

• Some gestures resemble depicting signs or depicting signs have gestural elements (Kendon 2004)

• **Pikes**: peaks of physical (including speech) activity (Loehr 2007)

• Gesture can contribute **aspects of meaning** which are not easily verbalised (Rowbotham et al. 2011)

• **Growth point**: common origin for imagery and categorial content (MacNeill 2005)
There is a continuum between co-speech gesture and sign language

Based on MacNeill (2005: 7-10)

<table>
<thead>
<tr>
<th></th>
<th>gesticulation</th>
<th>emblems</th>
<th>sign language</th>
</tr>
</thead>
<tbody>
<tr>
<td>speech?</td>
<td>present</td>
<td>optional</td>
<td>absent</td>
</tr>
<tr>
<td>linguistic properties</td>
<td>absent</td>
<td>some</td>
<td>present</td>
</tr>
<tr>
<td>conventionalised?</td>
<td>no</td>
<td>partly</td>
<td>fully</td>
</tr>
<tr>
<td>semiosis</td>
<td>global</td>
<td>segmented</td>
<td>segmented</td>
</tr>
<tr>
<td></td>
<td>synthetic</td>
<td>synthetic</td>
<td>analytic</td>
</tr>
</tbody>
</table>

- sign languages contain gesture
- co-speech gesture shares some similarities with sign language
Some gestures resemble depicting signs
(or: depicting signs have gestural elements)

Gestures specifying size, shape and structure of the cheese crates and how the cheeses were placed in them.
A. **Size-shape-specifier** gesture serving as a referent for the deictic pronoun ‘that’.
B. **Outline sketching** gesture, describing the shape of the ends of the crate.
C. Extended spread fingers moved laterally and horizontally depict the lateral horizontal arrangement of multiple long thin objects — the slats on the crate.
D. Spread hands perform 'object-placement' movements and thus indicate the relative position of the two cheeses packed in the crate.

Kendon (2004: 166)
## Talking about pain

### Table 4  Examples of semantic feature coding for gesture and speech

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>“Sort of lower stomach pains”</td>
</tr>
<tr>
<td>Hand moves repeatedly from left to right and back again across the lower stomach</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>“It feels like my head’s getting smaller and my brain’s getting bigger”</td>
</tr>
<tr>
<td>Hands held with palms facing each other and fingers outstretched, first moved towards each other and then out again, to demonstrate the idea of shrinking and swelling</td>
<td></td>
</tr>
</tbody>
</table>

also: quality, intensity, progression, cause, effects, presence

1.2. Speech perception is enhanced with visual information
(Visual) prosody

• Consensus: auditory signal takes precedence over visual signal, but visual signal improves intelligibility.

• The upper part of the face facilitates identification of prosodic prominence (Swerts & Krahmer 2008) and intelligibility of F0 (Beskow, Granström & House 2006)

• Questions and statements signalled bimodally (Srivasanen & Massaro (2003)

• Head movements improve intelligibility in synthetic speech (Munhall et al 2004)

• Utterance-finality detected better with audio and visual cues, despite individual variation (Barkhuysen, Krahmer & Swerts 2008)
2. The prosody-segment distinction is arbitrary
‘Segmental’ features in conversation

• Aspiration in Tyneside English delimits turns (Local, Kelly & Wells 1986)

• [ʔ̚, p̚, k̚] etc. mark turn holding in English (Local & Kelly 1986) and Finnish (Ogden 2001)

• Clicks to delimit sequence boundaries (Wright, 2005, 2011a & b)

• = non-lexical functions usually explained primarily in terms of intonation
Articulatory setting + prosodic features

- Repeats in other-initiated self-repair (Curl 2005)
- Turn-projection and talk-projection (Local & Walker 2012)
- Agreement and disagreement (Ogden 2006)
- Repair type affects phonetic design (Plug 2010, 2011)
- Complaints that seek affiliation vs. complaints that close the sequence (Ogden 2010)
- Intensification (Ogden 2012)
Phonetic resources for (dis-) agreeing (Ogden, JPrag 2006)

<table>
<thead>
<tr>
<th>Strong agreement or Disagreement</th>
<th>Weak agreement (+ disagreement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative to first assessment:</td>
<td>Relative to first assessment:</td>
</tr>
<tr>
<td>• wider pitch span</td>
<td>• narrower pitch span</td>
</tr>
<tr>
<td>• more dynamic contours</td>
<td>• less dynamic contours</td>
</tr>
<tr>
<td>• tenser articulations</td>
<td>• laxer articulations</td>
</tr>
<tr>
<td>• slower tempo</td>
<td>• faster tempo</td>
</tr>
<tr>
<td>upgrade</td>
<td>downgrade</td>
</tr>
<tr>
<td><em>it’s supposed to be really really pretty</em></td>
<td><em>and it’s a lot of hassle</em></td>
</tr>
<tr>
<td><em>oh it’s supposed to be gorgeous</em></td>
<td></td>
</tr>
</tbody>
</table>

1  ➔  ? deferred  
2  ➔  a response  | 1  ➔  ? deferred  
2  ➔  it’s a lot of effort  

Ward and Hirschberg, 1985; Kadmon, 2001; Ogden, JPrag 2006

Disagreement

Relative to first assessment:
• faster tempo
• narrower pitch span
• less dynamic contours
• laxer articulations

Strong agreement or Disagreement

Relative to first assessment:
• wider pitch span
• more dynamic contours
• tenser articulations
• slower tempo

Weak agreement (+ disagreement)


the position taken in 1

the turn is presented with
the heritage

It is supposed

oh it’s supposed to be really really pretty

it’s a lot of hassle

and it’s a lot of effort

It is quite a lot of effort

impressionistic records of the assessments in Fragment 13

speed is 4.7 and 7.7 syllables/s, respectively.

Impressionistic records of these utterances are given

relative to first assessment:
• faster tempo
• narrower pitch span
• less dynamic contours

It is quite a lot of effort

Relative to first assessment:
• wider pitch span
• more dynamic contours
• tenser articulations
• slower tempo

it’s a lot of hassle

Oh it’s supposed to be gorgeous

impressionistic records of the two utterances are given

Relative to first assessment:
• faster tempo
• narrower pitch span
• less dynamic contours
• laxer articulations
• faster tempo

it’s a lot of hassle

Oh it’s supposed to be gorgeous

Relative to first assessment:
• faster tempo
• narrower pitch span
• less dynamic contours

It is quite a lot of effort

Oh it’s supposed to be gorgeous

Relative to first assessment:
• wider pitch span
• more dynamic contours
• tenser articulations
• slower tempo

it’s a lot of hassle

Oh it’s supposed to be gorgeous

Relative to first assessment:
• faster tempo
• narrower pitch span
• less dynamic contours
• laxer articulations
• faster tempo
CH en_4822.924-961

01 A [speaking of which, I had the `rUssians come to the
02 `PARty on ‘thUrsday,
03 B <<laugh> oh ‘yEAh?>
04 → A <<all> we had> ?a `B:A:[SH.
05 B [oh my god
06 A you have never seen so much alcohol consumed b[y a] (*)(*)
07 B [re-]
08 r- re[all]y
09 A [(*)(*)]
10 unbelievable
11 B o[:h
12 A [they brought all this gin
13 B he[hehe
14 A [and vodka
15 B really <<laugh> gi[n>
16 A [they got
17 A <<len> `↑SCHNO::Ckere[d=> =they they we had a]
18 B [ oh my god ]
19 A `↑P:A:R↑ty he[re.
20 B [<<breathy> wo::w::>]
21 A dancing arou:nd; spilling stu::ff; still sticky everywhere
we had a **bash**

**full form:** [eɪ]

- **closure:** 325 ms

---

**glottalisation**

- **faster**

**voicing + closure:** 85 ms

---

**faster**

- **slow**
Summary


- ‘Prosodic’ + ‘segmental’ modifications work together

- Actions (like words) have exponents which can be stated in terms of the mutual co-ordination of phonetic parameters in time
3. Sounds that are ‘in’ speech but not ‘of’ it:

Clicks and percussives in English conversation

*Under review; e-mail for a sneak preview!*
Why study clicks?

• In English: non-phonemic, but demonstrably meaningful, sounds (‘sound objects’, Reber 2012)

• A better understanding of the resources available to speakers to manage talk and mark stance and affect

• Conventional accounts of paralinguistic form : meaning (e.g. biological codes, Gussenhoven 2004) are inadequate in accounting for clicks

• A better understanding of where the boundaries of language lie
Affective interpretations of clicks in English

(Wright 2005: 41)

‘Broad transcription’ of clicks in English

• [!] Tongue-tip click, generally slow, affricated release down the centre line of the vocal tract. ‘tsk tsk’, ‘tut’. Closure dental or alveolar.

• [ǁ] Tongue-tip click, lateral release. Closure ranges from dental to palatal (retroflex).
Turn-initiation
Pre-turn position. Marks incipient speakership.

ell sum04 cheese
01 P what’s on offer at the market then today
02 I !h↓ well we’ve got lots of cheeses
03 we’ve got h↓ pancakes
04 very h↓ traditional continental dishes
Word search
Mid-turn position. Marks continuing speakership.

Salford M & D 1328 secondary school

word search
01 M because there were one or two big houses
02 D yeah
03 M my (0.2) uhm (0.4)! secondary school
04 the girls' high school. that had been a big house.
Preface to ‘Resistance’
Marks incipient speakership.

CallHome 4861.60

01 Mom obviously I didn’t do a good enough job of raising you
02 Deb ➔ [!h↓] [oh s]top
03 that

Click + inbreath
Spoken part of the turn at a TRP
Clicks as metronomes.
TCU-initial on-beat click in overlap

CK/SW sympathy #63

1 A it’s just kinda dull
2 A God what a miserable miserable [weekend
3 B that’s a shame

Display of sympathy

Rhythmical negative assessment

Click + response token

GLOTTALIC AND VELARIC AIRSTREAMS

2 P what's on offer at the market then today

well we've got lots of different cheeses

1 I ((complains about new houses to be built in M—))

and you're unhappy about that

yes and so's a lot of people in M——

thank you

Councillor P——?

oh hello Imogen

the issue is, as you know, is the Government...
Anthony and Ray are talking about places they’d like to go. Anthony would like to go to Australia.

01 A  I'll just go on the topless beaches
02 R  yeah (..) you wish
03 A  no I know where they are
04 [ || || [ || ] ]
05 R  [<<laugh> * * * * * * * * * * * * ]
06 A  and titty bars
07 R  right, anyway
08 A  <<laugh>>-[---------------------------------------->]
09 R  [so where else would <<laugh> you wanna go>]
Salford A & R lateral click 1654 topless beaches
01 A  I'll just go on the topless beaches
02 R  yeah (...) you wish
03 A  no I know where they are
04    [ || || || || ]
05 R  [<<laugh> * * * * * * * * * * * * ]
06 A  and titty bars
07 R  right, anyway
08 A  <<laugh>-[-------------------------------------------------->]
09 R  [so where else would <<laugh> you wanna go>]
Clicks as embodied sound objects.
Display of affect

16. Virginia p.3 01:27

23 VIR: Can I please get that dress, please mom? Lemme g[et that-
24 MOM: [Dreh(ss)-?
25 VIR: >You know that [one-<
26 MOM: [OH VIRginia, we('ve) been through this
27 befo[RE, you'v[E, we('ve] been through this
28 ???: [huhh! ((laughter?))
29 VR?: [uhhh! ((“pained” sound))
30 MOM: =just wait a
31 in.
32 VIR: → (!]
33 (0.5)

01 PRU: I[t's s o| 02 MOM: [If you s|a
16. Virginia p.3 01:27

23 VIR: Can I please get that dress, please mom?
24 MOM: >You know that one-
25 VIR: > You know that [one--
26 MOM: [OH VIRginia]
27 befo[RE, you've got enough-
28 PRU: [hmmm! ((laughter?))
29 VIR?: [hmmm! (("pained" sound))
30 MOM: =just wait an' get- some of the new fa:ll stuff when it comes
31 in.
32 VIR: [!] (0.5)
33
01 PRU: [It's s
02 MOM: [If you
Head-turn

• Virginia’s click is accompanied by a head-turn away from Mom

• It is accompanied by other features which display her exasperation with Mom: nagging, a ‘pained’ sound, a complaint

• The click is but one part of how Virginia displays her stance towards Mom
and then I went over to the library to-uhm (0.6)

I can’t remember the name of the card now but to

to use other universit[y libraries

[oh a (sconal) card

yeah and I asked the woman at the library and she sort of said

yeah you can do that,
and then I went over to the library to—uhm (1)(2)(3)(0.6) see about that—uh
I can’t remember the name of the card now but to use other university libraries
[oh a (sconal) card]
yeah and I asked the woman at the library and she sort of said yeah you can do that,
Swallowing

- The click is the culmination of other physical activity: lip closure, swallow, release. Swallowing is compatible with rarefaction.

- The speaker swallows in a classical word-search environment.

- [swallow] + [click] are ‘in’ speech but not ‘of’ it.

- Visible to us, but R isn’t looking.
velaric ingressive click

Displaying a stance

+ [turn-initial particle: {oh, ah, ooh,...}]

voice quality; duration; intonation; pitch span; pitch range....

+ [loud in-breath]

Marking incipient speakership

[separation of articulators]

[ingressive airflow: {pulmonic, velaric}]

+[ə::], [m::]

Sequence management

[swallow]

[word+uh]

[in-breath]
‘In’ language but not ‘of’ it...

• Clicks (and percussives) have a regular, but complex, distribution

• Clearly they are not phonemic, but there is an argument that they are linguistic: regular distribution, part of bigger meaning-bearing prosodic constructions (l’arbitraire du signe)
Conclusions

- Sounds have linguistic (especially *lexical*) value. Many ‘paralinguistic’ phenomena display orderly linguistic and sequential patterning.

- Sound types = consonant, vowel, suprasegmental (under-theorised on IPA chart). Good evidence from interaction that features of speech work together in bundles; our task is to consider what bundles together, and how it works in time.

- Speech is made in the range: larynx to lips/nostrils. Increasing evidence from both lab speech and speech in its most natural setting that visual/gestural information is significant in the accomplishment of action, and in the perception of traditional categories.
Thanks!

University of York Research Priming Fund

For helpful discussions along the way:
Members of the Finnish Centre of Excellence in Intersubjectivity in Helsinki; Kearsey Cormier, Paul Drew, John Local, Dan Loehr, Emma Piggott, Jim Scobie, Rein Sikveland, Jane Stuart-Smith, Marilyn Vihman, Dom Watt, Melissa Wright