# Joint attention and play - a data-driven summary of longitudinal, multimodal, child-parent interactions

**Petter Kallioinen and Tove Gerholm**

**TAKE HOME MESSAGE:** The ability to initiate and engage in JA is more important than time spent in JA. Strict JA-focus blur the search for interactive means *necessarily* related to, or facilitating, language acquisition.

## RESEARCH QUESTIONS

**i)** How does JA relate to later language skills in a longitudinal data set? **ii)** Can JA be identified through a data-driven approach?

## RESULTS

**i:** JA was at 12 m o a related to productive morphology at 36 m o a, and between 12-18 m o a it was weakly correlated to SCDI at 36 m o a. There was no correlation between JA and receptive vocabulary (PPVT) at 33 m o a.

**ii)** JA could not be identified through a data-driven search based on the definition of triadic JA

**iii)** A frequent cluster of behaviors, Play-w-object, was identified and had stronger correlations to later language tests than JA. Gaze-at-object alone, had (at 12 moa), stronger correlation to later vocabulary than JA. JA was more predictive of morphology.

**FURTHER QUESTIONS**: What is Play-w-object? How is it important for language acquisition (if it is)?

**Data.** *Participants and recordings*. Annotated audio-visual data for 8 interactions/dyads for 14 children (7 girls) were selected from the MINT-data (MINT-project; MAW 2001.0070; VR 2018-01135), The child-parent dyads were recorded at 9, 12, 15, 18, 21, 24, 30, 36 months of age while interacting in a lab environment.

For each child-adult dyad, the starting and finishing times of each of the following behaviors were annotated: vocalization/verbalization, gaze, gesture, facial expression/mood, and touch. The current study is based on analyses lasting 4,4 to 13,3 minutes (M = 9.71, SD = 1,40).

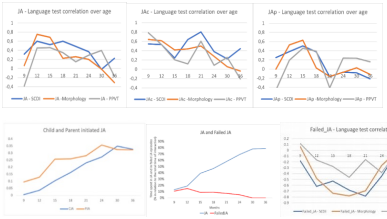
**Methodology.** *Definition JA* (Tomasello & Todd, 1983; Tomasello & Farrar, 1986). A joint attention episode is initiated when i) one of the participants initiates interaction around an object or event with the other; ii) both the child and the parent are focused on an object or event for at least 3 seconds; and iii) it is overtly expressed that the child is aware of the joint attention. JA was also divided by initiator to child-initiated JA (**JAc**) and parent-initiated JA (**JAp**).

*Failed joint attention* episodes (**Failed_JA)**, were defined as episodes where a child or parent attempted to establish joint attention but the interlocutor did not follow up, thus, *i* and sometimes *ii* above were present, but not *iii* (F. Eriksson, 2019).

Annotations in ELAN were exported and restructured into 63 subgroups of behaviors (e.g., parent's deictic gesture, etc.). Each subgroup of behavior was given a binary value per second, to reflect whether the behavior was observed during that specific one-second time-window. In the analysis, the percentage of time that each behavior occurred in relation to the total duration of the session was used to make comparisons feasible.

## ANALYSIS & RESULTS

**RQ1,** How does JA relate to later language skills in a longitudinal data set?

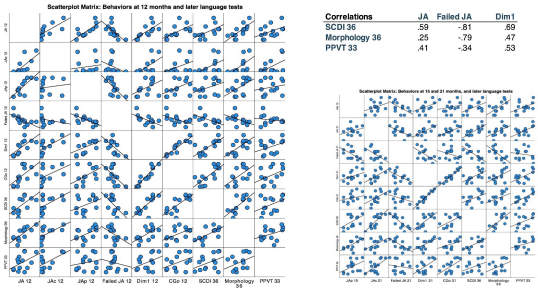


**RQ2,** Can JA be identified through a data-driven approach?

| | Dim1 | Dim2 | Dim3 | Dim4 | Dim5 | Dim5 top loading variables | Correlation |
|---|---|---|---|---|---|---|---|
| Eigenvalues | 3.4 | 1.7 | 1.6 | 1.2 | 1.1 | Child_gaze-at-object | 0.85 |
| % of variance | 25.9 | 13.3 | 12.3 | 9 | 8 | Child_action-gesture | 0.8 |
| Cumula. vari. | 25.9 | 39.2 | 51.4 | 60.4 | 68.5 | Parent_gaze-at-object | 0.77 |
| | | | | | | Parent_action-gesture | 0.67 |
| | | | | | | Parent_gaze-away | -0.67 |
| | | | | | | Child_gaze-away | -0.74 |

| Correlations | JA | Failed JA | Dim1 |
|---|---|---|---|
| JA | 1 | | |
| Failed JA | -.46 | 1 | |
| Dim1 | .28 | -.69 | 1 |

**Further investigations of Dim1.**



| Correlations | JA | Failed JA | Dim1 |
|---|---|---|---|
| SCDI 36 | .59 | -.81 | .69 |
| Morphology 36 | .25 | -.79 | .47 |
| PPVT 33 | .41 | -.34 | .53 |

**DISCUSSION**. The low number of participants precludes strong conclusions and testing. In the present study, children who are able to initiate JA early on are likely to also have better scores on later language tests. Further, the ability to perform JA is related to time spent in dyadic joint attention rather than triadic joint attention. An assumption is that some underlying ability or skill predicts the early JA-engagement and also the children's later language skills, as time spent in JA does not come through as important. What this ability is – and whether it can be addressed and traced through language studies at all – remains to be investigated.

We find a cluster of behaviors that occur repeatedly in the data (Dim1) and although this dimension shares some behaviors with JA, it is also clearly distinct, and patterned more closely with gaze behaviors and early child-initiated JA, and less with child-initiated JA at older ages, and not with mutual gaze or verbalizations. A hypothesis is that Dim1, which we have labelled play-w-object, is a Joint Engagement-behavior, potentially especially frequent in semi-naturalistic settings such as an interaction lab where parent and child are alone in a room equipped with age-adequate toys and the child has the parent's full attention regardless of interactive behaviors performed. The relation between JA and Dim1 is unclear, but a hypothesis would be that play-w-object is both a background to JA, and a behavioral cluster that – once JA is established as a means of interaction – continues to drive language acquisition as children's interests in objects, and actions on these objects, increases.