

## **Data, data documentation and analysis scripts for**

### *Multimodal character viewpoint in quoted dialogue sequences*

Kashmiri Stec, Martijn Wieling, Mike Huiskes & Gisela Redeker

Center for Language & Cognition  
University of Groningen, the Netherlands

Journal: **Glossa** (2017)

<https://doi.org/10.5334/gjgl.255>

## **Abstract**

We investigate the multimodal production of character viewpoint by performing complementary qualitative and quantitative analyses of two quoted dialogues, focusing on the storyteller's use of character viewpoint gestures, character intonation, character facial expression, spatial orientation and gaze. A micro-analysis revealed that the extent of multimodal articulation depends on (i) the quoted speaker, with different multimodal articulatory patterns found for quotes by the speaker's past self vs. a third person character, and (ii) the position of the quoted utterance within the quoted dialogue, with mid-dialogue utterances garnering less co-articulation than initial or final utterances within the quoted dialogue. We further investigate these observations using a quantitative approach, which was based on generalized additive modeling (GAM). The GAM analysis revealed different multimodal patterns for each quoted character, as indicated by the number of co-produced multimodal articulators. These patterns were found to hold regardless of the quote's position within the narrative. We discuss these findings with respect to previous work on multimodal quotation.

**Keywords:** multimodality, co-speech gesture, viewpoint, direct speech, quotation

## 1 Packages

```
library(mgcv)
```

```
R.Version()$version.string
```

```
## [1] "R version 3.1.2 (2014-10-31)"
```

```
packageVersion("mgcv")
```

```
## [1] 1.8.4
```

## 2 Dataset

```
load ('gam.suppl.csv', header=T)
```

Multimodal quotes (704 observations of 26 variables)

The dataset contains values for the following variables. Only data for speaker 12 (“Black”) is included. The file contains variables from several stages of the project (ELAN export, ELAN annotations, speaker variables and variables for analysis).

### (I) Variables from ELAN export

1. Item: Item number
2. Begin time (hh:mm:ss:ms): Time code for the start of the quote
3. Begin time (ss.msec): Time code for the start of the quote
4. End time (hh:mm:ss:ms) Time code for the end of the quote
5. End time (ss.msec): Time code for the end of the quote
6. Duration (hh:mm:ss:ms): Duration of the quote
7. Duration (ss.msec): Duration of the quote

### (II) Annotations made in ELAN

8. Speech: Transcription of the quote
9. Quotative: Classification of the utterance as a quote (yes), not a quote (no) or unsure (maybe)
10. Context: An identification of the quote context: single quote (also called Quote Island), quoted monologue (QM or QMI for initial QM utterance), or quoted dialogue (QD or QDIa, QDIb for the initial utterances in a QD sequence).
11. Character: An identification of the quoted character (A-E, speaker, addressee, speaker+addressee)
12. Intonation: The speaker produced character intonation during the quoted utterance (yes), did not (no) or unclear (maybe)
13. Gesture: The speaker produced a character viewpoint gesture (CVPT), another gesture (Other) or no gesture (NG).
14. Facial expression: The speaker displayed the quoted character’s facial expression (yes), did not (no) or unclear (maybe).
15. Gaze: The speaker looked away from the addressee (yes), maintained gaze with the addressee (no), looked away from the addressee after the quoted utterance started (late change) or the speaker’s gaze jumped around (quick shift).
16. Change: Classification of speaker’s body movements as horizontal, vertical, sagittal, none or unsure.

### (III) Speaker/File data

17. Speaker: Speaker ID number.
18. File: The filename of the narrative clip.

#### (IV) Variables created for analysis

19. IsIntonation: Binary variable indicating if the speaker used character intonation (1) or not (0). All unclear cases from #11 were coded as '0'.
20. IsFvpt: Binary variable indicating if the speaker produced character facial expression (1) or not (0). All unclear cases from #13 were coded as '0'.
21. Gaze\_Any: Binary variable indicating if the speaker made a “meaningful” use of gaze. Three codes from #14 were coded as 1 (yes, late change, quick shift). Maintaining gaze with the addressee (no) was coded as 0.
22. IsCVPT: Binary variable indicating if the speaker produced a manual character viewpoint gesture (1) or no (0). All Other gestures and no gestures from #12 were coded as 0.
23. Change\_Any\_Direction: A binary variable which indicates if the speaker moved. Values are based on codes from #15; Horizontal, Vertical and Sagittal were counted as (1) and no movement and unclear were counted as (0).
24. ArtCnt5 (Articulator Count 5): A variable which counts the number of 1's occurring in #19-23. Values range from 0 (no 1's, or no multimodal articulator use) to 5 (all 1's, or all multimodal articulators are used).
25. IsCH: 3-way variable indicating whether the quoted speaker is the Addressee (A), the Speaker (S) or a Character (CH). Re-coded from Character.
26. IsSeq: Binary variable indicating whether the quoted utterance is initial (Initial) or non-initial (Other). Re-coded from Context.

### 3 Codebook

#### *(I) Linguistic features*

**Quoted utterance** – Is the utterance a direct speech quote?

- Yes – it is.
- No – it is not.
- Unclear

**Quote context** – In which context does the quoted utterance occur?

- QI – Quote Island, or single quote
- QM(I) – Quoted Monologue (Initial utterance)
- QD(Ia, Ib) – Quoted Dialogue (Initial utterance by character a, b)
- QDM(Ia, Ib) – Quoted Dialogue with an AAB or ABB instead of ABAB.. pattern
- O - Other

**Quoted character** – Enter the quoted character. Please be consistent across the narrative (i.e. all “Sue said” quotes should be identified with the same letter, e.g. A.)

- Speaker – the speaker quotes themselves
- Addressee – the speaker quotes the addressee
- Speaker + Addressee – the speaker quotes themselves + the addressee (e.g. “We’d be all...”)
- Letters A-F – indexes which refer to unique characters which are quoted in the narrative (e.g. A always identifies quotes made by Sue, B always identified quotes made by William, etc.)

#### *(II) Multimodal features*

**Manual gesture** – CVPT – character viewpoint (McNeill 1992). Other – any gesture which is not a CPVT gesture. NG – no gesture.

**Facial expression** – Does the speaker’s face show aspects of the quoted speaker, e.g. affect such as fear, surprise, anger, joy? This should be an easy binary distinction. If you hesitate or are unsure, the correct choice is No. We want only to capture major facial expressions.

- Yes – the speaker’s face depicts aspects of the quoted speaker’s.
- No – it does not.

**Gaze** – Is the speaker looking directly at the addressee, or are the speaker and addressee making eye contact? Is the speaker looking somewhere else (not at the addressee)? Is the speaker’s gaze undirected and/or jumping all over the place?

- Yes – the speaker is gazing at the addressee
- No – the speaker is looking away from the addressee
- Late change – the speaker looks away from the addressee, but not on the left boundary of the quote

- Quick shift – the speaker’s gaze jumps all over the place

**Intonation** -- Does the speaker’s voice change to show aspects of the quoted speaker, e.g. changes in pitch, loudness, accent, or emotions such as joy, anger, confusion, etc.? This should be an easy binary distinction. If you hesitate or are unsure, the correct choice is No. We want only to capture major changes in intonation.

- Yes – the speaker’s voice depicts aspects of the quoted speaker’s.
- No – it does not.

**Body Orientation (Change)** – The perceived direction of the speaker’s movement during the quoted utterance.; “movement” can refer to the speaker’s head or torso, or where manual gestures are produced. We are interested in a holistic impression of the speaker’s “movement” during the entire utterance. Does it look like they’re basically staying neutral, moving forward, or up, or left?

- Horizontal – the speaker moves laterally (left, right).
- Vertical – the speaker moves vertically (up, down).
- Sagittal – the speaker moves forwards towards the addressee or backwards, away from them.
- None – the speaker makes no movement.
- ? – not sure.

## 4 Analysis and results

The file `sharing.R` contains code and instructions for running our models.

Our dataset is small; it takes less than 5 minutes to run each GAM.

In this section we show:

- GAM results for the narrative “Airports” using REML and ML methods. ML is more conservative and gives lower p-values. REML gives better smooths.
- GAM results for speaker 12 (narratives Airports and Philosophy) using REML and ML methods. We use Philosophy because it also has an almost equal number of quotes for the two most quoted characters (speaker and A). See Table A1 for quote comparisons in other narratives by speaker 12.

### Airports (only) REML

```
> m1b = gam(ISA ~ s(ArtCnt,k=3) ,data=tmp,family='binomial',method='REML')
> summary(m1b)
```

```
Family: binomial
Link function: logit
```

```
Formula:
ISA ~ s(ArtCnt, k = 3)
```

```
Parametric coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.09419    0.45672   0.206   0.837
```

```
Approximate significance of smooth terms:
              edf Ref.df Chi.sq p-value
s(ArtCnt)  1.901   1.99  7.493  0.0234 *
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
R-sq.(adj) = 0.404   Deviance explained = 40.3%
-REML = 16.312   Scale est. = 1           n = 37
```

```
> m2b = gam(IsSpeaker ~ s(ArtCnt,k=3) ,data=tmp,family='binomial',method='REML')
> summary(m2b)
```

```
Family: binomial
Link function: logit
```

```
Formula:
IsSpeaker ~ s(ArtCnt, k = 3)
```

```
Parametric coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.09419    0.45672  -0.206   0.837
```

```
Approximate significance of smooth terms:
              edf Ref.df Chi.sq p-value
s(ArtCnt)  1.901   1.99  7.493  0.0234 *
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
R-sq.(adj) = 0.404   Deviance explained = 40.3%
-REML = 16.312   Scale est. = 1           n = 37
```

## Airports (only) ML

```
> m1b = gam(IsA ~ s(ArtCnt,k=3) ,data=tmp,family='binomial',method='ML')
> summary(m1b)
```

```
Family: binomial
Link function: logit
```

```
Formula:
IsA ~ s(ArtCnt, k = 3)
```

```
Parametric coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.1046    0.4532   0.231   0.817
```

```
Approximate significance of smooth terms:
              edf Ref.df Chi.sq p-value
s(ArtCnt)  1.878  1.985  7.843  0.0196 *
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
R-sq.(adj) = 0.405   Deviance explained = 40.2%
-ML = 16.735   Scale est. = 1           n = 37
```

```
> m2b = gam(IsSpeaker ~ s(ArtCnt,k=3) ,data=tmp,family='binomial',method='ML')
> summary(m2b)
```

```
Family: binomial
Link function: logit
```

```
Formula:
IsSpeaker ~ s(ArtCnt, k = 3)
```

```
Parametric coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.1046    0.4532  -0.231   0.817
```

```
Approximate significance of smooth terms:
              edf Ref.df Chi.sq p-value
s(ArtCnt)  1.878  1.985  7.843  0.0196 *
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
R-sq.(adj) = 0.405   Deviance explained = 40.2%
-ML = 16.735   Scale est. = 1           n = 37
```