The influence of handedness and pointing direction on deictic gestures and speech interaction:

Evidence from motion capture data on Polish counting-out rhymes

Katarzyna Stoltmann [1,2], *Susanne Fuchs*[1]

[1] Leibniz Centre General Linguistics (ZAS), Germany
[2]Humboldt-Universität zu Berlin, Germany
stoltmann@zas.gwz-berlin.de, fuchs@zas.gwz-berlin.de

Abstract

Does handedness influence pointing gestures? Within the scope of our study, we investigated the influence of handedness, in this case dominant vs. non-dominant hand, on pointing gestures in Polish counting-out rhymes under two conditions: fast and normal speech. For this, pointing gestures of the index finger were recorded with a motion capture system. Speech acoustics was recorded simultaneously. We hypothesized that a detailed analysis of individual gestures would reveal different kinematic patterns for the dominant versus non-dominant hand. Moreover, we expected that pointing towards the addressee would differ from pointing towards the origo (the speaker), since the pointing hand is part of origo and different muscle synergies are involved. Results of our study revealed shorter duration of pointing gestures and higher peak velocities under time pressure. Higher velocity peaks were also found for pointing with the dominant hand in comparison to the non-dominant hand, in particular while pointing to the addressee. These findings suggest that hand gestures adapt to temporal constraints, and provide first insights that handedness and pointing direction have an impact on kinematic properties of deictic gestures.

Index Terms: pointing gestures, speech interaction, handedness, motion capture, Polish counting-out rhymes

1. Introduction

Eeny, meeny, miny, moe,
Catch a tiger by the toe.
If he hollers, let him go,
Eeny, meeny, miny, moe.

The rhyme above is one of the English counting-out rhymes, but similar rhymes exist in many languages. *Counting-out rhymes* are considered performative utterances. Their target is to establish a central player, usually to eliminate a person from a group as a *loser* or a *winner* of a game (Arleo 2009: 310). They are characterized by their specific regulatory function that makes them easy to recognize cross-culturally. Already in 2009, Arleo conducted an analysis on 1884 counting-out rhymes from 51 languages investigating the ablaut reduplication in those counting-out rhymes. Usually, these games are played using words and pointing gestures. At least two persons are required for the game.

In these games, *personal deixis* in the text as well as in the pointing gestures (pointing towards the speaker vs. addressee) are crucial. Included personal deixis and pointing gestures supplemented by game rules decide the game outcome.

Deixis [greek 'pointing', 'indicating'] was introduced by Bühler (1934) in linguistics. The meaning of deixis depends on the context. Personal deixis, which is used in the text of the counting-out rhymes here, can be represented by the first, second or third personal pronouns. Personal pronouns occur in almost all counting-out rhymes. Thereby, the first (the speaker) and second pronoun (the addressee) are deictic, the third one can be anaphoric or deictic (Consten 2003: 224). In our case, the third pronoun was also deictic. In our experiment the origo was verbally and gesturally attributed to the speaker (see Fricke 2002, 2014 for more details on *gestural origo*).

The meaning of deictic gestures (hereafter pointing gestures) depends on the context as well. Deictic gestures can for instance indicate real, implied or imaginary persons, objects, or directions. Moreover, they belong to their environment or 'gesture space', including their point of origin, and they can be performed with or without speech (see MODE (2012)). In the case of spoken counting-out rhymes, deictic gestures towards the speaker or the addressee imply the *winner* or *loser*, depending on the set of rules of the rhyme.

There are a few studies analyzing pointing gestures in detail which feature a large sampling frequency as well as a precise analysis of the 3D space.

Rochet-Capellan et al. (2007) analyzed deictic gestures in coordination with speech articulation in very controlled CVCV speech material. They showed that two jaw cycles is the maximum number of cycles that could be realized within a stroke without affecting the pointing duration. They discussed their results within the Sign and Speech Frame perspective, where the Sign Frame corresponds to the arm-hand-motor control and the Speech Frame to the opening and closing of jaw. The "Rendez-vous Frame" is the frame in which both are coordinated. Pfeiffer (2011: 112) investigated the precision of gestures in situations with and without speech production. In situations with speech production, participants put more effort into speech and used fewer gestures to describe a situation. In comparison, participants' pointing gestures were more accurate when performing similar tasks without speech.

Fuchs & Reichel (2016) recorded speech-gesture coordination in German counting-out rhymes. They showed that the relation between the number of syllables and pointing gestures is rather stable under time pressure, but to some

10.21437/AVSP.2017-5

extent speaker-specific. They also indicated that a fast speech rate not only affects speech, but also leads to a shortening in pointing gesture duration. Similarly, in our previous work (Stoltmann & Fuchs, 2017) we found rather stable relations between the number of syllables and pointing gestures for Polish counting-out rhymes. Moreover, we found higher velocities and shorter durations for fast speech in comparison to normal speech and no effect of handedness. We noted however, that handedness differences might be visible in a detailed analysis, because some first inspections seemed promising.

The relation between handedness and gestures is an important topic which has been well addressed in neuroscience, speech acquisition and evolution, but less so in the linguistic domain. In bimanual gestures, the dominant hand is the active one, whereas the non-dominant hand provides passive support. For instance, when pouring a beverage, the dominant hand pours and the non-dominant holds the cup (e. g. Cochet et al. 2011, Vauclair et al. 2005). Regarding pointing gestures, Cochet (et al. 2011) emphasized the right-sided asymmetry. Vauclair et al. (2009) investigated 123 infants and toddlers regarding their preference for pointing gestures. The results of the study indicated that all participant groups (right-handed, left-handed and ambidextrous) tended to use the right hand for pointing. Cochet et al. (2012) provided further evidence for hand preferences and lateralization of speech processing in human adults. However, their results did not confirm the hypothesis that the degree of hand preference differs between pointing gestures produced with speech and those without.

We wish to expand upon previous work on hand preferences and our own global results and investigate the kinematic properties of pointing gestures in Polish counting-out rhymes. We hypothesize that the dominant hand, which is more frequently used, would be faster. Moreover, pointing gestures directed away from the origo (i.e. from the speaker to the addressee) may differ in certain kinematic properties in comparison to those pointing towards the speaker, because the hands are connected to the origo and receive direct sensorimotor feedback. Moreover, moving the arm forward or backward requires different muscle synergies.

2. Experiment

2.1. Materials and Methods

2.1.1. Participants

9 healthy Polish native speakers, 7 females and 2 males, took part in this experiment. They were Erasmus students, who have lived abroad for no longer than 5 months. All were between 21 and 27 years old (mean 24.1 years). For their participation in this experiment, participants received 10€ compensation. They had to fill out a questionnaire based on the Edinburgh Handedness Inventory (Oldfield (1971)). According to this questionnaire, all attendees were right-handed (dominant hand).

2.1.2. Materials

Data were recorded by means of a motion capture system (OptiTrack, Motive Version 1.9.0) with 12 cameras (Prime 13) and a sampling rate of 200 Hz. Speech acoustic were recorded simultaneously with 44.1 kHz. The motion capture system was calibrated before each day and reached a precision of 0.4 mm in the 3D space. Figure 1 shows the experimental scenario. The participant stands to the right, wearing a black jacket as well as a headband and gloves for motion capturing. Five markers were attached to the jacket, three on the headband and two on each glove. On the headband, one marker was placed on the front (above the nose), one on the back and one on the right lateral side. On the jacket, one marker was placed at the height of the C7 vertebra, one at the right shoulder joint, one at the left shoulder joint, one at the right elbow, and one at the left elbow. On each glove, one marker was placed on the knuckle of the index finger and one at the wrist. Additional markers were attached to the nail of the left and right index fingers. All in all, we placed a total of 14 markers on each participant. The fifteenth and final marker was placed on the teddy bear's nose and was considered as a reference point. In the analyses we describe here, we focus only on x, y, and z motions of the index finger motion.

Figure 1 visualizes in detail the experimental scenario. The participant is playing the rhyme game with the teddy bear as a fictitious person using the left hand. The distance between the teddy bear and the participant was approximately 1.5m.
The left part of the figure presents a computer screen with subject's markers. The lower figure visualizes the x, y and z time series of the dominant index finger.

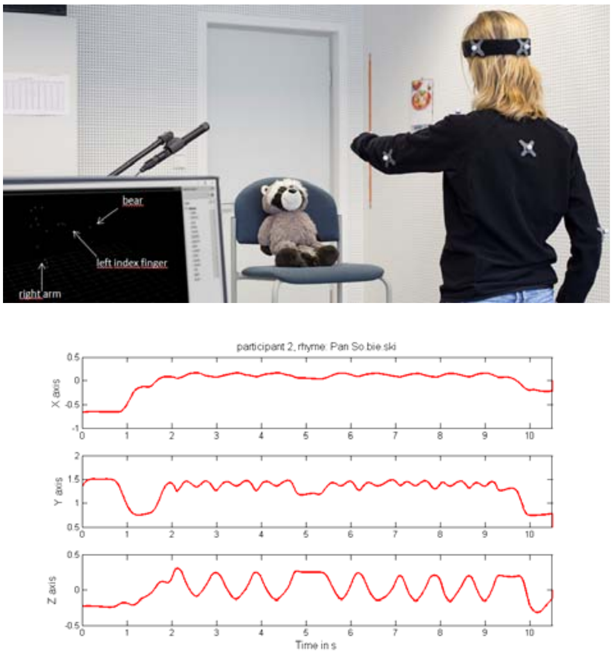

Figure 1: *Above: the experimental scenario, and below: x, y and z time series of the index finger.*

The experiment was split into three blocks:

- reading the rhymes (without pointing gestures)
- performing the rhymes naturally
    - using the dominant hand
    - using the non-dominant hand
- performing the rhymes fast
    - using the dominant hand
    - using the non-dominant hand.

The order of the blocks was the same for all participants, but the order of counting-out rhymes was randomized as well as whether they began with the dominant or non-dominant hand.

The following table presents the counting-out rhymes (1st column) in Polish orthography. We selected the counting-out rhymes from the following Polish webpage “wyliczanki”: http://wyliczanki.net/wyliczanki/Wyliczanki+do+wybierania[1]

Table 1: *Counting-out rhymes. Dots: syllable boundaries within words; Line breaks: prosodic boundaries; Number of syllables = 2nd column, Number of words =3rd column, the sum and ratio between syllables and words are displayed below the text of each rhyme.*

Orthographic representation of counting-out rhymes	No. of syllables	No. of words
Ent.li.czek – pent.li.czek,	6	2
czer.wo.ny sto.li.czek,	6	2
na ko.go wy.pa.dnie,	6	3
na te.go - bęc!	4	3
Ratio= 2	**20**	**10**
Raz, dwa, trzy,	3	3
wy.chodź ty,	3	2
jak nie ty, no to ty.	6	6
Ratio=1.1	**12**	**11**
Pan So.bie.ski miał trzy pie.ski,	8	5
czer.wo.ny, zie.lo.ny, nie.bie.ski.	9	3
Raz, dwa, trzy,	3	3
po te pie.ski i.dziesz ty.	7	5
Ratio=1.69	**27**	**16**
Bzy, bzy, bzy,	3	3
By.ły so.bie pszczół.ki trzy:	7	4
Ma.ja, Gu.cio, Kle.men.ty.na	8	3
I wy.cho.dzisz ty.	5	3
Ratio = 1.77	**23**	**13**
Wpa.dła bom.ba do piw.ni.cy,	8	4
na.pi.sa.ła na ta.bli.cy:	8	3
S. O. S. – głu.pi pies.	6	3
Tam go nie ma, a tu jest.	7	7
Ratio = 1.71	**29**	**17**
Tre.le.le.le, tre.le.le.le,	8	2
Zja.dłem dzi.siaj trzy mo.re.le.	8	4
Raz, dwa, trzy,	3	3
Dziś o.bia.du nie jesz ty!	7	5
Ratio = 1.86	**26**	**14**

2.1.3. Data pre-processing, gesture and speech annotation

The motion capture data were exported to the c3d format. Markers were renamed according to their anatomical position using the Biomechanical Toolkit (Barré & Armand, 2014). The velocity vector (v) of the x, y and z time series with a length from 1 to j was calculated as the central difference for index finger (equation 1).

v(j) = sqrt(((x(j+1)-x(j-1))/2)^2 + ((y(j+1)-y(j-1))/2)^2 + ((z(j+1)-z(j-1))/2)^2)

The velocity vector was saved in wav-file format and annotated together with the speech wav-file using Praat (version 6.0.26). In the speech wav-file, we manually labelled the on- and offset of the respective counting-out rhyme so as to have a frame of reference for the pointing gestures.

Index finger turning points were labelled as velocity minima only during the vocal production from the beginning to the end of the rhyme. Turning points which occurred before or after the rhyme were not taken into account. A stroke was defined as a movement between two successive velocity minima from the speaker pointing towards the teddy bear or back. We did not count strokes at the start and end of the counting-out rhyme, as these ones differ from all others since they start and end in a different rest positon. For each stroke, the duration, displacement and maximum velocity were calculated.

2.2. Statistical analysis

Because the dataset concerning each single stroke is unbalanced, linear mixed models were run using R (3.2.3.) [12] and the lme4 library [19]. In three models we considered the VELOCITY, AMPLITUDE and DURATION of each stroke as dependent variables, CONDITION (fast vs. normal speech), DIRECTION (speaker vs. teddy) and ARM (left vs. right) as fixed factors, and SPEAKER and RHYME as random effects as well as speaker-specific slopes. We considered t-values below and above |2| as significant.

2.3. Results

Results for all kinematic variables are displayed in Figures 2-4. Our sample consisted of n=3274 pointing gestures.

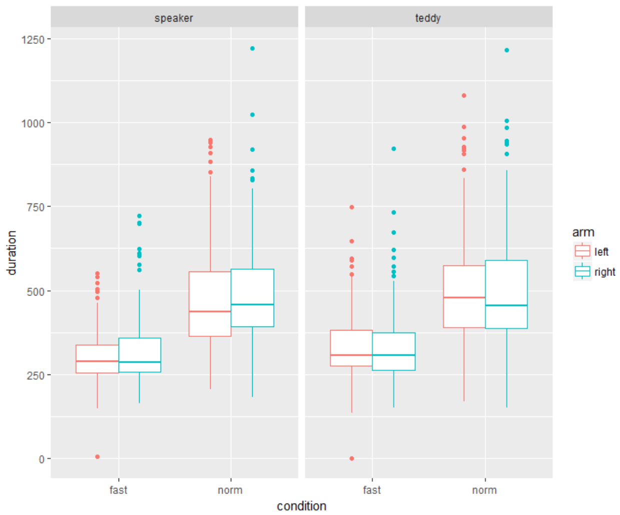

Figure 2: *Boxplots with stroke duration (in ms) for the fast and the normal condition (x-axis). Different box colours refer to the left (red) or right=dominant (green) arm. The left subplot corresponds to the pointing gestures towards the speaker themself (origo) while the right subplot corresponds to pointing gestures away from the speaker (teddy bear).*

The results reveal a clear effect of speech rate (i.e. condition) with longer gesture durations in normal than fast speech (t=-3.77), a rather weak effect for the arm with longer durations for the dominant arm in comparison to the non-dominant one, (t=-2.73), and an effect of direction (longer for pointing to the teddy bear than to the origo t=-2.86).

The peak velocities of the pointing gestures are higher for faster speech than for normal speech (t=3.8). The obtained

[1] Retrieved in April 2017.

results are comparable with the ones for duration: a higher velocity was found for the dominant right arm than for the non-dominant one (*t*=-2.8). Velocities were found to be higher when pointing towards the origo than towards the teddy bear (*t*= -3.0). Moreover, the right arm moved with a larger movement amplitude than the left arm (*t*=-2.65). Pointing towards the teddy bear was done with a larger movement amplitude in comparison to pointing toward the origo (t=-2.22), but CONDITION revealed significant results only in interaction with ARM as well as DIRECTION. Effects are generally weak and highly dependent on the random structure of the model. In a next step, further tests will be carried out.

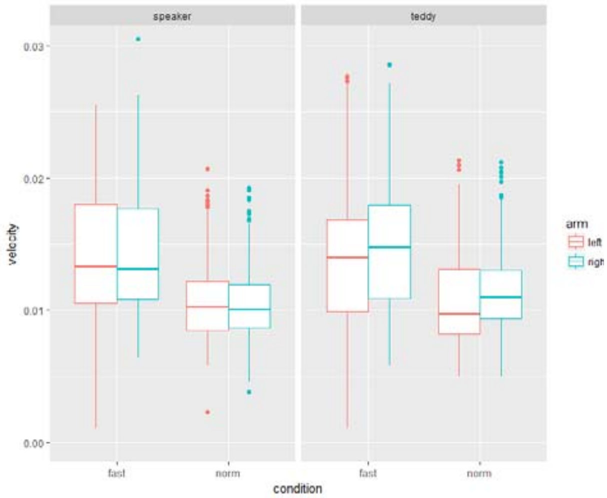

Figure 3: *Same as Figure 2, but for peak velocity.*

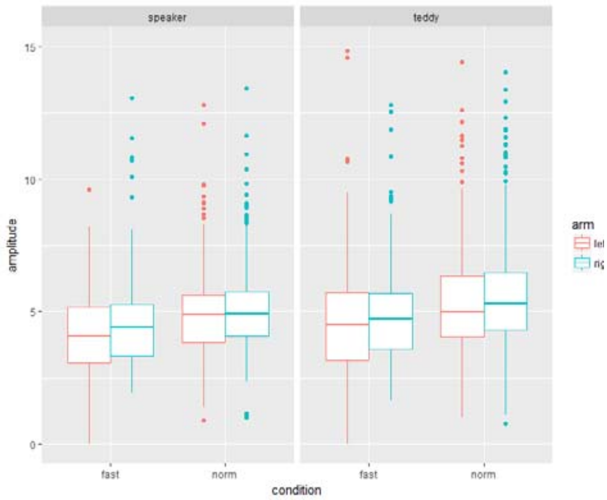

Figure 4: *Same as Figure 2, but for movement amplitude.*

2.4. Summary and discussion

Counting-out rhymes are an interesting research topic, because they occur in a variety of languages and the players combine speech production with deictic gestures. They are therefore an ecologically valid testbed for the study of speech and gesture coordination. Our study builds on previous work, but extends the analyses to a detailed investigation of each pointing gesture and its kinematic properties. The particular focus of this investigation lies in the potential differences of the dominant versus non-dominant hand. Since the dominant hand is more frequently used, and often associated with more fine-tuned control, we expected differences in the detailed analysis. Moreover, pointing direction (towards the origo, i.e. the speaker, or the addressee) might have an impact on gestural production, since the pointing hand is physically connected to the origo while it is not to the addressee. Our findings support the predictions. Speakers adapted to the different situational contexts and produced shorter and faster deictic gestures under temporal constraints. The pointing gesture for the dominant hand, in this case the right for all speakers, was slightly longer and faster. The velocity differences in handedness were particularly pronounced when speakers pointed in the direction of the addressee, but not to themselves.

In future work, we will further expand the experimental paradigm to more than two players of the game. So far our work is limited to an inanimate teddy bear which had the advantage of keeping the behavior of the addressee constant. However, in the next step, real interactions between players will be recorded.

3. Acknowledgements

This work was supported by a grant from the BMBF (Ministry of Education and Research, 01UG1411) as well as Leibniz society to both authors. We would like to thank Olivia Maky for her help in data preprocessing and annotation, Joerg Dreyer for technical support, Gediminas Schüppenhauer for the technical support in picture setting of the experimental scenario, our participants and Mark Tiede for scientific input. Thanks to Jolanda for providing her teddy bear Tom.

4. References

[1] A. Arleo, “Pif Paf Poof: Ablaut reduplication in children’s counting-out rhymes,” *in Towards a Typology of Poetic Forms: From linguistic structure to metrics and beyond*, pp. 307-323, 2009.

[2] A. Barré and S. Armand, “Biomechanical ToolKit: Open-source framework to visualize and process biomechanical data.” *Computer Methods and Programs in Biomedicine* vol. 114, no.1, pp. 80-87, 2014.

[3] P. Boersma and D.Weenink Praat: doing phonetics by computer [Computer program]. Version 6.0.26, 2017, retrieved 5 February 2017 from http://www.praat.org/

[4] K. Bühler, “Sprachtheorie. Die Darstellungsfunktion der Sprache” Jena: Verlag von Gustav, 1934.

[5] H. Cochet, M. Jover, and J. Vauclair, “Hand preference for pointing gestures and bimanual manipulation around the vocabulary spurt period”, in *Journal of experimental child psychology*, vol. 110 no. 3, pp. 393-407, 2011.

[6] H. Cochet and J. Vauclair. "Hand preferences in human adults: Non-communicative actions versus communicative gestures," *cortex*, vol. 48, no. 8, pp. 1017-1026, 2012.

[7] M. Consten, “Towards a unified model of domain-bound reference,” in *Deictic Conceptualisation of Space, Time and Person.* Amsterdam/Philadelphia: John Benjamins, pp. 223-248, 2003.

[8] E. Fricke, “Deixis, gesture, and embodiment from a linguistic point of view” in *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction (HSK 38.2).* Berlin, Boston: De Gruyter Mouton, pp. 1802–1823, 2014.

[9] S. Fuchs and U.D. Reichel, “On the relationship between pointing gestures and speech production in German counting out rhymes: Evidence from motion capture data and speech

acoustics," *in Proceedings of P&P 12,* München: LMU, 2016, pp. 1-4.

[10] R. C. Oldfield, "The assessment and analysis of handedness: The Edinburgh inventory," in *Neuropsychologia*, vol. 9, pp. 97– 113, 1971.

[11] T. Pfeiffer, "Interaction between speech and gesture: Strategies for pointing to distant objects", *Gestures in Embodied Communication and Human-Computer Interaction, 9th International Gesture Workshop, GW 2011*, E. Efthimiou and G. Kouroupetroglou, eds., Athens, Greece: National and Kapodistrian University of Athens, 2011, pp.109-112.

[12] R Core Team (2015). R: A Language and Environment for Statistical Computing [Computer program]. Version 3.2.3., retrieved October 2015 from https://www.R-project.org

[13] A. Rochet-Capellan, R. Laboissiere, A. Galvan, and J.L. Schwartz, "The speech focus position effect on jaw–finger coordination in a pointing task," *JSLHR 51.6*, pp. 1507–1521, 2008.

[14] K. Stoltmann and Fuchs S., "Syllable-pointing gesture coordination in Polish counting out rhymes: The effect of speech rate" in *Journal of Multimodal Communication Studies*, 2017, submitted.

[15] J. Vauclair, A. Meguerditchian, and W. D. Hopkins, "Hand preferences for unimanual and coordinated bimanual tasks in baboons (Papio anubis)," *Cognitive Brain Research* vol. 25, no. 1, pp. 210-216, 2005.

[16] J. Vauclair, Imbault J., "Relationship between manual preferences for object manipulation and pointing gestures in infants and toddlers." *Developmental science*, vol. 12, no. 6, pp. 1060-1069, 2009.

[17] *MODE. Glossary of multimodal terms. https://multimodalityglossary.wordpress.com/.* 2012 *Retrieved 040517.*

[18] Wyliczanki: *http://wyliczanki.net/wyliczanki/Wyliczanki+do+wybierania*

[19] D. Bates, M. Maechler, B. Bolker, B. and S. Walker. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-6. 2014. *http://CRAN.R-project.org/package=lme4*