Joint Attention and Language Development: A Meta-Analysis

By

Luke Armao

(Under the Direction of Drew Abney)

Abstract

Learning to coordinate attention with a social partner is a key developmental milestone for infants that usually occurs before the first year of life. This behavior reliably emerges before the onset of language acquisition and early engagement in joint attention (JA) behaviors has been shown to positively correlate with future vocabulary development. Co-occurrence of joint attention episodes and object naming events is a possible explanation for this relationship, but the precise mechanism of action remains an open question. Many experiments have been done to explore the association between JA and language development, but different researchers have used different operationalizations of JA in their coding protocols. To systematically explore the association between JA and language development and compare results reported using different operationalizations, the current study uses a structured literature search and meta-analysis of the relevant existing literature. The full analysis features 44 articles and 283 effect sizes representing 1370 participants.

Index Words: Joint Attention, Language Acquisition, Language, Developmental Psychology, Word Learning, Meta-Analysis

Joint Attention and Language Development: A Meta-Analysis

By

Luke Armao

AB Cognitive Science and Philosophy, University of Georgia, 2022

A Thesis Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2022

Joint Attention and Language Development: A Meta-Analysis

By

Luke Armao

Major Professor: Drew Abney

Committee: Janet Frick,

Neal Outland

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2022

Table of Contents

Chapters

5. Results

6. Discussion

CHAPTER 1

INTRODUCTION

The introduction of the term “joint attention” (JA) to the literature is often attributed to Jerome Bruner and Michael Scaife, whose early research showed that infants follow changes in adult gaze direction as early as the first year of life. Their initial experiment involved repeated gaze following trials, where an experimenter first makes eye contact with the infant before shifting their gaze 90 degrees. If the infant follows the experimenter's gaze and fixates on the same point as them for an uninterrupted 7 seconds, the interaction is considered a successful joint attention episode (Scaife & Bruner, 1975). In the years since this novel study, many efforts have been made to further our understanding of joint attention and the relation to other developmental processes. However, despite shared motivation and similarities between different research programs, there has been a notable lack of agreement on the way that “joint attention” should be operationalized (Gabouer & Bortfeld, 2021). Major differences between widely-used operationalizations tend to center around differences in gaze pattern, level of engagement of social partners, and initiation of the interaction, but all share a focus on the idea of “shared” or “coordinated” attention.

Much of the literature exploring joint attention has focused on its relationship to language learning, emphasizing its importance in helping infants learn to communicate (Brooks & Meltzoff, 2005; Abney et al., 2020). Although the mechanism by which JA contributes to language development has yet to be unveiled completely, JA is recognized as a crucial tool for helping listeners determine the referential intent of the speaker (Brooks & Meltzoff, 2008; Yu et al., 2021). For example, if a parent and child are jointly attending to a toy and the parent names the object ("car!", "duck!", "ball"!), the child can link the name to the object by virtue of their awareness of their shared attention to the object. Language acquisition cannot be fully explained by this process (especially the development of more advanced constructs like the use of abstract concepts). However, the co-occurence of joint attention and naming events has been shown to be a better predictor of language development than JA episodes alone (Yu et al., 2018), and this mechanism (although not in isolation) does seem to contribute to early word learning and provide scaffolding for future language acquisition.

St. Augustine of Hippo provides a beautiful illustration of this mechanism in recounting his own early language acquisition*: "When they named any thing, and as they spoke turned towards it, I saw and remembered that they called what they would point out by the name they uttered. That they meant this thing and no other was plain from the motion of their body, the glances of the eye… thus by constantly hearing words, as they occurred in various sentences, I collected gradually for what they stood "* (Augustine of Hippo, ~400CE). Although popularly criticized by Wittgenstein as portraying a simplistic view of language and the way it represents/attributes meaning to objects (Wittgenstein, 1953), these criticisms are primarily relevant with respect to fully developed language, and do not preclude Augustine's view from being an accurate description of early word acquisition (which is ultimately used to scaffold/provide a basis for later, more advanced natural language constructs). Later language development constructs like high level syntax or abstract concepts do certainly have more varied and complex mechanisms involved in their acquisition, but Augustine's recollection does a good job of illustrating a likely mechanism connecting JA to language acquisition in the earliest stages of development.This essentially involves the pairing of audio and visual inputs for the child- the repeated overlap of a visual input (a ball, for example) and a naming event ultimately leading to the child's mental mapping of the word/sound to the concept/object.

Referential ambiguity refers to uncertainty in the subject of a sentence. For example, if someone says to "go over there," without further specification, the "there" they are referencing could mean multiple different places. The ambiguity of the reference is generally filled in with some sort of context clue, such as the place they were looking or pointing when "there" was said, or from the previous topic of conversation. The role of joint attention in resolving referential ambiguity is fairly intuitive. If a parent-child dyad is playing with two toys (a ball and a car, for example), the referent of a naming event can be deciphered by identifying the object the speaker is "paying attention" to (in most cases, one would expect that the object they are attending to is the one they would name). Despite the value that this attentional monitoring of the speaker provides, recent work has shown that naming episodes are a stronger predictor of language development when co-occurring with "sustained attention" episodes (operationalized only according to infant gaze and involving consistent, uninterrupted attention to an object) than joint attention episodes (Yu et al., 2018).

Although research on the contribution of JA to language development is valuable simply by virtue of providing an increased understanding of early word acquisition, it has also provided important insights for addressing developmental gaps in children with disorders that tend to correlate with delayed or disrupted language acquisition (Adamson et al., 2020). Infants with autism spectrum disorder are particularly well represented in the joint attention literature, and many studies have been performed to examine differences in JA and subsequent language development between Autism Spectrum Disorder (ASD) and traditionally developing (TD) populations. These studies generally find that infants diagnosed with ASD lag behind their TD peers in JA skills, and these skills are predictive of future language development for both groups (Poon et al., 2013). Interestingly, a recent meta-analysis that sought to synthesize research studying this difference found that JA was an even more robust predictor of language development for ASD populations than TD populations (Bottema-Beutel, 2016). As research continues to explore exactly how early JA skills scaffold language development, findings can be used to inform interventions that help clinical populations fill the developmental gap. For example, a future where a child can wear glasses that recognize and name objects they look at is not so far off, and might be a genuinely beneficial supplement to delayed JA engagement.

CHAPTER 2

MAJOR OPERATIONALIZATIONS

Although widely studied and used in similar experimental contexts, JA has been measured in the existing literature using multiple different operationalizations. There are minor differences (such as time requirements for sustained gaze fixations) that occur across many studies using otherwise similar operationalizations, but the existing literature is also divided by major schools of thought/research largely based on the interpretation of the social intention involved in joint attention (for a recent theoretical discussion, see Gabouer & Bortfeld, 2021). Gabouer & Bortfeld (2021) explored the operationalization of joint attention and suggested a classification of major perspectives of joint attention as “associative” or “social.” Whereas social accounts essentially posit that the developmental advantages associated with JA require an active recognition of shared attention, associative accounts disregard this recognition in operationalizing JA (the sensorimotor information that they receive while looking at an object is - in an associative count - functionally the same whether they knew their attention was shared or not).

Because of the basic disagreement on the necessary components of joint attention, different operationalizations have been used by researchers with different theoretical motivations. Unfortunately, as JA is measured in different ways, findings cannot be reconciled for a greater understanding of “joint attention” as a whole until they are systematically compared. Ultimately, the systematic comparison of different operationalizations of JA and the effect their use has on experimental findings related to language development is the primary goal of this meta-analysis.

To carry out this study, the literature used for analysis was divided into two major groups that essentially parallel (with key differences) the associative/social dichotomy. On the "social side," studies included had to use the coding schema set out in the "Early Social Communication Scales (ESCS)" (Mundy et al., 2003) for assessment of JA variables. This schema (described in the next section) is used to code a standardized ~15-25-minute interaction between a child and researcher, with social cues and initiative intention as a central criteria of JA episodes. Studies using this coding scheme were chosen for this analysis because it is well-represented in the literature on JA and language development and essentially exhaustive of those studies using an associative account of JA that could reasonably be combined in a meta-analysis. The associative account of joint attention unfortunately does not have an overarching coding schema that is shared by the majority of its representatives. To ensure the group representing the "associative" approach in comparisons between social/associative operationalization groups had a standardized criteria for inclusion, their coding breakdown for JA scoring had to break down entirely to gaze pattern. For example, specifying simply that a JA episode is when the child/parent engage in a certain gaze pattern for a set amount of time.

2.1 ESCS and IJA/RJA

One of the most well represented schemas for studying joint attention was established by the Early Social Communication Scales (ESCS) (Mundy et al., 2003). The ESCS is a test that provides measures of important nonverbal communication skills that emerge between 8-30 months. Although it provides measures for other constructs such as "behavioral requests" and "social interaction behaviors," this exploration of the ESCS will focus on its operationalization of joint attention. The ESCS is typically administered in a standardized room, with an experimenter and child across a table from one another (parent can optionally be present, and the entire interaction is filmed for later behavioral coding). Toys that have been chosen for their likelihood to elicit social behavior (wind up toys, a balloon, a rolling car, etc.) are presented, along with questions and tasks chosen to create many ideal opportunities for JA episodes and other behaviors of interest (one included question, for example, is an open ended "what toy do you want to play with?").

In the ESCS, joint attention episodes are divided according to the social partner who initiated the bid. An episode is considered "initiating joint attention" (IJA) if the child initiates it, and "responding to joint attention" (RJA) if the episode is initiated by the experiment (or parent if they are present).

2.1.2 IJA

Behaviors coded as initiating joint attention included

1.The child making eye contact with the tester while touching a mechanical toy,

2. The child alternating gaze between the tester and an active object

3. The child pointing to an object (index finger extended apart from other fingers)

and 4. The child raising a toy towards a tester's face while looking at them to "show" the object.

Behaviors 1 and 2 are considered “low level” IJA behaviors, and 3 and 4 are considered “higher level” behaviors. If the child initiates a bid to their parent who is also in the room with the tester, the behavior will be counted as well.

2.1.3 RJA

Behaviors coded as responding to joint attention include:

1. Following the tester’s pointing with their gaze (lower level)
2. Following the tester’s line of regard (higher level). Trials for following line of regard are repeated for left, right, and behind trials.

RJA is certainly a more “associative” measure than IJA, but its inclusion of the pointing gesture as a central component criteria (and the inability to separate that data from gaze following in reported data) means that it is measuring a wholly different behavior than simple gaze patterns (the idea of a “response” is also indicative of the social context).

The IJA/RJA schema for studying joint attention has been used by many researchers since the manual was released, and has been employed often in studies examining the relationship between JA and language development (Markus et al., 2001; Galeote et al., 2020). For the context of the current study, the most important difference between the ESCS (“social”) treatment of joint attention and the “associative” treatment is the inclusion of behavioral and intention-related components to the criteria for the coding of joint attention. Although this does provide important context that gaze pattern studies miss, it also introduces subjective components to the coding pipeline (as these behavioral points are manually coded). Also, conclusions that can actually be drawn from the ultimately reported correlation are much more nebulous than a study correlating one specific gaze pattern because there are multiple behaviors that contribute to RJA and IJA. For example, pointing to an object and showing a parent an object are entirely different macro level behaviors that contribute to IJA.

2.2 Gaze Pattern

A recently developed school of research on joint attention has focused primarily on the gaze patterns involved in joint attention episodes. Scaife & Bruner's original operationalization might properly be thought to belong to this group, as it is defined in terms of gaze pattern alone and without attempting to discuss awareness of attention or the nature of the initiative bid. However, it has been shown that joint attention episodes occur quite often without gaze following, being prompted by other mechanisms such as hand-eye coordination on objects being engaged with (Yu & Smith, 2013). To account for JA episodes which may have no clear initiative or responsive bid and would be ignored by schema such as the ESCS, some researchers have focused on classifying joint attention episodes strictly in terms of the gaze patterns that generate joint attention. This type of operationalization is not wholly incompatible with those that came before it, but contemporary use of eye-tracking paradigms has allowed for coding engagement this way much more productively.

To categorize and compare the gaze patterns that make up dyadic joint attention episodes, a structured analysis was recently used to uncover a mutually-exclusive set of gaze patterns that generate coordinated attention bouts (Abney et al., 2020). This study made use of eye-tracking technology in a free-play experiment to break down dyadic attention episodes into seven mutually exclusive gaze patterns. In these gaze patterns parent and child can each have one of three types of gaze: face, object, or triadic. Face and object gaze are self explanatory, and a triadic gaze pattern consists of "continuous alignment toward the same object with looks to the partner's face for less than 5 seconds" (Abney et al., 2020). The 7 gaze patterns include 1. "mutual gaze", 2. "parent face, infant object", 3. "parent object, infant face", 4."parent triadic, infant triadic", 5. "parent triadic, infant object", 6. "parent object, infant triadic", and 7."parent object, infant object".

Longitudinal vocabulary development was correlated with time of engagement in each mutually exclusive gaze pattern. Importantly, significant differences were found between the correlations, indicating that different types of gaze patterns have different relationships to vocabulary development- and some seem to contribute to a greater degree. These differences justify the structure of the gaze pattern division used for this meta-analysis.

2.3 Supported/Coordinated Engagement

Following Scaife and Bruner's paper, Roger Bakeman and Lauren Adamson investigated JA in 1983, operationalizing it as episodes of shared attention to an object or person lasting more than three seconds- a change from Scaife & Bruner's 7 second attentional episodes (Bakeman & Adamson, 1984). In addition to changing the duration required of episodes, Bakeman & Adamson's operationalization took into account the infant's awareness of their parent's engagement. To this end, they divided joint attention into two similar constructs: "Passive Joint

Attention" and "Coordinated joint attention". In passive joint attention episodes, both the parent and infant are actively engaged with the same object, but the infant does not evidence awareness of their parent's presence. In coordinated joint attention episodes, the infant and parent are again both engaged with the same object, but the child actively acknowledges the parent's attention to them (for example by looking back and forth) (Bakeman & Adamson, 1984). Although classified in Gabour's dichotomy as a social account due to its interest in the difference between supported and coordinated attention (a difference essentially accounted for by intent but often coded according to child gaze pattern), studies that utilize this basic schema for operationalization may still technically be coded completely according to gaze pattern- in these cases, effect sizes were coded according to the respective gaze pattern they represent.

CHAPTER 3

MEASURES OF CHILD VOCABULARY

The studies of interest for this meta-analysis measure the relationship between joint attention by separately collecting measures of joint attention and child vocabulary development (either concurrently or longitudinally). Having discussed the different ways JA is often measured, attention will now be given to measurements of child vocabulary development.

Although many of the studies represented in this meta-analysis used different tests to quantify child vocabulary development, almost all of them employed either a form of communicative development inventory (CDI), Mullen Scales of Early Learning (MSEL) or Reynell Developmental Language Scale (RDLS). Each of these assessments provides a quantitative measure of child expressive vocabulary (the words that they produce on their own) and receptive vocabulary (the words they understand when heard). The tests do have some differences (outlined below), but are all parent reported measures of child vocabulary. Distribution of the tests among included studies was variable, so papers were used regardless of which test they employed- as long as the measure has been shown to have a strong correlation to scores on the other tests. All of the tests discussed here and employed by studies included in the quantitative meta-analysis have been shown to have strong correlation among resulting expressive/receptive language measures (Belteki, 2022),

3.1 CDI

CDI's were used in the majority of included studies (most commonly, the MacArthur-Bates CDI). CDI's are parent reported measures of child language comprehension which have been standardized and norm referenced using large samples of initial results (Hudry, 2021). Although the precise test structure differs slightly across age groups, the basic structure of a CDI involves a vocabulary checklist with words that a parent can indicate their child (1) understands or (2) understands and says. CDI's have been in use for 30 years, the first edition CDI manual having been published in 1992 (Fenson et al., 1992). Although it is certainly worth noting and considering the bias inherent in these tests due to their being parent-reported (it is not uncommon for parents of younger children to overreport their receptive language skills, for example), the tests have been shown to have strong convergent validity with more extensive laboratory tests of vocabulary development (Fenson et al., 2007). Furthermore, given the time and cost-effectiveness of CDI's, there are relatively few comparably feasible options for research of this kind and scale.

3.2 RDLS

The Reynell Developmental Language Scales were the second most represented measure of vocabulary development among studies included in the quantitative meta analysis. The RDLS were first made available in 1969 by Joan Reynell to provide quantitative measures of child expressive and receptive vocabulary skills, much like a CDI (Reynell, 1977). The receptive vocabulary scale focuses on nouns, verbs, prepositions, and following directions; the expressive vocabulary scale focuses on syntax, vocabulary breadth, and sentence content (Simmons, 2013).The RDLS are administered by a researcher and generally take less than 30 minutes to complete. Unlike the parent-reported CDI, the RDLS make use of objects and pictures that the child is asked to engage with (for example, they might be asked to pick a specific toy out of a group or to manipulate an object in a certain way). The tactile and visually-stimulating nature of the scales make them a popular choice for use in clinical populations, but CDI's are generally a better option for children younger than 20 months (Edwards et al., 1999). Although the test administration certainly differs from CDI's, experiments employing both of these measures of expressive/receptive vocabulary show high correlation between the measures, suggesting strong convergent validity (Stallings et al, 2000a; Stallings et al., 2000b) and justifying the choice of including effect sizes from studies using both measures within the same meta-analysis.

CHAPTER 4

THE PRESENT META-ANALYSIS

Since Scaife & Bruner's 1975 introduction of JA, a large amount of research has been done on the nature of JA, with many of the studies focusing on its effects on language development. However, there have been few attempts to synthesize this literature (a task that is all the more important due to the lack of uniformity between popular operationalizations). One notable meta-analysis was done in 2016 by Kristen Bottema-Beutel (Bottema-Beutel, 2016), with a specific emphasis on comparing JA & language development relations in typically developing (TD) infants and infants with ASD. Although this was certainly a significant contribution to the literature, the present study differs in that it (1). restricts its domain of analysis to only TD groups (2). includes important additions to the literature from 2016-2021 (nascent years for the gaze-pattern paradigm), (3). includes a study of longitudinal differences in the effect of interest, and (4). focuses its analysis on the differences between JA constructs that operationalize according to gaze pattern or the IJA/RJA schema.

Regardless of the way they operationalize JA, the studies of interest use the same general experimental paradigm for assessing the relationship between joint attention and language development. This involves an initial experimental session where JA is assessed (often free-play with gaze following or an ESCS setup), followed by either a concurrent or longitudinal measure of language development that utilizes a standardized and quantifiable measure of vocabulary that shows strong convergent validity with the measures of the other experiments (mostly CDI). The result obtained by this experiment is used as a proxy measurement for the "relationship between JA and language development," and is essentially a correlation measure that seeks to find a connection between early JA skills and later language development. The infant's propensity to engage in joint attention episodes (however they be defined) is being correlated with their scores on a parent reported test of expressive and receptive vocabulary development. To frame this more explicitly: a large effect size for a given experiment would indicate that the infants in the study with high engagement in JA episodes tended to have higher scores on the followup test of vocabulary development. Studies use this correlation to assess the relationship between JA behaviors and later language development.

4.1 Search Method

The search method for this meta-analysis involved querying the digital databases PubMed and Google Scholar. Pubmed was queried on 11/12/2021 using the following search string: ("joint attention" | "coordinated attention" | "mutual gaze") AND ("word learning" | "language acquisition" | "language development" | vocabulary). 188 results were scraped by manual entry for inclusion in the initial search results database. Google Scholar was queried on 11/12/2021 using the search string ("joint attention" | "coordinated attention" | "mutual gaze") AND ("word learning" | "language acquisition" | "language development") & cdi & (concurrent | longitudinal). 1,061 articles were scraped for inclusion in the initial search database by using the "Octoparse" web data scraping software. To ensure that important contributions to the literature were not excluded because they were not represented in the database search, a forward and backward literature search was also employed. These searches focused on checking the papers which cite or are cited by the major papers from the different groups of operationalizations identified after the database search.

4.2 Inclusion criteria

Selection criteria used to assess the eligibility of articles for inclusion in the Meta-Analysis were:

(a) Includes measure of receptive or expressive infant vocabulary

(b) Includes a quantifiable measurement of level of engagement in joint attention behaviors

(c) Report of a correlation (either longitudinal or concurrent) between joint attention variable and vocabulary variable. Correlation must be zero order, and either reported as a pearson correlation or a metric that can be converted to a Pearson correlation.

(d) Operationalization of joint attention used to quantify the episodes must be able to be defined in terms of explicit gaze patterns or IJA/RJA behavior that is measured using ESCS procedures

4.3 Exclusion criteria

Studies found in the initial search were excluded from the quantitative analysis if they:

(a) Lacked “TD” participants in the experiment

(b) Did not use original experimental result (i.e. were re-reporting or using data from past experiments)

Included below is a flow diagram that documents the steps involved in the search process. The flowchart was created following the guidelines of the PRISMA 2020 statement (Page et al., 2020). PRISMA is an established standard for reporting meta analysis results that encourages consistency and transparency in reporting results.

The provided flowchart walks through the major steps of the search process, noting the number of studies excluded from consideration in each round. Where the first round is more general (focus on checking abstracts and titles), the second round involves a precise search of papers to ensure they are topically eligible and have all the data necessary for inclusion. An enumerated list of the reasons for exclusion in the second round is included as well to offer insight into the process. Any further questions about this process are welcome, and the author can be contacted for more information.

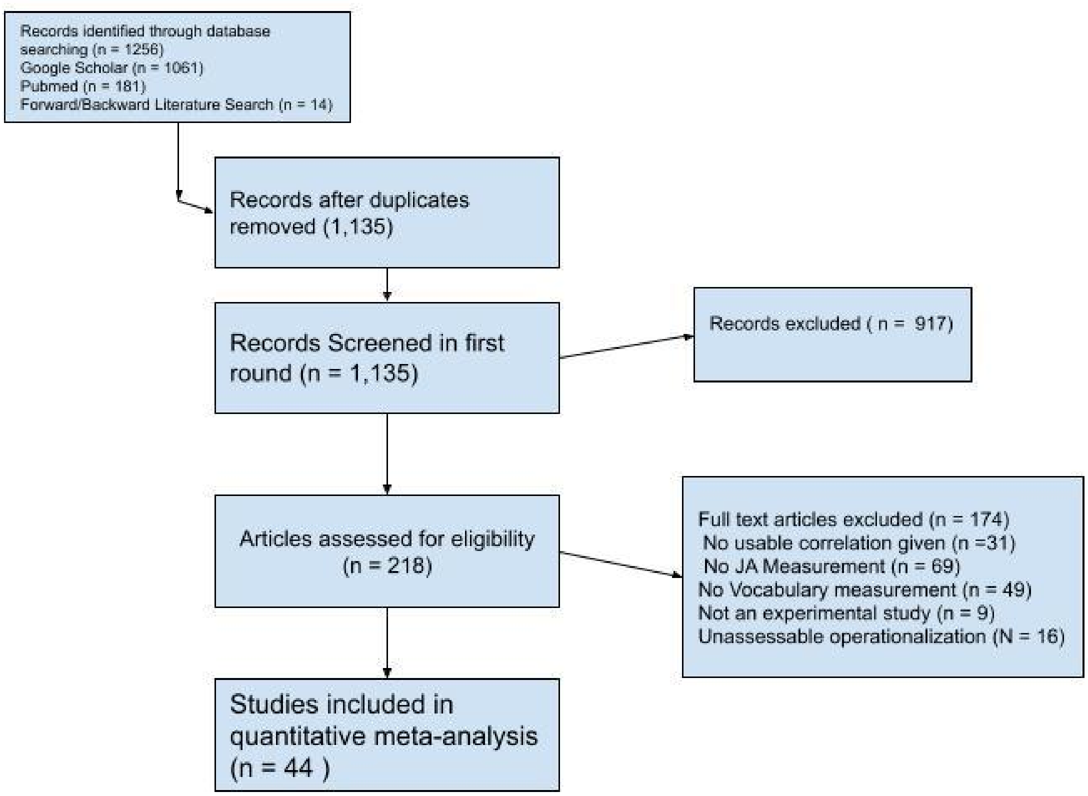

Figure 4.1: Prisma Flowchart Summarizing Search Strategy

4.4 Coding the Studies

Features collected from the included studies for analysis included the following:

1. Sample size
2. Age of JA measurement collection
3. Age of follow-up language assessment (concurrent or longitudinal)
4. JA operationalization
    - IJA/RJA: According to ESCS Schema if they explicitly mention it
    - Gaze Patterns: fitting into any of the previously mentioned gaze pattern groups. Studies were put here if the way JA was operationalized and coded/measured in the study relied primarily on gaze pattern without reference to subjective coding of intent. Did the best we could and if studies were ambiguous between two groups (such as specifying the child being triadic but the parent ambiguous between object or triadic), they were used in both groups for composite analysis,
5. Type of effect size (Spearman/Pearson/Tau)
6. Language variable type
    - Receptive
    - Expressive
7. Test used for language assessment (CDI, RDLS, etc.)
8. Value of the correlation for the association
9. Publication Status (published/unpublished)

Primary coding for the study was performed by the author, assisted by a team of undergraduate research assistants (trained by author and performing with over 90% agreement on a test subset before completing independent coding). After completing training, coders sorted the initial dataset as outlined in the PRISMA diagram, with 30% of the initial search results being sorted by two coders and used for testing reliability. By the end of the second round sort, coder's results shared 94% agreement, and all disagreement was resolved by consensus of the coding team.

Training for the extraction/coding of data from studies included in the quantitative analysis (those passing the sort rounds) was done by having coders train on a sample set of eight studies (36 effect sizes) which included studies representative of all major groups used for later analysis. Once able to complete the training set perfectly, studies were split among the coding team for data extraction. Ultimately, the data used for analysis was coded by the author, and the results obtained by the training team were used for confirmation when available and helpful.

All processing for this analysis was performed in R, primarily utilizing the dmetar and meta packages (Harrer et al., 2019; Balduzzi et al., 2019).This meta-analysis used Pearson's Product Moment Correlation (r) as its effect size of interest. To pool effect sizes between studies, correlations were converted to Fisher's z before pooling and converted back to Pearson's r for reporting (Borenstein et al., 2009).

CHAPTER 5

RESULTS

The structured literature search yielded a total of 283 effect sizes representing 44 papers and 1370 participants. Summary statistics for the papers included in the quantitative analysis are summarized in table 5.1. In this table, “Year” indicates year of publication, “N” indicates study sample size, “Effect size” indicates the number of separate effect sizes reported/included in the analysis, “conc/long” indicates concurrent, longitudinal or both, “Group” indicates whether the study operationalizes using IJA/RJA or gaze pattern (GP), and “age of JA” indicates the average age in months that the children were when the JA variable was collected (months separated by commas indicate JA was taken in each of the months). Months followed by a “-” indicate a range of more than 3 separate months where a JA variable was measured.

Table 5.1

Summary Characteristics of Studies in Analysis

Author	Year	N	Effect Sizes	Conc/ Long	Prod/ Comp	Group	Age of JA
Nagell	1996	24	70	Both	Both	GP	9-15
Saxon	1997	24	4	Long	Comp	GP	6,8
Saxon, Reilly	1998	60	2	Conc	Comp	GP	25
Morales, Mundy, Rojas	1998	20	4	Long	Prod	GP	6
Mundy, Gomes	1998	24	8	Long	Prod	RJA	6-24
Charman et al.	2000	13	4	Both	Both	GP	20
Morales, Mundy, Delgado	2000	22	16	Long	Prod	RJA	6-24
Markus et al.	2001	21	5	Both	Both	RJA	12,18
Slaughter et al.	2003	60	2	Conc	Both	GP	12
Rollins	2003	11	2	Long	Comp	GP	11
Mundy, Fox, Card	2003	32	4	Long	Prod	IJA/RJA	14,18
Namy, Nolan	2004	21	6	Both	Prod	GP	13, 8, 25
Fletcher, Perez, Hooper	2005	11	1	Conc	Comp	GP	11
Heimann et al.	2006	27	4	Conc	Both	IJA/RJA	16
Deák et al.	2007	33	2	Conc	Comp	GP	15,21
Mundy et al.	2007	52	4	Long	Both	IJA/RJA	12
Williams	2009	16	2	Conc	Prod	GP	34
Mundy, Block	2009	72	16	Long	Both	IJA/RJA	9-18
Tek	2010	18	5	Long	Comp	Both	21

Salley et al.	2011	52	1	Long	Prod	RJA	14
De Schuymer et al.	2011	60	8	Long	Both	IJA/RJA	9,14
Miller	2012	47	12	Both	Both	IJA/RJA	14,18
Vuksanovic, Bjekic	2013	25	10	Both	Both	GP	21,26,31
조윤정	2015	59	6	Long	Both	IJA/RJA	13
Brooks, Meltzoff	2015	27	1	Long	Prod	GP	11
Quinn	2016	52	4	Both	Both	GP	18.32
R Sperotto	2016	58	8	Long	Both	IJA/RJA	13,18
Cochet, Byrne	2016	14	4	Conc	Both	IJA/RJA	14
Okumura et al.	2017	37	3	Long	Prod	GP	9
AbdelAziz	2017	33	4	Long	Both	IJA/RJA	27
Parikh	2017	30	2	Long	Both	IJA	9
Edmunds et al.	2017	34	5	Both	Prod	RJA	12,15
Roemer	2018	14	2	Long	Comp	GP	12
Yu, Suanda, Smith	2018	26	2	Long	Comp	GP	9
Seager	2018	30	2	Conc	Comp	IJA/RJA	10
Seager et al.	2018	30	4	Conc	Both	IJA/RJA	10
Mason-Apps et al.	2018	32	8	Long	Both	IJA/RJA	10
Driggers-Jones	2019	88	2	Conc	Prod	GP	15
Kushner	2019	14	1	Long	Prod	GP	5
Abney et al.	2020	25	14	Long	Comp	GP	9
Galeote, Checa, Soto	2020	22	16	Long	Prod	RJA	6-24

Distribution statistics for variables used to separate groups in comparative analyses are provided in table 5.2. All groups are fairly well represented, although it is worth noting that RJA was represented by 97 effect sizes, while there were only 48 IJA effect sizes. Gaze pattern effect sizes were well represented, but due to uneven distribution and underrepresentation in some individual groups, the groups used for analysis were limited only to those that had enough samples for meaningful analysis and comparison with other groups.

Table 5.2

Effect Size Sample Distribution by Variable & Operationalization Group

Group	**N**	**Published**	**Concurrent**	**Long**	**Comp**	**Prod**
IJA	48	32	13	35	24	24
RJA	97	84	21	76	27	70
Gaze	138	127	37	101	70	68
Total	283	243	71	212	121	162

Although the compatible results and convergent validity of the CDI and RDLS tests used for vocabulary measurement were considered when designing this meta-analysis, the difference in effect size between these groups was tested to verify the relationship for our sample. A two-sample t-test of the groups using the different tests yielded a t-score of 1.23 ($p = .218$), showing that the groups are not significantly different from each other and reinforcing the decision to consider their results together for this meta-analysis.

5.1 IJA V.S. RJA

To compare IJA and RJA with respect to language development, effect sizes for all included studies that used the ESCS experimental schema were calculated and pooled according to their group. To visualize the results of this meta-analysis (as with the other groups), we use forest plots that compare the summary effect sizes of the selected groups alongside the pooled result of all the studies. Primary data provided in the forest plots includes the effect size by group (numerically, and converted back to Pearson's r once pooled) and the standard error (ultimately scaled by the sample size of the included studies). The effect size is also presented graphically, with each study's effect size plotted on a number line for parallel comparison to the other groups. Finally, a 95% confidence interval for the effect size is included, as well as a "weight %" that indicates how much each group's result contributes to the pooled total at the bottom. Also, an I^2 statistic is provided with each forest plot to assess heterogeneity in addition to any discussed Q statistics.

Past research directly comparing the two groups has found RJA to be a stronger predictor of future language than IJA, and the current study found their effect sizes to be and it is worth noting that both longitudinal groups have larger mean effect sizes than their concurrent counterpart groups. However, as their overlapping confidence intervals indicate- these differences are not significantly different from each other for either RJA (p = .17) or IJA (p = .91). Still, all longitudinal ESCS group effect sizes are significantly greater than 0 (results provided in table 5.3), providing continued evidence for the existence of a predictive relationship between early JA and later vocabulary development (at least in the ESCS group).

Results of the vocabulary type analysis are reported in figure 5.1, showing significant, positive effect sizes for all groups. Groups were relatively similar in this analysis, and RJA/IJA were found to not be significantly different from each other with respect to either comprehensive or productive vocabulary effect sizes (results in table 5.3).

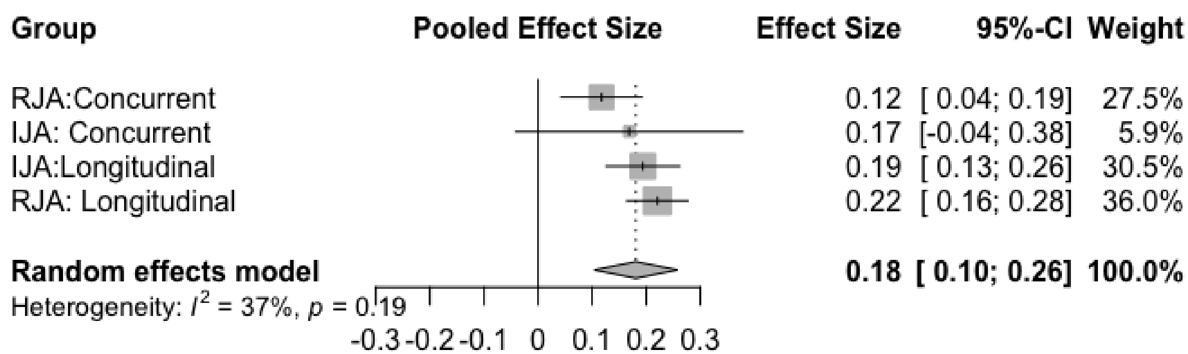

Figure 5.1: Forest Plot Comparing Pooled Concurrent & Longitudinal Effect Sizes by ESCS Group

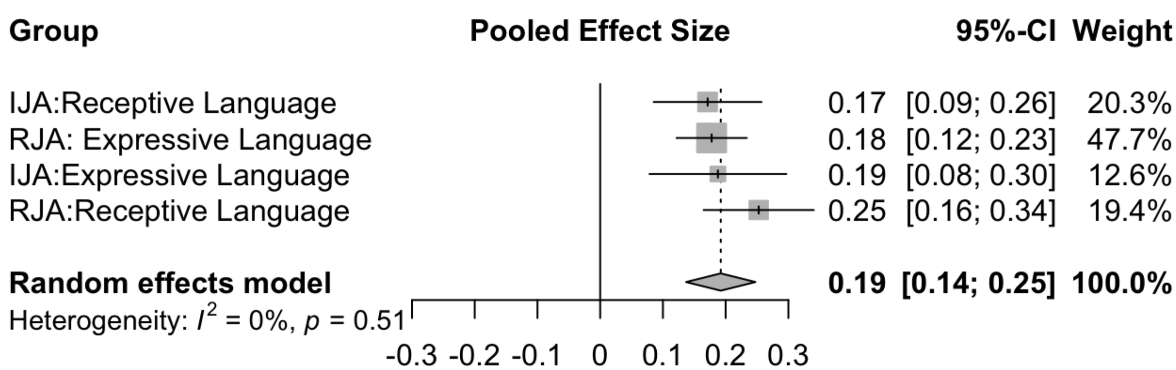

Figure 5.2: Forest Plot Comparing Pooled Productive/Comprehensive Vocabulary Effect Sizes by ESCS Group

Heterogeneity of these analyses was assessed using Cochrane's Q (Cochrane, 1950), which essentially measures heterogeneity between studies. This measure gives a quantitative description of how much the included groups differ from each other with respect to their distribution of effect sizes. This is a particularly interesting measure in a meta-analysis that is studying differences between groups, because very high heterogeneity can suggest that the groups may actually be studying different underlying effect sizes. In the context of this study, if groups have very high heterogeneity, they may be connected to different underlying mechanisms related to language development.

The random effects model used for analysis gave a Q statistic of 4.75 ($p = .19$) (between groups). Although there is some heterogeneity, it is not significant and can likely be reasonably attributed to a slight difference in longitudinal and concurrent effect sizes for this relationship, which will be explored in other analyses. Heterogeneity for the receptive/expressive vocabulary model was also very low, yielding a Q of 2.32 ($p = .51$), which suggests the underlying relationship being measured is similar between groups.

Results of significance tests for the effect sizes of ESCS groups are summarized in table 5.3. Besides concurrent IJA effect sizes, every group was significantly greater than 0 which supports past findings that this relationship does indeed exist. Also, the sample size of the IJA concurrent group was the smallest of all groups (n = 13), so it has a very high standard error and conclusions shouldn't be drawn based on this difference alone. Overall, RJA effect sizes were not significantly greater than IJA for any variable group.

Table 5.3

Results of T-Tests for Significance and Comparison of IJA and RJA Group Effect Sizes

Category	**RJA *t*-Statistic**	**IJA *t*-Statistic**	**2-Sample *t*-Statistic**
Prod. Vocabulary	5.511***	3.8741***	-0.726
Comp. Vocabulary	5.440***	3.542***	0.643
Longitudinal	6.8987***	5.9187***	0.0602
Concurrent	3.0297***	.224*	0.895
Total	7.4726***	***5.2827	-0.401

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

5.2 Gaze Pattern

For the analysis of studies that operationalize JA according to gaze pattern, some gaze patterns were excluded from analysis due to lack of representation in the literature. The gaze patterns ultimately compared in analysis included "parent triadic, child triadic" (triTri), "Parent object, child triadic" (objTri), and "parent object, child object" (objObj). Results of the meta-analysis of studies using gaze patterns are again presented in forest plots, separately comparing concurrent v.s. longitudinal studies effect sizes and productive/receptive vocabulary effect sizes.

Heterogeneity for both of the main gaze pattern analyses were rather high. The concurrent/longitudinal analysis reported a Q of 8.5(p = .13) which although not significant, is worth considering. This may, as suggested in the IJA/RJA analysis, point towards a difference in the actual effect size being reported between concurrent/longitudinal studies, and may also be affected by a difference between gaze pattern groups, an assumption that would be supported by the previously discussed findings in Abney et al., 2020. Regardless, the Q for this analysis was not actually significant, so we must rely on more study of the longitudinal and group differences to decide.

The heterogeneity of the vocabulary type analysis was more significant, coming out to a Q of 52.5 (p = .01). This is a very significant result, and suggests there may be a real difference in the effect of gaze-based JA on language development depending on the specific pattern and vocabulary type being developed. Given the size of this statistic and p-value, it is quite likely that there is a real difference, and the remainder of the analysis should pay special care to possible causes of this difference.

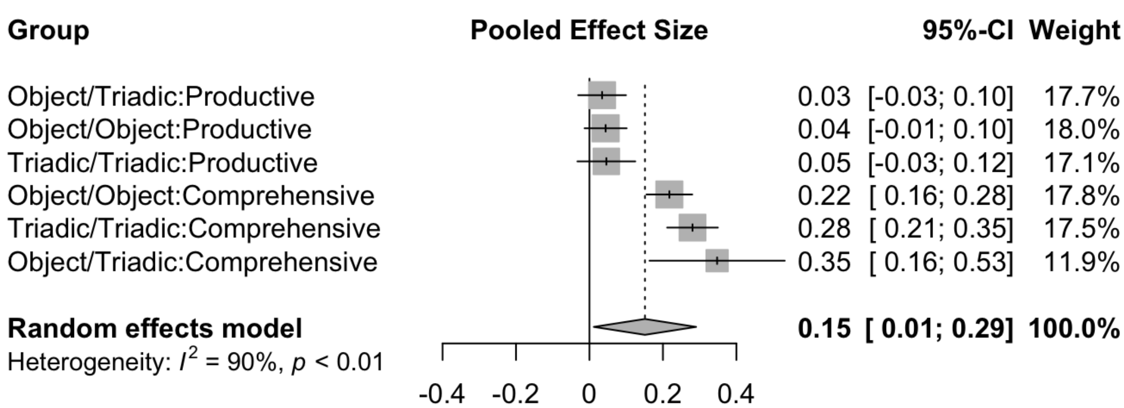

Figure 5.3: Forest Plot Comparing Pooled Productive/Comprehensive

Vocabulary Effect Sizes by Gaze Pattern Group

In the vocabulary group analysis, receptive vocabulary effect sizes were significantly greater than the expressive vocabulary effect sizes for all respective gaze pattern groups Object/Object ($t = 2.89$, $p < .01$), Object/Triadic ($t = 4.10$, $p < .01$), and Triadic/Triadic ($t = 4.63$, $p < .01$) . The gaze pattern groups did not, however, display significantly different effect sizes between categories. This was concluded after a pairwise t-test comparison of gaze pattern groups found a significant difference only between the "Object Object" group and the "Object Triadic" group ($t = 1.66$, $p = 0.099$). Interestingly, when this comparison was carried out again between the groups' productive and comprehensive vocabulary separately, the significance held for comprehensive vocabulary ($t = 1.72$, $p = 0.0903$), but the difference was not significant for productive vocabulary ($t = 0.27$, $p = 0.7842$). As these gaze patterns differ with respect to the infant's gaze, any difference in effect between patterns might theoretically mean the difference is related to the mechanism underlying JA's connection to language acquisition. Triadic gaze differs from object gaze in that it involves gaze shifts between a social partner and the object rather than sustained gaze on the object. The difference in these groups, then, essentially shows that the infant object gaze pattern is more connected to/predictive of future language development than a triadic gaze pattern. Although it is risky to draw any serious or broad conclusions about JA from just this analysis, it is worth noting that this difference runs counter to a "social" account of JA. Theoretically, a "social" account would expect the triadic gaze pattern to have a higher effect size as it allows the infant to orient themselves towards what their parent is looking at with certainty, and to be actively aware of the attention they share. However, this analysis shows that the object gaze pattern had a stronger effect. Although not necessarily compatible with a social account, an associative perspective neatly models this difference and is supported by it. In an associative model, the infant's sustained attention to a visual input is the most important factor in language acquisition (Yu et al., 2018). In theory, regardless of their recognition of the shared attention with their parent, hearing a reference (i.e. a naming episode) for an object will be most readily disambiguated and applied to where an infant is giving sustained attention to (looks back and forth to a parent reduces the quality and duration of the visual input around the naming event). Future studies on the relationship between JA and language development should focus on the difference between the child-triadic and child-object gaze patterns to more clearly understand the actual mechanism of acquisition.

Interestingly, the effect sizes for expressive vocabulary in gaze pattern groups was significantly ($p < .05$) smaller than that reported for both IJA and RJA expressive language effect sizes (RJA: $p = 0.019$, IJA: $p = 0.018$). Differences between the studies involved in these pooled effect sizes were considered, revealing that the gaze pattern effect sizes for the expressive vocabulary relationship are skewed towards younger ages and smaller age differences (the opposite is the case for expressive vocabulary). A meta regression was performed on the expressive vocabulary effect sizes to assess the moderating effect of age and longitudinal age difference, ultimately showing that longitudinal age difference was a significant positive moderator (regression coefficient β = .010**, standard error = .004, $p < 0.01$) for effect size.

The forest plot for the longitudinal/concurrent analysis (figure 5.4) seems to show clearly that longitudinal effect sizes are greater than concurrent sizes, but this is confirmed only with $p < .1$ significance for the "object object" and "triadic, triadic" groups. Significance remained at the .1 level for "object, triadic", was close to $p < .05$ ($p = .050$).

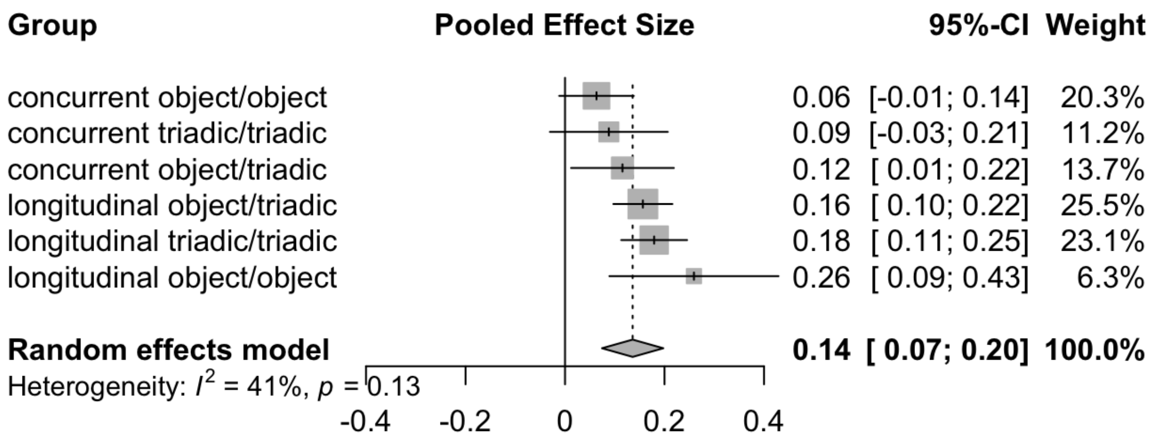

Figure 5.4: Forest Plot Comparing Pooled Longitudinal/Concurrent Effect Sizes by Gaze Pattern Group

Table 5.4

Results of T-Tests for Significance and Comparison of Gaze and ESCS Group Effect Sizes

Group	**Gaze Pattern**	**ESCS**	**2-Sample T-Test**
Productive Vocabulary	2.1151***	3.8741***	2.8066***
Comprehensive Vocabulary	6.7655***	3.5419***	0.10919
Longitudinal	6.072***	5.92***	-0.2524
Concurrent	1.8842**	1.8724*	2.9934***
Total	6.2201***	9.173***	1.6025

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

5.3 Longitudinal Relationship

To study differences in the relationship of JA on language development at different ages, multiple approaches were taken. First, longitudinal studies were separated into groups corresponding to age ranges (ranges chosen to maximize an even distribution of studies and roughly parallel average ages of important milestones in child language and motor development). Age was reflective of the age at which JA was tested in the included study, as the aim of this analysis was to study the age at which JA skills are most predictive of later language development.

The group of studies including children under eight months old had the largest effect size of the age groups (.54). Pairwise comparisons of age groups show that this group's effect size is significantly greater ($p < .01$) than all groups except the 19-23 month group ($p = .011$). Included in the <8 group are eleven effect sizes from five studies - and it is worth noting that ten of them took JA measurements at six months, and eight of them measured vocabulary over 14 months later. The analysis had a Q statistic of 247.9 ($p < .01$), indicating a high level of heterogeneity between groups. This suggests that there may indeed be different effect sizes being measured by these different groups, or raise the possibility that the effect of JA on language development is reduced or diluted by other newly developed and involved behaviors as age increases. To investigate the source of this heterogeneity, the moderating effect of age on the effect size of interest must be analyzed.

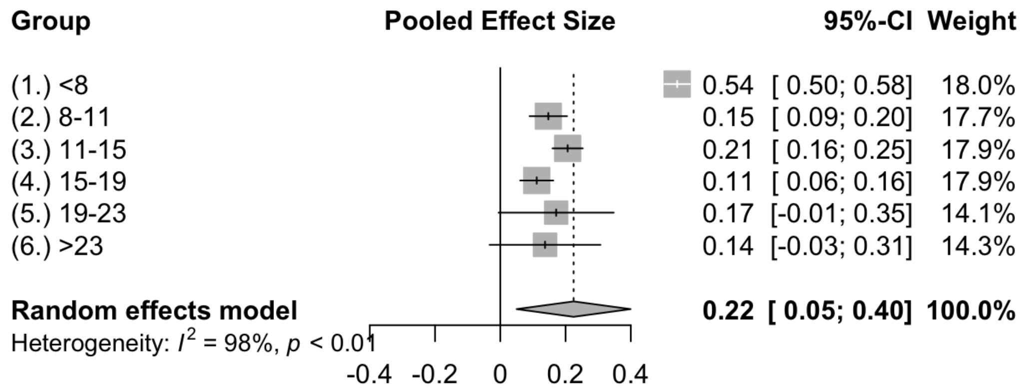

Figure 5.5: Forest Plot Comparing Pooled Reported Effect Sizes

by Age Group of JA Variable Measurement

To study the longitudinal relationship of joint attention to language development, central questions of interest are the effect of (1). The age of JA variable collection and (2). The length of the longitudinal interval between JA and vocabulary measurements on the found correlation. To visualize differences among the data, a bubble plot is used, mapping the age at which the JA variable is measured (x-axis) to the reported effect size (y-axis). In this plot, the relative size of the points corresponds to the length of the longitudinal scope of the experiment, and color corresponds to the operationalization group (yellow = gaze pattern, blue = IJA, red = RJA.

Figure 5.6 :Bubble Plot Displaying the Relationship Between Effect Size, Age at JA Variable measurement, and Experiment Variable Measurement Gap

To go a step further than visualization, a series of meta regressions were carried out to see how the longitudinal features of interest moderate our effect size of interest for each of the operationalization groups. The results of these regressions (coefficient and standard error of age at JA measurement/longitudinal experiment length when used in a regression of effect size) are summarized in Table 2. Statistically significant moderators included age of JA variable measurement in both the gaze pattern and ESCS groups, and the longitudinal length of the experiment when the groups of operationalization are considered together. The implications of these results are discussed in section 6.2.

Table 5.5

Results of Meta-Regressions on Group Effect Size Using Age Variables as a Moderator

Group	**Age Coef. β (SE)**	**Longitudinal Coef.β (SE)**
All Long. Effects	-0.005 (0.003)	0.005 (0.002)*
Gaze Pattern	-0.011 (0.040)**	0.003 (0.003)
ESCS	-0.014 (0.005)*	0.005 (0.003)

*p < 0.05, **p < 0.01, ***p < 0.001

Having established that age is a significant moderator of both of the operationalization groups but not of the combined group, an analysis was performed to examine differences in longitudinal effect sizes between the operationalization groups with respect to the age of the participants at the time of JA measurements. Four groups were created for comparison in a meta-analysis, including "old Gaze," "young Gaze," "old ESCS," and "young ESCS," where "young" indicates all effects sizes where participants were, on average, less than 13 months old, and "old" is all older group effect sizes. The "Gaze" or "ESCS" term indicates the operationalization group the effect sizes pooled for analysis are from. Results of this analysis are illustrated in figure 5.7. As this analysis' focus was essentially centered on the effect of age, and productive vocabulary is not well developed/measured until later ages, a comparison between old and young groups seems on more even grounds when compared with respect to comprehensive vocabulary. Figure 5.8 below presents this analysis, paralleling 5.7 before it.

Heterogeneity for both of these models was significant. The full vocabulary analysis gave a Q statistic of 9.35 ($p = .03$), and the comprehensive vocabulary analysis has a Q statistic of 14.36 ($p < 01$). This suggests that there could be a difference in the relationship/effect size being measured by the groups, and the increase in heterogeneity when productive vocabulary is removed suggests that much of the heterogeneity effect is coming from differences between comprehensive vocabulary measures in the group, giving further reason to explore these differences.

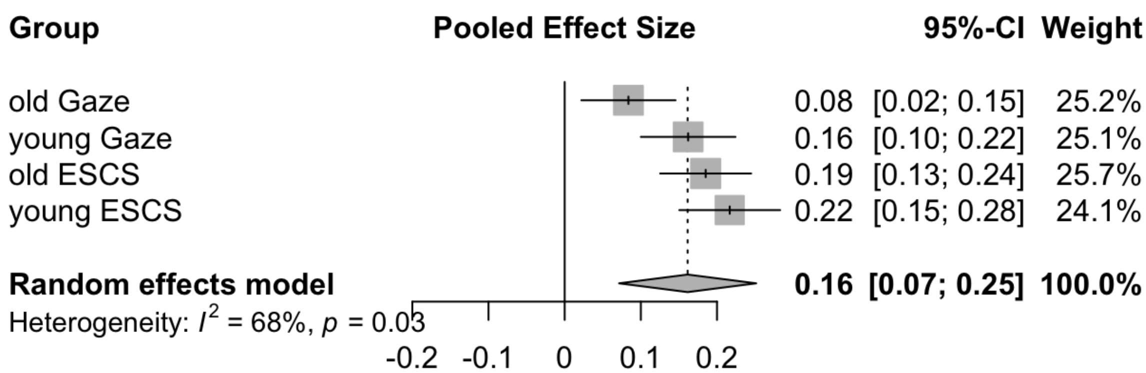

Figure 5.7: Pooled Results Comparing Old/Young Groups for all Vocabulary Types

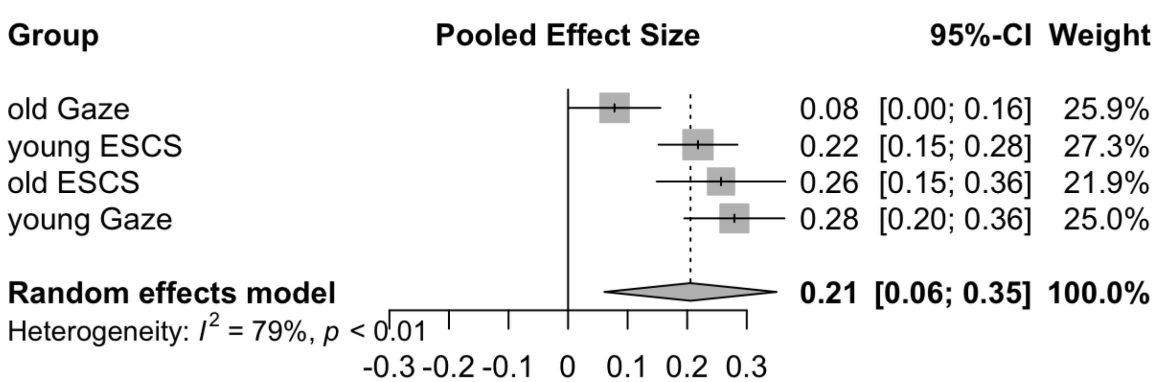

Figure 5.8: Pooled results for comprehensive vocabulary only

Once comprehensive vocabulary effect sizes were isolated and pooled, results changed in a few obvious ways, most notably greatly increasing the effect size of the young gaze pattern group. This change, alongside the young gaze effect size being significantly greater than the old gaze effect size (t = 2.71, p = 0.009), suggests that JA operationalized by gaze may be particularly related to comprehensive language development at a young age. That being said, the importance of this finding is in the difference between the age groups, not in the strength of the effect, which did not even differ significantly from the young ESCS group (t = -1.46, p = 0.148). Still, considering comparison of the old/young ESCS groups did not find any significant difference (t = -1.43, p = 0.16), the strength of the young gaze group in light of the weaker old gaze group merits further study.

As this difference seemed to arise when comprehensive vocabulary was considered on its own, differences between effect sizes for productive and comprehensive vocabulary were compared. For the gaze pattern group, comprehensive effect sizes were significantly greater than productive vocabulary (p = 0.00127), which was expected.

IJA productive vocabulary, however, was not significantly different than IJA comprehensive vocabulary (p = .983). This may perhaps be explained by the fact that IJA is coded to include more social behaviors which could be connected to productive vocabulary development. RJA still showed a significantly larger effect size for comprehensive vocabulary groups (t = 1.68, p = 0.10). As RJA is much closer to a gaze based operationalization than RJA, it makes sense that RJA would occupy the middle ground between the "social" IJA which shares a unique relationship with productive vocabulary, and "associative", gaze based JA which is more connected to comprehensive vocabulary.

5.4 Publication Bias

When carrying out a meta-analysis, it is important to assess the studies used for analysis for possible “publication bias.” Sometimes known as the “file drawer problem,” this source of this bias is essentially the fact that published studies are more likely to be statistically significant than unpublished studies (positive findings being more likely to be published) (Rosenthal, 1995).

Of the 283 effect sizes used for quantitative analysis, 40 are taken from unpublished literature. A meta regression on effect size using publication as a regressor finds that publication status is not a significant moderator of the data ($\beta = 0.032$, $SE = 0.041$). Although this does not guarantee a lack of publication bias, the inclusion of 40 unpublished effect sizes in the analysis makes it very unlikely that our results reflect publication bias if it is not a significant moderator of effect size.

Although the current study utilizes unpublished results, the possibility of publication bias existing in the published results when separated was considered as well. To do so, we made use of funnel plots to visualize bias, a standard method for assessing publication bias in meta-analyses (Page et al., 2020). These funnel plots graph studies with found effect size (using standardized mean difference as a summary effect size reflective of its Pearson correlation) on the x-axis, and standard error of a study on the y-axis. In practice, the points should generally form a symmetric funnel shape in the absence of publication bias: studies with lower standard error (larger studies) should have lower variability (close to true population effect size), pooling together at the top of the funnel. Small studies with high standard error are expected to have greater variability among effect sizes, filling out the larger bottom range of the funnel.

Three funnel plots were considered, reflecting (1) All published results, (2) ESCS, and (3) gaze pattern group effect sizes plotted separately. All plots are roughly symmetrical, and separating the operationalization groups did not significantly change results. In both groups, there is a very slight asymmetry caused by a few effect sizes that fall to the right of the funnel plot. These effects were examined individually, and it was noted that they were all studies whose JA measure was taken at 6 months of age. As the group of effect sizes including JA measured before 8 months had significantly greater effect sizes than other age groups, the minor asymmetry in the funnel plots caused by these points is more likely to be attributable to variability in effect size explained by age than publication bias.

To make a more quantitative assessment of asymmetry in the funnel plots, Egger's test (Egger, 1997) was used for each analysis. This test is essentially concerned with the intercept of the line that defines the funnel plot, reporting an intercept/β value that can be used to assess asymmetry in the plot. The results of Egger's test are not significant for any of the funnel plots representing the full ($\beta = .18$, $p = .54$) , ESCS ($\beta = .49$ $p = .27$), and Gaze ($\beta = .49$, $p = .33$) analyses. Although asymmetry is not always a perfect measure of publication bias, these results show there is definitely not significant asymmetry, and continue to add evidence that it is unlikely publication bias is affecting this meta-analysis.

To conduct a final test for the possibility of publication bias in the dataset used for this meta-analysis, Rosenthal's fail-safe number (Rosenthal, 1995) was calculated for the full dataset and found to be 4372 ($p < .01$). Rosenthal's fail safe number is calculated by finding the minimum number of additional studies with non-significant results that would be required to decrease this meta-analysis' findings to non-significance. Meta-analyses are thought to be robust if the failsafe number exceeds a critical value calculated by Rosenthal's formula of $5 * k + 10$, where k is the number of studies currently included. For the current meta-analysis of 44 studies, 4372 is certainly greater than the critical value of 450, verifying the robustness of this meta-analysis, and further reinforcing its lack of publication bias.

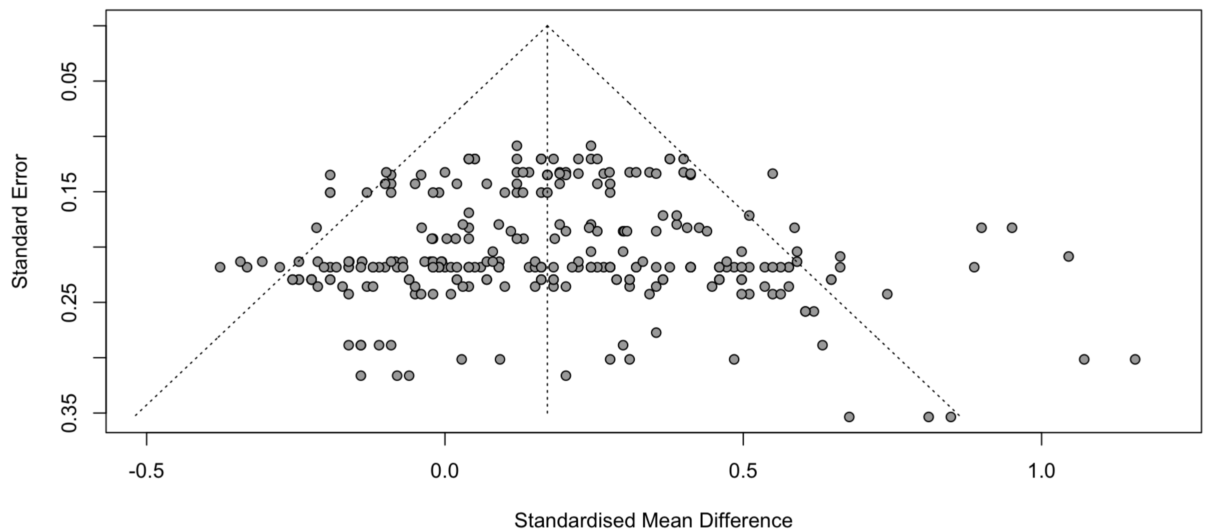

Figure 5.7: Funnel Plot Including All Effect Sizes

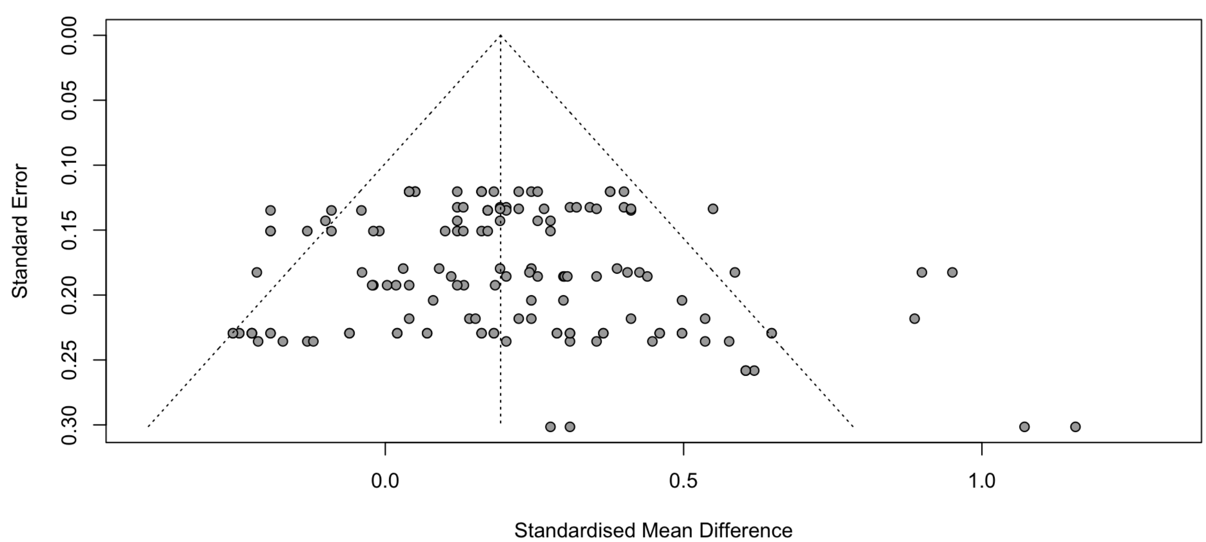

Figure 5.8: Funnel Plot Including All ESCS Group Effect Sizes

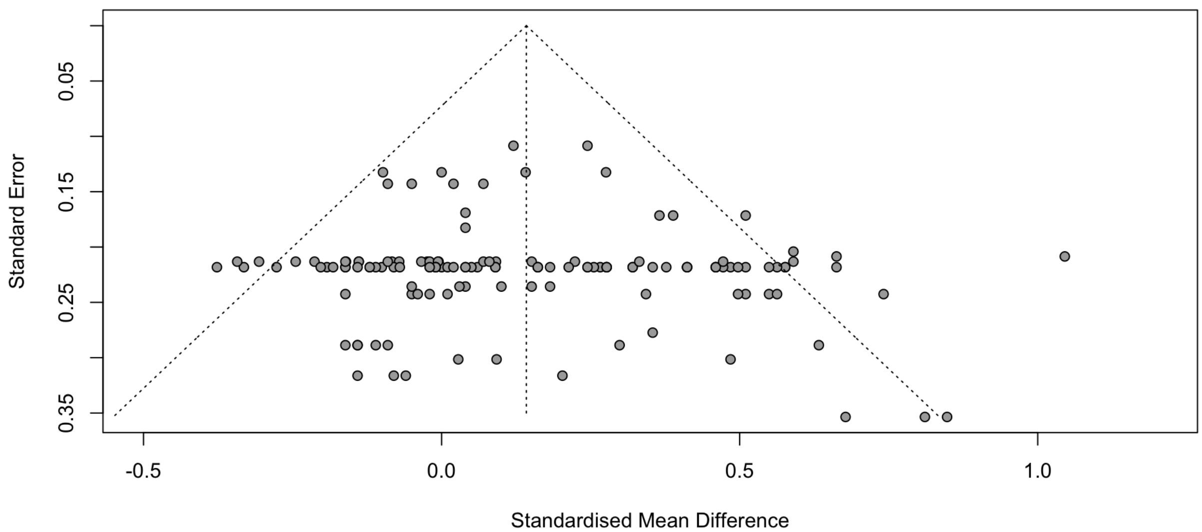

Figure 5.9: Funnel Plot Including All Gaze Pattern Group Effect Sizes

CHAPTER 6

DISCUSSION

Overall, the present meta-analytic study provides support for a weak to strong relation of early joint attention and later language development. Significant differences were found between effect sizes depending on the type of vocabulary measured and the group of JA variable operationalization. Longitudinal effect sizes were, in every compared case, greater than (not always significantly) concurrent effect sizes for otherwise identical groups. This generally supports the idea that JA scaffolds language in a developmental manner (rather than being related to language by way of a third moderator variable connected to JA and vocabulary). JA's effect on language development was also significantly moderated by age, decreasing as age increases. This relationship supports the conception of JA as an early behavior for scaffolding other social behaviors. As multiple social behaviors emerge which also contribute to the acquisition of language (i.e. drawing an example, or asking a question), JA's own effect is less dramatic.

This proposed model also explains the many results showing a significant difference in JA's effect on productive and comprehensive vocabulary development. Comprehensive vocabulary necessarily precedes productive vocabulary, so while the skill "catches up" over time, the effect of JA decreases. This chronological relationship does help make sense of the lower effect sizes for productive vocabulary in many groups, but it may also give insight into an important difference between IJA and other operationalizations. IJA is the most social and “subjective” of the schema addressed in this paper, but it does still have a comparable relationship to language as the other operationalizations, overall. However, it was unique in that it did not have a significant difference between productive and comprehensive vocabulary groups- which might suggest that IJA’s social dimension captures a dimension of behavior which is predictive of productive vocabulary and not captured by the other groups.

6.1 The Effect of Operationalization

Regardless of operationalization used, nearly all pooled effect sizes are statistically significant and positive, indicating that all operationalizations used do share a positive correlation with later language development. These results build on and agree with the results reported in Dr. Bottema-Beutel’s 2016 meta-analysis, indicating the additions to the literature in recent years have been consistent with the existing findings.

Between the ESCS groups, RJA measures were correlated more strongly with language development than IJA measures on average, but not significantly so. This is consistent with findings from studies that directly compare the groups and find RJA to be a superior predictor of language development (Sigman & McGovern, 2005; Mundy & Jarrold, 2010). Although this may indicate that the behavioral construct it measures is more advantageous than IJA for acquiring language, it may also be due related to the relatively more subjective nature of the coding schema for IJA bids, and the variability in environmental factors between studies which may influence propensity to engage in IJA (for example, included toys share categories like “wind-up” toys, but can differ in specific studies & more exciting toys in one study can bias the IJA measure with respect to studies with different toys). This is an issue for the IJA measure because they cannot be elicited in a controlled way like RJA bids (can be done a consistent number of times and ways across studies and participants).

Within the gaze pattern group, significant positive effect sizes were found for all gaze patterns considered in longitudinal analyses and with respect to receptive vocabulary. Although significant differences across gaze patterns were not identified in this meta-analysis, the general pattern makes an interesting contrast with the results on gaze pattern differences for language development correlation reported in Abney et al., (2020). Although the study reports a strong correlation between the "parent: triadic, child:object" gaze pattern reported and comprehensive language development which is supported by this meta-analysis, the correlations reported in the study for the other gaze patterns included in the quantitative meta-analysis are much smaller than the meta-analysis predicts. While this difference is open to interpretation (and within the realm of possibility given the confidence intervals involved in both studies), a likely explanation for the difference lies in the way the JA episode was motivated. Although all studies used for the gaze pattern analysis were chosen for their operationalization which is restricted to only gaze pattern criteria, the experimental manipulation involved may still affect features of the behavior which are also affected by gaze pattern differences. For example, in the case of the "parent:object, infant:object" gaze pattern, a study might code a JA episode as when the parent and child are looking at the same object for 3 seconds but involve trials that have a parent look at something to elicit a child's gaze. In contrast, the gaze patterns measured in Abney et. al.'s 2020 study were naturally occurring in free-flowing interaction. Although the gaze patterns are the same in both cases, their ontology is incompatibly different (although still agnostic to a child's recognition/intention with regard to attention and therefore still compatible with an associative perspective). Future research on JA that uses gaze pattern constructs should ideally take measurements in a context that is naturalistic to both parent and child to methodically study the differences that arise because of gaze pattern.

Ultimately, although there are important central differences between the major operationalizations, they have all been shown to be significant predictors of the effect size of interest, so the research into all groups is certainly worthwhile and justified. That being said, the scientific advantage conferred by using a gaze pattern schema is immense- primarily because it breaks down the independent variable into an objective behavior measure- which cannot be said for the subjective results of social accounts that ultimately end up going into quantitative analysis. For researchers to move any further towards knowledge of real causation rather than correlation, the systematic approach characterizing the gaze pattern operationalization is much more likely to lead to objective research progress. That being said, considering the ESCS is an important resource for use in joint attention research relating to autism, its value in research cannot be broken down simply to its ability to be objectively correlated to another variable. Rather, it should be considered a very valuable tool - just simply not for this context (one seeking a quantitative correlation measure with as little variability as possible).

6.2 The Effect of Age and Longitudinal Length of Experiment

Due to the nature and distribution of this data used for this meta-analysis (specifically with respect to many age points not represented in the continuous distribution, inconsistencies in longitudinal variable gaps for measurements at the same time point), a single pooled "longitudinal effect" is difficult to isolate (rather, different aspects of the longitudinal effect can be learned from different analyses).

Overall, when the effect sizes of the independent operationalization groups (ESCS & gaze pattern) are regressed on age of JA variable measurement, both groups are significantly moderated by age (both negatively). This suggests that for both groups of operationalizations, effect size is larger at younger ages- a finding that makes sense in context of the view that JA scaffolds later language development (Yu et al., 2018). As language and other social gestures are gradually acquired, the mechanisms for acquisition become more complicated (as well as less reliant on joint attention) and difficult to disentangle from our effect size of interest.

The longitudinal difference between the JA and vocabulary variable measurements was not a significant moderator of effect size for the independent operationalization groups, but it was a significant positive moderator for the model including all effect sizes, suggesting the longitudinal effect size tends to be larger over larger gaps of time (although the effect is small and warrants further consideration before being used to draw conclusions).

In a discussion of age's relation to our effect size of interest, it is worth noting that the different operationalizations of JA being compared differ a bit with respect to age- particularly in their lower bounds. Whereas the "social" ESCS is much less effective for use in children under the age of 8 months (basic social skills are necessary to fill criteria and to get any meaningful amount of occurrences within the 15-25 minute measurement administration. Studies of gaze pattern, meanwhile, do not (in theory) face restrictions on minimum age for use besides limitations/sensitivity of the eye-tracking hardware. Considering joint attention's effect on language development is greater at younger ages, it is imperative that a method which can explore the behavior from its earliest onset be employed for research into the correlation (especially if the goal is to ultimately explain the basic mechanism linking JA to language acquisition).

6.3 Limitations and Future Directions

To fully contextualize the results of this meta analysis, limitations inherent to the study and those studies used for the qualitative analysis should be considered. First of all, it is important to remember that the effect size of interest is Pearson's r- a correlative measure that can/should not be mistaken for a causal measure between joint attention and language development. Although it does seem intuitive that JA is directly involved in language development, the data that is collected in the general experimental paradigm cannot properly be used to deduce causation. Also relevant in a discussion of causation- there are multiple possible confounding variables which were not included in the analysis and likely explain some of the variation in results between studies (variables such as parent education, socio-economic status, race, etc. which were not collected/provided consistently enough in the studies used to include in analysis).

When comparing results from groups that use separate JA operationalizations, it is also important to consider the difference in the distribution of studies between groups. The groups of studies using gaze-pattern based operationalizations are particularly underrepresented (some groups excluded from group comparison analyses for this reason). Although pooling in a meta analysis synthesizes the results within a group (group size/distribution ultimately reflected in the associated error), more research utilizing these groups should ideally be done to create a more robust dataset for meta-analysis that gives more insight into differences across gaze patterns.

Most research focusing on JA in terms of only gaze pattern is relatively recent, as technological developments like eye-tracking systems have recently become more readily available and accurate (making coding results more efficient and objective) . Much of the category distribution issue in gaze pattern groups is due to a lack of papers that have explicitly employed these separate gaze categories, so a future meta-analysis should be performed when more experiments of this type have been carried out and reported. As this research is done, it is important that the gaze categories be meticulously coded according to a consistent and well described schema. Some of the included papers that use a gaze-pattern based operationalization give coding schemas that are ambiguous between more than one gaze pattern (i.e. says child looks at object but is not specific about parent being triadic or just object). If ambiguous but still explicitly based on gaze pattern, results like these were included in each of the groups their operationalization might fit into (an essential choice for addressing the group distribution issue - and fair because the ambiguous gaze pattern is still inarguably centrally captured by the provided operationalization.

Some other limitations to this research are due to the nature of the vocabulary measures taken. Most obviously: the vocabulary tests employed do not provide a full picture of a child's language development. As discussed, many of the values used are taken from parent- reported tests. Although the tests have been shown to have strong validity nonetheless, it is worth considering the inherent bias of these tests (bias also exists, of course, in those measures like the RDLS where an experimenter administers a test to the child in person- although valuable, the results from those ~30 minute interactions are not exhaustive). Still, any truly accurate measure of a child's language ability would be very difficult (if not impossible) to collect, so the tests do a great job, all things considered. That being said, many of the tests utilized focus mostly on metrics relating to vocabulary size (as opposed to other linguistic features such as sentence complexity). As vocabulary is certainly not the only skill involved in language acquisition, future studies should also consider the correlation between JA and the development of higher level language features (one might imagine that JA is primarily beneficial for the initial vocabulary development through naming episodes and less involved in syntactic development- more research will have to be done to know for sure).

Finally, some limitations to the results are due to the nature of the experimental paradigm. Generally done in a lab in fairly short time frames, these studies can rarely be rightfully called "naturalistic." (they take place in a lab setting whose appearance and structure may cause participants to act differently than they would in a normal setting). This is an important consideration for any experimental research in psychology, but it is especially worth considering in developmental populations. The lab setting can be unfamiliar and feel foreign to them, and things like the sheer novelty of the situation for an infant might really affect their gaze behaviors. For similar reasons, the time frame that the JA variable is collected should be carefully considered in future experiments as it is possible that results may change once the participants have gotten comfortable and used to any equipment they are wearing (like an eye-tracker).

6.4 Conclusion

The goal of this meta-analysis was to make a systematic study of the correlation between early joint attention skills and later language development, with particular attention to the comparison of effects between different common JA operationalizations. Although this correlation is widely studied, operationalizations are still not standardized. As significant differences exist among gaze patterns that can be used to measure JA, a systematic comparison of the operationalizations used in the existing literature is necessary for future research to most effectively study the phenomena they are interested in. Although a meta-analysis of the JA and language development literature was performed in 2016 (Bottema-Beutel, 2016), many important contributions to the literature (especially with respect to gaze pattern research thanks to development of eye-tracking technology) have been made. The findings of this meta-analysis parallel those of the 2016 paper with respect to IJA and RJA in TD populations, and additionally contrasts these findings with effects found by gaze pattern based operationalizations as well as adding a longitudinal analysis. Importantly, age at time of JA measurement was found to be a significant moderator of the effect size within the operationalization groups, suggesting the reported relationship decreases in strength as children get older. As the "associative" gaze pattern based analyses do not have prerequisite social skills for measurement of a JA variable, they do not face the limitations on measurements in very young children as those measurements in the ESCS do. Although both of the groups have proven to have significant & similar positive correlations with future language development, a gaze pattern based approach can provide a quantified JA variable at any age and is the suggested approach for progress to be made in the understanding of joint attention's role in language acquisition.

BIBLIOGRAPHY

AbdelAziz, A. M. (2017). *Joint attention, supported joint engagement and follow-in comments in the language acquisition process of typically developing children and children with autism: A longitudinal analysis*. [Unpublished doctoral dissertation]. University of Connecticut.

Abney, D. H., Suanda, S. H., Smith, L. B., & Yu, C. (2020). What are the building blocks of parent-infant coordinated attention in free-flowing interaction? *Infancy: The Official Journal of the International Society on Infant Studies*, *25*(6), 871–887. https://doi.org/10.1111/infa.12365

Akhtar, N., Dunham, F., & Dunham, P. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language*, *18*, 41–49. https://doi.org/10.1017/S0305000900013283

Akhtar, N., & Gernsbacher, M. A. (2007). Joint attention and vocabulary development: A critical look. *Language and Linguistics Compass*, *1*(3), 195–207. https://doi.org/10.1111/j.1749-818X.2007.00014.x

Augustine, S. (1950). The confessions of Saint Augustine (E. B. Pusey, Trans.). Peter Pauper Press. (Original work published ca. 400 CE).

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother–infant and peer–infant interaction. *Child Development, 55*(4), 1278–1289. https://doi.org/10.2307/1129997

Belteki, Z., Lumbreras, R., Fico, K., Haman, E., & Junge, C. (2022). The vocabulary of infants with an elevated likelihood and diagnosis of autism spectrum disorder: A systematic review and meta-analysis of infant language studies using the CDI and MSEL. *International Journal of Environmental Research and Public Health*, *19*, 1469. https://doi.org/10.3390/ijerph19031469

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley and Sons. https://doi.org/10.1002/9780470743386

Bottema-Beutel, K. (2016). Associations between joint attention and language in autism spectrum disorder and typical development: A systematic review and meta-regression analysis. *Autism Research*, *9*(10), 1021–1035. https://doi.org/10.1002/aur.1624

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, *35*(1), 207–220. https://doi.org/10.1017/s030500090700829x

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind.

Journal of Experimental Child Psychology, *130*, 67–78. https://doi.org/10.1016/j.jecp.2014.09.010

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, *15*(4), 481–498. https://doi.org/10.1016/S0885-2014(01)00037-5

Cochet, H., & Byrne, R. W. (2016). Communication in the second and third year of life: Relationships between nonverbal social skills and language. *Infant Behavior & Development*, *44*, 189–198. https://doi.org/10.1016/j.infbeh.2016.07.003

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, *37*(3/4), 256–266.

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, *30*(4), 352–366. https://doi.org/10.1016/j.dr.2010.10.001

de Barbaro, K., Johnson, C. M., Forster, D., & Deák, G. O. (2016). Sensorimotor decoupling contributes to triadic attention: A longitudinal investigation of mother-infant-object interactions. *Child Development*, *87*(2), 494–512. https://doi.org/10.1111/cdev.12464

Deák, G. O., Walden, T. A., Kaiser, M. Y., & Lewis, A. (2008). Driven from distraction: How infants respond to parents' attempts to elicit and re-direct their attention. *Infant Behavior & Development*, *31*(1), 34–50. https://doi.org/10.1016/j.infbeh.2007.06.004

Driggers-Jones, L. P. (2019). *A tactful conceptualization of joint attention: Joint haptic attention and language development*. [Unpublished doctoral dissertation]. East Tennessee State University.

Edmunds, S. R., Ibañez, L. V., Warren, Z., Messinger, D. S., & Stone, W. L. (2017). Longitudinal prediction of language emergence in infants at high and low risk for autism spectrum disorder. *Development and Psychopathology*, *29*(1), 319–329. https://doi.org/10.1017/S0954579416000146

Edwards, S., Garman, M., Hughes, A., Letts, C., & Sinka, I. (1999). Assessing the comprehension and production of language in young children: an account of the Reynell Developmental Language Scales III. *International Journal of Language & Communication Disorders, 34*(2), 151-171.

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997 Sep 13;315(7109):629-34. doi: 10.1136/bmj.315.7109.629

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories* (2nd ed.). Baltimore: Paul H. Brookes.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J., Bates, E., Hartung, J. P., et al. (1992). *MacArthur communicative development inventories: User's guide and technical manual* (1st ed.). Baltimore: Paul H. Brookes.

Fletcher, K., Perez, A., Hooper, C., & Claussen, A. (2005). Responsiveness and attention during picture‐book reading in 18‐month‐old to 24‐month‐old toddlers at risk.

Early Child Development and Care, *175*, 63–83. https://doi.org/10.1080/0300443042000230339

Gabouer, A., & Bortfeld, H. (2021). Revisiting how we operationalize joint attention. *Infant Behavior and Development*, *63*, 101566. https://doi.org/10.1016/j.infbeh.2021.101566

Galeote, M., Checa, E., & Soto, P. (2020). Joint attention and vocabulary development in toddlers with Down syndrome and their peers with typical development: The role of maternal interactive style. *Journal of Communicative Disorders, 84*, 105975. https://doi.org/10.1016/j.jcomdis.2020.105975

Gillespie-Lynch, K., Khalulyan, A., Del Rosario, M., McCarthy, B., Gomez, L., Sigman, M., & Hutman, T. (2015). Is early joint attention associated with school-age pragmatic language? *Autism: The International Journal of Research and Practice*, *19*(2), 168–177. https://doi.org/10.1177/1362361313515094

Harrer, M., Cuijpers, P., Furukawa, T. & Ebert, D. D. (2019). dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'. R package version 0.0.9000. URL http://dmetar.protectlab.org/.

Heimann, M., Strid, K., Smith, L., Tjus, T., Erik Ulvund, S., & Meltzoff, A. N. (2006). Exploring the relation between memory, gestural communication, and the emergence of language in infancy: A longitudinal study. *Infant and Child Development*, *15*(3), 233–249. https://doi.org/10.1002/icd.462

Hudry K. (2021). Communicative development inventories. *Encyclopedia of Autism Spectrum Disorders*, 1113-1120. https://doi.org/10.1007/978-3-319-91280-6_2009

Kushner, E. H. (2019). *An exploration of caregiver object labels to children at high vs. low risk for Autism Spectrum Disorder: The role of child factors and joint engagement context*. [Unpublished undergraduate thesis]. University of Pittsburgh.

Markus, J., Mundy, P., Morales, M., Delgado, C. E. F., & Yale, M. (2000). Individual differences in infant skills as predictors of child-caregiver joint attention and language. *Social Development*, *9*(3), 302–315. https://doi.org/10.1111/1467-9507.00127

Mason-Apps, E., Stojanovik, V., Houston-Price, C., & Buckley, S. (2018). Longitudinal predictors of early language in infants with Down syndrome: A preliminary study. *Research in Developmental Disabilities*, *81*, 37–51. https://doi.org/10.1016/j.ridd.2017.12.021

Miller, S. E. (2012). *The relationship between executive function and joint attention in the second year of life.* [Unpublished doctoral dissertation]. University of North Carolina at Greensboro.

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, *21*(2), 373–377. https://doi.org/10.1016/S0163-6383(98)90014-5

Mundy, P., Block, J., Delgado, C., Pomares, Y., Vaughan Van Hecke, A., & Parlade, M. V. (2007). Individual Differences and the Development of Joint Attention in Infancy. *Child Development*, *78*(3), 938–954. https://doi.org/10.1111/j.1467-8624.2007.01042.x

Mundy, P., Delgado, C., Goldstein, J., Parlade, M., Hogan, A., Seibert, J., & Mundy, D. (2003). *Early social communication scales (ESCS)*. *305*.

Mundy, P., Fox, N., & Card, J. (2003). EEG coherence, joint attention and language development in the second year. *Developmental Science, 6*(1), 48-54. https://doi.org/10.1111/1467-7687.00253

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, *21*(3), 469–482. https://doi.org/10.1016/S0163-6383(98)90020-0

Mundy, P., & Jarrold, W. (2010). Infant joint attention, neural networks and social cognition. *Neural Networks*, *23*(8), 985–997. https://doi.org/10.1016/j.neunet.2010.08.009

Nagell, K. M. (1996). *Joint attention and early communicative development in 9- to 15-month old infants* (Publication No. 9700305). [Doctoral dissertation, Emory University]. ProQuest Dissertations Publishing.

Okumura, Y., Kanakogi, Y., Kobayashi, T., & Itakura, S. (2017). Individual differences in object-processing explain the relationship between early gaze-following and later language development. *Cognition*, *166*, 418–424. https://doi.org/10.1016/j.cognition.2017.06.005

Page M. J., McKenzie J. E., Bossuyt P. M., Boutron I., Hoffmann T. C., Mulrow C. D., et al. (2020). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews, 10*(1), 1-11.

Page, M. J., Sterne, J. A., Higgins, J. P., and Egger, M. (2021). Investigating and dealing with publication bias and other reporting biases in meta-analyses of health research: A review. *Research Synthesis Methods*, *12*(2), 248-259.

Reynell, J. K., & NFER-Nelson Publishing Company. (1977). *Reynell developmental language scales*. Windsor, England: NFER-Nelson.

Roemer, E. J. (2018). *Exploring the role of eye contact in everyday interactions: Joint engagement, caregiver input, and language development in infant siblings of children with autism spectrum disorder* [Unpublished master's thesis]. University of Pittsburgh.

Rollins, P. R. (2003). Caregivers' contingent comments to 9-month-old infants: Relationships with later language. *Applied Psycholinguistics*, *24*(2), 221–234. https://doi.org/10.1017/S0142716403000110

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*(2), 183–192. https://doi.org/10.1037/0033-2909.118.2.183

Salley, B., Panneton, R. K., & Colombo, J. (2013). Separable attentional predictors of language outcome. *Infancy: The Official Journal of the International Society on Infant Studies*, *18*(4), 462–489. https://doi.org/10.1111/j.1532-7078.2012.00138.x

Saxon, T. F. (1997). A longitudinal study of early mother-infant interaction and later language competence. *First Language*, *17*(51), 271–281. https://doi.org/10.1177/014272379701705112

Saxon, T. F., & Reilly, J. T. (1998). Language competence and joint attention in mother‐toddler dyads. *Early Child Development and Care*, *142*(1), 33–42. https://doi.org/10.1080/0300443981420104

Seager, E. (2018). *Early intervention for infants with Down syndrome* [Unpublished doctoral dissertation]. University of Reading.

Seager, E., Mason-Apps, E., Stojanovik, V., Norbury, C., Bozicevic, L., & Murray, L. (2018). How do maternal interaction style and joint attention relate to language development in infants with Down syndrome and typically developing infants? *Research in Developmental Disabilities*, *83*, 194–205. https://doi.org/10.1016/j.ridd.2018.08.011

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attentIon in the infant. *Nature, 253*, 265-266. http://dx.doI.org/10.1038/253265a0

Simmons, E. S. (2013). Reynell developmental language scales. *Encyclopedia of Autism Spectrum Disorders*. https://doi.org/10.1007/978-1-4419-1698-3_985

Sigman, M., McGovern, C. W. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *J Autism Dev Disord 35*, 15–23. https://doi.org/10.1007/s10803-004-1027-5

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *The Journal of Genetic Psychology*, *164*(1), 54–71. https://doi.org/10.1080/00221320309597503

Stallings L. M., Gao S., & Svirsky M. A. (2000a). Assessing the language abilities of pediatric cochlear implant users across a broad range of ages and performance abilities. *The Volta Review*, *102*(4), 215–235.

Stallings, L. M., Kirk, K. I., Chin, S. B., & Gao, S. (2000b). Parent word familiarity and the language development of pediatric cochlear implant users. *Volta Review, 102*(4).

Tek, S. (2010). *A longitudinal analysis of joint attention and language development in young children with autism spectrum disorders* (Publication No. 3429221). [Doctoral dissertation, University of Connecticut]. ProQuest Dissertations Publishing.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454–1463. https://doi.org/10.2307/1130423

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*(12), 197–211. https://doi.org/10.1177/014272378300401202

Vaughan Van Hecke, A., Mundy, P. C., Acra, C. F., Block, J. J., Delgado, C. E. F., Parlade, M. V., Meyer, J. A., Neal, A. R., & Pomares, Y. B. (2007). Infant joint attention, temperament, and social competence in preschool children. *Child Development*, *78*(1), 53–69. https://doi.org/10.1111/j.1467-8624.2007.00985.x

Vuksanovic, J., & Bjekic, J. (2013). Developmental relationship between language and joint attention in late talkers. *Research in Developmental Disabilities*, *34*(8), 2360–2368. https://doi.org/10.1016/j.ridd.2013.04.017

Wittgenstein, L., & Anscombe, G. E. M. (1953). *Philosophical investigations*.

Yoon-jung, J. (2015). *The Relationship between CJA, IJA, RJA of Joint Attention and Language in Infancy* [Master's thesis, Seoul National University].

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE, 8*(11): e79659. https://doi.org/10.1371/journal.pone.0079659

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Yu, C., & Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Current Biology: CB*, *26*(9), 1235–1240. https://doi.org/10.1016/j.cub.2016.03.026

Yu, C., & Smith, L. B. (2017). Multiple sensory-motor pathways lead to coordinated visual attention. *Cognitive Science*, *41 Suppl 1*, 5–31. https://doi.org/10.1111/cogs.12366

Yu, C., Suanda, S. H., & Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science*, *22*(1), e12735. https://doi.org/10.1111/desc.12735

Yu, C., Zhang, Y., Slone, L. K., & Smith, L. B. (2021). The infant's view redefines the problem of referential uncertainty in early word learning. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(52), e2107019118. https://doi.org/10.1073/pnas.2107019118