

Data, data documentation and analysis scripts for

Multimodal character viewpoint in quoted dialogue sequences

Kashmiri Stec, Mike Huiskes & Gisela Redeker

Center for Language & Cognition
University of Groningen, the Netherlands

Journal: **Journal of Pragmatics** (accepted for publication July 2016)

Abstract

This study investigates how speakers of American English use multimodal articulation when quoting characters in personal narratives. We use the concept of role shift, adapted from signed languages, where it refers to a device used to represent one or more characters with one or more bodily articulators, to describe multimodal role shift practices. In a regression analysis, four bodily articulators were found to predict the impression of a role shift: character intonation, character facial expressions, character viewpoint gestures, and changes in body orientation; gaze was not a significant predictor. Most of the 704 quotes in our data are accompanied by activation of two or three articulators (55.3%) and very few (2.6%) occur without any of the articulators we have annotated. The extent of multimodal articulation depends on the type of quoted utterance: quotations of actual, witnessed speech events tend to garner fewer articulators than constructed ('fictive interaction') quotations. These findings demonstrate that speakers, like signers, use a range of bodily articulators when they take on another's role in quotation and thus underpin the importance of investigating the systematic use of the visual modality in quotation and, more generally, in ordinary interaction.

Keywords: co-speech gesture; viewpoint; quotation; direct speech; fictive interaction; multimodality

1 Packages

```
library(mgcv)  
library(Hmisc)
```

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

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

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

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

```
packageVersion("Hmisc")
```

```
## [1] 3.14.6
```

2 Dataset

```
load ('gam.supp.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. Qhead: The quoting predicate.
11. FI: Classification of the utterance as fictive interaction (1), normal quote (0) or unsure (0.5).
12. Role shift: The speaker was judged to role shift (1), not role shift (0) or unsure (0.5).
13. Intonation: The speaker produced character intonation during the quoted utterance (yes), did not (no) or unclear (maybe)
14. Gesture: The speaker produced a character viewpoint gesture (CVPT), another gesture (Other) or no gesture (NG).
15. Facial expression: The speaker displayed the quoted character’s facial expression (yes), did not (no) or unclear (maybe).
16. 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).
17. Change: Classification of speaker’s body movements as horizontal, vertical, sagittal, none or unclear.

(III) Speaker/File data

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

(IV) Variables created for analysis

20. IsFI: Binary variable indicating whether the quoted utterance is an instance of fictive interaction (1) or not (0). All unclear cases from #11 were coded as '0'.
21. IsRS: Binary variable indicating whether the speaker role shifted (1) or not (0). All unclear cases from #12 were coded as '0'.
22. IsIntonation: Binary variable indicating if the speaker used character intonation (1) or not (0). All unclear cases from #13 were coded as '0'.
23. IsFvpt: Binary variable indicating if the speaker produced character facial expression (1) or not (0). All unclear cases from #14 were coded as '0'.
24. Gaze_Any: Binary variable indicating if the speaker made a "meaningful" use of gaze. Three codes from #15 were coded as 1 (yes, late change, quick shift). Maintaining gaze with the addressee (no) was coded as 0.
25. IsCVPT: Binary variable indicating if the speaker produced a manual character viewpoint gesture (1) or no (0). All Other gestures and no gestures from #16 were coded as 0.
26. Change_Any_Direction: A binary variable which indicates if the speaker moved. Values are based on codes from #17; Horizontal, Vertical and Sagittal were counted as (1) and no movement and unclear were counted as (0).
27. ArtCnt5 (Articulator Count 5): A variable which counts the number of 1's occurring in #13-17. Values range from 0 (no 1's, or no multimodal articulator use) to 5 (all 1's, or all multimodal articulators are used).

3 Codebook

(I) Linguistic features

Quoted utterance – Is the utterance a direct speech quote?

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

Speech – If the utterance is a direct speech quote, transcribe it in full.

Quoting predicate – Enter the quoting predicate used. If none, write ‘bare’.

Fictive interaction – Is the utterance an instance of fictive interaction (Pascual 2002; 2014)? If so, Yes. If not, No. If unsure, Maybe.

(II) Multimodal features

Role Shift – Does the speaker appear to use their body to show aspects of the quoted speaker, e.g. what the quoted speaker was doing or feeling? This should be an easy binary distinction. If you hesitate or are unsure, the correct choice is No.

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

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 should take less than 2 minutes to run all analyses.

In this section we show:

- GAM results and code for the models presented in Section 4 of the paper (Tables 7 and 8)
- As stated in Footnote 3, the `glm` function in the `lme4` package is typically used for running regressions however its optimizers had difficulty converging (sometimes they would, sometimes they would not) so we used the `gam` function in the `mgcv` package instead.
- Presented below are results of GAM and GLMER analyses, with AIC differences showing a preference for the GAM method.

Table 7 Best-fit model for Role Shift

```
# glmer regression
```

```
summary(m1a <- glmer(IsRS ~ IsCVPT + IsIntonation + IsFvpt +  
                    + ChangeAnyDirection +  
(1|Speaker) +  
                    +  
(1|File), data=dat, family='binomial',  
                    +  
control=glmerControl(optimizer="bobyqa")))  
Generalized linear mixed model fit by maximum likelihood (Laplace  
Approximation) ['glmerMod']  
Family: binomial ( logit )  
Formula: IsRS ~ IsCVPT + IsIntonation + IsFvpt + ChangeAnyDirection +  
          (1 | Speaker) + (1 | File)  
Data: dat  
Control: glmerControl(optimizer = "bobyqa")
```

```
AIC      BIC    logLik deviance df.resid  
725.2    757.1   -355.6    711.2     697
```

Scaled residuals:

```
      Min       1Q   Median       3Q      Max  
-3.8846 -0.5615 -0.1379  0.5928  4.2235
```

Random effects:

```
      Groups   Name      Variance Std.Dev.  
File    (Intercept) 0.8312    0.9117  
Speaker (Intercept) 0.7310    0.8550  
Number of obs: 704, groups: File, 85; Speaker, 25
```

Fixed effects:

```
      Estimate Std. Error z value Pr(>|z|)  
(Intercept)      -2.9142    0.4243  -6.869 6.47e-12 ***  
IsCVPT             1.5308    0.2786   5.495 3.92e-08 ***  
IsIntonation       0.8189    0.2167   3.778 0.000158 ***  
IsFvpt             1.1900    0.2228   5.342 9.22e-08 ***  
ChangeAnyDirection 1.7451    0.3572   4.885 1.03e-06 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Correlation of Fixed Effects:

```

(Intr) IsCVPT IsIntn IsFvpt
IsCVPT      -0.196
IsIntonatin -0.213  0.056
IsFvpt      -0.155  0.046 -0.193
ChngAnyDrct -0.728  0.060 -0.042 -0.049

# glmer fit
somers2(fitted(m1a),dat$IsRS)
C          Dxy          n      Missing
0.8852214   0.7704428 704.0000000   0.0000000

# gam regression
summary(m1b <- gam(IsRS ~ IsCVPT + IsIntonation + IsFvpt +
                    + ChangeAnyDirection +
s(Speaker,bs='re') +
                    +
s(File,bs='re'),data=dat,family='binomial',method='ML'))

Family: binomial
Link function: logit

Formula:
IsRS ~ IsCVPT + IsIntonation + IsFvpt + ChangeAnyDirection +
s(Speaker, bs = "re") + s(File, bs = "re")

Parametric coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept)      -2.6300      0.4110  -6.399 1.56e-10 ***
IsCVPT            1.4696      0.2756   5.332 9.70e-08 ***
IsIntonation      0.7845      0.2156   3.638 0.000274 ***
IsFvpt            1.1629      0.2222   5.235 1.65e-07 ***
ChangeAnyDirection 1.5963      0.3546   4.501 6.76e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
edf Ref.df Chi.sq p-value
s(Speaker) 1.411e-04      1    0.0    0.918
s(File)    5.047e+01     84 127.6 2.11e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.417 Deviance explained = 39.6%
UBRE = -0.0059557 Scale est. = 1 n = 704

# gam fit
somers2(fitted(m1b),dat$IsRS)
C          Dxy          n      Missing
0.8914674   0.7829347 704.0000000   0.0000000

# AIC test
AIC(m1a) - AIC(m1b)
[1] 25.37005

```


Table 8 Best-fit model for Fictive Interaction

```
# glmer regression
summary(m2a <- glmer(IsFI ~ IsRS + Qhead_bare + (1|Speaker) + (1|File),
                    + data=dat,family='binomial',
                    control=glmerControl(optimizer="bobyqa")))
Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: IsFI ~ IsRS + Qhead_bare + (1 | Speaker) + (1 | File)
Data: dat
Control: glmerControl(optimizer = "bobyqa")

AIC      BIC    logLik deviance df.resid
793.4    816.1   -391.7   783.4     699

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.1732 -0.6498 -0.2610  0.6942  3.4470

Random effects:
    Groups Name          Variance Std.Dev.
File      (Intercept)  2.68      1.6372
Speaker   (Intercept)  0.31      0.5568
Number of obs: 704, groups: File, 85; Speaker, 25

Fixed effects:
    Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.3810     0.2986  -4.625 3.74e-06 ***
IsRS          0.4520     0.2188   2.066 0.03883 *
Qhead_bare    0.7093     0.2498   2.840 0.00451 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
    (Intr) IsRS
IsRS      -0.382
Qhead_bare -0.260  0.013

# glmer fit
somers2(fitted(m2a),dat$IsFI)
C          Dxy          n      Missing
0.8510973  0.7021946 704.0000000  0.0000000

# gam regression
summary(m2 <- gam(IsFI ~ IsRS + Qhead_bare + s(Speaker,bs='re') +
                  + s(File,bs='re'),
                  + data=dat,family='binomial',method='ML'))

Family: binomial
Link function: logit

Formula:
    IsFI ~ IsRS + Qhead_bare + s(Speaker, bs = "re") + s(File, bs = "re")
```

Parametric coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.4533	0.6797	-2.138	0.03249 *
IsRS	0.4287	0.2347	1.827	0.06777 .
Qhead_bare	0.7106	0.2705	2.627	0.00862 **

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

Approximate significance of smooth terms:

	edf	Ref.df	Chi.sq	p-value
s(Speaker)	1.962e-04	1	0.0	0.526
s(File)	7.042e+01	84	122.1	4.5e-05 ***

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

R-sq.(adj) = 0.289 Deviance explained = 34.3%
UBRE = 0.048654 Scale est. = 1 n = 704

gam fit

somers2(fitted(m2b),dat\$IsFI)

C	Dxy	n	Missing
0.8518637	0.7037274	704.0000000	0.0000000

AIC test

AIC(m2a) - AIC(m2b)

[1] 55.11028