

Fakultät für Kulturwissenschaften

Universität Paderborn

DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Philosophie (Dr. phil.)

**Enkodierungsstrategien und die Retention von Wort-Objekt-
Verknüpfungen im Fast-Mapping-Prozess von 10 Monate alten Kindern**

vorgelegt von

Eugenia Wildt

Paderborn, November 2019

Gutachterinnen:

Prof. Dr. Katharina J. Rohlfing, Universität Paderborn

Prof. Dr. Heike M. Buhl, Universität Paderborn

Inhaltsverzeichnis

1 Einleitung	1
2 Frühes Wortlernen im Zusammenhang mit Aufmerksamkeits- und Gedächtnisprozessen.....	4
2.1 Aufmerksamkeit.....	4
2.1.1 Formen der Aufmerksamkeit	4
2.1.2 Aufmerksamkeitssteuernde Mechanismen	8
2.2 Gedächtnisprozesse	11
2.2.1 Enkodierung.....	12
2.2.2 Retention und Konsolidierung	14
2.3 Wortlernen	16
2.3.1 Fast Mapping.....	18
2.3.2 Die Brücke vom Fast zum Slow Mapping.....	20
2.4 Das Referenzproblem.....	22
3 Enkodierungsstrategien im Fast-Mapping-Prozess	23
3.1 Methoden in der Spracherwerbsforschung.....	23
3.1.1 Untersuchung der Protosemantik anhand des Blickverhaltens	26
3.1.2 Preferential-Looking-Paradigma	28
3.2 Hypothesen einschränkende Prinzipien	30
3.2.1 Neuheitseffekt	32
3.2.2 Neuheitsgrad mehrerer Objekte	33
3.2.3 Perzeptuelle Aufmerksamkeitslenkung durch Top-down-Prozesse.....	34
3.2.4 Intermodal-Preferential-Looking-Paradigma.....	34
3.3 Assoziatives Wortlernen.....	36
3.3.1 Das Emergentist-Coalition-Modell	37
3.3.2 Perzeptuelle Objektsalienz.....	41
3.3.3 Perzeptuelle Aufmerksamkeitslenkung durch Bottom-up-Prozesse	42
3.3.4 Interactive-Intermodal-Preferential-Looking-Paradigma	43
3.4 Sozio-pragmatisches Wortlernen.....	45
3.4.1 Referentielles Verständnis	48
3.4.2 Loslösung von Gesichtern.....	49
3.4.3 Soziale Aufmerksamkeitslenkung.....	51
3.4.4 Social-Looking-Paradigma	53
3.5 Interaktionistisches Wortlernen	55
3.5.1 Alleinige und gemeinsame Spielsituationen	57
3.5.2 Reale Anwesenheit und Responsivität des Kommunikationspartners	60
3.5.3 Aufmerksamkeitslenkung in Interaktionen – Lernen durch Relevanz	62
3.5.4 Beobachtungsverfahren	64
3.6 Diskussion der Theorien und Forschungsmethoden.....	67
3.7 Ableitung der Fragestellung und Hypothesen für die vorliegende Studie	72

4 Untersuchung der frühen Enkodierungsstrategien – Experiment 1.....	75
4.1 Die Eye-Tracking-Methode	76
4.1.1 Messung der Fixationen und Sakkaden.....	76
4.1.2 Ablauf eines Eye-Tracking-Experiments	78
4.1.3 Möglichkeiten und Grenzen der Eye-Tracking-Methode	83
4.2 Mixed-Methods-Forschungsdesign	84
4.2.1 Studienteilnehmer	87
4.2.1.1 Rekrutierung	87
4.2.1.2 Stichprobenzusammensetzung	87
4.2.1.3 Ausschlusskriterien	89
4.2.1.4 Randomisierung der Variablen Geschlecht und Bedingung.....	92
4.2.2 Versuchsablauf.....	92
4.2.2.1 Aufwärmphase	95
4.2.2.2 Explorationsphase	95
4.2.2.3 Salienzphase.....	97
4.2.2.4 Trainingsphase	98
4.2.2.5 Testphase	101
4.2.3 Stimulusmaterial	102
4.2.3.1 Objekte.....	103
4.2.3.2 Pseudowörter.....	104
4.2.4 Kodierschema	105
4.2.4.1 Blickrichtung.....	106
4.2.4.2 Handlungsstrategien	107
4.2.4.3 Soziale Hinweise.....	108
4.2.4.4 Reliabilität der Kodierungen.....	109
4.3 Ergebnisse	111
4.3.1 Aufbereitung der Verhaltens- und Eye-Tracking-Daten	111
4.3.2 Auswahl der statistischen Testverfahren.....	112
4.3.3 Blickdauer während der Eye-Tracking-Salienzphase (Hypothese 1)	114
4.3.4 Blickdauer während der Benennungsphase (Hypothese 2)	115
4.3.5 Zusammenhang der Enkodierungsstrategien und der Retention von Wort- Objekt-Verknüpfungen (Hypothese 3)	118
4.3.5.1 Gruppenabhängiges Blickverhalten während der Spielphase.....	119
4.3.5.2 Gruppenabhängiges Blickverhalten während der Testphasen	121
4.3.5.3 Zusammenhang der Blickzeiten während der Spiel- und Testphase	127
4.3.5.4 Gruppenabhängige Handlungsstrategien während der Spielphase	129
4.3.5.5 Zusammenhang der Handlungsstrategien und der Eye-Tracking-Daten ..	130
4.3.6 Weiterführende Analysen.....	133
4.3.6.1 Soziale Hinweise während der Trainingsphase	133
4.3.6.2 Aufmerksamkeitsverlauf – Weighted Gaze Samples.....	134
4.4 Zwischendiskussion	136
4.5 Ableitung der Hypothese für ein Kontrollexperiment.....	142

5 Stabilität der Aufmerksamkeitsspanne – Kontrollexperiment 1	143
5.1 Methode	143
5.2 Ergebnisse der Eye-Tracking-Daten	145
5.2.1 Aufmerksamkeitsverlauf – Weighted Gaze Samples	145
5.2.2 Blickverhalten während der Salienzphase	146
5.2.3 Blickverhalten während der Testphasen	146
5.3 Zwischendiskussion	148
6 Konsolidierung der Wort-Objekt-Verknüpfung – Kontrollexperiment 2.....	150
6.1 Methode	151
6.2 Ergebnisse der Eye-Tracking-Daten	153
6.2.1 Blickverhalten während der Salienzphase – Tag 1	153
6.2.2 Blickverhalten während der Testphase – Tag 1	154
6.2.3 Blickverhalten während der Testphase – Tag 2	155
6.3 Zwischendiskussion	156
7 Gesamtdiskussion.....	160
7.1 Enkodierungsstrategien im frühen Alter	162
7.2 Retention und Konsolidierung der Wort-Objekt-Verknüpfung	169
7.3 Forschungsdesiderate der Studie und Ausblick auf künftige Forschung.....	172
8 Literaturverzeichnis.....	175
9 Anhang	196

Abkürzungsverzeichnis

AFC	Alternative Forced Choice
AOI	Area of Interest
DELFI	Denken, Entwickeln, Lieben, Fühlen, Individuell
ECM	Emergentist Coalition Model
EEG	Elektroenzephalogramm
ELAN	EUDICO Linguistic Annotator
ERP	Event-Related Potential
IBR	Initiating Behavior Regulation
IIPPL	Interactive-Intermodal-Preferential-Looking-Paradigma
IJA	Initiating Joint Attention
IPLP	Intermodal-Preferential-Looking-Paradigma
LENA	Language Environmental Analysis
LWL	Looking While Listening
KALPHA	Krippendorffs Alpha
MB-CDI	MacArthur–Bates Communicative Development Inventory
MTL	Medialer Temporallappen
N ₃ C	Novel-Name–Nameless-Category-Principle
PCCR	Pupil Centre Corneal Reflection
PEKIP	Prager Eltern-Kind-Programm
PLP	Preferential-Looking-Paradigma
RBR	Responding to Behavior Request
RJA	Responding to Joint Attention
ROI	Region of Interest
SLP	Social-Looking-Paradigma
SSW	Schwangerschaftswoche
WGS	Weighted Gaze Samples

Abbildungsverzeichnis

Abbildung 1. Zwei Areas of Interest aus der Studie der vorliegenden Arbeit.	79
Abbildung 2. Präsentierte Animation während der Fünf-Punkt-Kalibrierung.	80
Abbildung 3. Der Interaktionsbereich im SprachSpielLabor.	86
Abbildung 4. Der Eye-Tracking-Bereich im SprachSpielLabor.	86
Abbildung 5. Höchster erreichter Bildungsabschluss der Erziehungsberechtigten (Experiment 1)	89
Abbildung 6. Zeitachse der Salienzphase am Beispiel der Randomisierungsgruppe 1.....	97
Abbildung 7. Zeitachse der Testphase am Beispiel der Randomisierungsgruppe 1.....	102
Abbildung 8. Verwendete Stimuli: (1) Maiskolben, (2) Schachtelpuppe, (3) Schiff, (4) Klotz.....	104
Abbildung 9. Höchster erreichter Bildungsabschluss der Erziehungsberechtigten (Kontrollexperiment 1)	144
Abbildung 10. Höchster erreichter Bildungsabschluss der Erziehungsberechtigten (Kontrollexperiment 2)	152

Tabellenverzeichnis

Tabelle 1 Überblick der Blickmessungsmethoden in der Spracherwerbs- und Entwicklungsforschung.....	27
Tabelle 2 Ablauf der Experimente 1–3 der vorliegenden Studie.....	94
Tabelle 3 Die Kategorie Blickrichtung mit den jeweiligen Annotationen.....	107
Tabelle 4 Die Kategorie Handlungsstrategien mit den jeweiligen Annotationen.....	108
Tabelle 5 Die Kategorie Soziale Hinweise mit den jeweiligen Annotationen.....	109
Tabelle 6 Prozentuale Blickdauer der beiden Gruppen während der Benennungsphase im Vergleich zum Wahrscheinlichkeitswert von 20 %.....	116
Tabelle 7 Prozentuale Blickdauer während der Benennungsphase in Abhängigkeit von der Gruppenzuordnung.....	117
Tabelle 8 Prozentuale Blickdauer aller Probanden während der Benennungsphase im Vergleich zum Wahrscheinlichkeitswert von 20 %.....	117
Tabelle 9 Vergleich der prozentualen Blickdauer beider Gruppen zwischen der Benennungs- und Spielphase.....	120
Tabelle 10 Prozentuale Blickdauer während der Spielphasen in Abhängigkeit von der Gruppenzuordnung.....	121
Tabelle 11 Prozentuale Blickdauer aller Probanden zum Ablenker während der Eye- Tracking-Onset-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 %...	123
Tabelle 12 Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye- Tracking-Onset- und Offset-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 %.....	124
Tabelle 13 Prozentuale Blickdauer beider Gruppen zum Ablenker während der drei- sekündigen Eye-Tracking-Offset-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 %.....	126
Tabelle 14 Die Zeitdauer in Sekunden bis zur ersten Fixation eines der beiden Objekte während der Eye-Tracking-Offset-Testphasen.....	127
Tabelle 15 Durchschnittliche Anzahl der angewandten Handlungsstrategien während der Spielphase in Abhängigkeit von der Gruppenzuordnung.....	130
Tabelle 16 Prozentuale Blickdauer zum Ablenker während der Eye-Tracking- Testphasen in Abhängigkeit von der Subgruppen-Zuordnung.....	132
Tabelle 17 Durchschnittliche Anzahl der sozialen Hinweise während der Spielphase in Abhängigkeit von der Gruppenzuordnung.....	133
Tabelle 18 Weighted Gaze Samples aller Probanden (N = 66) in den Eye-Tracking- Tests.....	135
Tabelle 19 Weighted Gaze Samples aller Versuchsteilnehmer (N = 12) in den Eye- Tracking-Tests.....	145
Tabelle 20 Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye- Tracking-Onset-Testphasen im Vergleich zum Wahrscheinlichkeitswert von 50 %.....	147
Tabelle 21 Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye- Tracking-Offset-Testphasen im Vergleich zum Wahrscheinlichkeitswert von 50 %.....	147

Tabelle 22 Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye-Tracking-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 % (1. Erhebungstermin)	155
Tabelle 23 Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye-Tracking-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 % (2. Erhebungstermin)	156
Tabelle 24 Randomisierte Seite der Objekte [(1) Maiskolben und (2) Schachtelpuppe, (3) Schiff und (4) Klotz] im ersten und zweiten Versuch sowie die randomisierte Zuordnung der Objektnamen (Doma, Modi, Guse, Nebo) zu den Objekten in den acht Randomisierungsgruppen (Experiment 1)	197
Tabelle 25 Prozentuale Blickdauer zum Ablenker während der Offset-Testphasen im Vergleich zum Wahrscheinlichkeitswert von 50 % (Experimente 1–3).....	198

1 Einleitung

Noch bevor ein Kind seine ersten Worte produzieren kann, nimmt es an der Kommunikation mit seinen Mitmenschen aktiv teil und wird von diesen als ein Kommunikationspartner¹ behandelt. Am Anfang erfolgt der Austausch seitens des Kindes noch nicht durch die von einer Sprachgemeinschaft konventionalisierten Worte, sondern durch nonverbale Kommunikationsmittel: Indem das Kind zum Beispiel mit Blicken, Lächeln, Weinen oder Gesten auf die sprachlichen Äußerungen seiner Bezugspersonen reagiert, kann geschlussfolgert werden, ob es auf semantischer Ebene bereits Teile von Äußerungen versteht, auch, wenn es dieses Verständnis noch nicht sprachlich ausdrücken kann. Aus Studien mit Kindern vor dem ersten Lebensjahr geht hervor (z. B. Bergelson & Swingley, 2012; Nomikou, Rohlfing, Cimiano & Mandler, 2018), dass das rezeptive dem produktiven Wortwissen vorausgeht, da Kinder bereits alltägliche Wörter verstehen (vgl. Kapitel 2.3).

Im Zusammenhang mit den erwähnten Befunden zur frühkindlichen Semantik entsprang die Frage, wie der Wortlernprozess von Kindern abläuft. Zur Klärung dieser Frage galt zu untersuchen, auf welche Hinweise Kinder achten, um den Referenten eines neuen Wortes zu bestimmen. Im Hinblick auf die Lösung des Referenzproblems (vgl. Kapitel 2.4) gingen eine Reihe an Theorien hervor, die den frühen Wortlernprozess unterschiedlich erklären und im Verlauf der vorliegenden Arbeit vorgestellt werden. Im Allgemeinen unterscheiden sich die theoretischen Ansätze in den angenommenen Enkodierungsstrategien (vgl. Kapitel 3), die Kinder nutzen, um ein Wort und dessen Referenten aus der Menge an Reizeinflüssen zu identifizieren und miteinander zu verbinden.

Die vorliegende Dissertationsschrift knüpft an die bereits bestehenden Ansätze zur Lösung des Referenzproblems an und überprüft diese anhand einer neuen Studie (vgl. Kapitel 4). In dieser Studie wird jedoch nicht der gesamte Wortlernprozess, sondern der initiale Schritt, das Fast Mapping, fokussiert und mit den dabei ablaufenden Gedächtnisprozessen erklärt. Die Annahme der vorliegenden Dissertationsschrift leitet sich aus dem interaktionistischen Ansatz (vgl. Kapitel 3.5) ab, der besagt, dass 10 Monate alte Kinder solche Wort-Objekt-Verknüpfungen beibehalten, die sie im Rahmen einer wiederholenden, zielgerichte-

¹ Aus Gründen der einfacheren Lesbarkeit wird in der vorliegenden Arbeit das generische Maskulinum verwendet, aber selbstverständlich sind beide Geschlechter gleichermaßen gemeint.

ten Interaktion erfahren. Zur Überprüfung dieser Annahme werden zunächst die theoretischen Ansätze aufgegriffen und untersucht, welche Enkodierungsstrategien Kinder nutzen. Anschließend wird die Abruffähigkeit der enkodierten Wort-Objekt-Verknüpfung getestet. Im Übrigen ist das Alter von 10 Monaten für diese Arbeit insofern interessant, als Kinder in diesem Alter über ein frühes Wortverständnis besitzen, aber in der Regel noch nicht selber Wörter produzieren können, da dies etwa zwischen dem 9. und 18. Lebensmonat zu erwarten ist (Szagun, 2013, S. 74).

Zunächst wird in Kapitel 2 der Aufmerksamkeits-, der Gedächtnis- sowie der Mapping-Prozess beschrieben und miteinander in Zusammenhang gebracht. Anhand der unterschiedlichen Aufmerksamkeitsformen und -mechanismen (vgl. Kapitel 2.1) wird dargestellt, was die kindliche Aufmerksamkeit leisten muss, damit ein Reiz ins Gedächtnissystem aufgenommen wird. Daraufhin werden die Enkodierung und Konsolidierung erläutert, die zu den Gedächtnisprozessen zählen und für die Informationsaufnahme und -verarbeitung verantwortlich sind (vgl. Kapitel 2.2). Des Weiteren werden die zwei Schritte des Wortlernens, das Fast und das Slow Mapping, erläutert (vgl. Kapitel 2.3) und abschließend das Referenzproblem (vgl. Kapitel 2.4) beschrieben. Diese Konzepte bilden nicht nur die theoretische Grundlage für die nachfolgenden theoretischen Ansätze, sondern werden auch in der Studie dieser Abhandlung aufgegriffen.

Kapitel 3 eröffnet die bestehende Forschungsliteratur zum frühen Wortlernprozess, indem zunächst die in diesem Forschungsfeld angewandten klassischen Forschungsmethoden vorgestellt werden. Darauf folgen vier unterschiedliche Wortlerntheorien – die Hypothesen einschränkende Prinzipien, die assoziative, die sozio-pragmatische sowie die interaktionistische Theorie –, die anhand von bisher durchgeführten Studien exemplarisch vorgestellt werden. An die Studien anknüpfend wird auf die Aufmerksamkeitsform eingegangen, die der jeweiligen Sichtweise nach für den Wortlernprozess notwendig ist. Des Weiteren wird für die ersten drei Theorien jeweils ein Blickmessungsparadigma vorgestellt, um die unterschiedlichen Paradigmen des Blickmessungsverfahrens sowie deren Vor- und Nachteile zu verdeutlichen. Im Zusammenhang mit der interaktionistischen Wortlerntheorie wird das Beobachtungsverfahren als charakteristische Forschungsmethode beschrieben. Für die

Studie der vorliegenden Arbeit wurden das Beobachtungs- und das Eye-Tracking-Verfahren² zu einer Mixed-Methods-Methode verbunden, um Rückschlüsse auf den Zusammenhang der Enkodierungsstrategien und der Retention von Wort-Objekt-Verknüpfungen zu ziehen.

Die in Kapitel 3 beschriebenen Theorien werden hinsichtlich der daraus abgeleiteten Forschungsfrage und Hypothesen (vgl. Kapitel 3.7) in dem ersten Experiment der vorliegenden Arbeit überprüft (vgl. Kapitel 4). Zunächst wird das methodische Vorgehen beschrieben. Anschließend werden die Ergebnisse der statistischen Auswertungen vorgestellt und in einer Zwischendiskussion mit den aufgestellten Hypothesen in Zusammenhang gebracht.

In Kapitel 5 und 6 folgen zwei Kontrollexperimente, deren Bedarf sich aus den Befunden des ersten Experiments ergeben hat. In dem ersten Kontrollexperiment (vgl. Kapitel 5) wurde die Durchführung des Eye-Tracking-Verfahrens modifiziert, um die visuelle Aufmerksamkeit der Probanden länger aufrechtzuhalten. Das zweite Kontrollexperiment (vgl. Kapitel 6) geht über den Gedächtnisprozess der Enkodierung hinaus und bezieht die Konsolidierung mit ein, um die Ergebnisse der Blickdaten aller drei Experimente zu erklären. Auch die Befunde der jeweiligen Kontrollexperimente werden anschließend diskutiert (vgl. Kapitel 5.3 und 6.3).

Den Schluss der vorliegenden Dissertationsschrift bildet die Gesamtdiskussion (vgl. Kapitel 7), in der die Ergebnisse der gesamten Studie im Zusammenhang mit den Gedächtnisprozessen erklärt werden. Zunächst werden in Kapitel 7.1 die Befunde hinsichtlich der aufgestellten Forschungsfrage, welche Hinweise Kinder als frühe Enkodierungsstrategien im Fast-Mapping-Prozess hinzuziehen, diskutiert. Den Ergebnissen nach werden sowohl saliente, im Sinne der assoziativen Wortlerntheorie, als auch soziale Hinweise, im Sinne des sozio-pragmatischen Ansatzes, von Kindern im ersten Lebensjahr zur Enkodierung im Fast-Mapping-Prozess genutzt. Die Annahme, dass eine im interaktiven Lernformat enkodierte Wort-Objekt-Verbindung von Kindern rezeptiv abgerufen wird, kann jedoch nicht unterstützt werden. Dieser Befund wird in Kapitel 7.2 in Bezug auf die Retention und Konsolidierung näher diskutiert. Anschließend werden die Desiderate der vorliegenden Forschungsarbeit aufgezeigt, der Erkenntnisgewinn für die bestehende Forschung erläutert sowie weiterführende Fragestellungen und Ideen zur künftigen Forschung vorgeschlagen, um die Erkenntnisse über den Wortlernprozess von Kindern im ersten Lebensjahr zu erweitern.

² Eye Tracking ist eine Technologie, mit der Augenbewegungsdaten erfasst werden.

2 Frühes Wortlernen im Zusammenhang mit Aufmerksamkeits- und Gedächtnisprozessen

Beim Wortlernen sind Wahrnehmungsprozesse und kognitive Mechanismen beteiligt, die eine Aufnahme von Informationen ins Gedächtnissystem sowie deren kognitive Weiterverarbeitung ermöglichen. Dieses Kapitel beinhaltet das übergreifende Ziel zu beschreiben, wie neue Gedächtniseinträge entstehen, und fungiert damit als theoretische Grundlage für das nachfolgende Kapitel über Wortlernetheorien (vgl. Kapitel 3). Dabei werden bisherige Forschungsansätze zu Aufmerksamkeits-, Gedächtnis- und Mapping-Prozessen vorgestellt und zueinander in Bezug gesetzt. So lässt sich im Zuge dessen verdeutlichen, wie die verschiedenen Prozesse miteinander zusammenhängen und weshalb alle drei in Bezug auf das frühe Wortlernen berücksichtigt werden sollten.

2.1 Aufmerksamkeit

In der Psychologie wird der Begriff *Aufmerksamkeit* als visuelle oder auditive Wahrnehmung eines Stimulus und das gleichzeitige Ausblenden anderer Stimuli beschrieben, die sowohl bewusst als auch unbewusst erfolgen kann (Posner, 1980; Posner, Snyder & Davidson, 1980). Da Aufmerksamkeit notwendig ist, damit Informationen in weitere kognitive Verarbeitungsprozesse gelangen, hat sie eine große Bedeutung für das Wortlernen. Allerdings gibt es nicht nur eine, sondern verschiedene Formen der Aufmerksamkeit, die unterschiedliche Funktionen im Wortlernprozess ausführen. Welche Rollen die verschiedenen Aufmerksamkeitsformen dabei spielen, wird im nächsten Unterkapitel beleuchtet. Im darauffolgenden Unterkapitel werden die aufmerksamkeitssteuernden Mechanismen *Bottom-up* und *Top-down* beschrieben, die darüber entscheiden, was in den Aufmerksamkeitsfokus gelangt.

2.1.1 Formen der Aufmerksamkeit

Eine wesentliche Form der Aufmerksamkeit ist die *Selektion*. Unter dem Begriff *selektive Aufmerksamkeit* wird die Fähigkeit verstanden, seine Beachtung einer bestimmten Sache zu widmen, während andere Objekte oder Geschehen ausgeblendet werden (James, 1890,

S. 404). William James (1890, S. 402) verdeutlicht die Rolle der selektiven Aufmerksamkeit folgendermaßen:

Millions of items of the outward order are present to my senses which never properly enter into my experience. Why? Because they have no interest for me. My experience is what I agree to attend to. Only those items which I notice shape my mind – without selective interest, experience is an utter chaos.

Demzufolge ist das Selektieren eine wichtige Fähigkeit, da der Mensch von einer Masse an Informationen überflutet werde, die jedoch nicht alle auf einmal verarbeitet werden können. Eine Überlastung des Systems, die durch eine parallele Informationsverarbeitung entstehen würde, verhindert ein Informationsfilter, auch als *serieller Flaschenhals* (engl.: serial bottleneck) bezeichnet, der nur bestimmte Informationen in die Gedächtnisprozesse gelangen lässt (Anderson, 2007, S. 91–92). Die „Flaschenhälse“ können dabei für verschiedene Modalitäten auftreten, d. h., dass sowohl auditive als auch visuelle Informationen gefiltert werden können, bevor sie den Verarbeitungsprozessen unterliegen. Die Fähigkeit, eine selektierte Informationsauswahl zu treffen, sei für junge Kinder schwierig, nehme aber laut Lane und Pearson (1982) mit höherem Alter zu. Dies wird damit begründet, dass die „Filter“ anfangs noch nicht so weit entwickelt seien wie bei älteren Kindern (Hagen, 1967).

Für den Wortlernprozess spielt die selektive Aufmerksamkeit insofern eine wichtige Rolle, als neue Wörter und Referenten aus einer Flut einströmender Reize selektiert werden, die in die Gedächtnisprozesse eintreten sollen. Dabei können auch bestimmte auditive Reize, beispielsweise besondere Sprachlaute, oder visuelle Reize, zum Beispiel bestimmte Objekteigenschaften, näher in den Fokus rücken.

Als eine weitere Aufmerksamkeitsform ist die *obligatorische Aufmerksamkeit* aufzuführen, unter der man die Fähigkeit bzw. Reife versteht, den Fokus der Aufmerksamkeit eigenwillig zu steuern, um sich beispielsweise von einem aufmerksamkeitslenkenden Reiz zu lösen (engl.: attention disengagement) und auf Dinge in ihrem Umfeld zu achten, auf die andere Personen referieren. Diese Art der kontrollierten Aufmerksamkeit scheint sich, ebenso wie die selektive Aufmerksamkeit, mit dem Alter allmählich zu entwickeln. Yurovsky und Frank (2017) zeigten in ihrer Studie, dass es jüngeren Kindern im Vergleich zu älteren Kindern schwerfällt, ihre Aufmerksamkeit von einem Gesicht des Sprechers zu lösen, um der Blickrichtung dieser Person zum Zielobjekt zu folgen. Doch mit zunehmendem Alter lernen sie nicht nur, sich von Gesichtern loszulösen und referierende Objekte zu fokussieren, sondern werden auch schneller darin (Frick, Colombo & Saxon, 1999). Aktuelle Befunde zeigen (Niedźwiecka, Ramotowska & Tomalski, 2017), dass eine wechselseitige,

visuelle Aufmerksamkeit (engl.: mutual gaze) in frühen Interaktionen zur Entwicklung der kindlichen Aufmerksamkeitskontrolle beiträgt. Für den Spracherwerb ist die obligatorische Aufmerksamkeit insofern wichtig, als Kinder in der Lage sind, sich in Wortlernsituationen nicht von aufmerksamkeitslenkenden, irrelevanten Reizen ablenken zu lassen.

Erlangen Kinder zusätzlich die Fähigkeit, einen Referenten mithilfe von Hinweisen seitens Erwachsenen zu selektieren, wird diese Form der Aufmerksamkeit als *geteilte Aufmerksamkeit* (engl.: joint attention) bezeichnet (Bremner & Slater, 2008, S. 317–355). Anders ausgedrückt, wird die ursprüngliche *Dyade* – eine Zweierbeziehung zwischen Mensch und Mensch oder Mensch und Objekt – durch einen dritten Referenten zu einer *Triade* (Bakeman & Adamson, 1984) ergänzt. Das dabei resultierende, referentielle Dreieck (Müller, Pradeu & Schäfer, 2004, S. 217–237) stellt die Koordination in dieser Konstellation dar. Somit ist das höchste Level der Aufmerksamkeitskoordination die triadische Interaktion, in der Kinder nicht nur in einer dyadischen, wechselseitigen Kommunikation stehen, sondern auch *über* andere Dinge und Handlungen *mit* einem Gesprächspartner kommunizieren. In der Forschungsliteratur (Tomasello, 1995; Tomasello, Carpenter, Call, Behne & Moll, 2005) wird berichtet, dass Kinder zwischen dem 9. und 12. Lebensmonat ihre Aufmerksamkeit zwischen Objekten und Gesprächspartnern teilen können.

Mundy et al. (2007) untersuchten responsive und initiierte Verhaltensformen von Erwachsenen und Kindern im Alter von 9 bis 18 Monaten hinsichtlich der Frage, wie sie die geteilte Aufmerksamkeit koordinieren. Daraus ergaben sich vier entwicklungsbedingte Stufen der Aufmerksamkeitskoordination, zu denen beispielsweise RJA (Responding to Joint Attention) zählt, worunter die Forscher eine Reaktion des Kindes auf die sozialen Achtungshinweise von Erwachsenen, wie zum Beispiel Gesten oder die Blickrichtung verstehen. Eine weitere Form ist IJA (Initiating Joint Attention), bei der Kinder selbst die Initiative ergreifen und durch Gesten oder ihre Blickrichtung den Kommunikationspartner auf etwas aufmerksam machen. IBR (Initiating Behavior Regulation) wird ebenfalls mit Blicken oder Gesten kommuniziert, um eine Hilfestellung des Gegenübers zu erhalten. RBR (Responding to Behavior Request) ist die Reaktion, in Form einer Hilfestellung, die auf solche Signale folgt.

Der geteilte Aufmerksamkeitsfokus gilt als Schlüssel für ein erfolgreiches Wortlernen (Moore, 2014), da Kinder in einer sozial koordinierten Interaktion nicht nur lernen, auf einen Referenten zu achten, welcher für beide Gesprächspartner von Interesse ist, sondern sich auch über die Wortbedeutung dieses Referenten auszutauschen (Bruner, 1983, S. 68–77). Um auf einen Referenten in einen gemeinsamen Aufmerksamkeitsfokus hinzuweisen und

damit die Wort-Objekt-Verknüpfungen für Kinder zugänglicher zu gestalten, nutzen Erwachsene multimodale Hinweise, indem sie beispielsweise Sprache mit Gesten kombinieren, die vorteilhaft in Bezug auf die Sprachentwicklung wirken (Zukow-Goldring, 1996). Studien belegen den Zusammenhang zwischen RJA im Alter von 6, 8, 10, 12 und 18 Monaten und der späteren Entwicklung des Wortschatzes von Kindern (Morales et al., 2000). Tomasello und Farrar (1986) erklären diesen Zusammenhang so, dass in freien Spielsituationen zwischen Müttern und ihren Kindern, Mütter mehr Sätze produzieren, häufiger kommentieren und die Konversationen länger sind, wenn ein gemeinsamer Aufmerksamkeitsfokus gegeben ist. Zudem ergänzen Tomasello und Farrar (1989), dass das Eingehen des Erwachsenen auf einen bereits vom Kind hergestellten Aufmerksamkeitsfokus eine positive Auswirkung auf das spätere Vokabular hat.

Allerdings reicht es oft nicht aus, die Aufmerksamkeit auf etwas zu lenken bzw. zu richten, damit wahrgenommene Reize erfolgreich verarbeitet werden. Hierzu bedarf es neben der visuellen Orientierung noch das Aufrechterhalten der Aufmerksamkeit (engl.: *sustained attention*). Aus der bisherigen Forschung mit Kindern im Vorschulalter (Ruff & Lawson, 1990) ist bekannt, dass die Fähigkeit, denselben Aufmerksamkeitsfokus über eine längere Zeitspanne aufrechtzuhalten, mit dem Alter zunimmt. In weiteren Studien (Lawson, Parrinello & Ruff, 1992; Ruff & Lawson, 1990) stellte sich heraus, dass nicht nur das Alter, sondern auch die Art der Darbietung eines Objektes mit der Aufmerksamkeitsspanne zusammenhängt. Demnach können Mütter, indem sie die Objekteigenschaften in gemeinsamen Interaktionen mit ihren 12 Monate alten Kindern hervorheben, derer Aufmerksamkeitsdauer expandieren (Lawson, Parrinello & Ruff, 1992, S. 226). Zudem wurde festgestellt, dass sich die Aufmerksamkeitszunahme vor allem bei den Kindern zeigte, die eine generell kurze Aufmerksamkeitsspanne hatten. In eine ähnliche Richtung geht auch die Studie von Wass, Clackson und Kollegen (2018), die 12 Monate alte Kinder aus einer Bedingung, in denen sie alleine spielten („Solo-Play“) mit Kindern aus einer Bedingung, in der sie mit einem Experimentator spielten („Joint-Play“), untersuchten (vgl. Kapitel 3.5.1). Sie fanden heraus, dass die kindliche Aufmerksamkeitsdauer stark von dem Verhalten des Erwachsenen in einer gemeinsamen Interaktion abhängt. Um genauer zu sein, schauten die Probanden das Zielobjekt länger an, wenn dieses in ein gemeinsames Spiel eingebunden wurde, als in einer Situation, in der beide mit dem Objekt alleine spielten.

Die Auffassung, dass die geteilte und aufrechterhaltende Aufmerksamkeit nicht unabhängig voneinander zu funktionieren scheinen, wird insbesondere in der Studie von Yu und

Smith (2016) erkennbar, in der die Forscher den Zusammenhang zwischen den beiden Aufmerksamkeitsformen untersuchten. In der Studie (Yu & Smith, 2016) wurde die andauernde Aufmerksamkeitsspanne von 10- und 12-Monatigen in einer dyadischen Interaktion gemessen, wobei sich folgendes gezeigt hat: Je länger die Kinder den gemeinsamen Aufmerksamkeitsfokus mit der Bezugsperson teilten, umso länger konnten sie ihre Aufmerksamkeit zum Objekt aufrechterhalten. Das bedeutet, dass die Aufmerksamkeitsdauer durch die Länge der geteilten Aufmerksamkeit beeinflusst wurde. Des Weiteren stellten Yu und Smith (2016) fest, dass die Probanden das Zielobjekt sogar dann weiterhin fixierten, wenn die erwachsene Person nach einiger Zeit ihre Blickrichtung änderte. Hinsichtlich dieser Ergebnisse schließen sich Yu und Smith anderen Forschern (z. B. Csibra, 2010) insofern an, als das Erkennen und Verstehen einer kommunikativen, sozialen Intention anfangs nicht ausschlaggebend für die Etablierung einer geteilten Aufmerksamkeit sei und schlagen daher vor, dass die *geteilte Intentionalität*³ (nach Tomasello, Carpenter, Call, Behne & Moll, 2005) als Erklärung für die Entwicklung der geteilten Aufmerksamkeitsform, die letztlich mit dem Spracherwerb zusammenhängt, überdacht werden sollte. Darüber hinaus sei aber auch die aufrechterhaltende Aufmerksamkeit nicht nur eine Folge der Selbstkontrolle, die sich altersabhängig entwickelt. Die Ergebnisse sprechen vielmehr dafür, dass gemeinsame soziale Interaktionen das Lernen fördern, weil die kindliche Aufmerksamkeit zu einem Referenten durch aufmerksamkeitslenkende, referentielle Hinweise in einer Interaktion gefestigt wird (Yu & Smith, 2016, S. 5).

2.1.2 Aufmerksamkeitssteuernde Mechanismen

Wie bereits im vorherigen Kapitel aufgeführt wurde, ist die selektive Aufmerksamkeitsform entscheidend dafür, was in die Informationsverarbeitungsprozesse des Gehirns aufgenommen wird. Die klassische Sichtweise auf die kognitive Informationsverarbeitung geht auf Atkinson und Shiffrin (1968, 1971) zurück. Ihre Theorie stellt ein Modell der Gedächtnissysteme dar, welches das Langzeit-, Kurzzeit- bzw. Arbeitsgedächtnis umfasst. Hiernach wird ein Reiz zunächst durch die Sinne wahrgenommen: visuell, auditiv oder haptisch. Dabei verweilen die Informationen im sensorischen Gedächtnis nur wenige Sekunden. Wird die Aufmerksamkeit auf die wahrgenommene Information weiterhin gerichtet, so gelangt

³ *Shared Intentionality* (Tomasello et al. 2005, S. 676): Die Theorie der geteilten Intentionalität beruht auf der Ansicht, Menschen seien biologisch derart ausgestattet, dass sie von Anfang an in sozialen, geteilten Handlungen mit ihren Mitmenschen teilnehmen. Allerdings erfordern solche Interaktionen nicht nur das Verstehen von Intentionen und Zielen anderer, sondern auch die Bereitschaft bzw. Motivation sich darüber auszutauschen.

diese in das Arbeitsgedächtnis, oder auch Kurzzeitgedächtnis genannt, das etwas mehr Kapazität hat als das sensorische Gedächtnis. Diese Information kann kognitiv weiterverarbeitet werden, indem sie weiter in das Langzeitgedächtnis gelangt und damit langfristig abgespeichert wird. Allerdings läuft die Informationsverarbeitung nicht ausschließlich sequenziell ab, sondern es findet eine Interaktivität der Systeme statt, wodurch zwei Aufmerksamkeitsprozesse – der Bottom-up- und Top-down-Prozess – hervorgerufen werden. Diese zwei Mechanismen werden im Folgenden näher beschrieben.

Wird eine Information, aufgrund von besonderen physischen Eigenschaften sensorisch erfasst und dann in die weiteren Systeme geleitet, spricht man von einem daten- bzw. reiz-gesteuerten Prozess. Diese Art von Aufmerksamkeitslenkung wird als bottom-up oder auch als *exogen* bezeichnet, da dieser Wirkungsprozess automatisch „von unten nach oben“ als eine schnelle Reaktion auf externe Einflüsse, meist aufgabenunabhängig, auftritt (Itti & Koch, 2001, S. 194). In erster Linie rufen perzeptuell saliente, d. h. ins Auge springende (lat., *salire*: springen), Informationen den Bottom-up-Prozess hervor. Dieser Auswahlprozess scheint evolutionär bedingt und lebensnotwendig zu sein, damit eine Situation schnell erfasst und bewertet werden kann, um anschließend und situationsbedingt handeln zu können. Die Selektion eines Reizes kann zum Beispiel durch einen abrupten Anstieg in der physikalischen Stärke eines Stimulus (grelle Farbe, lauter Ton, schnelle Bewegung) erfolgen.

Da ein Stimulus auf unterschiedliche Arten perzeptuell salient sein kann, wurde der Begriff *Salienz* versucht mithilfe von Modellen näher zu beschreiben und zu konkretisieren. So stellten beispielsweise Itti und Koch (1998) Salienz-Charakteristika in einem Modell anhand von sogenannten Salienzkarten dar, welche sich durch die Aktivierung von linearen Filtern ergeben. Diese lineare Filterung erfolgt durch Farbe, Intensität/Kontrast und Orientierung. Mithilfe der Salienzkarten konnten die salienten Stimuli aus visuellen Szenen aufgrund ihrer physikalischen Stärke (Farbe, Geräusch) gefiltert werden, sodass auf Grundlage dessen ermittelt werden konnte, welcher Stimulus die Aufmerksamkeit am wahrscheinlichsten auf sich lenkt. Der Interpretation der Salienzkarten nach, wird eine Information als salient definiert, wenn sie auf irgendeine Art und Weise aus dem Kontext hervorsteht (Itti & Koch, 2001, S. 202), wofür auch der umschreibende Begriff *Local Feature Contrast* verwendet wird (Nothdurft, 2000, S. 3181). Ein roter Mantel steche beispielsweise nur dann heraus, wenn dieser von schwarzen Mänteln umgeben sei (Itti & Koch, 2001, S. 194). Somit ist die perzeptuelle Salienz nicht nur von Merkmalen eines Objektes, sondern auch vom jeweiligen Umfeld des Objekts abhängig. Außer Acht gelassen wurde bei den Salienzkarten

jedoch der Aspekt, dass nicht nur perzeptuelle, sondern auch soziale Reize die Aufmerksamkeit steuern können. Des Weiteren wird bei den Salienzkarten nicht berücksichtigt, dass die Blickrichtung auf Dinge im Umfeld ebenso von individuellen, persönlichen Interessen oder sogar vom Vorwissen geprägt sein können.

Studien zur frühkindlichen visuellen Wahrnehmung zeigen, dass Kinder von Geburt an Präferenzen bzw. Wahrnehmungsneigungen (engl.: biases) haben, die ihre Aufmerksamkeit lenken. Es wurde festgestellt (z. B. Adams, 1987; Adams & Courage, 1995), dass bereits Neugeborene zwischen farbigen und farblosen Objekten differenzieren können. Ab dem dritten Monat zeigen sie eine Präferenz für die Farben Rot und Gelb im Vergleich zu Blau und Grün (Adams, 1987; Franklin, Bevis, Ling & Hurlbert, 2010). In einer anderen Studie (Fantz, 1963) zeigte sich, dass Säuglinge nicht nur den Unterschied zwischen homogen farbigen und gemusterten Flächen erkennen, sie zeigten außerdem eine Präferenz für das Muster, indem sie dieses länger anschauten. Andere Studien haben die kindliche Wahrnehmung von statischen und dynamischen Stimuli untersucht und herausgefunden, dass fünf Monate alte Kinder eine Präferenz für dynamische im Vergleich zu statischen Gesichtern zeigen (Wilcox & Clayton, 1968). In einem anderen Experiment (Nagata & Dannemiller, 1996) mit 14 Wochen alten Kindern wurden rote und grüne Stimuli präsentiert, von denen sich einer bewegt hat. Bei dieser Kontrastierung von Farbe und Bewegung zeigte sich, dass für Säuglinge die Stimuli-Farbe eine wichtige Rolle dabei spielt, ob sie den bewegenden Stimulus identifizieren. Den Probanden war es nicht möglich einen sich bewegenden, grünen Ziel-Stimulus in einem Mix aus grünen und roten Stimuli zu finden, während es ihnen bei einem roten Ziel-Stimulus gelang. Kurzgefasst nehmen Säuglinge visuelle Stimuli dann als salient wahr, wenn sich diese bewegen, eine auffällige Farbe haben oder gemustert sind.

Im Gegensatz zum Bottom-up- ist der Top-down-Prozess *konzeptgesteuert* (Connor, Egeth & Yantis, 2004; Itti & Koch, 2001). Dieser wird auch als *endogen* bezeichnet, da dieser Wirkungsprozess „von innen heraus“ entsteht und nach außen wirkt. Das bedeutet, dass das Langzeitgedächtnis die Aufmerksamkeit auf einen Referenten lenkt, der von dem sensorischen System erfasst werden soll, wodurch eine Interaktivität zwischen den Gedächtnissystemen herrscht. Das Ziel kann beispielsweise die Suche nach einem bekannten Objekt sein, wobei das Wissen über die Objektfarbe genutzt werden kann, um es zu selektieren. Genauso gut kann die Person aber auch ein neues Objekt ausfindig machen, indem das Merkmal „bekannte Farbe“ als Ausschlusskriterium verwendet wird. In diesen Fällen wird die

Aufmerksamkeit top-down, durch Wissen, Vorerfahrung und Erwartungen gesteuert (Ansorge & Leder, 2017, S. 9)

An dieser Stelle soll abschließend erwähnt werden, dass eine solch klare Differenzierung der beiden Prozesse wie in dem Beispiel, sich in der Realität jedoch schwierig gestaltet. Denn auch ein vermeintlich bottom-up-gesteuerter Aufmerksamkeitsprozess, wie beim Neuheits-Effekt (vgl. Kapitel 3.2.1), wird ebenso top-down gesteuert, indem auf vorhandenes Objektwissen zurückgegriffen wird.

2.2 Gedächtnisprozesse

Neben den aufmerksamkeitssteuernden Mechanismen sind auch die in den Gedächtnissystemen ablaufenden *mnemonischen* Prozesse (griech., *mneme*: Gedächtnis) (Gruber, 2018, S. 5) bedeutsam, um Erkenntnisse darüber zu gewinnen, wie neue Gedächtniseinträge entstehen. Diese Prozesse lassen sich in drei Phasen unterteilen: die Enkodierung, Retention/Konsolidierung und den Abruf. Einen erhellenden Beitrag lieferte Wojcik (2013) zu der Bedeutung dieser Prozesse für den Bereich der Spracherwerbsforschung: Einerseits schaffte sie in ihrem Überblicksartikel die Brücke zwischen den Befunden aus der Gedächtnisforschung und den Erkenntnissen aus Wortlernstudien. Andererseits kritisierte sie, dass, obwohl der Worterwerb ein Prozess ist, der das Gedächtnis betrifft, viele Studien die Gedächtnisprozesse nicht berücksichtigen oder die betroffenen Prozesse als solche nicht benennen:

Yes, children must first map labels onto referents, but they must also encode, consolidate, and retain these representations. The process of how infants and young children encode, store, and retrieve representations has been studied rigorously for half a century, and yet this research has rarely been used to inform our study of word learning (Wojcik, 2013, S. 10).

Infolgedessen kommt es zu einer uneinheitlichen Verwendung von Begriffen und zu unterschiedlichen Schlussfolgerungen in Studien, wenn beispielsweise von *Wortlernen* die Rede ist, aber lediglich die Referentenauswahl während der Enkodierungsphase untersucht wurde (Wojcik, 2013, S. 2). Im Folgenden werden die Gedächtnisprozesse vorgestellt und mit Beispielen aus Wortlernstudien veranschaulicht. Für die vorliegende Arbeit steht insbesondere der Enkodierungsprozess im Vordergrund.

2.2.1 Enkodierung

Die Enkodierung (engl.: encoding) ist ein Erlebniseindruck, der im Gedächtnis eine erste Gedächtnisspur (Engramm⁴) hinterlässt (Wojcik, 2013). Es handelt sich bei diesem Prozess nicht nur um eine Aufnahme eines Reizes als *Code* (Gruber, 2018, S. 6), sondern auch um einen Abgleich und Verknüpfung des Codes mit bereits vorhandenen mentalen Repräsentationen (S. 70). Solch ein Code kann von sensorischer, phonologischer und/oder semantischer Natur sein (S. 6). Zum Beispiel kann ein *Tennisball*, der als sensorischer Reiz visuell registriert und ins Arbeitsgedächtnis aufgenommen wurde, während des Enkodierungsprozesses mit dem bereits bekannten *Fußball* in Verbindung gebracht werden. Welche Stimuli letztendlich enkodiert werden, hängt davon ab, worauf die selektive Aufmerksamkeit (vgl. Kapitel 2.1.1) fällt.

Aus der bisherigen Forschung ist bekannt, dass Kinder mit zunehmendem Alter neue Reize schneller enkodieren (Wojcik, 2013, S. 5). Hinzu kommt, dass ältere Kinder nicht nur schneller in der Enkodierung sind, sondern auch weniger Unterstützung dabei benötigen im Vergleich zu jüngeren Kindern (Horst & Samuelson, 2008; Vlach & Sandhofer, 2012). Während ältere Altersgruppen selbstständig kontrollieren können, was und wie lange sie etwas enkodieren, haben jüngere Kinder hingegen Schwierigkeiten ihre Aufmerksamkeit kontrolliert zu steuern.

Daher sollten Erwachsene junge Kinder in Wortlernsituationen vor allem in der Selektion und Aufrechthaltung ihrer Aufmerksamkeit unterstützen, damit ihnen die Enkodierung neuer Wort-Referent-Verknüpfungen gelingt. Die Enkodierung eines *Zielobjektes* kann unterstützt werden, indem Erwachsene den Referenten zum Beispiel mit Gesten hervorheben oder damit handeln (Deák, Krasno, Jasso & Triesch, 2018). Ebenso unterstützend auf die Enkodierung kann es wirken, indem das Objekt durch Hochhalten oder Beleuchtung als salient markiert wird (Axelsson, Churchley & Horst, 2012). Zusätzlich kann die Salienz der anderen Objekte, die von dem Referenten ablenken, abgeschwächt werden, indem man die Objekte verdeckt oder gar ganz aus dem Sichtfeld wegräumt. Die Enkodierung des *Zielwortes* kann dadurch gefördert werden, indem Erwachsene zum Beispiel eine kindgerichtete Sprache (engl.: infant-directed speech) nutzen (Csibra, 2010; Fernald, 1985; Golinkoff, Can, Soderstrom & Hirsh-Pasek, 2015) und das Zielwort mehrmals wiederholen (Vlach & Sandhofer, 2012). Neben den aufgeführten Strategien, die Erwachsene nutzen, um die Aufmerksamkeit von Kindern auf eine Wort-Objekt-Verbindung zu lenken und diese zu halten,

⁴ „Engram: the physical record of a memory item in the brain; a memory trace“ (Dudai, 2012).

haben Kinder ebenso ihre eigenen Strategien, um Informationen über das Wort, den Referenten und die Verknüpfung zwischen den beiden Komponenten zu enkodieren (Wojcik, 2013, S. 1).

Laut des *Emergentist-Coalition-Modells* (vgl. Kapitel 3.3.1) nutzen Kinder perzeptuelle, soziale und linguistische Faktoren als Enkodierungsstrategie, allerdings zu unterschiedlichen Zeitpunkten ihrer Entwicklung (Hollich et al., 2000, S. 30):

Attentional, social, and linguistic factors all play a necessary role in children's first word acquisition [...] We would emphasize, however, that although these mechanisms are available from the start, they might not be exploited to the same degree at different points in development.

Diesem Modell zufolge wird die Aufmerksamkeit von Kindern in ihrer ersten Entwicklungsphase von perzeptuell salienten Dingen in ihrer Umwelt gelenkt. In der zweiten Phase erkennen sie, dass auch soziale Hinweise für die Enkodierung hinzugezogen werden können. Schließlich werden in der dritten Phase auch linguistische Achtungssignale genutzt. Obwohl das Modell aufgrund dieses stufenweisen Entwicklungsverlaufs kritisch zu betrachten ist (für eine kritische Reflexion vgl. Kapitel 3.6), bleibt dennoch unbestritten, dass Kinder sich im Laufe der Zeit in all ihren Fähigkeiten weiterentwickeln und deshalb nicht dieselben Enkodierungsstrategien nutzen. Hierfür können zwei Studien als Beispiel hinzugezogen werden. In der Studie von Carey und Bartlett (1978) wurden drei- und vierjährige Kinder mit der Aussage „the chromium tray, not the blue one“ instruiert, ein Zielobjekt zu identifizieren. Die Versuchsteilnehmer waren erfolgreich in dieser Aufgabe, da sie durch die Enkodierungsstrategie des *lexikalischen Kontrasts* die Schlussfolgerung ziehen konnten, auf welches Objekt referiert wurde. Dies war möglich, weil sie bereits das Farbwort *Blau* kannten und dadurch das blaue Objekt als Referent ausschließen konnten. Während ältere Kinder ihr vorhandenes Wortwissen nutzen können, welches bei jüngeren Kinder noch nicht in diesem Maße vorhanden ist, greifen die jüngeren deshalb auf andere Strategien zurück.

Abschließend ist zu betonen, dass eine erfolgreiche Enkodierung für die weitere Informationsverarbeitung essentiell ist. Aus der Gedächtnisforschung ist bekannt (Anderson & Schooler, 1991; Rubin & Wenzel, 1996; White, 2001; Wixted, 2004), dass Informationen im Laufe der Zeit exponentiell vergessen werden. Munro, Baker, McGregor und Arciuli (2012) stellten fest, dass dies auch auf den Wortlernprozess zutrifft und der größte Verlust des Wortwissens bereits in den ersten Minuten stattfindet. Die Phase des Enkodierens funktioniert wie ein „Flaschenhals“, durch den nicht alle Informationen durchkommen. In der Studie von Vlach und Sandhofer (2012) stellte sich außerdem heraus, dass das Einsetzen

von Gedächtnisstützen während der Enkodierungsphase dem Vergessensprozess entgegensteuert. Kinder, die keine Gedächtnisstützen während der Enkodierung erhielten, vergaßen im Laufe der Zeit die neuen Wörter (vgl. Horst & Samuelson, 2008). Mit drei Gedächtnishilfen (Gesten, Salienz, Wiederholung) konnten die Wort-Objekt-Verknüpfungen hingegen auch noch nach einem Monat erfolgreich abgerufen werden.

2.2.2 Retention und Konsolidierung

Nach der Enkodierung folgt in dem mnemonischen Prozess die Aufrechterhaltung und Verfestigung des neu erworbenen Wissens. In der Retentionsphase (lat., *retinere*: zurückhalten) wird die enkodierte Information zunächst im Arbeitsgedächtnis bzw. Kurzzeitgedächtnis aufrechterhalten. Diese zwischenzeitige Speicherung der Informationen erfolgt im Hippocampus, welcher im Bereich des medialen Temporallappens (MTL) liegt und aufgrund seiner Form seinen Namen (lat., *hippocampus*: Seepferdchen) erklärt (Gruber, 2018, S. 64). Die Dauer, wie lange eine Information im Gedächtnis aufrechterhalten werden kann, wird als *Retentionsintervall* bezeichnet (S. 70–73). Erst im nächsten Schritt, der Konsolidierungsphase (lat., *consolidare*: zusammenfügen), werden die zunächst noch vom Hippocampus abhängigen Gedächtnisinhalte nach und nach neuronal verfestigt (Dudai, 2004, 2012; Gruber, 2018, S. 6), indem sie in die äußere Hirnrinde, den Neocortex, übertragen werden.

Das Vorhandensein eines schnellen und eines langsamen Lernsystems hat zwei wesentliche Funktionen: Mithilfe des schnellen Mechanismus können neue Inhalte leicht „aufgeschnappt“ und im Hippocampus abgelegt werden. Diese Zwischenspeicherung von neuem Wissen dient dazu, dass es nicht zu einer Überschreibung von vorhandenem Wissen mit neuen Informationen im Langzeitgedächtnis kommt. Erst durch den langsamen Prozess werden die „abgelegten“ Inhalte allmählich in den Neocortex überspielt und somit verfestigt (Dudai, 2004, S. 62). Wäre statt der zweien nur das schnelle Lernsystem vorhanden, dann würde beispielsweise die bereits bestehende mentale Repräsentation eines Apfels bei jedem neu wahrgenommenen Apfel überschrieben werden. Die Überschreibung von alten Gedächtniseinträgen durch neue, wodurch verhindert wird, dass ein Eintrag nachhaltig abgespeichert werden kann, wird als *katastrophaler Interferenzeffekt* (engl.: *catastrophical inference*) genannt (Henderson, Weighall, Brown & Gaskell, 2012, S. 684).

Das beschriebene duale Gedächtnissystem (engl.: *dual memory system*) wurde ebenfalls für die Verarbeitung von neuen Wörtern untersucht. Hierzu sollten in der Studie von Henderson und Kollegen (2012) 7–12 Jahre alte Kinder über das Vorhandensein einer Pause am Ende eines Kunstwortes entscheiden (*Pause detective paradigm*), welches sich lediglich

in der letzten Silbe von einem echten Wort unterschied (z. B. **cathed_ruke* und *cathed_ral*). Die Forscher stellten fest, dass Pausen schneller erkannt wurden, wenn sich das Kunstwort noch nicht im Langzeitgedächtnis, sondern im „Zwischenspeicher“ befand. Sobald aber das Pseudowort in das bestehende mentale Lexikon integriert wurde, entstanden Konkurrenzefekte zwischen dem Kunstwort (**cathed_ruke*) und dem bereits bekannten Wort (*cathed_ral*), da die Probanden mehr Zeit brauchten, um eine Pause zu erkennen. Durch das Überspielen des Gedächtnisinhalts in den Neocortex befanden sich letztlich zwei ähnliche Wörter in demselben Gedächtnissystem und konkurrierten deshalb miteinander, was dazu führte, dass die Pause langsamer erkannt wurde (engl.: Lexical Competition Effect). Die Verzögerungen in der Pausenerkennung werden als Hinweise auf „Störungen“ wie folgt interpretiert: Wenn eine verstärkte lexikalische Aktivierung vorhanden ist, können weniger Ressourcen der Pausenerkennung zugeschrieben werden. Anders ausgedrückt, entsteht ein Wettstreit zwischen zwei ähnlich klingenden Wörtern erst, nachdem ein neues Wort ins Langzeitgedächtnis integriert wurde und sich dadurch das neuronale Netzwerk verändert hat.

Des Weiteren verdeutlichte die Studie von Henderson et al. (2012), dass für den Konsolidierungsprozess von neu erworbenem Wissen – in diesem Fall von neuen Wörtern in ein existierendes Vokabular – der Schlaf und nicht die Zeit allein eine große Bedeutung hat. Da die 7–12 Jahre alten Probanden zu unterschiedlichen Zeitpunkten getestet wurden, konnte festgestellt werden, dass die neu trainierten Wörter erst nach dem Schlaf erfolgreich abrufen werden konnten. Weitere Studien fanden heraus, dass sogar schon „Nickerchen“, d. h. kleinere Schlafeinheiten, direkt nach dem Lernen die Konsolidierung von neuen Wörtern fördern (Axelsson, Swinton, Winiger & Horst, 2018; Horváth, Myers, Foster & Plunkett, 2015). Schlaf ist einerseits wichtig, damit relevante Informationen stabilisiert werden, und andererseits, damit irrelevante Informationen aussortiert und vergessen werden können (Gruber, 2018, S. 79). Kurzgefasst: Im wachen Zustand werden entweder neue Informationen enkodiert oder vorhandene Gedächtnisinhalte abgerufen, wobei die Schlafphase eine optimale Bedingung bietet, encodierte Informationen zu konsolidieren (Rasch & Born, 2013, S. 683).

Infolge der Konsolidierung können Gedächtnisinhalte auch nach einem längeren Zeitraum erfolgreich abgerufen werden. Die Abruffähigkeit (engl.: retrieval) kann entweder rezeptiv anhand der Rekognition oder expressiv anhand der Reproduktion überprüft werden (vgl. Kapitel 3.1). Bei der ersteren der beiden Vorgehensweisen handelt es sich um die Wiedererkennung, zum Beispiel die eines zuvor gelernten Wortes, während die Reproduktion eine Wiedergabe dieses Wortes erfordert. Des Weiteren kann die Wiedergabe entweder frei sein (engl.: free recall) oder durch Hinweisreize unterstützt werden (engl.: cued recall)

(Gruber, 2018, S. 7). Da der Fokus dieser Arbeit hauptsächlich den Enkodierungs- und zum Teil den Konsolidierungsprozess betrifft, wird im Folgenden nicht näher auf den Abruf eingegangen.

2.3 Wortlernen

Die zentrale Aufgabe von Kindern im Spracherwerbsprozess besteht darin eine *Referenz*, d. h. den Bezug von sprachlichen Zeichen auf Dinge in der Welt, herzustellen (Vater, 2012, S. 1). Das Schema einer Referenz wird in einem semiotischen Dreieck deutlich, welches aus dem Zeichenausdruck, dem bezeichneten Referenten und der mentalen Einheit eines Zeichenbenützers besteht (Rohlfing, 2019, S. 215; Rohlfing, 2013, S. 33). An dem semiotischen Dreieck ist charakteristisch, dass kein unmittelbarer Bezug des Zeichenausdrucks zu dem bezeichneten Gegenstand besteht, sondern die Vermittlung der Wortbedeutung⁵ über die Vorstellung eines Begriffs in der mentalen Repräsentation des Zeichenbenützers erfolgt. Referiert eine Person auf einen Gegenstand oder eine Handlung mit einem Symbol (dem Wort), schwingt mit dem lautsprachlich realisierten Zeichenausdruck gedanklich eine Bedeutung des Zeichens mit. Somit handelt es sich bei einer Wortbedeutung um eine mentale Repräsentation, die mit der Wortform zusammenhängt. Diese Repräsentation kann selbst dann bestehen, wenn das sprachliche Zeichen noch nicht produziert werden kann (Bloom, 2001).

Im Erwerb einer neuen Wortbedeutung bewältigt das Kind eine Reihe an Aufgaben: Es ordnet das neue Wort einem Referenten oder einem Ereignis zu, indem es die lautlichen Einheiten aus dem Sprachstrom extrahiert und das Objekt bzw. das Ereignis findet, worauf sich das Wort bezieht. Dabei füllt es die Verbindung mit einem Inhalt, welcher im Idealfall der mentalen Repräsentation des Zeichenbenützers nahekommt.

Das bedeutet, dass Kinder einem Referenten am Anfang nicht denselben Bedeutungsinhalt zuweisen wie Erwachsene. Bruner (1983, S. 68) macht diesen Aspekt mit dem Beispielwort *Elektrizität* deutlich: Eine Mutter, die Physikerin ist, erklärt ihrem vierjährigen Sohn die Gefahr eines elektrischen Schlags. Da sich eine Wortbedeutung nicht immer unmittelbar auf die Wirklichkeit bezieht oder sogar abstrakt sein kann, wird der Bedeutungsinhalt des Kindes dem eines Erwachsenen anfangs womöglich nicht entsprechen. Doch auch, wenn das frühe semantische Wissen über Elektrizität nicht direkt vollständig ist, wurde eine erste Vorstellung der Wortbedeutung fundamentierte und kann im Laufe der Zeit weiter ausgebaut

⁵ Die Termini „Referenz“ und „Bedeutung“ sind nicht gleichzusetzen, da ersteres sich auf außersprachliche Referenten bezieht und letzteres einen innersprachlichen Bezug des Zeichens meint (Vater, 2012, S. 3).

werden. Eine Annäherung an den Bedeutungsinhalt eines Erwachsenen wird erreicht, wenn Kinder nach und nach semantische Merkmale hinzuaddieren (Szagun, 2013, S. 137–138, siehe auch Rohlfing, 2019, S. 219). Als Beispiel kann das Wort *Hund* hinzugezogen werden, welchem zunächst das semantische Merkmal *Vierbeiner* zugeordnet und im Laufe der Zeit mit weiteren spezifischen Merkmalen, wie *bellt* angereichert wird.

Die ersten Wortbedeutungen entstehen noch bevor Kinder ihre ersten Wörter, sogenannte Protowörter (Kauschke, 2012, S. 53), um das erste Lebensjahr herum anfangen zu produzieren (Rohlfing, 2019, S. 147; Szagun, 2015, S. 74). Um nachzuweisen, ob ein frühes Wortverständnis vorhanden ist, werden Verständnistests anhand von Blickzeitmessungen durchgeführt (vgl. Kapitel 3.1.1), indem üblicherweise die Assoziation⁶ zwischen dem Wort und einem Objekt getestet wird. Zeigen die Probanden eine längere Blickdauer zum Zielobjekt im Vergleich zu einem anderen, gleichzeitig präsenten Objekt, während sie das Zielwort hören, dann wird dies als Indiz für eine rudimentäre Form des Wortverständnisses gesehen.

Tincoff und Jusczyk (1999, 2012) deckten auf, dass Kinder im Alter von sechs Monaten bei den Wörtern *mommy* und *daddy* die Bilder länger anschauen, auf denen ihre Eltern zu sehen sind, im Vergleich zu Bildern, auf denen fremde Frauen und Männer abgebildet sind. In weiteren Studien (Bergelson & Swingley, 2012, 2015) wurde herausgefunden, dass bei Kindern ab dem sechsten Lebensmonat bereits Assoziationen von alltäglichen Wörtern aus der Kategorie *Körper* und *Essen* und Bildern von den jeweiligen Objekten vorhanden sind. Nomikou und Kollegen (2018) weisen darauf hin, dass neben der Wortklasse der Nomina 10 Monate alte Kinder auch eine mentale Vorstellung von Verben, die ihnen im Alltag begegnen haben, beispielsweise *bauen*. Auf Basis der eben aufgeführten Studien kann geschlossen werden, dass Kinder Wörter nicht nur mit Objekten, sondern auch mit Menschen und Funktionshandlungen assoziieren.

Allerdings wurde in den meisten Studien zum frühen Wortverständnis hauptsächlich die Verknüpfung von Wort-Objekt-Verknüpfungen untersucht. Um nicht nur zu erfahren, ob und welche Wörter verstanden werden, sondern auch, welchen Bedeutungsinhalt die ersten Wörter tragen, sollten neben Nomen-Objekt-Verknüpfungen auch Assoziationen mit anderen Wortklassen (z. B. mit Verben, vgl. Nomikou et al., 2018; mit Adjektiven, vgl. Forbes & Plunkett, 2019) untersucht werden. Neben der Frage, was Kinder bereits verstehen, wird

⁶ Eine *Assoziation* ist nicht nur die Erinnerung an ein Objekt, sondern auch an alles, was damit zusammenhängend erfasst wurde, beispielsweise seine Funktion oder der Name (Rohlfing, 2019, S. 102). Sie ist zunächst situativ gebunden, kann aber durch wiederholten Gebrauch auch auf neue Situationen übertragen werden (S. 208).

auch die Forschungsfrage diskutiert, *wie* erste Wortbedeutungen gelernt werden. Hinsichtlich dieser Frage werden im weiteren Verlauf dieser Arbeit Theorien beschrieben, die unterschiedliche Erklärungsansätze aufgestellt haben wie Kinder Wörter mit Referenten verknüpfen (vgl. Kapitel 3).

Der Prozess der Zuordnung wird auch als *Mapping* bezeichnet (Kauschke, 2012, S. 43) und besteht aus zwei wesentlichen Phasen – dem Fast und Slow Mapping –, auf die in den folgenden Unterkapiteln eingegangen wird.

2.3.1 Fast Mapping

Der Begriff *Fast Mapping* (Carey, 1978, 2010) bezeichnet einen rudimentären Lernprozess, bei dem auf die Schnelle einem neuen Wort ein Referent zugeordnet wird (Horst & Samuelson, 2008; Markson & Bloom, 1997). Da ein Kind bei dem ersten Aufeinandertreffen eines Referenten und eines neuen Wortes vor zahlreichen Möglichkeiten steht, wie sich der Referent und die phonologische Form zueinander verhalten, werden in diesem ersten Schritt Hinweise aus dem gegebenen Kontext hinzugezogen, um sich die erste Wortbedeutung zu erschließen (Carey, 1978). Hierzu zählt neben der Berücksichtigung der linguistischen Informationen – wie zum Beispiel der phonetischen Aspekte, Syntax oder der semantischen Kategorie – auch die Berücksichtigung des nicht-linguistischen Kontextes (Heibeck & Markman, 1987; Horst, Samuelson, Kucker & McMurray, 2011). Nach der Referentenzuordnung ist die Wort-Referent-Verknüpfung weder vollständig noch fest im Gedächtnis verankert.

An dieser Stelle soll der Zusammenhang zwischen Fast Mapping und Enkodierung verdeutlicht werden, da beide Prozesse zwar eng miteinander zusammenhängen, aber nicht ein und dasselbe sind. Während es sich bei der Enkodierung um einen kognitiven Informationsverarbeitungsprozess handelt, den man auch als Umwandlung einer aufgenommenen Information in einen neuronalen Code beschreiben könnte, versteht man unter Fast Mapping den Prozess des Zusammenbringens zweier Komponenten, wobei eine erste mentale Repräsentation der Wort-Objekt-Verknüpfung entsteht. Aufgrund der damit verbundenen Funktionen kann die Enkodierung als Voraussetzung für das Fast Mapping gesehen werden. In diesem Prozess werden die ersten Informationen über das Wort und das Objekt enkodiert, um diese anschließend im Fast-Mapping-Prozess miteinander in Verbindung zu bringen. Horst und Samuelson (2008, S. 129) beschreiben den Fast-Mapping-Prozess, dem sie die

Enkodierung unterordnen, folgendermaßen:

To learn the word *kitty* from the sentence “Look at the kitty,” the child must (a) segment the target word from the speech stream, (b) find the referent of the novel word in the current scene that typically contains many possible referent objects, (c) encode the novel word form (e.g., the individual phonemes and their sequence), (d) encode something about the referent (e.g., where it was, its color, its shape, what it was doing), and (e) store the encoded information in such a way that the different pieces are linked and can be retrieved at a later point in time when the child needs to find the kitty, recognize the name, or produce the name.

Eine zentrale Fast-Mapping-Studie ist die von Carey und Bartlett (1978), in der sie die Fähigkeit der schnellen Wort-Referent-Zuordnung von drei- und vierjährigen Kindern untersuchten. Die Versuchsteilnehmer hörten das neue Farbwort *chromium* in dem Satz „Can you give me the chromium tray, not the red one, the chromium one“. Indem die Kinder den Referenten auswählten (engl.: referent selection⁷), demonstrierten sie, dass sie sich den Referenten für das Farbwort *chromium* richtig erschließen konnten. Dieses Experiment brachte die Erkenntnis, dass Kinder schnelle Wort-Objekt-Verknüpfungen mit nur wenig Vorwissen bilden können. Heibeck und Markman (1987) gingen in einer weiteren Studie der Frage nach, ob neben Farbwörtern auch das Lernen neuer Wörter für Form und Beschaffenheit von Objekten möglich sei. Sie zeigten, dass sowohl Zweijährige als auch ältere Kinder in gleichem Maße die Fähigkeit besitzen, bereits bei der ersten Auseinandersetzung mit einem Objekt und dessen Namen eine schnelle Zuordnung herzustellen. Die wesentliche Erkenntnis dieser Studie war, dass Fast Mapping scheinbar in jeder Domäne, d. h. seien es Farbwörter oder Wörtern für Objektformen, möglich ist. Aus weiteren Studien zum Fast Mapping ging hervor, dass sich 30 Monate alte Kinder bis zu sechs neue Wörter innerhalb einer Sitzung merken können (Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992) und eine schnelle Wort-Objekt-Zuordnung bereits unter drei Sekunden bei 2-Jährigen gelingt (Spiegel & Halberda, 2011).

Während in den referierten Studien Kinder als „word-learning wizards“ (Thom & Sandhofer, 2009, S. 466) dargestellt werden, die neue Wörter regelrecht aufsaugen, zeigen andere Studien widersprüchliche Ergebnisse (Rice, 1980, S. 46–83). Demnach brauchen 2- bis 3-jährige Kinder hunderte Wiederholungen, um drei Farbwörter zu lernen (vgl. Rice,

⁷ „Referent selection/object selection task: A method to assess children’s language knowledge or word construal. Children are asked to find a target object from an array of objects by following a verbal prompt such as ‘find the apple’“ (Golinkoff, et al., 2013, S. 333).

1980, Figure 1, S. 73). In genaueren Untersuchungen stellte sich heraus, dass die unterschiedlichen Ergebnisse hinsichtlich der Fast-Mapping-Leistung mit der Erfahrung, die Kinder in einer bestimmten Domäne bereits besitzen, zusammenhängen (Gershkoff-Stowe & Hahn, 2007). Das bedeutet, dass Kinder, die bereits viele Farbwörter in ihrem Wortschatz haben, schneller neue Wörter aus derselben Domäne lernen, als Kinder, die wenige oder noch gar keine Farbwörter kennen (Thom & Sandhofer, 2009).

2.3.2 Die Brücke vom Fast zum Slow Mapping

An die erste Phase des Fast Mappings schließt eine zweite Phase, das *extended mapping* (Carey & Bartlett, 1978; Vlach & Sandhofer, 2012) bzw. Slow Mapping (Kucker & Samuelson, 2011; Swingley, 2010) an. Slow Mapping ist ein Prozess, bei dem die erste, rudimentäre Wort-Referent-Zuordnung im Laufe der Zeit mit weiterem Wissen angereichert und langfristig abgespeichert wird. Wie bereits in dem Kapitel 2.2.2 beschrieben wurde, existiert ein schnelles und ein langsames Lernsystem. Das heißt, sofern aufgegriffene Wörter nicht direkt wieder nach der Enkodierungsphase vergessen werden, gelangen sie in den Konsolidierungsprozess und werden nach und nach in das bestehende mentale Lexikon integriert, um den katastrophalen Interferenzeffekt zu vermeiden (Henderson et al., 2012). Während das Fast Mapping ein schneller Prozess ist, der allerdings eine schwache Wort-Objekt-Verknüpfung hervorbringt, ist der Slow Mapping Prozess zwar langsam, aber er führt zu einer langfristigen, stabilen Wort-Objekt-Verknüpfung.

Ein kritischer Aspekt in Bezug auf die bisherige Forschung zum Mapping ist, dass zahlreiche Studien zum Fast Mapping existieren (vgl. Kapitel 2.3.1), es aber nur wenige Studien zum Slow Mapping gibt, die auch die Konsolidierung der Wort-Referent-Verbindung mitberücksichtigt haben (z. B. Goodman, McDonough & Brown, 1998; Henderson et al., 2012; Markson & Bloom, 1997; Munro, Baker, McGregor & Arciuli, 2012; Woodward, Markman & Fitzsimmons, 1994). Ohne einen ausreichenden zeitlichen Abstand zwischen der Trainings- und Testphase bleibt jedoch unklar, ob eine im Fast-Mapping-Prozess aufgegriffene Wort-Objekt-Verknüpfung auch langfristig im Gedächtnis bestehen bleibt. Aus diesem Grund ist hervorzuheben, dass der Begriff Fast Mapping kein Synonym für das Wortlernen, sondern lediglich der erste Schritt im Wortlernprozess ist.

Vergleicht man die bisherigen Forschungsergebnisse zum Mapping-Prozess, so fällt auf, dass in einigen Studien die Wort-Referent-Verbindungen nach einer Woche (Waxman & Booth, 2000) und sogar nach einem Monat (Carey & Bartlett, 1978; Markson & Bloom, 1997) beibehalten wurden, wohingegen in anderen Studien der Objektname bereits nach fünf

Minuten vergessen wurde (Horst & Samuelson, 2008). Motiviert durch die unterschiedlichen Erkenntnisse wurde in der weiterführenden Forschung die Frage untersucht, wie eine zunächst schwache Verknüpfung, robust und langfristig in dem mentalen Lexikon verankert wird: Wie gelingt also die Brücke vom Fast zum Slow Mapping?

Dieser Frage gingen Horst und Samuelson (2008) in ihrer Studie mit 24 Monate alten Kindern nach. Dabei stellten sie fest, dass Kinder in der Referentenauswahl eine gute Leistung erbringen, aber bereits nach fünf Minuten das Wort weder produktiv noch rezeptiv abrufen konnten. Aus diesen Ergebnissen folgerten sie, dass, obwohl das Fast Mapping gelungen war, das Slow Mapping scheiterte. Auch ein mehrmaliges Benennen des Referenten während der Trainingsphase reichte nicht aus, um die Enkodierung der Wort-Objekt-Verknüpfung soweit zu verstärken, dass die Verknüpfung langfristig beibehalten wird. Erst in einer weiteren Kontrollstudie, nachdem das Zielobjekt durch Hochhalten und Zeigegesten aus dem Kontext hervorgehoben wurde, konnten die Probanden das Wort nach der fünfminütigen Pause rezeptiv abrufen. Auch weitere Studien (z. B. Vlach & Sandhofer, 2012) bestätigen die gewonnenen Erkenntnisse von Horst und Samuelson (2008), dass ohne zahlreiche Gedächtnisstützen eine aufgegriffene Wort-Objekt-Verknüpfung schnell wieder vergessen wird. Daraus wurde geschlossen, dass während der Enkodierungsphase das Wort und dessen Referent hervorgehoben werden müssen, damit diese auch nach der Enkodierung beibehalten werden. Das bedeutet, je tiefer die semantische Repräsentation des Wortes ist, umso erfolgreicher kann das Wort abgerufen werden (Capone & McGregor, 2004, 2005).

Neben der Verstärkung einer Wort-Referent-Verknüpfung *während* des Fast-Mapping-Prozesses kann auch die Erfahrung *vor* dem Fast Mapping eine Rolle dabei spielen, ob ein Wort beibehalten wird. Kucker und Samuelson (2011) belegten, dass das Wortlernen bereits vor dem Fast Mapping stattfindet, indem sie die Probanden vor dem Training entweder mit Objekten oder mit Wortformen familiarisierten. Es stellte sich heraus, dass die Familiarisierung mit Objekten den 24 Monate alten Kindern geholfen hat einen neuen Objektamen aufzuschnappen und diesen auch noch nach fünf Minuten abzurufen. Die Gruppe, die der Wortform familiarisiert wurde, konnte den neuen Objektamen nach fünf Minuten allerdings nicht beibehalten. Des Weiteren zeigten Studien (Rice, 1980; Gershkoff-Stowe & Hahn, 2007; Thom & Sandhofer, 2009), dass auch das Wortwissen in einer Domäne Einfluss darauf hat, ob ein neues Wort erfolgreich gelernt wird oder nicht. So haben beispielsweise Thom und Sandhofer (2009) in ihrer Studie 16–26 Monate alten Probanden über einen längeren Zeitraum entweder zwei, vier oder sechs Farbwörter trainiert. Anschließend wurde getestet, ob sie ein komplett neues, nicht-trainiertes Farbwort, der richtigen Farbe zuordnen

konnten. Es stellte sich heraus, dass die Probanden, die nach dem Training die meisten Farbwörter in ihrem Lexikon hatten (6 Farben), die neuen, nicht-trainierten Farbwörtern besser lernten, als diejenigen, die im Training weniger Farbwörter gelernt hatten (2–4 Farben).

2.4 Das Referenzproblem

Aus der bisherigen Forschungsliteratur ist bekannt, dass Kinder schon im Alter von sechs Monaten, d. h. noch bevor sie ihre ersten Wörter äußern können, über eine frühe Semantik (Protosemantik) verfügen (vgl. Kapitel 2.3). Doch bisher ist umstritten, wie dieses frühe Wortverständnis bei Kindern im ersten Lebensjahr genau entsteht. Da die Entstehung der Referenz eine wesentliche Aufgabe im Wortlernprozess darstellt, wird in den vorgestellten theoretischen Ansätzen dieser Arbeit insbesondere die Frage nach der Lösung des Referenzproblems, welches im Folgenden näher erklärt wird, diskutiert.

In Anbetracht der Referentenauswahl besteht die herausfordernde Aufgabe eines Kindes darin, herauszufinden, auf welchen Referenten sich die empfangene sprachliche Einheit bezieht, da im Alltag meistens nicht nur ein, sondern mehrere Objekte im visuellen Umfeld liegen, die als Referenten in Frage kommen. Aber auch, wenn ein Sprecher auf einen bestimmten Referenten verweist, bleibt dennoch die Frage offen, ob sich das neue Wort auf das ganze Objekt oder lediglich auf einen Teil dessen bezieht.

Zur Verdeutlichung dieses Problems wird in der Forschungsliteratur auf das Quinean conundrum, ein von Quine (1960) formuliertes, philosophisches Gedankenexperiment, referiert. Es handelt von einem Wissenschaftler, der in ein fremdes Land reist, um die Sprache der Einheimischen zu erforschen. Dabei ereignet sich eine Situation, in der einer der Einheimischen das Wort *gavagai* ruft und gleichzeitig auf einen vorbeihüpfenden Hasen zeigt. Der Forscher übersetzt das Wort *gavagai* zunächst als *Hase*, obwohl die tatsächliche Übersetzung aus dem gegebenen Kontext nicht abzuleiten ist. Es ist also unklar, ob mit *gavagai* der Hase selbst, seine Ohren oder vielleicht nur dessen Farbe gemeint ist.

Mit diesem Gedankenexperiment intendiert Quine (1960), dass es für einen Lerner zahlreiche Möglichkeiten gibt, worauf sich ein sprachlicher Ausdruck beziehen könnte und stellte damit die Frage, wie dieses Referenzproblem gelöst wird. Auch in der Spracherwerbsforschung wurde dieser philosophische Gedanke aufgegriffen und motivierte Forscher dazu zu untersuchen, wie Kinder das Referenzproblem lösen und welche Hinweise sie nutzen, um einen Referenten zu identifizieren. In Bezug auf diese Fragestellungen, werden im weiteren Verlauf dieser Arbeit vier unterschiedliche Wortlerntheorien vorgestellt.

3 Enkodierungsstrategien im Fast-Mapping-Prozess

Nach einer Einführung in die Forschungsmethoden von Wortlernstudien werden in den nächsten Kapiteln der vorliegenden Arbeit vier Theorien vorgestellt, die das von Quine (1960) beschriebene Gedankenexperiment in Bezug auf das Referenzproblem im Wortlernprozess aufgreifen: Woher weiß ein Kind worauf ein Sprecher mit einem neuen Wort referiert? Diese Wortlerntheorien, bestehend aus dem assoziativen Erklärungsansatz, der Lerntheorie basierend auf Hypothesen einschränkenden Prinzipien, der sozio-pragmatischen Theorie sowie der interaktionistischen Sichtweise, unterscheiden sich hauptsächlich in ihren Annahmen über die Enkodierungsstrategien, die Kinder nutzen, um ein Wort mit einem Referenten zu verknüpfen. Als Enkodierungsstrategien bezeichnet Wojcik (2013, S. 1) Mechanismen, die dem Lerner helfen, das Zielwort aus dem Sprachstrom zu filtern und dafür einen Referenten zu finden.

Das Ziel der folgenden Kapitel dieser Arbeit besteht darin, eine Auswahl an bestehenden Enkodierungsstrategien aus verschiedenen Perspektiven zu reflektieren. Hierfür wird jede Wortlerntheorie zunächst anhand von zwei Studien näher erklärt. Danach schließt ein weiteres Kapitel an, welches die Aufmerksamkeitsform und den aufmerksamkeitslenkenden Mechanismus mit der jeweiligen Wortlerntheorie in Zusammenhang bringt. Außerdem wird jeder theoretische Ansatz mit einem Kapitel zur Forschungsmethode abgerundet, die charakteristisch für die jeweilige Theorie ist bzw. in der referierten Studie angewandt wurde.

3.1 Methoden in der Spracherwerbsforschung

Die Untersuchung der kindlichen Semantik ist mithilfe unterschiedlicher Methoden bei Kindern verschiedener Altersstufen möglich, vorausgesetzt, dass sie den kognitiven und motorischen Fähigkeiten der Probanden entsprechen. Das Ziel dieses Kapitels ist, Möglichkeiten aufzuzeigen, um den Forschungsgegenstand *Sprache* im Kindesalter zu messen. In den darauffolgenden Unterkapiteln zu den Wortlerntheorien wird ein besonderes Augenmerk auf die grundlegenden Paradigmen der Blickmessungsmethode gelegt.

Möchte man Sprache messen, so unterscheidet man in erster Linie zwischen *impliziten* und *expliziten* Forschungsmethoden (Bjorklund & Causey, 2018, S. 95). In einem expliziten Messverfahren kommuniziert ein Proband seine Antworten bewusst, entweder verbal oder nonverbal. Dies ist möglich, indem entweder das rezeptive oder das produktive Wortwissen angesprochen wird (S. 357). Bei ersterem muss der Proband, nachdem er ein neues Wort

hört, die Auswahl eines Referenten beispielsweise durch eine Zeigegeste oder das Greifen danach kenntlich machen (Golinkoff, Ma, Song & Hirsh-Pasek, 2013, S. 333; Hallé & Boysson-Bardies, 1994). Diese Art der Methode, in der das Antwortverhalten, zum Beispiel in Form von Gesten, „erzwungen“ wird, ist in Wortlernstudien auch als Forced Choice Task⁸ (Kucker & Samuelson, 2011, S. 5) bekannt. Als Beispiel für das explizite Messverfahren soll das Experiment 1A der Studie von Horst und Samuelson (2008) beschrieben werden, da hier sowohl das rezeptive als auch das produktive Wortwissen überprüft wurde:

In der Studie (Horst & Samuelson, 2008, Experiment 1A) wurden 24 Monate alte Kinder im Training und im Test mit mehreren Objektsätzen konfrontiert. Jedes Set bestand aus drei unterschiedlichen Objekten, von denen zwei bereits bekannt und ein drittes neu war. Die Anzahl von drei Objekten sollte die Wahrscheinlichkeit des zufälligen Handelns möglichst gering zu halten. Im ersten Test (Referent Selection Trial), welches das rezeptive Wortwissen anspricht, wurde jedes Objektsatz zweimal präsentiert und erst nach einem der beiden familiären Objekte (z. B. „Can you get the *duck*?“) und dann nach dem neuen Objekt gefragt (z. B. „Can you get the *cheem*?“). Nach diesen Aufforderungen sollten sich die Probanden für eines der drei Objekte entscheiden. Tatsächlich gelang es den Probanden das neue Wort dem korrekten Referenten zuzuordnen. Nach einer fünf-minütigen Pause wurden die Probanden nochmals getestet, um zu überprüfen, wie robust die „aufgeschnappte“ Wort-Objekt-Verbindung war. Hierfür wurde im zweiten Test (Production Task) das produktive Wortwissen überprüft, indem die Kinder nach dem Objektnamen gefragt wurden (z. B. „What’s this? What is this called?“). Allerdings konnte keines der Kinder das zuvor aufgeschnappte Wort verbal abrufen. Der dritte Test (Retention Trial) war ebenso aufgebaut wie der erste Test, allerdings lagen hier die Ergebnisse der Referentenauswahl signifikant unter dem Wahrscheinlichkeitswert. Aus dem ersten Experiment dieser Studie konnte geschlossen werden, dass Kinder eine Wort-Objekt-Verknüpfung erfolgreich „fast-mappen“, aber schon nach wenigen Minuten weder rezeptiv noch produktiv abrufen können.

Angeichts der Ergebnisse, dass auch in den anschließenden Experimenten der produktive Abruf scheiterte und der rezeptive Abruf erst durch zusätzliche Gedächtnisstützen gelang (Horst & Samuelson, 2008, Experiment 2), kann aus der Studie die Schlussfolgerung gezogen werden, dass das Testen des produktiven Wortschatzes für 24 Monate alte Kinder

⁸ *Alternative Forced Choice Task* (AFC) ist ein „erzwungenes“ Aufgabenformat, bei dem die Probanden neben dem Zielobjekt noch weitere potentielle Referenten vor sich liegen haben, die als Ablenker dienen. Die Probanden hören ein Zielwort und müssen aus der Objektauswahl einen Referenten aussuchen, indem auf ihn zeigen oder nach ihm greifen. Fragebögen mit vorgegeben Antwortmöglichkeiten zählt ebenso zur AFC-Aufgabe.

schwierig ist. Ähnliche Befunde werden auch in der Studie von Booth, McGregor und Rohlfing (2008) berichtet. Die Durchführung von produktiven Tests sei sogar mit 28-monatigen Kindern schwierig, da die damit verbundenen Ergebnisse zeigten, dass nur wenige Probanden ein neu gelerntes Wort produktiv abrufen konnten. Aus diesen Studien lässt sich schließen, dass der Abruf von produktivem Wortwissen für Kinder unter 28 Monaten kaum gelingt und daher das rezeptive Wortwissen geeigneter zu sein scheint. Hinsichtlich der Testverfahren von rezeptivem Wissen sollte beachtet werden, dass Kinder im Durchschnitt zwischen dem 9. und 12. Lebensmonat anfangen, deiktische Gesten zu produzieren (Capone & McGregor, 2004; Carpenter, Nagell & Tomasello, 1998; Leung & Rheingold, 1981; Rohlfing, Grimminger & Lüke, 2017). Daher wäre diese Art des expliziten Messverfahrens erst für Kinder ab dem ersten Lebensjahr zuverlässig.

Die zuvor genannten Ergebnisse zeigen, dass eine andere Forschungsmethode als das explizite Verfahren notwendig ist, wenn sprachliche Fähigkeiten bei Kindern vor dem ersten Lebensjahr untersucht werden sollen. Aus diesem Grund werden implizite Verfahren herangezogen, um das Wortwissen zu testen, welches ein Kind weder sprachlich noch gestisch äußern kann. Implizite Verfahren (Bjorklund & Causey, 2018, S. 95) erfassen automatische Verhaltensweisen, die keine direkte oder bewusste Kommunikation von Wissen erfordern. Durch die Messung der kindlichen Reaktion auf sprachliche oder visuelle Stimuli können Rückschlüsse über Präferenzen, Differenzierungen und auch über das frühe Sprachverständnis gezogen werden.

Ein Beispiel für ein implizites Messverfahren ist die Frequenz des Saugverhaltens (engl.: high amplitude sucking procedure) (Kauschke, 2012, S. 18). In einer Studie (DeCasper & Spence, 1986) sollten werdende Mütter in den letzten Wochen ihrer Schwangerschaft regelmäßig eine bestimmte Passage eines Textes vorlesen. Nach der Geburt lasen die Mütter ihren neugeborenen Kindern wieder dieselben sowie auch neue Passagen vor. Die Forscher stellten mithilfe eines „Nucklexperiments“ fest, dass sich das Saugverhalten während des Vorlesens der bekannten Textpassagen im Vergleich zu neuen Textpassagen unterschied. Neben der Erkenntnis, dass die auditive Kognition pränatal geformt werden kann, zeigt diese Studie auch, dass die Fähigkeit zur auditiven Differenzierung bereits bei Neugeborenen gemessen werden kann. Neben der Saugfrequenz kann bei Säuglingen auch die Richtung ihrer Kopfbewegungen (Fernald, 1985; Nelson et al., 1995), die Blickrichtung und die Blickdauer (Fantz, 1958, 1964) gemessen werden. Als eine bewährte Methode hat sich vor allem die Blickmessung erwiesen, die im Folgenden expliziert werden soll.

3.1.1 Untersuchung der Protosemantik anhand des Blickverhaltens

Die Messung des Blickverhaltens ist eine Methode, die für Kinder unterschiedlichen Alters geeignet ist. Diese Methode wird sowohl in Studien mit Neugeborenen (Aslin & Smith, 1988; Atkinson, 2000) als auch mit Kindern bis zum vierten Lebensjahr (Yurovsky & Frank, 2017) eingesetzt. Doch in der Regel sind in den meisten Wortlernstudien, die die Blickmessungsverfahren nutzen, die Probanden zwischen 6 und 12 Monaten alt (Bergelson & Swingley, 2012; Nomikou et al., 2018; Pruden, Hirsh-Pasek, Golinkoff & Hennon, 2006). Dies hat den Grund, dass zu diesem Zeitpunkt das Seh- und Hörvermögen bereits entwickelt ist, wohingegen (Zeige-) Gesten und Sprache erst ab dem ersten Lebensjahr genutzt werden. Da die rezeptiven und produktiven Modalitäten zum direkten Ausdruck der Referentenauswahl noch nicht soweit ausgereift sind, dass man sie als einen zuverlässigen Messgegenstand hinzuziehen kann, wird das rezeptive Wissen anhand des Blickmessungsverfahrens erfasst.

Noch vor einigen Jahren wurde das Blickverhalten durch Beobachtung der Probanden und das gleichzeitige Stoppen der Zeiten gemessen (z. B. Fantz, 1958). Mit dem wachsenden technischen Fortschritt kann heutzutage das Blickverhalten mithilfe von Videokameras aufgezeichnet und die Videodaten anschließend von Kodierern hinsichtlich der Blickrichtung und Blickdauer kodiert werden. Als eine noch modernere und präzisere Methode gilt jedoch das Eye Tracking (Gredebäck, Johnson & von Hofsten, 2009). Doch abgesehen davon, ob die Augenbewegungen auf Video aufgezeichnet, manuell kodiert oder mit Eye-Tracking-Systemen automatisch erfasst werden, müssen die Grenzen dieser Methode bedacht werden. Deshalb soll an dieser Stelle auf die verdeckte und offene Aufmerksamkeit (engl.: covert bzw. overt attention) hingewiesen werden.

Die verdeckte Aufmerksamkeit ist nicht messbar, da sie von der Blickrichtung unabhängig ist. Vielmehr richtet das Gehirn die Aufmerksamkeit auf periphere Zonen, die nicht im visuellen Bereich liegen. Auf eine verdeckte Aufmerksamkeit kann eine offene folgen, indem sich die Probanden zum Zielstimulus wenden (Carrasco, 2011). Für Blickmessungsmethoden kann lediglich die offene Aufmerksamkeit genutzt werden, um anhand von *Fixationen* und *Sakkaden* (vgl. Kapitel 4.1.1) Rückschlüsse über die kognitiven Prozesse ziehen zu können.

In den folgenden Unterkapiteln werden das *Preferential Looking Paradigma* sowie die Abwandlungen dieser Methode vorgestellt. Diese und auch weitere Methoden zur Erfassung des Blickverhaltens bei Kindern, auch vor dem ersten Lebensjahr, werden in einer eigenen Übersichtstabelle (vgl. Tabelle 1, S. 27–28) dargestellt.

Tabelle 1

Überblick der Blickmessungsmethoden in der Spracherwerbs- und Entwicklungsforschung

Methode	Erklärung	Untersuchungsgegenstand	Fachliteratur
Habituation/Dishabituation	Die Gewöhnung der Probanden an einen visuellen Stimulus bis deren Aufmerksamkeit (gemessen anhand der Fixationsdauer) sinkt (Habituation). Austausch des bekannten durch einen neuen Stimulus, mit der Annahme, dass die Aufmerksamkeit steigt (Dishabituation).	<ul style="list-style-type: none"> - Fähigkeit zur Diskriminierung zweier Stimuli - Gedächtnisleistung - Kategorisierungsfähigkeit - Präferenz für neue Stimuli 	Bornstein (1985); Bornstein & Benasich (1986); Fantz (1964)
Head Turn Preference	Die Probanden hören Lautsequenzen abwechselnd aus zwei Lautsprechern aus zwei Richtungen. Die Annahme lautet, dass Kinder länger in die Richtung der präferierten Lautsequenzen schauen.	<ul style="list-style-type: none"> - Phonologische Bewusstheit der Muttersprache 	Jusczyk, Hohne & Baumann (1999)
Switch Task	Die Methode gleicht der Habituation/Dishabituation mit dem Unterschied, dass auditive Stimuli verwendet werden. Die Probanden sitzen vor einem Bildschirm auf dem ein buntes Muster zu sehen ist, während sie ein und dieselbe Lautsequenz hören bis sie sich abwenden. Dann wird eine neue Lautsequenz abgespielt.	<ul style="list-style-type: none"> - Diskriminierung zweier auditiver Stimuli - Wahrnehmung von phonetischen Veränderungen 	Werker & Stager (1997)
Preferential-Looking-Paradigma (PLP)	Zwei visuell unterschiedliche, <i>statische</i> Stimuli werden ohne linguistischen Input nebeneinander präsentiert. Die Annahme ist, dass Kinder länger zu dem präferierten Stimulus schauen. Die Augenbewegungen werden entweder auf Video aufgezeichnet oder mit einem Eye-Tracking-System automatisch erfasst.	<ul style="list-style-type: none"> - Visuelle Präferenzen - Wissen über die Umwelt (familiäre vs. neue Objekte) 	Fantz (1958, 1961)
Split-Screen-Preferential-Looking-Paradigma	Eine abgewandelte Form des PLP. Im Gegensatz zum PLP sind die Stimuli <i>dynamisch</i> . Zwei Videos werden gleichzeitig auf einem Bildschirm abgespielt werden. Die Videos können linguistischen Input enthalten.	<ul style="list-style-type: none"> - Sprachverständnis - Fast Mapping - Entwicklung des Erwerbs von Nomen und Verben 	Hirsh-Pasek & Golinkoff (1996)
Intermodal-Preferential-Looking-Paradigma (IPLP)	Eine weitere Variante des PLP: Die Probanden sehen zwei statische oder dynamische Stimuli entweder auf einem Split Screen oder auf zwei Bildschirmen, während sie einen sprachlichen Input hören. Das Video enthält eine <i>Pre-</i> und <i>Post-Naming</i> -Phase. Die Annahme ist, dass Kinder sich in der Post-Naming-Phase schneller zu dem Stimulus orientieren, das zur Lautsequenz passt.	<ul style="list-style-type: none"> - Sprachverständnis - kreuzmodales Matching von Wort und Referent - grammatikalisches Verständnis - Wahrnehmung von phonologischen Veränderungen 	Bergelson & Swingley (2012); Hirsh-Pasek & Golinkoff (1996); Swingley & Aslin (2002); Tincoff & Jusczyk (1999, 2012)

Tabelle 1

Überblick der Blickmessungsmethoden in der Spracherwerbs- und Entwicklungsforschung (Fortsetzung)

Methode	Erklärung	Untersuchungs-gegenstand	Fachliteratur
Primed IPL	Es wird untersucht wie das lexikalische Netzwerk von Kindern organisiert ist und wie semantisches Priming bzw. die Anbahnung eines Reizes die darauffolgende Referentenauswahl unterstützt.	- semantisches Netzwerk	Styles & Plunkett (2009)
Looking While Listening (LWL)	In allen PLPs wird entweder die proportionale Blickzeit, die Gesamtblickdauer oder die durchschnittliche Blickdauer zu beiden Stimuli berechnet. Diese Parameter sind in der LWL-Methode zweitrangig. Hier steht die Entwicklung des Blickverlaufs im Fokus. Die Kodierung der Augenbewegung erfolgt direkt mit dem Wort-Onset.	- Sprachverarbeitungsprozess und damit verbundene Unterschiede, z. B. hinsichtlich des Alters - Dauer bis zur Fixation des Zielobjekts nach dem linguistischen Input	Fernald et al. (2008); Swingley & Aslin (2000, S. 158)
Interactive IPLP	Zwei Stimuli, die sich in ihrer perceptuellen Salienz unterscheiden, werden anstatt auf einem Bildschirm, auf einem schwarzen Brett fixiert. Der Experimentator referiert mit sozialen Signalen (linguistischer Input und Blickrichtung) zu einem der beiden Stimuli.	- Ressourcen, die Kinder im Wortlernprozess nutzen (perzeptuelle Salienz, soziale Hinweise)	Hollich et al. (2000); Pruden et al. (2006)
Mixed-Methods-Paradigma	Die in der vorliegenden Arbeit verwendete Methode setzt sich aus einer modifizierten Variante des IPLPs und dem Beobachtungsverfahren zusammen. Während des Beobachtungsverfahrens wird das interaktive Verhalten zweier Personen während einer Wortlernsituation auf Video aufgezeichnet. Im anschließenden Eye-Tracking-Test wird das enkodierte Wissen überprüft.	- Ressourcen, die Kinder im Wortlernprozess nutzen (vorhandenes Wortwissen, perzeptuelle Salienz, soziale Hinweise, Objekthandlungen) - Zusammenhang der Wortlernsituation und der Behaltensleistung eines Wortes	Wildt et al. (2019); Wildt & Rohlfing (2018)

3.1.2 Preferential-Looking-Paradigma

Robert Fantz entwickelte eine Blickmessungsmethode für Säuglinge, die nun als Preferential-Looking- (PLP) oder Visual-Preference-Paradigma bezeichnet wird, und leistete somit eine bedeutsame Pionierarbeit für weitere Paradigmen. In seinen Studien nutzte Fantz (1958, 1961) das PLP, um einen Einblick in die visuelle Verarbeitung der Umwelt bei Säuglingen zu erhalten und mögliche Präferenzen für bestimmte, visuelle

Stimuli bei Neugeborenen untersuchen. Die perzeptuellen Fähigkeiten der Kinder sowie ihr vorhandenes Wissen über die Umwelt kann anhand des Blickverhaltens gemessen werden, indem man den Probanden zwei visuelle Stimuli nebeneinander präsentiert. Dabei wird die Blickrichtung und die Blickdauer zu beiden Stimuli als abhängige Variable gemessen: Liegt eine längere Blickzeit zu einem der Stimuli bzw. eine Differenz in der Blickdauer zwischen den beiden Stimuli vor, spricht dies dafür, dass die Probanden die visuellen Reize voneinander unterscheiden können und sogar eine Präferenz für eines davon haben. In diesem Kontext wird der Begriff *Präferenz* nicht nur als besonderes Interesse oder Vorliebe für etwas aufgefasst werden, sondern gilt auch als Indiz dafür, dass Kinder bereits sehr früh visuelle und/oder auditive Reize voneinander unterscheiden können und deren Blickverhalten deshalb nicht auf einen Zufall zurückzuführen ist.

In seinen Studien mit Säuglingen nutzte Fantz das *Looking Chamber*, eine kleine Kabine, in die Säuglinge reingelegt wurden und an der Decke angebrachte, visuelle Stimuli sehen konnten. Dabei beobachtete ein Experimentator die Probanden durch einen Schlitz und stoppte die Zeit, sobald die Kinder zu einem der jeweiligen Stimuli schauten. In der Tat zeigte sich in Fantz' Studien (1958, 1961), dass Neugeborene eine Neigung für Gesichter im Vergleich zu anderen Formen haben sowie die Fähigkeit besitzen, verschiedenen Musteroberflächen, zum Beispiel zwischen schwarz/weiß-gemusterten Oberflächen und einfarbigen Oberflächen, zu differenzieren. Auch eine Präferenz für reichhaltige Muster im Vergleich zu einfarbigen Flächen hat Fantz anhand einer längeren Blickdauer der Säuglinge feststellen können.

Zunächst wurden den Säuglingen ausschließlich visuelle Reize ohne auditive Stimuli präsentiert. Doch diese Methode wurde von Spelke (1979) in einer Studie modifiziert, in der visuelle und auditive Stimuli kombiniert präsentiert wurden. Dabei wurden 4-Monatigen zwei dynamische Stimuli nebeneinander gezeigt, während gleichzeitig ein auditiver Stimulus ablief, welcher allerdings nur mit einem der beiden visuellen Reize kongruierte. Mit dieser Studie kam die Erkenntnis, dass Säuglinge visuelle und auditive Stimuli korrekt miteinander in Verbindung bringen, indem sie auf den Rhythmus sowie die Synchronität achten. Aus diesem Studiendesign entsprang die Vorstellung, dass auch das Sprachverständnis auf diese Weise untersucht werden könnte (Golinkoff et al., 2013). Somit geht der Ursprung der im Folgenden beschriebenen Paradigmen auf das ursprüngliche Preferential-Looking-Paradigma zurück.

3.2 Hypothesen einschränkende Prinzipien

Von den für die vorliegende Arbeit ausgewählten Wortlerntheorien wird als erstes der theoretische Ansatz der Hypothesen einschränkenden Prinzipien bzw. Constraints (z. B. Golinkoff, Mervis & Hirsh-Pasek, 1994; Markman, 1990; Markman & Wachtel, 1988; Markman, Wasow & Hansen, 2003; Merriman, Bowman & MacWhinney, 1989; Mervis & Bertrand, 1994) vorgestellt. Die Begriffe *Constraints* und *Prinzipien* werden meist synonym verwendet, doch um bei einer einheitlichen Benennung zu bleiben, wird im folgenden Verlauf dieser Arbeit die Bezeichnung *Prinzipien* gebraucht. Darunter wird die Einschränkung von möglichen Hypothesen, worauf sich ein neues Wort beziehen könnte, verstanden. Im Grunde handelt es sich um verschiedene Strategien, die ein Kind situationsbedingt verwendet (Golinkoff et al., 1994, S. 128), um die große Anzahl an möglichen Referenten, vor welchen es in einer Wortlernsituation steht, zu reduzieren und somit das Referenzproblem (vgl. Kapitel 2.4) zu lösen.

Hypothesen einschränkende Prinzipien haben gemeinsam, dass sie bereichsspezifisch (engl.: domain-specific) sind. Demnach verfügen Kinder über ein sprachspezifisches Wissen, welches sie nutzen, um das Referenzproblem zu lösen (z. B. He & Arunachalam, 2017). Dies gelingt ihnen, indem sie Strategien, wie z. B. das Prinzip der Ausschließbarkeit (engl. mutual exclusivity) anwenden. Samuelson und McMurray (2016, S. 4) kritisieren an der Theorie der Hypothesen einschränkenden Prinzipien, dass diese spezialisierten Strategien auf ganz bestimmte Probleme in Wortlernsituationen zugeschnitten seien. Zudem seien die Prinzipien statisch, da die Frage nach dem Entwicklungsprozess dieser Mechanismen kaum erforscht sei. Auch Golinkoff und Kollegen (1994, S. 126) kritisieren an den bisherigen Studien, dass die Prinzipien zwar einzeln untersucht wurden, aber ein einheitliches Modell, welches alle zusammenfasst und deren Entwicklungsverlauf aufzeigt, fehlt. Daher fassten sie in ihrem Modell insgesamt sechs Prinzipien zusammen und ordneten diese zwei Entwicklungsstufen zu. Während die erste Stufe grundlegende Prinzipien betrifft, die bei Kindern spätestens bis zum Ende des ersten Lebensjahres vorhanden sind (S. 150), enthält die zweite Stufe Prinzipien, die erst ab dem zweiten Lebensjahr verwendet werden.

Eine solcher Hypothesen einschränkenden Prinzipien ist zum Beispiel das des ganzen Objekts (engl.: whole object assumption) (Markman & Wachtel, 1988). Bei dieser Strategie zur Hypothesenreduzierung beziehen Kinder ein neues Wort auf ein ganzes Objekt, anstatt auf seine Einzelteile. Doch für dieses Prinzip wurden Ausnahmen festgestellt, denn es galt nur dann, als die Probanden ein neues und ein familiäres Objekt sahen. So schlossen die

Kinder zunächst das bekannte Objekt als Referent aus und bezogen dann das neue Wort auf das neue, namenlose, *ganze* Objekt. In einem zweiten Experiment der Studie von Markman und Wachtel (1988, S. 128–134) wurde eine weitere Bedingung geschaffen, in der die Kinder zwei familiäre Objekte sahen und ein neues Wort hörten. In diesem Fall wurde das neue Wort mit einem Teil und nicht mit dem ganzen Objekt verknüpft, beispielsweise mit der Flosse und nicht mit dem Fisch. Ein weiteres Prinzip ist das der Ausschließbarkeit (engl.: *mutual exclusivity*) (Halberda, 2003; Markman, 1989, 1990; Markman & Wachtel, 1988), laut dem Kinder ein neues Wort einem neuen anstatt einem familiären Objekt zuordnen. Allerdings gibt es unterschiedliche Annahmen darüber, worauf das Prinzip der Ausschließbarkeit beruht.

Eine Erklärung (Markman, 1989, 1990; Markman & Wachtel, 1988) ist die, dass Kinder annehmen, zwei Wörter können nicht genau dasselbe bezeichnen und würden ein Wort nur einem einzigen Objekt zuordnen, um lexikalische Überlappungen zu vermeiden. Allerdings berichtet Szagun (2013, S. 150) von Studien, die zeigen, dass Kinder diese Prinzipien oftmals selber verletzen und auch zwei Bezeichnungen für ein und dasselbe Objekt zulassen. Eine weitere Erklärung des Mutual-Exclusivity-Prinzips ist das N₃C (Novel-Name–Nameless-Category-Principle, vgl. Golinkoff et al., 1994, S. 142; Mervis & Bertrand, 1994). Demnach assoziieren Kinder neue Wörter mit namenlosen Objekten und greifen somit auf ihr vorhandenes lexikalisches Wissen zurück, um sich die Bedeutung neuer Wörtern zu erschließen (Houston-Price, Caloghiris & Raviglione, 2010). Im Gegensatz zu dieser Perspektive gibt es noch eine alternative Erklärung des Ausschließbarkeitsprinzips. Laut dieser Erklärung beruht die Referentenauswahl nicht auf der Namenlosigkeit, sondern auf dem Grad der Neuheit eines Objektes (Bion et al., 2013; Mather, 2013; Mather & Plunkett, 2010, 2012; Merriman, Marazita & Jarvis, 1995).

Im Folgenden soll das hypotheseneinschränkende Prinzip der Neuheitserkennung (engl.: *feeling of novelty*) als frühe Strategie zur Referentenauswahl anhand der Studien von Mather und Plunkett (2010, 2012) vorgestellt werden. Der Neuheits-Effekt bezieht sich auf die Präferenz für etwas Neues und wird meist im Rahmen von Habituations- (vgl. Tabelle 1, S. 27–28) oder Preferential-Looking-Experimenten genutzt, um neben der Strategie zur Referentenauswahl zu untersuchen, ob Kinder Objekte voneinander unterscheiden und kategorisieren können.

Die beiden Studien von Mather und Plunkett (2010, 2012) wurden für die vorliegende Arbeit als relevant erachtet, weil das Prinzip der Ausschließbarkeit, beruhend auf der Objektneuheit, bereits mit 10 Monate alten Kinder empirisch untersucht wurde. Außerdem

wurde in den beiden Studien ein Blickmessungsparadigma durchgeführt, welches charakteristisch für die Wortlernertheorie der Hypothesen einschränkenden Prinzipien ist.

3.2.1 Neuheitseffekt

Die Studie von Mather und Plunkett (2010) hatte zum Ziel zu untersuchen, welchen Ursprung das Prinzip der Ausschließbarkeit hat und ob diese Strategie auch von Kindern vor ihrem ersten Lebensjahr genutzt wird. Die Forschungsfrage lautete, ob sich die Anzahl der potenziellen Referenten reduziert, indem Objekte, deren Namen den Kindern bekannt sind, ausgeschlossen werden oder beurteilen Kinder vielmehr nach der Neuheit eines Objekts?

In ihrem ersten Experiment wurden 10 Monate alten Probanden zwei Objekte nebeneinander auf einem Bildschirm präsentiert. Während eines davon den Kindern bekannt war (z. B. ein Auto oder ein Ball), war das zweite Objekt für die Kinder neu. Gleichzeitig, als die Objekte auf dem Bildschirm präsentiert wurden, hörten die Probanden zusätzlich noch einen auditiven Stimulus. Dabei ertönte entweder ein familiäres Wort (Familiar Label Trial), ein neues Wort (Novel Label Trial) oder eine aufmerksamkeitslenkende Phrase ohne ein Zielwort (Control Trial). Die Forscher nahmen an, dass während des Novel Label Trials die Aufmerksamkeit der Kinder zum neuen Objekt steigt, sobald sie das neue Wort hören.

Zur Überprüfung ihrer Hypothese wurden über dem Bildschirm zwei Kameras angebracht, die das Blickverhalten der Kinder während des Experiments aufzeichneten. Die Auswertung der Blickdaten zeigte, dass das neue Wort tatsächlich eine längere Blickzeit zum neuen Objekt erweckte als in den anderen Bedingungen. Zudem wurde mithilfe eines Elternfragebogens⁹ das Wortverständnis für die sechs Stimuli überprüft und festgestellt, dass die Probanden die Namen für die familiären Objekte nicht kannten. Aus diesen Ergebnissen wurde geschlossen, dass Kinder bereits im Alter von 10 Monaten das Ausschließbarkeitsprinzip anwenden und dieses nicht auf dem vorhandenen Wortwissen, sondern auf der Neuheit von Objekten beruht.

Um den Einfluss des Wortwissens der familiären Objekte auf das im ersten Experiment resultierende Blickverhalten komplett ausschließen zu können, wurde ein zweites Experiment durchgeführt, in dem den Probanden zwei familiäre Objekte nebeneinander präsentiert und gleichzeitig einer der Objektamen genannt wurde. Da die 10-Monatigen in

⁹ Es wurde eine britische Adaption des MacArthur–Bates Communicative Development Inventory (MB-CDI) (Hamilton, Plunkett & Schafer, 2000) verwendet.

diesem Experiment kein Wortverständnis für die familiären Objekte zeigten, konnte bestätigt werden, dass die Kinder die Namen für die familiären Objekte nicht kannten. Somit unterstützt dieser Befund die Ergebnisse des ersten Experiments, dass die Referentenauswahl nicht aufgrund des Wortwissens für die familiären Objekte, sondern aufgrund der Objekt-Neuheit getroffen wurde.

3.2.2 Neuheitsgrad mehrerer Objekte

Anlehnend an die vorangehende Studie (Mather & Plunkett, 2010) führten Mather und Plunkett (2012) eine neue Studie mit 22 Monate alten Kindern durch. In dieser Studie bestand das primäre Ziel darin, den Ursprung des Ausschließbarkeitsprinzips anhand eines modifizierten Studien-Designs erneut zu überprüfen. Treffen Kinder eine Referentenauswahl aufgrund der Neuheit eines Objektes oder aufgrund des bestehenden lexikalischen Wissens?

Hierfür wurden statt zwei (vgl. Mather & Plunkett, 2010), drei unterschiedliche Objekte gleichzeitig auf einem Bildschirm präsentiert. Eines der Objekte sowie dessen Name war den 22 Monate alten Kindern bereits bekannt, wohingegen das zweite für die Probanden vollkommen neu und namenlos war. Mit dem dritten Objekt wurden die Kinder bekannt gemacht, sodass es familiarisiert, aber dennoch namenlos war. Während die drei Objekte präsentiert wurden und gleichzeitig ein neues Wort ertönte, wurde das Blickverhalten der Kinder mithilfe von zwei Kameras aufgezeichnet und anschließend ausgewertet, wohin und wie lange sie vor und nach dem Wort schauen.

Die Auswertung ergab, dass sobald die Probanden ein neues Wort hörten, sie den neuen, namenlosen Stimulus am längsten fixierten. Aufgrund dessen, dass die Aufmerksamkeit nur zum neuen-unbenannten Objekt anstieg, aber nicht zum familiären-unbenannten, bestand die Vermutung, dass nicht die „nameability“-Fähigkeit (S. 1160) allein für die Referentenauswahl verantwortlich war. Diese Annahme wurde in einem zweiten Experiment genauer überprüft, indem die bekannten-benannten Objekte (Schuh und Auto) ausgeschlossen wurden, sodass sich die Stimuli nur noch in dem Grad der Neuheit unterschieden. Übrig blieben somit die neuen-namenlosen und die familiären-namenlosen Objekte. Es wurde angenommen, dass die 22 Monate alten Kindern das neue Objekt, welches noch keinen Namen hat, als Referent auswählen, wenn das Ausschließbarkeitsprinzip auf der Objekt-Neuheit beruht. Im Vergleich zum ersten Experiment zeigte sich in dem zweiten Experiment, dass die visuelle Aufmerksamkeit zum neuen-namenlosen Objekt erst in dem dritten Versuch zunahm.

Als Fazit aus den beiden Experimenten zogen Mather und Plunkett (2012), dass Kinder im Alter von 22 Monaten Objekte als potenzielle Referenten ausschließen, indem sie nach dem Grad derer Neuheit urteilen. Das heißt, dass das neuste von allen Objekten als Referent für das neue Wort ausgewählt wird. In Bezug auf die Ergebnisse des zweiten Experiments erklären die Autoren, dass der Neuheits-Effekt im zweiten Experiment erst nach dem dritten Versuch auftrat, weil dieser Zeit braucht, um sich zu entfalten. Sofern aber familiäre-benannte Objekte dabei sind, kann durch den Abruf des lexikalischen Wissens die Neuheit zusätzlich verstärkt werden, das wiederum in einer schnelleren Referentenauswahl resultiert.

3.2.3 Perzeptuelle Aufmerksamkeitslenkung durch Top-down-Prozesse

Anknüpfend an die zuvor beschriebenen Studien von Mather und Plunkett (2010, 2012), ist der Neuheitseffekt ein gutes Beispiel für die top-down geleitete Aufmerksamkeit (vgl. Kapitel 2.1.2), da die visuelle Präferenz für neue Objekte auf einem bereits existierenden Wissen basiert. Das bedeutet, der Aufmerksamkeitsanstieg zu einem neuen Objekt nur dann gelingt, wenn der neue Stimulus aus dem familiären Kontext hervorsteht und dadurch die kindliche Aufmerksamkeit auf sich zieht. Die Studien von Mather und Plunkett (2010, 2012) zeigen, dass sogar junge Kinder über solch kognitive Fähigkeiten verfügen, um ihr Objektwissen aus dem Langzeitgedächtnis abzurufen und dadurch familiäre Objekte als potenzielle Referenten für neue Wörter auszuschließen.

Dem Neuheitseffekt zufolge entscheidet das vorhanden Objektwissen über die kindliche selektive Aufmerksamkeit, sodass keine sozialen Hinweise notwendig sind, um eine Referentenauswahl zu treffen. Daher wird in der Theorie der Hypothesen einschränkenden Prinzipien der sozialen Aufmerksamkeitsform, wie der Joint Attention, keine Funktion zugeschrieben.

3.2.4 Intermodal-Preferential-Looking-Paradigma

Anknüpfend an die Methode der Studien von Mather und Plunkett (2010, 2012) soll das Intermodal-Preferential-Looking-Paradigma (IPLP) vorgestellt werden, welches auf das klassische Preferential-Looking-Paradigma (vgl. Kapitel 3.1.2) zurückgeht. Eines der Ziele des IPLP ist es zu messen, ob Kinder bereits ein Sprachverständnis haben, noch bevor sie ihre ersten Wörter überhaupt äußern können.

Dieses Paradigma wurde zum ersten Mal in einer Studie mit 16-Monatigen durchgeführt, um deren Verständnis von Verben und Nomen zu untersuchen (Golinkoff, Hirsh-Pasek, Cauley & Gordon, 1987). Die Versuchsteilnehmer saßen auf dem Schoß ihrer Erziehungsberechtigten, während vor ihnen auf zwei separaten Bildschirmen zwei visuelle Stimuli präsentiert wurden, beispielsweise ein Schuh und ein Boot. Zusätzlich hörten die Kinder einen linguistischen Stimulus, z. B. „Wo ist der Schuh?“. Diese Situation wurde mit einer Videokamera, die zwischen beiden Bildschirmen platziert war, aufgezeichnet. Die längere Fixationsdauer sowie eine schnellere visuelle Orientierung zum Schuh als zum Boot zeigte, dass die Versuchsteilnehmer den visuellen Stimulus mit dem Wort korrekt verknüpft haben. In weiteren Studien stellte sich außerdem heraus, dass auch jüngere Kinder ein frühes Wortverständnis zeigten. In Studien mit 6-Monatigen, konnten die Versuchsteilnehmer die Wörter „mommy“ und „daddy“ mit dem jeweiligen Foto ihres Elternteils in Zusammenhang bringen, wenn dieses gleichzeitig mit einem Bild einer fremden Person gezeigt wurden (Tincoff & Jusczyk, 1999, 2012). Aber auch viele andere Wörter, die im Alltag häufig auftauchen, konnten Kinder im Alter von 6 bis 8 Monaten den richtigen Bildern zuordnen (Bergelson & Swingley, 2012, 2015).

In weiterführenden Studien wurde das IPLP modifiziert (vgl. Tabelle 1, S. 27–28), um Antworten auf neue Forschungsfragen zu finden. Abhängig von den Forschungsfragen ergaben sich unterschiedliche Varianten des IPLPs, so wie das LWL (Looking While Listening) Paradigma (Fernald, Zangl, Portillo & Marchman, 2008), das „primed“ IPL (Styles & Plunkett, 2009) und das IPLP mit drei, anstatt zwei Objekten (Mather & Plunkett, 2012). Die LWL-Methode testet nicht nur das, was Kinder bereits wissen, sondern auch den Verarbeitungsprozess von Wörtern, indem die Augenbewegungen kodiert werden. Das „primed“ IPL-Paradigma betrachtet, wie das lexikalische Netzwerk von Kindern organisiert ist und wie semantisches Priming bzw. die Anbahnung eines Reizes die darauffolgende Referentenauswahl unterstützt.

Zusammengefasst können mithilfe des IPLP nicht nur das frühe Sprachverständnis, sondern auch tiefere, kognitive Prozesse untersucht werden, die für die Referentenauswahl bei Kindern verantwortlich sind. Außerdem werden alle aufgeführten IPL-Paradigmen ohne das Hinzuziehen der sozialen Komponente durchgeführt. Es wird darauf zurückgegriffen, was das Kind schon weiß (bekannte Wörter oder Objekte) und wie es sich aus dem vorhandenen Wissen neues Wissen erschließt.

3.3 Assoziatives Wortlernen

Laut der assoziativen Sichtweise existiert das Referenzproblem lediglich in der Erwachsenenperspektive und nicht in der des Kindes (Samuelson & McMurray, 2016). Diese Annahme wird von Studien bekräftigt (Pereira, Smith & Yu, 2014; Yoshida & Burling, 2012), die mithilfe von mobilen Eye Trackern das visuelle Blickfeld von Kindern analysierten und dabei herausfanden, dass Kinder meistens nur ein bis zwei Objekte in ihrem visuellen Blickfeld haben. Dies hat unter anderem auch damit zu tun, dass Kinder zu diesem Zeitpunkt noch nicht laufen können, und im Sitzen bzw. beim Krabbeln nur eine bestimmte Perspektive haben. Darüber hinaus vermeiden Erwachsene das Referenzproblem für ihre Kinder, indem sie Dinge benennen, die die Kinder in dem Moment bereits fixieren (Tomasello & Todd, 1983).

Aus der assoziativen Perspektive (Golinkoff & Hirsh-Pasek, 2006; Plunkett, 1997; Samuelson & Smith, 1998; Smith, Colunga & Yoshida, 2010) entscheiden beim frühen Wortlernen nicht sprachspezifische (engl.: domain-specific), sondern allgemeine, kognitive Mechanismen (engl.: domain-general) darüber, was in den Aufmerksamkeitsfokus des Kindes gelangt. Dabei dementieren die Befürworter des assoziativen Ansatzes, dass Kinder von Anfang an mit einem spezifischen Wissen in eine Wortlernsituation treten, denn für die frühe Wortlernfähigkeit seien einfache kognitive Prozesse verantwortlich (Samuelson & McMurray, 2016). Der frühe Wortlernprozess verlange keine spezifische, kognitive Kompetenz, sondern laufe auf den „unteren“ Ebenen der Kognition ab (Hollich et al., 2000, S. 12–13, S. 85; Plunkett, 1997; Samuelson & Smith, 1998; Smith, Jones & Landau, 1996). Das bedeutet, dass zur Referentenauswahl für ein neues Wort Mechanismen verantwortlich sind, für die weder Hypothesen einschränkende Prinzipien (Hollich et al., 2000, S. 12; Samuelson & McMurray, 2016), noch soziale, referentielle Hinweise notwendig sind (Houston-Price, Plunkett & Duffy, 2006; Schafer & Plunkett, 1998). Hollich und Kollegen (2000, S. 86) argumentieren, dass ein assoziativer Wortlernprozess voraussetzen würde, dass nicht nur Wörter, sondern auch Musiknoten und Geräusche mit Objekten verknüpft werden müssten. In der Tat fanden sie heraus, dass 12 Monate alte Kinder nicht nur Wörter, sondern auch mit dem Mund erzeugte Geräusche mit neuen Objekten verknüpften (Hollich et al., 2000, S. 89; konträr zu den Ergebnissen von Marcus, Fernandes & Johnson, 2012).

Studien nach erfolgt das frühe Wortlernen, indem neue Wörter mit auffälligen Objekten aus dem visuellen Blickfeld verknüpft werden (Brandone, Pence, Golinkoff & Hirsh-Pasek, 2007; Moore, Angelopoulos & Bennett, 1999; Pruden et al., 2006; Werker,

Cohen, Lloyd, Casasola & Stager, 1998). Aus den Ergebnissen der Studie von Houston-Price et al. (2006) geht hervor, dass außerdem die Konstanz von salienten Reizen im Fast-Mapping-Prozess bei 15-Monatigen eine wichtige Rolle spielt. Das heißt, dass ein Objekt nicht nur in der Trainingsphase, sondern auch in der Testphase salient sein muss, um als Referent erkannt zu werden.

Namy (2012) argumentiert, dass die assoziative Wortlernstrategie die erste im Wortlernprozesses sei. Andere Strategien, wie das Hinzuziehen von sozio-pragmatischen und linguistischen Hinweisen nutzen Kinder zwar auch, aber erst ab dem 12. Lebensmonat (Hollich et al., 2000). So zeigen beispielsweise die Experimente von Hollich et al. (2000), dass Kinder im Alter von 12 Monaten zwar auf das mögliche referentielle Verhalten des erwachsenen Gegenübers achten, um das Referenzproblem zu lösen, aber erst ab dem 19. Monat soziale Hinweise zuverlässig nutzen, um neue Wörter mit Referenten zu verknüpfen. Für das Argument hinsichtlich der ersten Wortlernstrategie spricht außerdem, dass ein frühes Wortverständnis bei Kindern bereits im Alter von sechs Monaten zu beobachten ist (Bergelson & Swingley, 2012; Nomikou et al., 2018; Tincoff & Jusczyk, 1999, 2012), obwohl aus der Forschung hervorgeht, dass soziale Hinweise frühestens ab dem ersten Lebensjahr als Informationsquelle beachtet werden (Baldwin et al., 1996; Hollich et al., 2000; Moore et al., 1999; Woodward et al., 1994). Auf die assoziative Sichtweise deutet außerdem der Befund, dass autistische Kinder, die Schwierigkeiten damit haben referentielle Intentionen von Sprechern zu erkennen, dennoch die Sprache lernen (Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff & Tager-Flusberg, 2007; Preissler & Carey, 2005).

Im Folgenden werden zwei Studien (Hollich et al., 2000; Pruden et al., 2006) vorgestellt, in denen 10 bis 24 Monate alte Kinder getestet wurden. In diesen Studien wurden sowohl saliente als auch soziale Hinweise in eine Wortlernsituation einbracht, um deren Einfluss auf das Wortlernen zu untersuchen.

3.3.1 Das Emergentist-Coalition-Modell

Hollich und Kollegen (2000, S. 70) führten eine Reihe an Experimenten (1–10) durch, um den Fast-Mapping-Prozess von 12, 19 und 24 Monate alten Kindern zu beleuchten. Laut des von ihnen aufgestellten Emergentist-Coalition-Modells (vgl. Kapitel 2.2.1 und Kapitel 3.6) ziehen Kinder zu unterschiedlichen Zeitpunkten andere Hinweise für das Wortlernen hinzu. Im Grunde besteht das ECM aus drei Phasen, in denen die Faktoren perzeptuelle, soziale und linguistische Signale unterschiedlich stark gewichtet werden. In der ersten Phase, die

ungefähr bis zum Ende des ersten Lebensjahres andauert, entstehen Wort-Objekt-Verknüpfungen zwischen neuen Wörtern und auffälligen Objekten, die einen visuell „anspringen“. Soziale Informationen werden in dieser Phase nicht beachtet. Erst in der zweiten Phase, d. h. ab dem 12. Monat, erkennen Kinder, dass soziale Achtungssignale eine referentielle Bedeutung tragen und fangen an, diese zur Lösung des Referenzproblems zu nutzen. Ab dem 24. Lebensmonat ziehen Kinder nicht nur soziale, sondern auch linguistische Einheiten hinzu, um sich eine Wortbedeutung zu erschließen. Im Folgenden sollen die Studienergebnisse von Hollich et al. (2000) beschrieben werden. Da der Fokus der vorliegenden Arbeit auf der Entstehung von ersten Wort-Objekt-Links liegt, wird insbesondere auf die erste Phase des ECMs eingegangen.

In den ersten beiden Experimenten der Studie (Hollich et al., 2000) wurden den 12, 19 und 24 Monate alten Probanden zwei Objekte präsentiert, die sich in ihrer perzeptuellen Salienz unterschieden. In jeder Altersgruppe wurden die Teilnehmer zwei Gruppen zugeordnet: Die Kinder aus der „Coincidental“-Bedingung hörten den Namen für das saliente Objekte, wohingegen in der „Conflict“-Bedingung der Experimentator auf ein „langweiliges“¹⁰ Objekt referierte. Dabei stellten die Forscher fest, dass, obwohl die 19- und 24-Monatigen der Blickrichtung des Experimentators zum langweiligen Objekt während des Trainings in der Conflict Condition folgten, nur die älteste Altersgruppe das Wort mit dem langweiligen Objekt im Test verknüpfte und gleichzeitig den salienten Stimulus ignorieren konnte. Bei den 12-Monatigen fand man keinen Unterschied im Blickverhalten zwischen den beiden Bedingungen, da sowohl in der Coincidental- als auch in der Conflict-Bedingung das saliente Objekt während der Salienz-, Trainings- und Testphase stets länger fixiert wurde (S. 48). Für die beiden älteren Probandengruppen wurde das Fazit gezogen, dass sie sowohl die Blickrichtung als auch die perzeptuelle Salienz als Achtungshinweise zur Referentenauswahl beachten. Letztendlich schafften es aber nur die 24 Monate alten Kinder, sich von der Objektsalienz loszulösen und soziale Hinweise für das Wortlernen zu nutzen. Da bei der jüngsten Probandengruppe beobachtet wurde, dass sie stets eine Präferenz für das saliente Objekt zeigte, wurde vermutet, dass sie der Objektsalienz, im Vergleich zur Blickrichtung, Vorrang gewähren (Hollich et al., 2000, S. 50–51). Es blieb allerdings unklar, ob die längere Blickdauer ein Indiz dafür war, dass das Zielwort mit dem salienten Objekt verknüpft wurde oder ob das gemessene Blickverhalten vom Zielwort unabhängig war.

¹⁰ „Perceptual salience was manipulated by using two objects of unequal salience. The ‘interesting’ object was brightly colored and had moving parts. The ‘boring’ object was dull beige and did not have moving parts“ (Hollich et al., 2000, S. 42).

Aus diesem Grund wurden das Blickverhalten der 12 Monate alten Kinder in einem weiteren Experiment (3) nochmal untersucht. Dabei wurde die ursprüngliche Methode der beiden ersten Experimente verändert, indem neben dem ursprünglichen Test, der nach dem Zielwort fragte, zwei zusätzliche Tests („New Phase“ und „Recovery“) eingebaut wurden. Die Annahme lautete, dass eine Wort-Objekt-Verknüpfung erfolgreich gelernt wurde, wenn die Probanden im ersten Test das Zielobjekt am längsten anschauen. Da in der New-Testphase ein komplett neues Wort fällt, wurde angenommen, dass die Probanden daraus schließen können, dass es sich um das andere, unbenannte Objekt handelt und werden dieses länger fixieren. In der Recovery-Testphase, in der wieder das Zielwort fällt, sollte eine längere Blickzeit zum Zielobjekt vorliegen. Die Ergebnisse (S. 57) zeigten, dass die Probanden der Coincidental-Bedingung im ersten Test, in dem das saliente Objekt benannt wurde, wie erwartet das saliente Zielobjekt am längsten fixierten. In der New Phase nahm die Blickzeit zum salienten Objekt ab und war nicht mehr signifikant. In der Recovery-Testphase zeigte sich wieder eine längere Blickzeit zum Zielobjekt.

Dieses Blickverhalten deuten die Autoren als Hinweis darauf, dass die Probanden das Wort mit dem salienten Zielobjekt verknüpft haben und in der New Phase erkannten, dass nach einem anderen Referenten gefragt wird, wodurch sich die Blickzeit reduzierte. In der Conflict-Bedingung fand sich keine signifikant höhere Blickdauer zum salienten Objekt während der Trainingsphase, allerdings war die Fixationsdauer zum salienten Ablenker in allen Testphasen signifikant höher im Vergleich zum Wahrscheinlichkeitswert von 50 %.

Aus diesen Ergebnissen schlussfolgerten Hollich et al. (2000), dass junge Kinder sowohl auf soziale als auch auf perzeptuell saliente Hinweise achten, aber noch nicht in der Lage sind zu bewerten, welche dieser Informationsquellen die verlässlichere ist. Somit besteht die Herausforderung am Anfang des Wortlernprozesses für junge Kinder darin, zwischen der perzeptuellen Salienz und den sozialen Hinweisen zu entscheiden. Das Ergebnis, dass die Conflict-Gruppe zwar die Blickrichtung des Experimentators während des Trainings berücksichtigt haben, die Präferenz für das saliente Objekt im Test aber dennoch bestand, erklären Hollich und Kollegen (2000, S. 61) wie folgt:

The failure of 12-month-olds to show any evidence of word learning in the conflict condition of Experiment 3 suggests that the cue of perceptual salience and the cue of social eye gaze must both "point" to the same object. When they do not, word learning becomes very difficult for these infants, who are unsure of which cue are more reliable in the word learning situation.

Demnach würde das Wortlernen für 12-monatige Kinder am besten funktionieren (S. 59), wenn Erwachsene darauf achten, solche Objekte benennen, die bereits im Aufmerksamkeitsfokus von Kindern liegen (vgl. Dunham, Dunham & Curwin, 1993).

Die folgenden Experimente (4–11) von Hollich et al. (2000, vgl. Tabelle 4, S. 65) waren motiviert durch die bisherigen Erkenntnisse über die Wortlernstrategie von 12-Monatigen. Ihre weiterführenden Fragestellungen lauteten: Welches Blickverhalten zeigen Kinder, wenn sich zwei Objekte, die ihnen gezeigt werden, in ihrer Salienz nicht unterscheiden? Und: Welche sozialen Hinweise sind für 12-Monatige in diesem Kontext für ein Wort-Objekt-Mapping notwendig (S. 62 ff.)? Zur Untersuchung dieser Fragen wurde die perzeptuelle Objektsalienz eliminiert, um lediglich die Wirkung der sozialen, referentiellen Hinweise zu untersuchen.

In dem Experiment 4 stellte sich bereits heraus, dass die Blickrichtung allein nicht ausreicht, damit 12-Monatige ein neues Wort einem Referenten zuordnen, auch wenn sich die Objekte in ihrer Salienz nicht unterscheiden. Aus diesem Grund wurde in den weiteren Experimenten die Blickrichtung mit weiteren sozialen Hinweisen kombiniert. Erst die Kombination aus vielfältigen sozialen Hinweisen führte zu einer längeren Blickzeit zum Zielobjekt im Training und im Test (vgl. Experiment 7, 8, 10). Solche Zielobjekt-Präsentationen bestanden aus der Blickrichtung zum Zielobjekt, der Objektmanipulation (Drehen des Zielobjektes) und dessen Benennung. Daraus wurde deutlich, dass junge Kinder, im Vergleich zu älteren Kindern, unterschiedliche referentielle Hinweise zum Objekt sowie genügend Auseinandersetzung mit dem Zielwort (mehrfache Benennung oder längere Benennungszeit) benötigen, um den Referenten sowie das Zielwort zu enkodieren und miteinander zu verknüpfen.

Die wesentliche Erkenntnis der Studie von Hollich et al. (2000) ist, dass der Fast-Mapping-Prozess bei 12-, 19- und 24-Monatige unterschiedlich abläuft. Experiment 1 und 2 zeigten, dass 24-Monatige sich von der Objektsalienz nicht ablenken lassen und die Blickrichtung des Gegenübers für das Fast-Mapping ausreicht. 19-Monatige sind hingegen sensibel für die Blickrichtung des Kommunikationspartners, aber erachten diese noch nicht als sichere Quelle, sondern achten vielmehr auf die Objektsalienz. Bei 12 Monate alten Kindern ist die Blickrichtung als Informationsquelle noch gar nicht etabliert und muss daher mit anderen sozialen Hinweisen kombiniert werden, damit ein Referent in den Aufmerksamkeitsfokus gelangt. Hollich et al. (2000, S. 83) weisen aber darauf hin, dass die Aufmerksamkeit

von 12-Monatigen zum Zielobjekt während der Trainingseinheiten kein verlässlicher Indikator dafür sei, dass Kinder dieses Objekt im darauffolgenden Test als Referent auswählen: „Referential sensitivity during labeling is not the same thing as referential understanding“.

3.3.2 Perzeptuelle Objektsalienz

Pruden und Kollegen (2006) veröffentlichten eine Studie, in der sie die Befunde von Hollich et al. (2000) hinsichtlich der unterschiedlichen Gewichtung perzeptueller, sozialer und linguistischer Hinweise zu verschiedenen Alterszeitpunkten aufgriffen. Sie nahmen an, dass für Kinder im Alter von zehn Monaten die Objektsalienz für das Wortlernen relevanter sei als soziale Hinweise. Während ab dem 12. Lebensmonat Kinder sensibel seien für soziale Hinweise als Informationsquelle, sei vor dem ersten Lebensjahr diese Sensibilität noch gar nicht vorhanden.

Der Ablauf dieser Studie ähnelte der Studie von Hollich und Kollegen (2000). Den Probanden wurden jeweils zwei Objekte gleichzeitig gezeigt, die sich in ihrem Grad der Salienz unterschieden. Während die salienten Objekte in ihrer Beschaffenheit und Farbe auffällig waren, sahen die anderen zwei Objekte damit verglichen langweilig aus. Während der Trainingsphase stand ein Experimentator hinter dem Objektset und benannte entweder das saliente (Coincidental Condition) oder das langweilige Objekt (Conflict Condition). Der Konflikt in der zweiten Bedingung bestand darin, dass die Kinder die Herausforderung meistern mussten, ihre Aufmerksamkeit von dem auffälligen Objekt zu lösen und stattdessen den langweiligen Referenten zu fixieren. In der Coincidental-Bedingung wies der Experimentator mit seiner Blickrichtung auf das saliente Objekt als Referenten hin, sodass die Salienz und der soziale Hinweis übereinstimmten. Ebenso wie in der Studie von Hollich und Kollegen (2000) bestand auch der Test aus drei Testphasen, dem „Original Label“, „New Label“ und „Recovery Trial“.

Aus den Ergebnissen ging hervor, dass die 10 Monate alten Studienteilnehmer in beiden Bedingungen das neue Wort mit einem salienten Objekt verknüpften. Das bedeutet, dass in der Coincidental-Bedingung der entstandene Wort-Objekt-Link zwar korrekt war, aber es in der Conflict-Bedingung zum „Mismatching“ kam. Obwohl in dieser Gruppe das langweilige Objekt benannt wurde, wurde das neue Wort dennoch dem salienten Objekt zugeordnet. Pruden und Kollegen (2006, S. 274) schlussfolgern aus diesen Ergebnissen, dass die erste Wortlernstrategie assoziativ sei. Demnach handeln Kinder perzeptionsgeleitet und assoziieren neue Wörter, die sie hören, mit dem salientesten Objekt aus dem visuellen Blickfeld und ignorieren die sozialen Hinweise.

In den beiden beschriebenen Studien (Hollich et al., 2000; Pruden et al., 2006) liegt der Fokus auf dem Lernen von Wort-Objekt-Verknüpfungen, d. h. auf Nomen. Eine weiterführende Studie (Brandone et al., 2007) hat die assoziative Wortlerntheorie auf das Lernen von Verben ausgeweitet hat. Auch hier wurde die Wirkung der Salienz im Vergleich zu sozialen Hinweisen auf das Wortlernen untersucht, wobei die Salienz als eine Handlung, aus der ein Resultat hervorgeht, operationalisiert wurde. In der Studie stellte sich heraus, dass das ECM nicht nur auf Nomen, sondern auch auf das Lernen von Verben zutrifft. Anhand von Blickdaten konnte festgestellt werden, dass 21–24 Monate alte Kinder Verben perzeptionsgeleitet lernen und sozialen Hinweisen erst zwischen dem 33. und 36. Lebensmonat mehr Gewichtung als der Salienz beimessen. Das bedeutet, dass der Übergang von der ersten zur zweiten Phase im ECM im Erwerb von Verben später stattfindet, als im Erwerb von Nomen.

3.3.3 Perzeptuelle Aufmerksamkeitslenkung durch Bottom-up-Prozesse

Die in den Kapiteln 2.1.1 und 2.1.2 bereits eingeführten Aufmerksamkeitsformen und aufmerksamkeitslenkenden Mechanismen sollen nun in Zusammenhang mit der assoziativen Wortlerntheorie gebracht werden. Aus den Studien (Hollich et al., 2000; Pruden et al., 2006) ging hervor, dass Kinder im ersten Lebensjahr sozialen Hinweisen keine Beachtung schenken, da sie diese noch nicht als Informationsquelle wahrnehmen. Das bedeutet, dass die geteilte Aufmerksamkeitsform (Joint Attention) für den assoziativen Wortlernprozess keine Rolle spielt. Stattdessen werde die frühkindliche Aufmerksamkeit stark von salienten Stimuli gesteuert, was in Bezug auf die Aufmerksamkeitsform bedeutet, dass eine nicht-strategische, selektive Aufmerksamkeit (Smith et al., 1996, S. 145) bestimmt, welche Wort-Referent-Verbindung in den weiteren Wortlernprozess gelangt. Da die selektive Aufmerksamkeit von der perzeptuellen Salienz bestimmt wird, ist der hierfür verantwortliche Mechanismus demnach bottom-up-gesteuert. Des Weiteren verdeutlichen die Studien auch, dass die obligatorische Aufmerksamkeit im frühen Alter noch nicht so weit entwickelt ist, dass Kinder ihre Aufmerksamkeit intrinsisch von einem salienten Objekt lösen können.

Den vorgestellten Studien ist außerdem zu entnehmen, dass die perzeptuelle Salienz, die auf für den Bottom-up-Prozess verantwortlich ist, unterschiedlich operationalisiert wurde (für einen Überblick vgl. Wildt, Rohlfing & Scharlau, 2019). In der Studie von Hollich et al. (2000, S. 45) und Pruden et al. (2006) wurden Objekte als salient bezeichnet, die farblich aufmerksamkeitslenkend waren und entweder ein Geräusch machten oder sich bewegten. Zudem wurden sie zusammen mit langweiligen Alltagsgegenständen präsentiert, zum

Beispiel einem Flaschenöffner, wodurch sie noch auffälliger wirkten. In anderen Studien zeichnete sich die Salienz eines Objektes dadurch aus, dass das jeweilige Objekt größer als die anderen Objekte war (Pereira et al., 2014; Smith et al., 1996) oder im Blickfeld der Kinder dominierte. In wiederum anderen Experimenten wurde die Salienz eines Objektes hervorgehoben, indem es sich auf einer Drehscheibe drehte (Moore et al., 1999) oder sich auf dem Bildschirm bewegte (Houston-Price et al., 2006), während das daneben präsentierte Objekt statisch war. Eine weitere Operationalisierung der Objektsalienz ist die des Beleuchtens. So haben beispielsweise Axelsson und Kollegen (2012) das Zielobjekt in seiner Salienz verstärkt, indem es von unten beleuchtet wurde. Eine noch stärkere Salienz erreichten sie, indem sie gleichzeitig die zwei anderen Objekte mit einer Box abdeckten.

Yu und Smith (2012, S. 258) argumentieren, dass ein bottom-up-gesteuerter, sensorisch „sauberer“ Input zum frühen Wortlernen führe, da ein Referent durch das intrinsische Interesse ausgewählt werde und dadurch in den Aufmerksamkeitsfokus rücke. Demnach hängen das kindliche Interesse und das Lernen zusammen, weil durch das Interesse die Aufmerksamkeit auf etwas gerichtet wird. Auch Pruden et al. (2006, S. 276) plädieren dafür, dass 10-Monatige rein nach ihrem eigenen Interesse handeln und deshalb neue Wörter mit salienten Objekten verknüpfen, anstatt auf soziale Hinweise zu achten. Das wiederum bedeutet, dass Kinder meist selber solche Momente schaffen, indem ein bestimmtes Objekt, aufgrund seiner Salienz ins Blickfeld rückt. Aus assoziationistischer Perspektive sei es deshalb hilfreich, genau in solchen Momenten ein Objekt zu benennen, in denen die Aufmerksamkeit von Kindern bereits auf dieses gerichtet ist, um das Wort-Objekt-Mapping zu fördern (Dunham et al., 1993; Hollich et al., 2000; Pereira et al., 2014; Yu & Smith, 2012).

3.3.4 Interactive-Intermodal-Preferential-Looking-Paradigma

Eine modifizierte Variante des PLPs (vgl. Kapitel 3.1.2) und des IPLPs (vgl. Kapitel 3.2.4) ist das Interactive-Intermodal-Preferential-Looking-Paradigma (IIPLP), welches in zahlreichen Wortlernstudien (z. B. Brandone et al., 2007; Golinkoff et al., 2013; Pruden et al., 2006; Hollich et al., 1998; Hollich et al., 2000, S. 30–40) genutzt wurde. Im Folgenden sollen die Unterschiede zu den bisherigen Paradigmen dargestellt und damit auch die Vorteile des IIPLP verdeutlicht werden. Anschließend wird der Ablauf der Methode beschrieben.

Während das PLP visuelle Präferenzen und die Differenzierungsfähigkeit erfasst, geht das IPLP einen Schritt weiter und misst die kindliche Sensibilität für das lexikalische und syntaktische Wissen, indem nicht nur visuelle, sondern auch linguistische Stimuli präsentiert

werden. Mithilfe dieser Methode fand man unter anderem heraus, dass Kinder vor der Äußerung ihrer ersten Wörter über ein großes Sprachverständnis verfügen (Bergelson & Swingley, 2012, 2015; Golinkoff et al., 1987; Tincoff & Jusczyk, 1999, 2012). Diese Erkenntnis führte zu der weiterführenden Forschungsfrage, wie Kinder zu diesen ersten Wortbedeutungen kommen. Somit unterscheidet sich das IIPLP von den vorherigen Methoden insofern, als es nicht nur untersucht, *was* Kinder bereits verstehen, sondern auch *wie* sie dieses Wissen erwerben.

Das IIPLP ist auf Grundlage des ECMs entstanden, das besagt, dass Kinder bis zum 12. Lebensmonat lediglich visuell saliente Stimuli für das Fast Mapping nutzen und erst später auch soziale sowie linguistische Stimuli als zuverlässige Quellen erkennen und auf diese zurückgreifen. Der Vorteil des IIPLPs besteht vor allem in seiner flexiblen Anwendungsweise, da es nicht nur für unterschiedliche Altersstufen geeignet ist, sondern auch eine Variation der unabhängigen Variablen erlaubt (Hollich et al., 2000, S. 32). Die Methode der Blickmessung ist sowohl bei jungen Kindern im Alter von 10 Monaten als auch bei älteren Kindern ab dem 24. Lebensmonat möglich. In Bezug auf die Variation von Variablen kann beispielshalber die Salienz unterschiedlich operationalisiert worden. Dasselbe gilt aber auch für soziale Hinweise, die auf verschiedene Arten, z. B. durch die Blickrichtung, Zeigegesten oder Objektmanipulation, ausgeführt werden können. Charakteristisch für das IIPLP sind außerdem die realen Objekte und ein Experimentator, die eine „alltägliche“ Lernsituation schaffen (S. 32), aber nichtsdestotrotz eine hohe experimentelle Kontrolle gewähren.

Für die methodische Durchführung wird ein Tisch aufgestellt, auf dem ein schwarzes Brett aufgebaut wird (S. 34). Auf diesem Brett werden jeweils zwei Objekte präsentiert, von denen aber nur eins vom Experimentator benannt wird. Währenddessen sitzt das Kind auf dem Schoß eines Elternteils, ungefähr 75 cm von dem Tisch entfernt. Hinter dem Tisch, auf dem sich die beiden Stimuli befinden, steht eine Trennwand, hinter der eine Videokamera versteckt ist. Sie dient dazu, das Blickverhalten der Probanden aufzuzeichnen. Außerdem wird ein Spiegel hinter den Probanden aufgestellt, worin sich die Objekte spiegeln und dadurch auch auf dem Videomaterial zu sehen sind. Dies ermöglicht den Kodierern zu bestimmen, auf welches der beiden Objekte die Probanden geschaut haben. Die Probanden sitzen in allen Phasen auf dem Schoß der Bezugsperson und beobachten das Geschehen vor ihnen, während ihr Blickverhalten auf Video aufgezeichnet wird.

Das klassische IIPLP besteht aus der „exploration“, „salience“, „training“ und „test phase“ (Hollich et al., 2000, S. 34). Während der Explorationsphase dürfen die Kinder zunächst jedes der beiden Objekte nacheinander für 26 Sekunden explorieren. Hiernach folgen

zwei Salienzphasen, während der sich der Experimentator hinter dem schwarzen Brett versteckt und fragt „Eve, look at the board! What do you see?“ (vgl. Figure 4 in Hollich et al., 2000, S. 35). Diese Phase gilt als „Baseline“, wodurch das allgemeine Interesse in beide Objekte gemessen werden soll, um zu bestimmen, ob eines davon tatsächlich salienter ist als das andere. Hiernach folgt die Trainingsphase, während der Experimentator, hinter dem schwarzen Brett stehend, mit sozialen Signalen auf eines der beiden Objekte referiert und es gleichzeitig mehrmals benennt. Insbesondere die Trainingsphase kann in ihrer Durchführung von Studie zu Studie stark variieren, je nachdem welche Altersstufe und welche sozialen Signale als referentielle Hinweise überprüft werden sollen. Anschließend versteckt sich der Experimentator wieder komplett hinter dem Brett und fragt das Kind: "Eve, where's the modi? Do you see the modi?". Dabei wird kodiert und ausgewertet, auf welches der beiden Objekte die Kinder länger schauen, da eine längere Blickzeit als Indiz für das Wortverständnis gedeutet wird.

Bevor diese Methode eingesetzt wurde, um den Mapping-Prozess von neuen Wörtern zu untersuchen, führten Hollich und Kollegen (2000, S. 39) eine Pilotstudie mit 12, 19 und 24 Monate alten Kindern und mit familiären Wörtern durch. Alle Altersgruppen demonstrierten eine signifikant längere Blickzeit zu dem korrekten Referenten, sobald sie das Zielwort hörten. Den Ergebnissen der Testphase zufolge, hat sich das Studiendesign mit den soeben beschriebenen Phasen als eine geeignete Methode herausgestellt, um das frühe Sprachverständnis zu testen.

3.4 Sozio-pragmatisches Wortlernen

Die bisher vorgestellten Ansätze der Hypothesen einschränkenden Prinzipien (vgl. Kapitel 3.2) und des assoziativen Wortlernens (vgl. Kapitel 3.3) erklären das frühe Wortlernen als einen mechanistischen Prozess (Plunkett, 1997; Samuelson & Smith, 1998; Smith et al., 1996; Spencer, Perone, Smith & Samuelson, 2011), für den keine sozialen Hinweise notwendig sind. Andere Forscher sind wiederum der Auffassung, dass das Wortlernen in einem sozio-kommunikativen Kontext stattfindet (Tomasello, 2001). Dem sozio-pragmatischen Ansatz nach entstehen neue Wort-Referent-Verknüpfungen, indem Kinder soziale Hinweise von anderen Personen beachten (Carpenter et al., 1998; Tomasello, 1999).

Tomasello (2000, S. 404) kritisiert an der assoziativen Sichtweise, dass die sozio-pragmatischen Fähigkeiten der Kinder komplett außer Acht gelassen werden, obwohl Studien zeigen, dass Kinder von Geburt an sensibel dafür sind. Des Weiteren argumentiert er, dass

auch die Hypothesen einschränkenden Prinzipien als Wortlernstrategie überflüssig seien, da Kinder durch die Partizipation in einer triadischen Interaktion einen Referenten bestimmen können (S. 402). Tomasello (2000, S. 404) weist außerdem darauf hin, dass die Prinzipien lediglich erklären können wie Nomen gelernt werden. Wie das Lernen von Verben und Adjektiven abläuft, werde aus diesem Ansatz kaum untersucht. Weiterhin kritisiert er an dem assoziativen sowie am „Constraints“-Ansatz, dass diese nicht auf die Frage eingehen, warum Kinder ab dem Ende des ersten bzw. Anfang des zweiten Lebensjahres anfangen ihre ersten Wörter zu produzieren, wenn doch Wortbedeutungen durch mechanistische Gedächtnisprozesse schon viel früher gelernt werden (S. 405). Die sozio-pragmatische Theorie begründet diesen Zeitpunkt als Startschuss für das Wortlernen damit, dass sich erst im Alter von 9–12 Monaten die Fähigkeit zur geteilten Aufmerksamkeit (Joint Attention) entwickelt (S. 406). Ab diesem Zeitpunkt fangen Kinder an, andere Menschen als intentional handelnde Wesen zu begreifen (vgl. Woodward, 2003), die linguistische Symbole nutzen, um in Situationen, in denen der Aufmerksamkeitsfokus auf ein Objekt von dem Erwachsenen und dem Kind geteilt wird, Wissen zu vermitteln.

Anhand einiger Beispiele aus entwicklungspsychologischen Studien soll verdeutlicht werden, dass Kinder von Geburt an soziale Wesen sind, die auf andere Menschen reagieren und diese als intentional handelnd ansehen. Dabei soll nicht nur auf die visuelle, sondern auch auf die auditive Wahrnehmung von sozialen Signalen eingegangen werden. Die Forschung auf dem Gebiet der frühkindlichen Perzeption hat gezeigt, dass Kinder ein besonderes Interesse für Gesichter (Maurer, 1985) bzw. gesichtsähnliche Formen (Fantz, 1963) und insbesondere für Augen mitbringen.

Die Präferenz für Augen zeigte sich beispielsweise in einer Studie von Farroni, Csibra, Simion und Johnson (2002) mit 17 Neugeborenen (2–5 Tage alt), denen zwei Bilder nebeneinander präsentiert wurden. Auf dem ersten Bild war ein menschliches Gesicht abgebildet, dessen Augen bzw. Blick geradeaus auf den Säugling gerichtet waren bzw. war, während die Augen auf dem zweiten Bild in eine andere Richtung schauten. Die Forscher fanden heraus, dass Neugeborene sensibel für den direkten Augenkontakt waren, da sie signifikant länger das Gesicht fixierten, welches sie direkt anschaute. Ein zweites Experiment von Farroni und Kollegen (2002) unterstrich durch ERP-Untersuchungen¹¹ die vorherigen Er-

¹¹ Unter ERP (ereigniskorrelierte Potenziale, engl.: event-related potential) versteht man Wellenformen, die im Elektroenzephalogramm (EEG) aufgezeichnet werden und stark mit kognitiven Prozessen korrelieren (Müller, 2013).

gebnisse. Den Befunden nach verarbeitet das Gehirn von vier Monate alten Säuglingen, Gesichter mit direktem Augenkontakt am besten. Einer Studie (Reid et al., 2017) zufolge, bestehe eine Präferenz für gesichtsähnliche Formen bereits pränatal. Hierfür wurden 4D-Ultraschall-Untersuchungen angewandt, mithilfe derer die Forscher herausfanden, dass Föten sich häufiger zu einem, durch drei Lichtpunkte dargestellten Gesichtsmuster, als zu einem „umgedrehten“ Gesichtsmuster drehen. Aus der aktuellsten Forschungsliteratur geht hervor, dass Kinder bereits im Alter von vier Monaten auf Augen, die in Richtung eines Objekts schauen, mit einer stärkeren Hirnaktivität (ERP) reagieren als auf ein augenähnliches Schachbrett-Muster (Michel et al., 2019). Darüber hinaus ruft direkter Augenkontakt in einer realen Interaktion zwischen neun Monate alten Kindern und ihren Bezugspersonen eine erhöhte Hirnaktivität bei Kindern hervor (Höhl et al., 2014) im Vergleich zu Interaktionen, in denen kein direkter Augenkontakt aufgebaut wurde.

Säuglinge zeigen nicht nur für den direkten Augenkontakt eine Empfindsamkeit, sondern auch für die Blickrichtung anderer (Muir, Hains, Cao & D’Entremont, 1996). Bereits ab dem dritten Monat sind sie in der Lage, den Blicken anderer Personen zu einem Objekt zu folgen (Gredebäck, Astor & Fawcett, 2018; Hood, Willen & Driver, 1998; Morales, Mundy & Rojas, 1998; Senju & Csibra, 2008; Szufnarowska, Rohlfing, Fawcett & Gredebäck, 2015a; Woodward, 2003). Yurovsky und Frank (2017) weisen darauf hin, dass es jedoch für junge Kindern schwierig ist, sich von dem Gesicht eines Sprechers zu lösen und der Blickrichtung auf ein Zielobjekt zu folgen. Diese Fähigkeit entfalte sich aber im Laufe der Entwicklung, indem Kinder lernen die gemeinsamen Aufmerksamkeitsbezüge zu koordinieren (vgl. Bruner, 1983).

Neben den sozialen Sensibilitäten für visuelle Reize haben Kinder auch Präferenzen für bestimmte auditive Signale. Da bereits pränatal Außengeräusche gedämpft wahrgenommen werden, wird argumentiert, dass ein Kind nicht als „unbeschriebenes Blatt“ zur Welt kommt, sondern dadurch gewisse Neigungen für bestimmte Laute mitbringt (DeCasper & Spence, 1986). Zum Beispiel wird die Sprache der Mutter gegenüber anderen weiblichen Stimmen präferiert (DeCasper & Fifer, 1980). In der Studie von DeCasper und Spence (1986) zeigte sich, dass Neugeborene in der Lage sind, sprachliche Muster, die sie pränatal gehört haben, wiederzuerkennen. Darüber hinaus können Neugeborene ihre Muttersprache nicht nur von anderen Sprachen zu unterscheiden (Mehler et al., 1988), sondern präferieren diese sogar (Moon, Cooper & Fifer, 1993). Einer weiteren Studie zufolge schreien Neugeborene in der Satzmelodie der Muttersprache (Mampe, Friederici, Christophe & Wermke, 2009). Eine ebenso beachtliche Fähigkeit bei Neugeborenen ist die

Präferenz für die menschliche Sprache im Vergleich zur Vokalisation von Affen (Vouloumanos, Hauser, Werker & Martin, 2010) sowie die Präferenz für Wörter gegenüber anderen Tönen (Vouloumanos & Werker, 2004, 2007).

Neben den Experimenten, die die Sensibilität von Kindern für soziale Hinweise belegen, gibt es auch eine Reihe an Studien (Akhtar & Tomasello, 2000; Baldwin, 1991a, 1993; Booth et al., 2008; Esseily & Fagard, 2013; Tomasello, 2000; Wu, Tummeltshammer, Gliga & Kirkham, 2014), die die Wirkung von sozialen Hinweisen auf das Wortlernen getestet haben. Im Folgenden werden zwei sozio-pragmatische Studien zum Wortlernen vorgestellt werden. Die Studie von Baldwin (1991) wurde ausgewählt, da sie das klassische methodische Vorgehen in einer sozio-pragmatischen Lernsituation darstellt. Die Studie von Yurovsky und Frank (2017) soll als Beispiel für eine aktuelle Studie dienen, in der die Eye-Tracking-Methode angewendet wurde.

3.4.1 Referentielles Verständnis

Baldwin (1991, S. 877) thematisierte in ihrer Studie das Problem, dass zwar erforscht wäre, dass Kinder im Alter von 9 bis 12 Monaten in der Lage sind, sozialen Hinweisen (z. B. der Blickrichtung oder Zeigegesten) zu folgen, aber dennoch nicht eindeutig klar sei, ob sie solch nonverbalen Hinweise auch zum Fast Mapping nutzen. Ihre Annahme lautete, dass durch die Beachtung von sozialen Hinweisen korrekte Wort-Objekt-Zuordnungen entstehen und falsche Wort-Objekt-Verknüpfungen vermieden werden können.

Für die Studie wurden 16–19 Monate alte Kinder entweder der *Follow-in-Labeling*- oder der *Discrepant-Labeling*-Bedingung zugeteilt. Während mit der *Following-in-Labeling*-Bedingung gezeigt werden sollte, dass soziale Hinweise für das Wortlernen relevant sind, sollte mithilfe der *Discrepant-Labeling*-Bedingung eingeräumt werden, dass das frühe Wortlernen assoziativ abläuft, indem Kinder lediglich auf die temporäre Kontinuität von Wörtern und Objekten achten.

In der Trainingsphase saßen die Probanden mit einem Experimentator am Tisch, während dieser entweder ein Objekt benannte, welches die Probanden in dem Augenblick fixierten (*Follow-in Labeling Condition*) oder eins, welches nicht im visuellen Blickfeld des Kindes lag (*Discrepant Labeling Condition*). In der ersteren Bedingung holte der Experimentator zwei neue Objekte aus einer Box heraus und legte diese auf den Tisch. Danach wurde eines der beiden Objekte wieder zurück in die Box gepackt. Der Experimentator nahm das übrig gebliebene Objekt in die Hand und zeigte dem Kind, wie man dieses Objekt manipu-

lieren kann, woraufhin er es dem Kind reichte. In dem Moment, in dem das Kind das Spielzeug anschaute, benannte der Experimentator dieses. Dabei schaute der Experimentator im Wechsel entweder das Objekt oder das Kind an. Im Gegensatz zu dieser Bedingung schaute der Experimentator die Probanden der Discrepant-Labeling-Gruppe nicht an, sondern schaute mit gesenktem Blick in die Box auf das weggelegte Spielzeug. Durch einen Seitenspiegel beobachtete der Experimentator das Blickverhalten des Kindes. Er wartete ab, bis das Kind das auf dem Tisch liegende Objekt fixierte und äußerte dann erst den Objektnamen. In der anschließenden Testphase hörten die Probanden das Zielwort und wurden aufgefordert auf den Referenten zu zeigen. Zur Kontrolle wurde neben der Verständnisfrage, zusätzlich eine Präferenzfrage gestellt (z. B.: „Zeig auf das Objekt, das dir gefällt“).

Der gesamte Ablauf des Experiments wurde auf Video aufgezeichnet und kodiert. Die Auswertungen ergaben, dass die Follow-in-Labeling-Gruppe eine korrekte Referentenauswahl im Verständnistest getroffen hat. Zudem wählten die Probanden, denen Verständnisfragen gestellt wurden, häufiger den Referenten aus, im Vergleich zu den Kindern, denen eine Präferenzfrage gestellt wurde. Die Probanden der Discrepant-Labeling-Bedingung, in der der Experimentator ein Objekt benannte, welches nicht im visuellen Blickfeld des Kindes lag, verknüpften das neue Wort nicht mit dem zu sehenden Objekt, sondern mit dem, welches in der Box lag. Aus diesen Ergebnissen wurde geschlossen, dass Kinder im dem Alter von 16 Monaten von sozialen Hinweisen Gebrauch machen, um neuen Wörtern Referenten zuzuordnen. Sind die sozialen Hinweise jedoch nicht kongruent mit dem sichtbaren Objekt, dann vermeiden sie es, falsche Verknüpfungen zu bilden.

Neben der vorgestellten Studie gibt es weitere Wortlernstudien (z. B. Baldwin, 1993), die die Wirkung von anderen sozio-pragmatischen Hinweisen auf eine ähnliche Weise überprüft haben. So haben beispielsweise Booth und Kollegen (2008) nicht nur den Einfluss der Blickrichtung auf das Wortlernen, sondern in Kombination mit Gesten mit 28–31 Monate alten Kindern getestet.

3.4.2 Loslösung von Gesichtern

In Bezug auf das ECM (vgl. Kapitel 3.3.1), stellen Yurovsky und Frank (2017) die Frage, wie der in dem Modell beschriebene Wechsel vom assoziativen zum sozio-pragmatischen Wortlernen in der Entwicklung zustande kommt. Schreiben ältere Kinder ab einem bestimmten Zeitpunkt in ihrer Entwicklung sozialen Hinweisen eine größere Bedeutung zu als sali-

enten Stimuli oder reduziert sich der Einfluss der perzeptuellen Salienz im Laufe der Entwicklung? Als alternative Erklärung bringen sie ein, dass die Rolle der sozialen Hinweise zunehme, während gleichzeitig die Bedeutung der Objektsalienz abnehme.

Yurovsky und Frank (2017) untersuchten diese Fragen, indem sie zwei Wortlern-Experimente durchführten, an denen Kinder zwischen dem ersten und vierten Lebensjahr teilnahmen. Die von einer Experimentatorin durchgeführte Lernsituation wurde auf einem Bildschirm präsentiert, während das Blickverhalten der Kinder von einem Eye Tracker erfasst wurde. Für das erste Experiment wurden zwei Objekte ausgesucht, die auf einem Tisch vor der Experimentatorin lagen, während sie nur eines davon mehrmals benannte. Somit sammelte der Eye Tracker Blickdaten zu drei Regions of Interests bzw. Areas of Interests (ROIs/AOIs)¹² (Experimentatorin, Zielobjekt, Distraktor), anhand derer, Rückschlüsse über die Wahrnehmung sozialer Hinweise und die Referentenauswahl gezogen werden sollten. Es stellte sich heraus, dass während dieser Trainingsphase alle Kinder dem Blick der Experimentatorin hin zum Referenten gefolgt sind. Auch in den darauffolgenden Tests konnten die Probanden das Zielwort dem richtigen Referenten zuordnen, indem sie ihr Wortverständnis durch eine längere Blickdauer zum Zielobjekt demonstrierten. Der Unterschied zwischen den Altersgruppen äußerte sich dadurch, dass das Blickverhalten der älteren Kinder zum Referenten präziser war im Vergleich zu den jüngeren Probanden. Ein weiterer Befund war, dass alle Altersgruppen die meiste Zeit der Trainingsphase das Gesicht der Experimentatorin anschauten. Die jüngste Altersgruppe hatte sogar Schwierigkeiten sich von ihrem Gesicht zu lösen.

In einem zweiten Experiment wurden zwei Objekte präsentiert, die sich in der Salienz unterschieden. In diesem Experiment gehörten die Probanden entweder der *salient* Gruppe oder der *nonsalient* Gruppe an. Während in der salient-Bedingung, die Experimentatorin das saliente Objekt benannte, war in der nonsalient-Bedingung das weniger saliente Spielzeug das Zielobjekt. Somit hatte dieses Studiendesign Ähnlichkeit zur Studie von Pruden et al. (2006), in der die Probanden entweder der Coincidental- oder Conflict-Bedingung zugeteilt wurden. Aber auch in dem zweiten Experiment identifizierten alle Altersgruppen sowohl aus der Salient- als auch aus der nonsalient-Bedingung während der Trainingsphase den richtigen Referenten anhand der Blickrichtung der Experimentatorin. Doch es stellte sich heraus, dass die Probanden im Test, wenn nach dem neuen Zielwort gefragt wurde, schneller darin waren, den Referenten zu identifizieren, wenn dieser salient war.

¹² ROIs/AOIs sind Markierungen eines oder mehrerer Bereiche eines visuellen Stimulus im Eye-Tracking-System, für welchen die Blickdaten erfasst werden sollen (vgl. Kapitel 4.1.2).

Auf der einen Seite lässt sich festhalten, dass alle Altersgruppen soziale Hinweise des Gegenübers zur Lösung des Referentenproblems hinzuziehen, sodass die soziale Sensibilität auch bei der jüngsten Altersgruppe vorhanden ist. Dies interpretieren die Autoren als Beweis dafür, dass das Wortlernen ein fundamental sozialer Prozess sei. Auf der anderen Seite gelingt die Referentensuche in allen Altersstufen besser, wenn der Referent perzeptuell salient ist. Angesichts der Ergebnisse aus den beiden Experimenten distanzieren sich Yurovsky und Frank (2017) von dem ECM als hybrides Modell, bei dem soziale Hinweise und Salienz zu unterschiedlichen Zeitpunkten zum Wortlernen beitragen. Sie argumentieren, dass es sich nicht um eine Neugewichtung perzeptueller und sozialer Hinweise handelt, sobald Kinder soziale Signale als eine zuverlässige Informationsquelle beachten. Die Tatsache, dass die jüngste Zielgruppe Schwierigkeiten hatte sich von vom Gesicht der Experimentatorin zu lösen, weist darauf hin, dass es sich stattdessen um eine Veränderung der allgemeinen kognitiven Prozesse handelt, die Kinder dazu befähigt, ihre Aufmerksamkeit zu kontrollieren (vgl. Kapitel 2.1.1).

3.4.3 Soziale Aufmerksamkeitslenkung

Laut der assoziativen Wortlerntheorie wird eine Referenz gelernt, indem Kinder ihre selektive Aufmerksamkeit auf Stimuli richten und diese mit neuen Wörtern verknüpfen. Dabei kann die Auswahl des Referenten durch Hypothesen einschränkende Prinzipien oder durch die Präferenz für saliente Objekte erfolgen. Im Gegensatz zu diesen Ansätzen ist der sozio-pragmatische Ansatz dadurch geprägt, dass Kinder für den Erwerb einer neuen Wortbedeutung soziale Achtungshinweise berücksichtigen.

Für das sozio-pragmatische Wortlernen spielt insbesondere die geteilte Aufmerksamkeit (vgl. Kapitel 2.1.1) eine zentrale Rolle (Akhtar & Tomasello, 2000; Bruner, 1983; Bruner, 1975; Tomasello, 2000). Die Fähigkeit zur Joint Attention gilt sogar als ein Meilenstein in der Sprachentwicklung und könne bereits bei Kindern vor dem ersten Lebensjahr beobachtet werden (Liszkowski, 2015). Studien belegen, dass Kinder neue Wörter zuverlässiger mit Objekten verknüpfen, wenn die Benennung des Objektes in dem Moment erfolgt, wenn beide Kommunikationspartner ihren Aufmerksamkeitsfokus auf den Referenten richten (Baldwin, 1991; Tomasello & Farrar, 1986).

Die Aufmerksamkeitslenkung des Kindes auf ein externes Objekt, auf welches der Sprecher referieren möchte, gelingt dadurch, dass der Kommunikationspartner dem Kind signalisiert, dass er eine wichtige Information (z. B. ein neues Wort) vermitteln möchte. Dieses Signal an den Adressaten wird durch sogenannte *ostensive Signale* überbracht, wodurch

schlussendlich ein gemeinsamer Aufmerksamkeitsfokus hergestellt werden kann. Diese Bezeichnung geht ursprünglich auf Sperber und Wilson (1986) zurück und wurde in der aktuellen Forschungsliteratur in Bezug auf das sozio-pragmatische Wortlernen bei Kindern neu aufgegriffen (Akhtar & Tomasello, 2000; Csibra, 2010; Csibra & Gergely, 2006, 2009; Gredebäck et al., 2018; Senju & Csibra, 2008; Szufnarowska, Rohlfsing, Fawcett & Gredebäck, 2015b; Wu et al., 2014). Unter einer *Ostension* bzw. ostensivem Verhalten verstehen Sperber und Wilson (1986, S. 49–50) einen Hinweis, der einem Empfänger signalisiert, dass er von einem Sprecher adressiert wird und dass der Sprecher zudem auf ein bestimmtes Phänomen referieren möchte. Für eine Kommunikation sind ostensive Signale von Bedeutung, weil der Zuhörer durch sie die Erwartung erzeugen könne, dass die Information für ihn relevant sei (S. 50).

Im Kontext des Wortlernens werden ostensive Signale als solche definiert, die von einem Sprecher an einen Empfänger ausgesendet werden, um diesem offensichtlich zu vermitteln, dass der Sprecher eine kommunikative Absicht hat. Weiterhin erläutert Csibra (2010, S. 144), dass Kinder zwar für ostensive Signale empfänglich sind, diese aber gewisse Kriterien erfüllen müssen:

There are at least three types of stimuli that satisfy these criteria: direct gaze that generates eye contact, the special intonation pattern used with infants, which is called infant-directed speech or ‘motherese’, and contingent reactivity to the infant’s behaviour in a turn-taking manner.

Laut der *natural pedagogy* Theorie¹³ nach Csibra und Gergely (2009) sind Kinder sensibel für direkten Blickkontakt, kindgerichtete Sprache (engl.: infant-directed speech, motherese) und kontingente Reaktionen der Erwachsenen. Diese Sensibilitäten wurden bereits in einer Reihe an Studien untersucht. Zum Beispiel fand man heraus, dass 6-Monatige in der Lage sind, den Augenbewegungen des Experimentators zu folgen, sofern die Aufmerksamkeit vorher durch ostensive Signale erlangt wurde (Senju & Csibra, 2008). Andere Studien sprechen allerdings gegen die Theorie der natürlichen Pädagogik (nach Csibra & Gergely, 2009), indem sie zeigen, dass Kinder der Blickrichtung des Gegenübers nach allen aufmerksamkeitserregenden Handlungen folgen (z. B. sich schütteln), ohne vorher direkten Augenkontakt aufgebaut zu haben (Gredebäck et al., 2018; Szufnarowska et al., 2015).

¹³ Der Theorie der natürlichen Pädagogik zufolge ist der Sprecher darauf „ausgerichtet“, relevante Informationen zu vermitteln. Der Empfänger ist hingegen mit einer Sensibilität für spezifische pädagogische Hinweise ausgestattet, die es ihm ermöglicht, die Informationen des Sprechers aufzunehmen. Dadurch, dass Menschen schon früh über ein solches „Kommunikationssystem“ verfügen, ist es ihnen möglich, Wissen schnell und effizient zu erwerben.

In der Forschungsliteratur (Axelsson et al., 2012; Horst & Samuelson, 2008) findet sich außerdem die Bezeichnung *Ostensive Naming* (ostensives Benennen), die an dieser Stelle ebenfalls erläutert werden soll, da dieser Aspekt auch in der Studie der vorliegenden Arbeit aufgegriffen wird (vgl. Kapitel 4.2.2.4). Axelsson und Kollegen (2012) operationalisierten den Begriff *Ostensive Naming* durch eine Hervorhebung des Zielobjektes sowie das gleichzeitige Abdecken der Ablenker, während der Objektbenennung. Die Hervorhebung erfolgte entweder durch die Beleuchtung des Objekts oder durch eine Zeigegeste. Die Forscher plädieren dafür, dass das ostensive Benennen die Wort-Referent-Enkodierung bei Kindern unterstützt und sich in einem besseren Abruf im Test äußert. Dies erklären sich Axelsson et al. (2012) dadurch, dass aufgrund dieser Manipulation der Referent hervorsticht und gleichzeitig der Einfluss der potenzielle Ablenker geschwächt wird.

Abschließend lässt sich die sozio-pragmatische Sichtweise dahingehend zusammenfassen, dass ostensive Mittel den Spracherwerb unterstützen, indem sie die Aufmerksamkeit von Kindern erlagen, wodurch sichergestellt wird, dass beide Partner für eine Informationsvermittlung bereit sind (Csibra, 2010, S. 145).

3.4.4 Social-Looking-Paradigma

Das methodische Verfahren, welches in der Studie von Yurovsky und Frank (2017) durchgeführt wurde, wird von den Autoren als IPLP bezeichnet. Allerdings unterschieden sich das IPLP in dieser Studie von dem IPLP, welches beispielsweise in den Studien von Mather und Plunkett (2010, 2012) angewandt wurde (vgl. Kapitel 3.2.1, 3.2.2, 3.3.4). Der größte Unterschied in dem methodischen Vorgehen besteht darin, dass in der Studie von Yurovsky und Frank (2017) die Blickrichtung der Probanden nicht nur zu den Objekten, sondern auch zur Experimentatorin berücksichtigt wurde, um Rückschlüsse über den Einfluss sozialer Informationen auf das Wortlernen zu ziehen. Dieser Unterschied blieb in der Verwendung der Bezeichnung IPLP in der bisherigen Forschungsliteratur jedoch unberücksichtigt. Um diesen Unterschied zu verdeutlichen, wird in dem folgenden Verlauf der vorliegenden Arbeit die neue Bezeichnung *Social-Looking-Paradigma* (SLP) geführt. Mit einer anderen Bezeichnung soll der Unterschied in dem methodischen Vorgehen des IPLPs aufgezeigt und der Aspekt der sozialen Beachtung verdeutlicht werden.

Das Intermodal-Preferential-Looking- und das Social-Looking-Paradigma haben gemeinsam, dass beide zum Ziel haben das Wortverständnis von Kindern zu erfassen, indem die Probanden visuellen und linguistischen Input erhalten und dabei deren Blickverhalten

erfasst wird. Im IPLP wird die Blickdauer des Kindes lediglich zu den beiden Objekten gemessen, die sich, im Vergleich zum Interactive IPLP, nicht in ihrer perzeptuellen Salienz unterscheiden. Auf diese Weise kann das Wortverständnis getestet werden. Der ausschlaggebende Unterschied des SLP zum IPLP besteht darin, dass sich die Kinder im SLP, ähnlich wie beim IIPLP, in einer *triadischen* Interaktion (Kind, Experimentator, Objekt) befinden. Das bedeutet, dass der auditive Input nicht aus Lautsprechern ertönt, sondern von einem Experimentator ausgeht, der zusätzlich mit sozialen Hinweisen auf ein Zielobjekt referiert. Der Experimentator sitzt zwischen zwei Objekten und schaut eines der Objekte an (und/oder verweist mit weiteren sozialen Signalen darauf), während er es benennt.

Im Interactive IPLP weist ein Experimentator auf einen Referenten hin, während ein zweites, unbenanntes Objekt als salienter Ablenker dient. Dadurch soll erfasst werden, ob die Blickdauer zu dem Objekt länger ist, welches sozial hervorgehoben wurde oder zu dem, das perzeptuell salient ist. Sowie beim IIPL besteht der Vorteil des SLP vor allem darin, dass die sozialen Hinweise variiert werden können, um deren Einfluss auf das Wortlernen zu vergleichen. Im Unterschied zu dem IIPLP wird im SLP die Blickzeit sowohl zu den Objekten, also auch zum Experimentator gemessen, wodurch Rückschlüsse über die Berücksichtigung von sozialen Hinweisen im Fast-Mapping-Prozess gezogen werden. Dabei ist nicht nur von Interesse, *ob* die Probanden soziale Hinweise beachten, sondern auch *welche* für das Wortlernen genutzt werden. Im SLP wird der Input des Experimentators meistens auf Video aufgezeichnet und den Probanden auf einem Monitor präsentiert (Gliga & Csibra, 2009; Gredebäck et al., 2018; Houston-Price et al., 2006; Senju & Csibra, 2008; Szufnarowska et al., 2015; Wu et al., 2014; Yurovsky & Frank, 2017).

Als Beispiel für das SLP soll das bereits in Kapitel 3.4.2 erwähnte Studiendesign von Yurovsky und Frank (2017) beschrieben werden. In ihrer Studie wurden Videos gezeigt, auf denen zwei Objekte sowie eine Experimentatorin zu sehen war. Die Experimentatorin schaute zuerst das Kind an und begrüßte es. Daraufhin wandte sie ihren Kopf zu einem der beiden Objekte und benannte dieses drei Mal. Während dieser Lernphase wurde das Blickverhalten der Kinder zu drei AOIs ausgewertet – zur Experimentatorin und zu den beiden Objekten. Im Anschluss an diese Lernphase folgte die Testphase mit insgesamt drei Tests, in denen zwei Objekte auf dem Bildschirm gezeigt wurden. Zunächst sahen die Probanden im ersten Test zwei familiäre Objekte (z. B. Hund und Buch) und hörten das jeweilige Zielwort. Auf diese Familiarisierungsphase folgte eine Testphase, in der die zuvor in der Lernphase präsentierten Objekte wieder zu sehen waren. Daraufhin wurden zwei Tests durchgeführt, und zwar ein Novel- und ein Mutual-Exclusivity-Test (S. 4). Während beim ersteren

die Probanden nach dem Objekt gefragt wurden, welches in der Lernphase benannt wurde, hörten sie beim letzteren ein komplett neues Zielwort. Daraufhin sollten die Kinder anhand des Prinzip der Ausschließbarkeit das andere Objekt als Referenten identifizieren können.

3.5 Interaktionistisches Wortlernen

Bisher wurden drei Ansätze vorgestellt, die unterschiedliche Positionen einnehmen, wie Kinder das Referenzproblem lösen. Während die Theorie der Hypothesen einschränkenden Prinzipien (vgl. Kapitel 3.2) und der assoziative Ansatz (vgl. Kapitel 3.3) das Wortlernen als einen mechanistischen Prozess beschreiben, für den einfache kognitive Leistungen ausreichen, plädiert die sozio-pragmatische Sichtweise (vgl. Kapitel 3.4) dafür, dass erst soziale Signale die jungen Lerner dazu bewegen, auf einen bestimmten Referenten in dem visuellen Umfeld zu achten.

Im Folgenden soll der interaktionistische Ansatz vorgestellt werden, der davon ausgeht, dass Kinder von Geburt an bestimmte Kompetenzen auf der kognitiven und sozialen Ebene mitbringen, die sich erst in Interaktionen mit ihren Bezugspersonen entfalten (vgl. Klann-Delius, 2008). Um diese Kompetenzen weiterzuentwickeln, ist es allerdings notwendig, dass sich die Bezugsperson an den Entwicklungsstand des Kindes anpasst und das Kind durch einen wechselseitigen Austausch an der Kommunikation mitwirken lässt (S. 144 ff.). In diesem Kontext wird der Bezugsperson eine größere Bedeutung für den Wortlernprozess beigemessen, als im sozio-pragmatischen Ansatz, in dem die Rolle des Erwachsenen darin besteht, die geteilte Aufmerksamkeit zu koordinieren (Rohlfing, 2019, S. 217). Laut dem interaktionistischen Ansatz helfen soziale Hinweise Kindern zwar aus einer Flut von Sineiseflüssen den richtigen Referenten zu extrahieren, eine geteilte Aufmerksamkeit jedoch führe nicht zum Erwerb einer neuen Wortbedeutung, sofern diese nicht in einem *pragmatischen Rahmen* (engl.: pragmatic frames) durch bestimmte Zielhandlungen etabliert wird (Rohlfing, 2019, S. 224; Rohlfing, Wrede, Vollmer & Oudeyer, 2016; Wildt et al., 2019), „weil sie für die Zwecke einer Interaktion geformt wird“ (Rohlfing, 2019, S. 217).

Die sogenannten Pragmatic Frames werden an anderen Stellen der Forschungsliteratur auch als *Interactional Formats* (Bruner, 1983, S. 120) oder *Semantic Frames* (Geeraerts, 2006, S. 373 ff.) genannt. Bruner (1983, S. 120) beschreibt die Interaktionsformate als eine

Art Hilfsmittel, welches durch wiederkehrende, zum Teil auch spielerische Handlungsabläufe (z. B. „Guck-Guck-Spiele“; engl.: peekaboo¹⁴) die Interaktion von Kindern und ihren Bezugspersonen forme (Nomikou, Leonardi, Radkowska, Rączaszek-Leonardi & Rohlfing, 2017), um ihnen den Eintritt in den Spracherwerb zu erleichtern. Solche Interaktionsstrukturen etablieren sich durch die wiederkehrenden Handlungen beider Personen.

Dem interaktionistischen Ansatz nach sind Kinder mit vorsprachlichen Fertigkeiten ausgestattet (Klann-Delius, S. 146–148), an denen die Bezugspersonen anknüpfen können und sie dazu befähigen neuen Input aus den Interaktionen aufzunehmen. Zentral an dem interaktionistischen Ansatz ist, dass solche frühen Formate aus zielgerichteten Handlungen bestehen (Rohlfing et al., 2016), die von beiden Interaktanten koordiniert werden (S. 2). Zunächst sind es eng festgelegte Interaktionsmuster, die in einem bestimmten Kontext auftauchen und in denen das Kind und dessen Kommunikationspartner feste Rollen einnehmen. In ganz frühen Interaktionen übernimmt die Bezugsperson üblicherweise alle „Jobs“, also unterschiedliche zielgerichtete Handlungen, um eine Referenz aufzulösen (Heller & Rohlfing, 2017). Nach und nach entstehen daraus feste Routinen, bis das Kind in der Lage ist, einige dieser „Jobs“ selbst zu übernehmen oder die Rollen sogar getauscht werden können. Diese Routinen sind dann nicht mehr wie vorher nur auf eine bestimmte Handlung bezogen, sondern können sich auch auf andere Situationen ausweiten. Entscheidend an diesem Ansatz ist, dass das Kind dadurch keine passive Rolle mehr einnimmt, indem es beispielsweise nur beobachtet, sondern durch solche Formate aktiv in die kommunikative Handlung eingebunden wird.

Säuglinge sind keinesfalls passive Wesen, die nur Input von ihren Mitmenschen erhalten, sondern sind ebenso von Geburt an in der Lage Reize und Reaktionen auszusendenden, z. B. durch Schreien, auf die Erwachsene wiederum eingehen können (Lebovici, 1990, S. 85–87). *Reziproke*, d. h. wechselseitige Interaktionen, wurden bereits bei nur wenig Wochen alten Säuglingen mit ihren Bezugspersonen beobachtet; dabei wurde festgestellt, dass bereits zu diesem Zeitpunkt Interaktionen koordiniert ablaufen (Trevvarthen, 2011). Das bedeutet, dass der Input nicht nur seitens des Erwachsenen kommt, sondern auch ein Säugling multimodal, beispielsweise durch Blickkontakt oder rhythmische Arm- und Beinbewegungen, in eine Interaktion tritt. Trevvarthen (2011) bezeichnet dies als *Protokonversation*. Auch

¹⁴ Darunter wird eine spielerische Handlung mit Kindern verstanden, bei der die Bezugsperson ihr Gesicht verdeckt (z. B. mit ihren Händen) und mit der Äußerung „Guck-Guck“ bzw. „Boo“ wieder aufdeckt. Bruner (1983, S. 45) verdeutlicht: „They are games that are constituent by language and can exist only where language is present“.

weitere Studien (Nomikou & Rohlfsing, 2011; Nomikou, Rohlfsing & Szufnarowska, 2013) belegen, dass Mütter, die während einer Wickelsituation beobachtet wurden, das Verhalten ihrer Kinder tatsächlich aufgreifen und dementsprechend darauf reagieren. Dabei sind sie besonders sensibel hinsichtlich der Konzentrationsrichtung ihrer Kinder und passen ihr Verhalten daran an, damit sie wahrgenommen werden.

Im Folgenden sollen zwei Studien vorgestellt werden, die drei für eine Interaktion wichtigen Aspekte – gemeinsame Handlungen, Anwesenheit der Bezugsperson, kontingente Kommunikation – untersucht haben. Die Studie von Wass, Clackson und Kollegen (2018) wurde ausgewählt, da sie den interaktionistischen Ansatz stützt, indem sie auf die Wirkung des gemeinsamen Handelns zweier Personen in einer triadischen Interaktion auf die Aufmerksamkeitsspanne von Kindern eingeht. Die zweite Studie (Troseth et al., 2018) wird für die vorliegende Arbeit aus zweierlei Gründen als relevant erachtet:

Zum einen wird in der Studie von Troseth et al. (2018) die Rolle der realen Anwesenheit der Bezugsperson bzw. des Experimentators für die frühkindliche Lernleistung thematisiert. In einer Reihe an sozio-pragmatischen Studien wurde der soziale und linguistische Input des Experimentators auf Video aufgezeichnet und den Kindern auf einem Bildschirm präsentiert. Allerdings wurde in solchen Studien nicht darauf eingegangen, was es ausmacht, die Anwesenheit einer realen Person durch eine Videoaufzeichnung auszutauschen, und ob es dabei zu Unterschieden im Lernprozess kommt. Zum anderen gehen Troseth und Kollegen (2018) auf die Bedeutung der kontingenten Kommunikation für das Wortlernen bei Kindern ein. Wenn ein Experimentator sein Skript „aufsagt“ bzw. ein Lern-Video abgespielt wird, geht damit einher, dass die Interaktion nicht kontingent verläuft, da der Erwachsene seine Reaktion weder inhaltlich noch zeitlich an das Verhalten des Kindes anpassen kann. In solchen Situationen kann ein Experimentator auf die spontanen Handlungen und Reaktionen des Kindes nicht responsiv handeln.

Als Erstes wird die Studie von Wass, Clackson und Kollegen (2018) beschrieben, die die interaktionistische Sichtweise hinsichtlich der gemeinsamen Objekthandlungen stützt. Anschließend wird auf die Studie von Troseth et al. (2018) eingegangen.

3.5.1 Alleinige und gemeinsame Spielsituationen

Laut dem bisherigen Forschungsstand zur Koordination der geteilten Aufmerksamkeit von Kindern und Erwachsenen (Yu & Smith, 2016) ist bekannt, dass Erwachsene die Blickdauer ihrer Kinder auf ein Objekt verlängern können, indem sie dieses Objekt selber visuell fixieren (vgl. Kapitel 2.1.1). Wass, Clackson und Kollegen. (2018) kritisieren allerdings an der

Studie von Yu und Smith (2016), dass eine Kontrollbedingung fehlte, in der sich das Kind mit dem Spielzeug nicht in einer Interaktion mit der Bezugsperson, sondern alleine beschäftigt. Aus diesem Grund war nicht eindeutig geklärt, ob tatsächlich aufgrund der gemeinsamen Handlung die Aufmerksamkeit in einer triadischen, sozialen Interaktion tatsächlich höher ist als in einer nicht-sozialen Situation.

Wass, Clackson und Kollegen (2018) modifizierten die Studie (Yu & Smith, 2016) durch ein Within-Subjects-Design¹⁵, indem sie dieses durch eine Kontrollbedingung (Solo-Play-Bedingung) erweiterten. Somit nahmen 12 Monate alte Probanden sowohl in einer gemeinsamen Spielsituation (Joint Play) als auch an einer Bedingung teil, in der sie alleine mit den Objekten spielten (Solo Play). Wass, Clackson et al. (2018) gingen davon aus, dass auch in dieser Studie die Probanden eine längere Blickdauer zum Referenten in der Joint-Play-Bedingung demonstrieren würden. Doch durch die Modifizierung der Originalstudie sollten nicht nur die Ergebnisse von Yu und Smith (2016) überprüft, sondern auch einer weiterführenden Fragestellung nach der Ursache der Aufmerksamkeitslenkung nachgegangen werden. Die Forschungsfrage lautete: Ist die Blickdauer zum Referenten in einer Joint-Play-Bedingung aufgrund der endogenen Aufmerksamkeit höher oder aufgrund dessen, dass es in einer sozialen Interaktion mehr Möglichkeiten gibt, die exogene Aufmerksamkeit zu verstärken?

Während des Experiments saßen die 12-monatigen Probanden in der Joint-Play-Bedingung ihrer Bezugsperson gegenüber und spielten gemeinsam mit einem Objekt. Dabei wurde das Geschehen auf zwei Videokameras aufgezeichnet und anschließend kodiert. In der Solo-Play-Bedingung wurde eine Trennwand auf den Tisch zwischen die beiden Interakten gestellt, die den gegenseitigen Blickkontakt ermöglichte, die Kinder aber nicht das Objekt auf der Seite der Bezugsperson sehen können. In dieser Bedingung spielten die beiden Personen für sich allein mit einem Objekt, während sie auch hier auf Video aufgezeichnet wurden. In beiden Bedingungen wurde die Bezugsperson darauf hingewiesen, dass das Spielen, soweit es möglich ist, im Stillen durchgeführt werden soll, um den sprachlichen Einfluss auf die kindliche Aufmerksamkeit ausschließen zu können. Die Auswertung der Blickdaten beider Gruppen zeigte, dass die Einjährigen in der Joint-Play-Situation das vorgelegte Objekt länger fixierten als im Solo-Play-Kontext und untermauerten somit den Befund von Yu und Smith (2016).

¹⁵ In einem Within-Subjects-Design (oder Repeated-Measures-Design) läuft ein und dieselbe Probandengruppe die unterschiedlichen Bedingungen/Ausprägungen einer unabhängigen Variable nacheinander durch (Field, 2009, S. 15).

Um der zweiten Forschungsfrage nachzugehen, ob der endogene oder exogene Aufmerksamkeitsprozess für die längere Blickdauer verantwortlich sei, wurden weitere Auswertungen durchgeführt. Eine solche Auswertung betraf beispielsweise die Aufmerksamkeitsträgheit (engl.: *attentional inertia*)¹⁶. Es wurde angenommen, dass die Aufmerksamkeit endogen gesteuert sei, wenn das Blickverhalten kaum variiere und die Aufmerksamkeitsträgheit somit hoch ist. Wenn die Aufmerksamkeit exogen gesteuert sei, dann müsse das Blickverhalten schneller variieren und damit auf eine niedrige Aufmerksamkeitsträgheit hindeuten. Es stellte sich heraus, dass die Ergebnisse auf die letztere Erklärung hindeuten, da der Blick der Kinder nicht nur auf einen Stimulus fixiert war, sondern sich schnell veränderte.

Die Ergebnisse dieser Studie bestätigen die bisherigen Befunde von Yu und Smith (2016) insofern, als in einer Triade die visuelle Aufmerksamkeit in der Tat länger aufrechterhalten wird als in einer nicht-sozialen Interaktion. Daran anknüpfend, erklären Wass, Clackson und Kollegen (2018, S. 12) diese Ergebnisse mit der exogenen Aufmerksamkeitslenkung, da Erwachsene in gemeinsamen Interaktionen die Aufmerksamkeit von jungen Kindern, z. B. durch die Erhöhung der Objektsalienz lenken.

Die vorgestellten Studienergebnisse sind für die vorliegende Arbeit insoweit relevant, als sie den Einfluss von *gemeinsamen* Spielsituationen auf die Aufmerksamkeit von Kindern hervorheben. An dieser Stelle ist allerdings hinzuzufügen, dass es sich nicht um eine Wortlernstudie handelt, sodass noch nichts über die Enkodierungsstrategie der Objekthandlungen im Fast-Mapping-Prozess bekannt ist. Das heißt, es ist unklar, ob die erhöhte Aufmerksamkeit, aufgrund von gemeinsamen Objekthandlungen, auch mit einem besseren Wortlernen einhergeht. Des Weiteren geht aus der Studie nicht hervor, welche Handlungen die Dyaden gemeinsam durchgeführt haben und welche Objekthandlungen die Kinder der Solo-Play-Bedingung angewandt haben. Aus der interaktionistischen Sichtweise wäre es daher interessant zu erfahren, ob die erhöhte Aufmerksamkeit in der Joint-Play-Bedingung aus einer zielgerichteten Handlung resultierte, in die die Kinder eingebettet wurden.

Eine aktuelle Studie von Wass, Noreika und Kollegen (2018) unterstützt die zuvor beschriebenen Forschungsergebnisse von Wass, Clackson et al. (2018) insofern, als die Kinder während gemeinsamer Spielphasen aufmerksamer einen Referenten explorierten als in

¹⁶ Dieses Phänomen wurde bei Kindern in Bezug auf das Fernsehen entdeckt, indem beobachtet wurde, dass mit einer zunehmenden Sehdauer sich die Abwendung vom Fernseher reduziert (Richards & Anderson, 2004). Das heißt, je länger man sich mit etwas beschäftigt, desto robuster wird die Aufmerksamkeit zu diesem Stimulus.

alleinigen Spielphasen. Die Studie (Wass, Noreika et al., 2018) bietet außerdem Einblicke in die Hirnaktivitäten von 12 Monate alten Kindern und ihrer Bezugspersonen während gemeinsamer Interaktionen. Dabei zeigen Wass, Noreika und Kollegen (2018), dass die Aufmerksamkeitsspanne zu dem Objekt, welches die Kinder fokussierten, in Zusammenhang mit einer erhöhten Hirnaktivität der Bezugsperson steht. Daraus schlussfolgerten die Autoren, dass diese Ergebnisse die Rolle des sozialen Partners in einer sozialen Interaktion verdeutlichen.

3.5.2 Reale Anwesenheit und Responsivität des Kommunikationspartners

Die Studie von Troseth et al. (2018) basiert auf bisherigen Studien, die in Bezug auf das Lernen von Bildschirmen aussagen, dass Kinder von einer realen Situation im Vergleich zu einer Fernsehsituation mehr profitieren. So schreiben beispielsweise Anderson und Pempek (2005, S. 511):

Even if very young children are sensitive to the comprehensibility of TV programs, it is another question as to whether they can or do learn anything from them. The totality of results thus far suggests that very young children learn less from television than from equivalent real-life experiences. We refer to this as the video deficit.

Das erwähnte Video-Defizit wird damit erklären, dass in „On-Screen“-Paradigmen interpersonelle Hinweise fehlen, wodurch die Kommunikation nicht-kontingent abläuft und deshalb als quasi-soziale Situation gilt (Troseth et al., 2018). Mit diesem Hintergrund entstand die Frage, ob eine On-Screen-Situation, in der der Experimentator responsiv auf die Handlungen der Versuchsteilnehmer reagiert, das Video-Defizit beheben könnte. Daher entwickelten Troseth und Kollegen (2018) eine Studie, in der sie 24 sowie 30 Monate alte Kinder in vier unterschiedlichen Wortlern-Szenarien untersuchten. Diese unterschieden sich darin, dass der Experimentator entweder responsiv oder nicht-responsiv reagierte und dass die Wortlernsituation entweder in realer Anwesenheit des Experimentators oder auf Video ablief.

Zunächst spielen die Versuchsteilnehmer in einer *Warm-up*-Phase mit dem Objekt-Set, bestehend aus einem neuen und einem familiären Objekt. Darauf folgte eine *Practice*-Phase, in der den Kindern die Aufgabe nähergebracht werden sollte, indem die beiden Objekte dem Kind entgegengehalten wurden und es aufgefordert wurde, eines der beiden in eine Box zu legen. Haben die Versuchsteilnehmer das Spiel verstanden, dann folgte die *Labeling*-Phase. Hierfür wurde ein Objekt-Set aus zwei neuen Objekten gebildet, von denen nur eines benannt werden sollte. Beide Objekte wurden den Kindern zunächst gezeigt und

danach in zwei separate Behälter gelegt. Während das Zielobjekt in einen undurchsichtigen Behälter gelegt und danach benannt wurde, wurde das zweite neue Objekt in einen transparenten Behälter platziert und blieb namenlos. Somit bestand für die Versuchsteilnehmer die Herausforderung darin, den sozialen Signalen des Experimentators zu folgen, sich nicht von dem sichtbaren Objekt ablenken zu lassen und den Namen für das nicht-sichtbare Objekt zu lernen.

Die unterschiedlichen Bedingungen zeigten sich in der Art der Benennung. In der Bedingung *Responsive Live* blieb der Experimentator dem Kind gegenüber weiterhin sitzen und benannte das Zielobjekt aus der undurchsichtigen Box. Dabei verhielt sich der Experimentator responsiv, indem er das Kind anschaute, namentlich adressierte, lächelte und dessen Aufmerksamkeit steuerte, wenn es von anderen Reizen abgelenkt war. In der *Unresponsive-Video*-Bedingung wurde die Labeling-Phase zuvor auf Video aufgezeichnet und den Kindern auf einem Bildschirm abgespielt, sodass die Kommunikation nicht-responsiv ablief. In der *Unresponsive-Live*-Bedingung war der Experimentator zwar präsent, verhielt sich aber ebenfalls nicht-responsiv, indem er lediglich zu festgelegten Zeitpunkten lächelte und die Versuchsteilnehmer anschaute. Dadurch verlief die Kommunikation ebenfalls nicht-kontingent. Die *Responsive-Video*-Bedingung ermöglichte einen responsiven Videochat, in denen der Experimentator aus dem Nebenzimmer dazu geschaltet wurde und mit dem Kind genauso wie in der *Responsive-Live*-Bedingung interagierte, nur über einen Bildschirm.

Anschließend wurden die Versuchsteilnehmer getestet, ob sie das neue Wort gelernt haben. Hierfür folgte eine *Comprehension*-Phase, die das rezeptive Wortwissen überprüfte. Hierfür streckte der Experimentator dem Kind die beiden Objekte entgegen und forderte es auf, das Zielobjekt in einen Behälter und das namenlose Objekt in den anderen Behälter zu legen. Dasselbe wurde nochmal mit einem anderen Objektset durchgeführt, welches sich von dem ersten Set in der Farbe unterschied, um zu testen, ob Kinder das vorhandene Wortwissen auf ein ähnliches Objekt übertragen können.

Die Ergebnisse des Verständnistests dieser Studie zeigten, dass sowohl die 24 Monate alten Kinder als auch die 30 Monate alten Kinder das neue Wort in der *Responsive-Live*-Bedingung gelernt haben. Des Weiteren lernten die 30-Monatigen Kinder das Zielwort auch in der *Unresponsive-Live*-Bedingung. Allerdings konnte das Wortlernen bei beiden Altersgruppen weder in der responsiven noch in der nicht-responsiven Videobedingung verzeichnet werden. Aufgrund dieser Ergebnisse nehmen die Autoren an, dass das „Video-Defizit“ insbesondere bei jungen Kindern auftritt, weil sie Schwierigkeiten haben Objekte und Handlungen, die auf einem Bildschirm gezeigt werden auf die reale Welt zu übertragen. Daher

sei es laut Troseth und Kollegen (2018, S. 8) wichtig, dass Erwachsene in solchen Situationen anwesend sind und den jungen Kindern helfen, die Verbindung zwischen der digitalen und der realen Welt herzustellen, indem sie die notwendigen Hinweise geben, die Kinder brauchen, damit eine Referenz entsteht.

3.5.3 Aufmerksamkeitslenkung in Interaktionen – Lernen durch Relevanz

In Bezug auf die perzeptionsgeleitete Wortlerntheorie merkt Rohlfing (2019, S. 101–104) kritisch an, dass die perzeptuelle Aufmerksamkeitslenkung nicht von der sozialen Interaktion losgelöst betrachtet werden kann, da die perzeptuelle Salienz eines Referenten häufig erst durch die Interaktion erhöht wird, indem die Bezugsperson beispielsweise das Objekt bewegt. Eine klare Trennung von perzeptuellen, sozialen und linguistischen Hinweisen, so wie sie im ECM vorgeschlagen wird, trifft daher nicht immer zu. Diesen Punkt unterstützen aktuelle Studien, die zeigen, dass Bezugspersonen in Interaktionen mit Kindern in der Tat aufmerksamkeitslenkende Handlungen ausführen (Deák et al., 2018), zum Beispiel Zeigegesten und Objektmanipulationen, wodurch die Objektsalienz verstärkt und die exogene Aufmerksamkeit der Kinder angesprochen wird (Wass, Clackson et al., 2018). Dadurch wird den jungen Lernern erleichtert, den Referenten, auf den sich das neue Wort bezieht, zu bestimmen.

Der interaktionistische Ansatz unterscheidet sich vom sozio-pragmatischen Ansatz, indem argumentiert wird, dass die Koordination der geteilten Aufmerksamkeit für das frühe Wortlernen nicht ausreicht. Auch wenn es Kindern gelingt, eine richtige Referentenauswahl zu treffen und den Aufmerksamkeitsfokus mit der Bezugsperson zu teilen, so bedeutet es nicht, dass dieser Objektname gelernt wird (Rohlfing et al., 2016, S. 1). Da Sprache und Handlung eng miteinander verbunden sind, sind Kinder im frühen Wortlernprozess auf mehr als auf den gemeinsamen Aufmerksamkeitsfokus angewiesen, und zwar auf zielgerichtete, gemeinsame Handlungen (engl.: joint action).

Sprache kann nicht nur bestimmte Handlungsaspekte hervorheben, was bereits in frühen Mutter-Kind-Interaktionen deutlich wird (Nomikou & Rohlfing, 2011), ebenso kann eine Handlung die Bedeutung eines sprachlichen Ausdrucks in den Vordergrund bringen (Rohlfing, 2013, S. 126). Die Sichtweise, dass eine passive Lernsituation, z. B. in Form von Beobachtungen von Handlungen, nicht ausreicht, um ein Wort zu lernen, wird beispielsweise von einer Wortlernstudie mit 2–3-jährigen Kindern untermauert (Gampe, Brauer & Daum, 2016). Während es 36-monatigen Probanden gelingt, ein neues Verb lediglich durch die Beobachtung einer Handlung, die sprachlich begleitet wurde, zu lernen, müssen jüngere

Kinder im Alter von 24 und 30 Monaten die beobachtete Handlung eigenständig imitieren, um das neue Wort zu lernen.

Rohlfing und Kollegen (2016, S. 2) erklären die Bedeutung der Joint Actions für den frühen Spracherwerb damit, dass zielgerichtete, gemeinsame Handlungen neuen Wörtern erst eine Bedeutung geben, die für sie relevant seien: „Hence, children will pick up this word because it is being used for a particular purpose that is relevant to them, and because it is uttered hundreds of times to express activities in which it is involved.” Der Begriff *Relevanz* wurde im Kontext der Kommunikation bereits von Wilson und Sperber eingeführt, die ihn wie folgt definieren (2012, S. 62):

We characterise *relevance* as a property of inputs to cognitive processes which makes them worth processing. [...] At each point of our cognitive lives, there are many more potential inputs available than we can actually process: for example, we perceive many more distal stimuli than we can attend to, and have many more memories than we can reactivate at a single time. Efficiency in cognition is largely a matter of allocating our processing resources so as to maximise cognitive benefits. This involves processing inputs that offer the best expected cost/benefit ratio at the time.

An dem Relevanzprinzip ist charakteristisch, dass das menschliche, kognitive System darauf ausgerichtet ist, automatisch und mit einem möglichst geringen Aufwand, potenziell relevante Reize aufzunehmen (S. 63).

Das Relevanzprinzip wurde beispielsweise in einer Studie mit Kindern im Alter von 4 Jahren getestet (Henderson, Sabbagh & Woodward, 2013). Die Probanden wurden zwei Gruppen zugeordnet, entweder der *Nearby Condition* oder der *Faraway Condition*, in denen ein neues Wort für ein neues Objekt genannt wurde. Zusätzlich wurde die Information hinzugefügt, dass das Objekt entweder für Kinder aus der Nachbarschaft oder für Kinder, die weit weg wohnen, besonders ist. Die Annahme, dass die Probanden nur den Namen für das relevante Objekt lernen, d. h. welches von den Kindern aus der Nachbarschaft gekauft wird, wurde bestätigt. Henderson et al. (2013) schlussfolgern aus den Ergebnissen, dass Kinder nicht nur gute, sondern auch selektive Wortlerner sind, da sie die Informationen nach ihrer Relevanz filtern. Da diese Studie mit 4-jährigen Kindern durchgeführt wurde, bleibt dennoch unklar, ob das Relevanzprinzip auch auf den Wortlernprozess bei jüngeren Kindern zutrifft.

Bloom et al. (1993) haben sich mit dieser Frage beschäftigt, indem sie eine Langzeitstudie mit Kindern ab dem 9. Lebensmonat bis zum 2. Lebensjahr durchgeführt haben. Zunächst stand eine andere Frage im Fokus: Machen Nomen die größte Wortkategorie aus, die die Kinder in dieser Zeitspanne lernen? Die Überprüfung dieser Annahme wurde aus bisherigen Ansätzen hergeleitet, die dafür plädieren, dass der frühe Wortschatz von Kindern aus

einer Mehrzahl an Nomen besteht (vgl. Rohlfing, 2019, S. 150–151). Dies rühre daher, dass Nomen keine syntaktischen Zusammenhänge fordern und auch dann verständlich sind, wenn sie alleine stehen (Bloom et al., S. 433). Eine weitere Erklärung für die Annahme, dass Nomen im frühen Wortschatz von Kindern dominieren, lautete, dass Kinder eine Reihe an Hypothesen einschränkenden Prinzipien, z. B. Whole Objects Assumptions oder Principle of Object Scope, in eine Wortlernsituation mitbringen, die auf den Erwerb von Nomen, aber nicht auf Verben ausgerichtet sind.

In ihrer Langzeitstudie stellen Bloom und Kollegen (1993) allerdings fest, dass die Wortklasse der Nomina nicht den größten Teil des Wortschatzes, sondern ungefähr nur ein Drittel bei Kindern ausmacht, was bedeutet, dass die genannten Prinzipien in den meisten Fällen angewandt wurden. Bei vielen Wörtern trat sogar der Fall auf, dass diese keiner konkreten Wortart zugeordnet werden konnten und daher als z. B. „social words“, „event words“ oder „other“ (S. 445) kodiert wurden. Auch jüngere Studien zeigen, dass Kinder bereits im Alter von 10 Monaten nicht nur Nomen, sondern auch alltägliche Verben in ihrem rezeptiven Wortschatz haben (Nomikou et al., 2018).

Bloom und Kollegen (1993) schlussfolgerten, dass allgemeinere Annahmen oder Prinzipien notwendig sind als das Ausschließbarkeitsprinzips, um den frühen Wortlernprozess erklären zu können, weil die Hypothesen einschränkenden Prinzipien lediglich eine Schlussfolgerung aus der Sicht der Erwachsenen sind, die versuchen das Referenzproblem logisch zu lösen. Daher argumentieren Bloom et al. (1993) für das Relevanzprinzip zur Erklärung des Wortlernprozesses. Sie erklären (S. 447), dass im geteilten Aufmerksamkeitsfokus zwischen einem Kind und einem Erwachsenen ein Wort für das Kind erst dann relevant wird, wenn es mit dem, was in den Gedanken des Kindes mitschwingt, zu tun hat. Somit hängt die Relevanz eines Wortes von den Gedanken des Kindes über den geteilten Aufmerksamkeitsfokus ab. Demnach lernen Kinder das, was sie als relevant empfinden.

3.5.4 Beobachtungsverfahren

Da laut der interaktionistischen Sichtweise das Verhalten von Bezugspersonen eine wichtige Rolle für das frühe Wortlernen bei Kindern spielt, werden Verfahren durchgeführt, die vor allem den Zusammenhang zwischen dem Verhalten der Bezugsperson und des Kindes untersuchen. Die Methoden der Eltern-Kind-Beobachtungen unterscheiden sich zu den bisher vorgestellten, methodischen Verfahren hinsichtlich dessen, dass diese nicht nur das Blickverhalten der Kinder messen, sondern multimodale Handlungen beider Interakten einbezie-

hen, zum Beispiel Sprache und Gesten. Im Folgenden werden drei Varianten solcher Beobachtungsverfahren – *freie Interaktion*, *interaktive Alltagsaktivität* und die *aufgabenorientierte Interaktion* – vorgestellt (Kauschke, 2012; Rohlfing, 2013, 2019).

Das Ziel der freien Interaktion besteht darin, die Spontansprache des Kindes (Kauschke, 2012, S. 9) und/oder seine multimodale Verhaltensweise zu beobachten. Hierfür werden Kinder mit ihren Bezugspersonen in einer Spielsituation auf Video aufgezeichnet, wonach die Daten kodiert und anschließend ausgewertet werden. Dabei können aus solch einem Datensatz Rückschlüsse über den Zusammenhang zwischen dem elterlichen und dem kindlichen Verhalten gezogen werden. Es kann sich zum Beispiel um das gestische Verhalten von Kindern und ihren Bezugspersonen handeln (z. B. Grimminger, 2017).

Neben der freien Interaktion, können auch bestimmte Mutter-Kind-Interaktionsformate vorgegeben werden. Die Voraussetzung solcher interaktiven Alltagsaktivitäten besteht darin, dass es sich dabei um Beschäftigungen handelt, der alle Dyaden im Alltag nachgehen (z. B. Mahlzeit einnehmen, Wickeln). Da dadurch der Handlungsspielraum eingeschränkter ist als in einer freien Interaktion, sind die Daten vergleichbarer. Auch bei dieser Methode werden die dyadischen Interaktionen auf Video aufgezeichnet und *offline*, d. h. anschließend kodiert. Damit solche Interaktionen analysiert und miteinander verglichen werden können, werden Kategorien erstellt. Als Beispiel soll das Verfahren aus der Studie von Nomikou und Rohlfing (2011) vorgestellt werden:

Für diese Studie wurde die alltägliche Situation des Wickelns ausgewählt, bei denen Mütter mit ihren drei Monate alten Kindern in ihrem Zuhause gefilmt wurden. Das Ziel dieser Arbeit bestand darin, das an das Kind gerichtete Verhalten zu beobachten. Hierfür wurden der sprachliche Input sowie die Körperbewegungen der Mütter kodiert. Dabei wurde festgestellt, dass Mütter den sprachlichen Input mit ihren Handlungen kombinieren, um bestimmte Aspekte ihrer Äußerungen hervorzuheben. Für weitere Analysen wurden die Mutter-Kind-Interaktionen während derselben Aktivität über die Zeit, als die Kinder im Alter von 3, 6 und 8 Monaten waren, untersucht. Die Auswertungen ergaben, dass sich das Blickverhalten der Kinder verändert hat (Nomikou et al., 2013). Mit dem Alter wurden sie aktiver, ihre Blickdauer zur Mutter reduzierte sich, während sie mehr auf externe Objekte im Umfeld achteten. Die Mütter zeigten, dass sie die Kinder als aktive Interaktionspartner wahrnehmen und ihre Handlungen berücksichtigen, indem sie ihr Verhalten an das Blickverhalten der Kinder anpassten.

Ein interaktives Verhalten, welches aber eine begrenzte Möglichkeit an Aktivitäten zulässt, kann auch entstehen, indem die Bezugsperson eine bestimmte Aufgabe erhält, die

sie mit dem Kind erfüllen soll. Solch eine Aufgabe könnte beispielsweise lauten, dass die Bezugsperson das Kind instruieren soll ein bestimmtes Ziel zu erreichen (Grimminger, Rohlfing & Stenneken, 2010). Handlungen in einer Interaktion können aber auch durch eine Reihe an Spielen, wie zum Beispiel „Guck-Guck-Spiel“, begrenzt werden (Bruner, 1983; Nomikou et al., 2017). Solche Aktivitäten erfordern nicht nur sprachlichen Input seitens der Erwachsenen, sondern auch eine soziale Koordination der Handlungen beider Interaktanten. Eine weitere Möglichkeit für eine aufgabenorientierte Interaktion stellt das gemeinsame Betrachten von Bilderbüchern dar (Heller & Rohlfing, 2017; Yont, Snow & Vernon-Feagans, 2003), um das Verhalten von Dyaden in einem bestimmten Kontext zu beobachten.

Obwohl das Beobachtungsverfahren darauf hinzielt, einen Einblick in die natürliche Eltern-Kind-Interaktion zu schaffen, sollte die Generalisierung der Daten auf Alltagssituationen kritisch betrachtet werden (Rohlfing, 2013, S. 28). Der Nachteil dieses Verfahrens besteht darin, dass die Daten lediglich einen momentanen Ausschnitt in einer bestimmten Situation repräsentieren. Daher ist unklar, ob die Daten das tatsächliche Verhalten von Kindern wiedergeben, wenn sie unbeobachtet sind. Es sollte deshalb berücksichtigt werden, dass Kameras und die Anwesenheit der Experimentatoren einen Einfluss auf die Bezugsperson und auf das Kind haben könnten.

Das LENA¹⁷ System bietet die Möglichkeit, nicht nur einen momentanen Ausschnitt, sondern den tatsächlichen Input, dem ein Kind im Alltag ausgesetzt ist, zu erfassen. Es handelt sich um einen kleinen Audio-Recorder, der an das Kind befestigt wird und im Alltag getragen werden kann. Das Vorteilhafte an diesem System ist, dass es bis zu 16 Stunden Audiomaterial aufnehmen kann und somit den sprachlichen Input über den Tag hinweg erfasst. Nach den Aufzeichnungen werden die Daten auf einen Computer übertragen und hinsichtlich der Menge und der Art des Inputs, den die Kinder im Laufe eines Tages erhalten haben, analysiert. Zwar unterscheidet das System automatisch die einzelnen Sprecher voneinander, aber es bleibt nicht aus, dass für tiefergehende Analysen das Audiomaterial kodiert werden muss. Trotz der technischen Weiterentwicklung stellen die zeitintensiven Aufnahmen, Kodierungen und Auswertung der Daten dennoch einen Schwachpunkt des LENA Systems dar. Des Weiteren beschränkt sich LENA auf die Erfassung des sprachlichen Inputs, sodass der multimodale Input unberücksichtigt bleibt. In aktuelleren Studien wird daher das LENA System mit anderen Verfahren, z. B. mit der Eye-Tracking-Methode kombiniert, um auch die visuelle Wahrnehmung miteinzubeziehen (Odean, Nazareth & Pruden, 2015).

¹⁷ Language Environmental Analysis; siehe: <http://lenafoundation.staging.wpengine.com>

3.6 Diskussion der Theorien und Forschungsmethoden

Die Lösung des Referenzproblems (Quine, 1960) bei Kindern ist eine in der Spracherwerbsforschung umstrittene Frage, wenn es darum geht zu klären, welche Informationen Kinder nutzen, um den Referenten zu identifizieren. Man spricht von den sogenannten Enkodierungsstrategien (Wojcik, 2013), die Kinder anwenden, um für ein neues Wort einen Referenten zu finden. Die Enkodierung wird für das Fast Mapping vorausgesetzt, da zunächst erste Informationen über das Wort und das Objekt enkodiert werden müssen, um diese anschließend im Fast-Mapping-Prozess miteinander in Verbindung zu bringen.

Im bisherigen Verlauf der vorliegenden Arbeit wurden vier Wortlernetheorien vorgestellt, die sich nicht nur in ihren Sichtweisen auf die Enkodierungsstrategien, die Kinder im frühen Fast-Mapping-Prozess anwenden, unterscheiden, sondern auch hinsichtlich dessen, welche Form der Aufmerksamkeit hierfür notwendig ist. Auch die Forschungsmethoden, darunter auch die Blickmessungsparadigmen, die in den Studien der jeweiligen Wortlernetheorie durchgeführt wurden, sind zu differenzieren. Sowohl die theoretischen Ansätze als auch die damit verbundenen Forschungsmethoden werden nun im Folgenden zusammengefasst und kritisch diskutiert.

Die Theorie der Hypothesen einschränkenden Prinzipien (z. B. Golinkoff et al., 1994; Markman, 1990; Markman & Wachtel, 1988; Markman et al., 2003; Merriman et al., 1989; Mervis & Bertrand, 1994) besagt, dass Kinder mit einem spezifischen Wissen ausgestattet sind (Markman, 1994), um Annahmen darüber treffen zu können, worauf sich ein neues Wort bezieht. Dadurch, dass Kinder mit solchen Prinzipien in eine Wortlernsituation treten, können sie die Referentenauswahl reduzieren. Da in der Forschungsliteratur bereits zahlreiche Prinzipien aufgestellt wurden, musste für die vorliegende Arbeit die Auswahl der Prinzipien auf eines begrenzt werden, welches als Beispiel für das frühe Worterlernen dienen sollte. Hierfür wurde das Prinzip der Ausschließbarkeit, basierend auf dem Neuheits-Effekt ausgewählt. Laut den Ergebnissen der vorgestellten Studien (Mather & Plunkett, 2010, 2012), können Kinder bereits im Alter von 10 Monaten „fast-mappen“, indem sie familiäre Objekte als potentielle Referenten ausschließen und sich zum visuell neusten Objekt orientieren, sobald sie ein neues Wort hören. Die Überprüfung dieser Enkodierungsstrategie führten Mather und Plunkett (2010, 2012) mithilfe des IPLP (vgl. Kapitel 3.2.4) durch, indem zwei bzw. drei Objekte präsentiert wurden, die sich in ihrem Neuheitsgrad unterschieden, und die Probanden gleichzeitig einen linguistischen Input erhielten. Durch die Auswertung

der Blickdauer konnte bestimmt werden, welches Objekt aufgrund einer längeren Blickdauer als Referent für ein neues Wort ausgewählt wurde.

In Bezug auf die Theorie der Hypothesen einschränkenden Prinzipien sind drei Aspekte kritisch zu erwähnen, die die Sicht auf den Spracherwerb verengen: Als erster Kritikpunkt ist aufzuführen, dass dieser theoretische Ansatz nur dann sinnvoll ist, wenn keinerlei soziale Hinweise genutzt werden. Dies ist allerdings in einer alltäglichen Situation von Kindern und ihren Bezugspersonen eher unüblich. Daher ist an dieser Sichtweise zu bemängeln, dass die Rolle der Bezugspersonen und deren sozialen Hinweise in der Interaktion mit Kindern zur Erklärung des Wortlernprozesses nicht hinzugezogen werden. Der zweite Kritikpunkt lautet, dass die Hypothesen einschränkenden Prinzipien lediglich den Erwerb von Objektnamen betreffen, und daher unklar ist, wie dann der Rest des Wortschatzes, z. B. Verben (Nomikou et al., 2018), von Kindern erworben wird (vgl. Kapitel 3.5.3, S. 65–66). Als abschließender Punkt ist anzumerken, dass das Wortlernen durch die Hypothesen einschränkenden Prinzipien hauptsächlich mit 2-, 3- und 4-Jährigen überprüft wurden. Daraus gewinnt man zwar die Erkenntnis, dass ältere Kinder mit einem bereits vorhandenen Wortwissen die Prinzipien anwenden können und erfahre auch, *was* sie über das Wort hinaus lernen (z. B. dass es sich auf ein ganzes Objekt bezieht). Doch die Frage, *wie* Kinder vor dem ersten Lebensjahr, ohne ein großes Wort- oder Objektwissen zu besitzen, ihre ersten Wörter lernen, wird dabei nicht ausreichend erklärt (Bloom et al., 1993, S. 434).

Schließlich ist anzumerken, dass die vorgestellten Studien (vgl. Kapitel 3.2.1 und 3.2.2) dafür plädieren, dass 10 Monate alte Kinder das Prinzip der Ausschließbarkeit basierend auf der Objektneuheit nutzen. Es bedarf jedoch weiterer Forschung, die das Prinzip der Ausschließbarkeit, hinsichtlich der „Nameability“-Fähigkeit mit jungen Kindern untersucht, da aktuelle Studien zeigen, dass bereits 6-Monatige alltägliche Wörter verstehen. Somit ist noch die Frage offen, ob Kinder entweder ihr vorhandenes Wortwissen, die Objektneuheit oder die Kombination aus beidem als Enkodierungsstrategie nutzen.

Als zweites wurde in der vorliegenden Arbeit die assoziative Wortlerntheorie beschrieben. Befürworter dieses Ansatzes (z. B. Hollich et al., 2000; Pruden et al., 2006) plädieren dafür, dass ein Wort in einer temporären Kontiguität erworben wird, indem es auditiv aufgenommen und zugleich ein salienter Referent visuell wahrgenommen wird. Aus ihren Studienergebnissen schlussfolgern Pruden et al. (2006), dass 10-Monatige sozialen Hinweisen keine Bedeutung zuschreiben (S. 274), sondern ihre selektive Aufmerksamkeit ausschließlich auf perzeptuell saliente Stimuli richten, wodurch das frühe Wortlernen perzepti-

onsgeleitet abläuft. Rohlfing (2013, S. 74) weist jedoch kritisch darauf hin, dass die Schlussfolgerung nicht lauten sollte, dass soziale Hinweise in der ersten Entwicklungsphase komplett ignoriert werden, sondern vielmehr „[...] dass Kinder im Alter von 10 Monaten der perzeptuellen Auffälligkeit Vorrang gewähren und entsprechend der Blick eines Sprechers für ihr Wortlernen kaum relevant ist.“

Mit der Studie knüpfen Pruden und Kollegen (2006) außerdem an das ECM nach Hollich et al. (2000) an (vgl. Kapitel 3.3.1), laut welchem bis zum ersten Lebensjahr die Enkodierungsstrategie von Kindern darin besteht, das salienteste Objekt aus dem visuellen Umfeld als Referenten auszuwählen und erst ab dem zweiten Lebensjahr soziale Signale als sichere Informationsquelle zu beachten. Rohlfing (2013, S. 95–96, vgl. auch Rohlfing, 2019, S. 32–33) kritisiert jedoch am ECM, dass dort die Lösung des Referenzproblems im Fokus stehe und andere Aspekte im Zusammenhang mit der Sprachentwicklung außer Acht gelassen werden. Hollich und Kollegen (2000) führen beispielsweise Hinweise auf, die Kinder in den drei Entwicklungsphasen nutzen, um das Referenzproblem zu lösen, erklären aber nicht, *wie* es zu dieser Veränderung kommt. Des Weiteren wird im ECM die Rolle der sozialen Sensibilität, die bei Kindern von Beginn an vorhanden ist, in der ersten Entwicklungsphase nicht berücksichtigt.

Als Forschungsmethode wurde in den Studien von Hollich et al. (2000) sowie von Pruden et al. (2006) das Interactive IPLP (vgl. Kapitel 3.3.4) angewandt, um die Gewichtung der sozialen und perzeptuellen Stimuli im frühen Wortlernprozess anhand der Blickdauer miteinander zu vergleichen. Allerdings ist an der methodischen Durchführung der Studie von Pruden et al. (2006) kritisch anzumerken, dass soziale Hinweise lediglich als Blickrichtung bzw. Kopfbewegung operationalisiert wurden, obwohl die Experimente von Hollich und Kollegen (2000) zeigten, dass bei 12 Monate alten Kindern die Blickrichtung des Experimentators zum Referenten nicht ausreicht, um ein Wort zu lernen. Ein weiterer Kritikpunkt ist, dass die Blickrichtung lediglich zu den Objekten, aber nicht zum Experimentator ausgewertet wurde. Daher ist zum einen unklar, ob die Probanden die Veränderung der Blickrichtung überhaupt wahrgenommen haben (Rohlfing, 2019, S. 103). Zum anderen kann die Aussage „Data from the conflict condition, however, allow us to conclude that infants ignored the social cues“ (Pruden et al., 2006, S. 274) aufgrund der nicht berichteten Blickdaten nicht konkret belegt werden.

Insgesamt ist an dem assoziativen Ansatz zu bemängeln, dass auf die Rolle der sozialen Partner nicht ausreichend eingegangen wird und diese Theorie für eine vollständige Erklärung zur Entstehung der frühkindlichen Semantik daher nicht ausreicht. Rohlfing (2019,

S. 104) fasst kritisch zusammen, dass Kinder ohne den sozialen Aspekt eine Menge an „sinnfreien Assoziationen“ lernen würden und dies sei, wie Studien zeigen (z. B. Baldwin, 1991), nicht der Fall. Des Weiteren zeigt eine aktuellere Studie (Yurovsky & Frank, 2017), dass nicht nur junge, sondern auch ältere Kinder perzeptuell saliente Objekte im Vergleich zu langweiligen präferieren (vgl. Kapitel 3.4.2). Daraus geht die Frage hervor, ob die Strategie zur Referentenauswahl nicht altersabhängig, wie in dem ECM dargestellt, sondern vielmehr kontextabhängig sei. Daher bedarf es weiterer Forschung, die untersucht, unter welchen Umständen Kinder vor dem ersten Lebensjahr auf welche Hinweise – perzeptuell saliente oder soziale – achten.

Bei der an dritter Stelle beschriebenen Theorie dieser Arbeit (vgl. Kapitel 3.4) handelt es sich um den sozio-pragmatischen Ansatz, mit welchem die Annahme verbunden ist, dass der Mensch von Geburt an ein soziales Wesen ist, welches für die sozialen Signale seiner Mitmenschen sensibel sei. Kinder nutzen soziale Signale nicht nur, um den Referenten für ein neues Wort zu bestimmen, sondern auch, um falsche Wort-Objekt-Verknüpfungen zu vermeiden (Baldwin, 1991). Jungen Kindern fiel es sogar schwerer als älteren Kindern, sich von Gesichtern zu lösen, um einem referentiellen Hinweis zu folgen (Yurovsky & Frank, 2017). Die Befürworter des sozio-pragmatischen Ansatzes heben die Form der geteilten Aufmerksamkeit besonders hervor, die für die frühe Sprachentwicklung von Bedeutung sei.

Der Unterschied zu den bisher vorgestellten Forschungsparadigmen der assoziativen Sichtweise und der Hypothesen einschränkenden Prinzipien besteht darin, dass die Blickrichtung der Probanden nicht nur zu den Objekten, sondern zum Kommunikationspartner berücksichtigt wird, um Rückschlüsse über den Einfluss sozialer Informationen auf das Wortlernen zu ziehen. Daher wurde in der vorliegenden Arbeit die Bezeichnung *Social-Looking-Paradigma* eingeführt (vgl. Kapitel 3.4.4), um den Unterschied zu dem IPLP zu verdeutlichen sowie den Aspekt der sozialen Beachtung hervorzuheben. Mit dem Fortschritt der Technik geht einher, dass in aktuellen sozio-pragmatischen Studien auf reale Interaktionen verzichtet wird und den Kindern stattdessen Videoaufzeichnungen vom Experimentator auf dem Bildschirm präsentiert werden.

Damit verbunden ist am theoretischen Ansatz und auch am Paradigma kritisch anzumerken, dass eine passive Lernsituation hergestellt wird, in der ein Kind durch Beobachtung eine neue Wort-Objekt-Verknüpfung lernen soll. Dies spiegelt allerdings nicht die natürliche Situation wieder, die Kinder im Alltag erleben, wenn sie mit ihren Mitmenschen interagieren. Daher sollte die „soziale“ Wortlernsituation, die auf einem Bildschirm präsentiert wird, kritisch betrachtet werden, da aktuelle Studien (z. B. Troseth et al., 2018) zeigen, dass vor

allem junge Kinder eher von einem realen Kommunikationspartner lernen, anstatt von einem, der digital übertragen wird. Darüber hinaus geht aus der Studie von Troseth und Kollegen (2018) hervor, dass ein responsives Verhalten der Bezugsperson, welches in einer echten Interaktion realisiert wird, im frühen Alter für das Wortlernen wichtig ist.

Abschließend ist noch ein bedeutsamer Aspekt aufzuführen, der in weiterführenden Studien untersucht werden sollte, um die Forschung hinsichtlich der Rolle von sozio-pragmatischen Hinweisen für das frühe Wortlernen voranzutreiben: Studien zeigen, dass Kinder bereits im Alter von 6 Monaten sozialen Hinweisen, in Form von Blickrichtung, folgen können (z. B. D'Entremont et al., 1997) und diese Fähigkeit sogar mit der späteren Sprachentwicklung zusammenhängt (z. B. Morales et al., 1998). Des Weiteren wurde festgestellt, dass ein frühes Wortverständnis von Nomen (Bergelson & Swingley, 2012) und Verben (Nomikou et al., 2018) im Alter von 6 Monaten vorhanden ist. Trotz dieser Befunde, die sich auf ein Wortverständnis von Kindern bereits vor dem ersten Lebensjahr beziehen, wurde der sozio-pragmatische Ansatz bisher nur in Studien mit Kindern ab dem 12. Lebensmonat untersucht. Daher bleibt weiterhin unbeantwortet, ob Kinder schon vor dem ersten Lebensjahr sozio-pragmatischen Hinweisen (z. B. der Blickrichtung oder Gesten) einer Bezugsperson folgen, um im gemeinsamen Aufmerksamkeitsfokus neue Wörter zu lernen, und diese im anschließenden Test auch abrufen können.

An vierter Stelle wurde der interaktionistische Ansatz des Wortlernens vorgestellt (vgl. Kapitel 3.5). Diese Sichtweise unterscheidet sich von den zuvor beschriebenen Positionen hinsichtlich dessen, dass nicht nur dem Aufmerksamkeitsfokus, sondern insbesondere den gemeinsamen, zielführenden Handlungen mit einem anwesenden Interaktionspartner eine wesentliche Bedeutung für das Wortlernen bei Kindern beigemessen wird (Rohlfing et al., 2016). Demnach entstehen Wort-Referent-Verbindungen nicht aus dem Willen des Kindes heraus, neue Wörter zu lernen, sondern, weil Wörter als sprachliches Werkzeug in Interaktionen genutzt werden und sie für die Handlungen des Kindes relevant sind (Wildt et al., 2019).

Die Bedeutsamkeit eines realen Kommunikationspartners während einer Interaktion wurde an dem Beispiel der Studie von Troseth et al. (2018) dargestellt. Sie zeigten, dass es nicht ausreicht, wenn 2-Jährige an einer On-Screen-Wortlernsituation teilnehmen. Um ein Wort erfolgreich lernen zu können, brauchen sie einen anwesenden Interaktionspartner, der kontingent auf ihr Verhalten reagiert. In Bezug auf die wesentliche Rolle von gemeinsamen Handlungen zeigen aktuelle Studien mit 12 Monate alten Kindern und ihren Bezugspersonen

(Wass, Clackson et al., 2018; Yu & Smith, 2016), dass gemeinsame Spielsituationen die exogene Aufmerksamkeit der Kinder zu Referenten lenken und aufrechterhalten.

Im Vergleich zu den bisherigen Forschungsmethoden, die das Blickverhalten hauptsächlich anhand unterschiedlicher Preferential-Looking-Paradigmen gemessen haben, wird in interaktionistischen Studien das multimodale Verhalten der Dyaden erfasst. Hierfür interagiert eine Bezugsperson bzw. ein Experimentator mit dem Kind, während die Interaktion auf Video aufgezeichnet und mit einem Annotationsprogramm kodiert wird.

An dem theoretischen Ansatz sowie an der Forschungsmethode ist kritisch anzumerken, dass aus bisherigen Studien zwar hervorgeht, dass gemeinsame (Objekt-) Handlungen die kindliche Aufmerksamkeitspanne zu einem Objekt verlängern (z. B. Wass, Clackson et al., 2018), aber noch weiterführende Studien fehlen, die dies im Kontext des Wortlernens untersuchen. Somit ist auch noch unzureichend erforscht worden, welches interaktive Verhalten sich positiv auf das Fast Mapping auswirkt.

Aufgrund der unterschiedlichen Ansätze, verfolgt die vorliegende Arbeit das Ziel, die vier Theorien in einer Studie zusammen zu bringen, um untersuchen zu können, auf welche Enkodierungsstrategien Kinder im Alter von 10 Monaten zurückgreifen. Darüber hinaus soll unmittelbar nach der Enkodierung der rezeptive Abruf überprüft werden, um Rückschlüsse über die Retention der enkodierten Wort-Objekt-Verbindung ziehen zu können. Neben der Zusammenfügung der Theorien in einer Studie, wird außerdem das Eye-Tracking- und das Beobachtungsverfahren zu einer Mixed-Methods-Methode kombiniert.

3.7 Ableitung der Fragestellung und Hypothesen für die vorliegende Studie

Die aus der assoziativen und sozio-pragmatischen Theorie abgeleitete Annahme der vorliegenden Dissertationsschrift lautet, dass sowohl die Objektsalienz als auch soziale Hinweise von 10 Monate alten Kindern als Enkodierungsstrategien genutzt werden. Allerdings könne eine Wort-Objekt-Verknüpfung nur dann rezeptiv abgerufen werden, wenn diese durch zielorientierte Handlungen in einer gemeinsamen Interaktion enkodiert wurde. Diese Annahme wird von der interaktionistischen Sichtweise gestützt, laut der Kinder neue Wörter lernen, die für sie relevant sind (vgl. Kapitel 3.5.3). Die Relevanz der Wort-Objekt-Verknüpfung wird dem Kind im Rahmen von zielführenden Handlungen in einer Interaktion aufgezeigt.

In Bezug auf die formulierte Annahme sollen im ersten Experiment der vorliegenden Arbeit die bestehenden theoretischen Ansätze aufgegriffen und untersucht werden, ob sich

die Versuchsteilnehmer für die perzeptuelle Objektsalienz oder für soziale Signale, in Form der Blickrichtung und der Ausführung einer Objektfunktion, als Enkodierungsstrategie entscheiden, wenn beide Hinweise gleichzeitig präsent sind. Anknüpfend an der interaktionistischen Sichtweise, nimmt die Hälfte der Versuchsteilnehmer zusätzlich an einer gemeinsamen, zielführenden Interaktion teil, in der die Objektfunktion gemeinsam mit der Experimentatorin ausgeführt wird. Anschließend gilt zu überprüfen, nach welcher Enkodierungsstrategie die Wort-Objekt-Verknüpfung rezeptiv abgerufen werden kann. Abhängig davon, ob den Versuchsteilnehmern der Abruf gelingt oder nicht, kann das Prinzip der Ausschließbarkeit, basierend auf der Nameability-Fähigkeit, als weitere Enkodierungsstrategie untersucht werden.

Während die erste Hypothese anhand des Beobachtungsverfahrens kontrolliert wird, betrifft die zweite Hypothese die Eye-Tracking-Methode und die dritte Hypothese deren Zusammenhang. Die im Folgenden aufgestellten Hypothesen werden mithilfe der Mixed-Methods-Methode überprüft, indem zwei unterschiedlichen Verfahren kombiniert werden.

Hypothese 1: Perzeptuelle Salienz ist aufmerksamkeitslenkend.

Als erste Hypothese wird aufgestellt, dass 10 Monate alte Kinder eine visuelle Präferenz für das perzeptuell salienteste Objekt aus ihrem visuellen Umfeld haben. Dieser Hypothese nach sollten die Eye-Tracking-Daten, die sich aus der Salienzphase der vorliegenden Arbeit ergeben, die Ergebnisse der Studie von Pruden et al. (2006) sowie von Hollich et al. (2000) und damit den assoziativen Ansatz untermauern, dass junge Kinder ihre selektive Aufmerksamkeit auf physikalisch auffällige Objekte richten. Es wird angenommen, dass die visuelle Präferenz durch eine längere Blickdauer zum salienten Objekt als zu einem weniger salienten Objekt messbar sein wird.

Hypothese 2: Soziale Hinweise dienen als Informationsquelle zur Referentenauswahl.

Als zweite Hypothese wird formuliert, dass 10 Monate alte Kinder soziale Hinweise als Informationsquelle hinzuziehen, um das Referenzproblem zu lösen. In Bezug auf die Blickdaten, die aus der Experimentator-Kind-Beobachtungsmethode resultieren, wird angenommen, dass die Versuchsteilnehmer das weniger-saliente Objekt, welches aber in einer Trainingsphase durch soziale Hinweise hervorgehoben wird, länger anschauen als das saliente Objekt. Somit knüpft die zweite Hypothese an die sozio-pragmatische Sichtweise an, dass selbst junge Kinder soziale Hinweise von Kommunikationspartnern als Enkodierungsstrategie nutzen.

Damit geht auch einher, dass diese Hypothese in Kontrast zum assoziativen Ansatz und damit auch zu den Ergebnissen der Testphasen aus der Studie von Pruden et al. (2006) steht. Wird die zweite Hypothese nicht bestätigt, weil die Versuchspersonen anstelle des Zielobjekts das saliente Objekt länger anschauen, dann spricht es für die assoziative Sichtweise, laut der Kinder die Objektsalienz als Enkodierungsstrategie nutzen und dabei soziale Hinweise nicht beachten.

Hypothese 3: Es besteht ein Zusammenhang zwischen der Enkodierungsstrategie während der Trainingsphase und der Retention der Wort-Objekt-Verbindung in der Testphase.

Neben dem Ziel zu untersuchen, welche Enkodierungsstrategien Kinder im Alter von 10 Monaten nutzen, soll außerdem der Zusammenhang zwischen den Lernformaten und der Retention, gemessen anhand des Blickverhaltens, der enkodierten Wort-Objekt-Verknüpfung überprüft werden. Da die Annahme der vorliegenden Arbeit lautet, dass eine Wort-Objekt-Verknüpfung nur dann beibehalten werden kann, wenn sie durch zielorientierte Handlungen in einer gemeinsamen Interaktion enkodiert wird, werden 10 Monate alte Kinder zwei Gruppen zugeordnet, die an zwei unterschiedlichen sozialen Lernformaten teilnehmen (Wildt & Rohlfing, 2018; Wildt et al., 2019). In einer Gruppe können die Versuchspersonen die Beobachtung von sozialen Hinweisen, im Sinne der sozio-pragmatischen Theorie, als Enkodierungsstrategie nutzen und anschließend mit dem Objekt selber spielen. In der anderen Gruppe folgt auf die Beobachtung der sozialen Hinweise eine gemeinsame Handlung, die als Enkodierungsstrategie im Sinne des interaktionistischen Ansatzes gilt.

Als dritte Hypothese wird abgeleitet, dass die zwei verschiedenen Lernformate in Zusammenhang mit dem Blickverhalten der beiden Gruppen im anschließenden Eye-Tracking-Test stehen. Anknüpfend an diese Hypothese stellt sich die Frage, ob und welche der Enkodierungsstrategien zu einer Retention und somit dem rezeptiven Abruf der Wort-Objekt-Verknüpfung führt.

4 Untersuchung der frühen Enkodierungsstrategien

– Experiment 1

Die Studie der vorliegenden Arbeit besteht aus insgesamt drei Experimenten, von denen die letzten beiden als Kontrollexperimente durchgeführt wurden. Alle Daten für die gesamte Studie wurden im Zeitraum von Januar 2017 bis April 2019 von der Autorin der vorliegenden Arbeit im SprachSpielLabor¹⁸ der Universität Paderborn erhoben. Im Rahmen des ersten Experiments wurden Kinder im Alter von 10 Monaten hinsichtlich ihrer Enkodierungsstrategien im Fast-Mapping-Prozess untersucht. Anschließend bestand das Ziel darin zu untersuchen, ob die Probanden die enkodierte Wort-Objekt-Verbindung beibehalten können.

Dem Experiment liegt ein *Between-Subjects*-Studiendesign (engl.: *between-groups*) zugrunde, da zwei Gruppen – die Joint-Engagement- und Self-Exploring-Gruppe – bestehend aus unterschiedlichen Versuchsteilnehmern miteinander verglichen wurden (Field, 2009, S. 15). Mithilfe der Gruppenaufteilung wurde überprüft, wie sich die Ausprägungen der unabhängigen Variable (gruppenabhängiges Wortlernformat) auf die abhängige Variable (rezeptives Wortwissen) auswirken. Dabei wurde das Wortverständnis durch die Messung der Blickdauer zu den jeweiligen Stimuli mithilfe eines Eye Trackers operationalisiert.

Die Besonderheit des angewandten Mixed-Methods-Forschungsdesigns besteht darin, dass zwei Methoden – das Beobachtungs- und das Eye-Tracking-Verfahren – kombiniert wurden. Zum einen wurde das qualitative Beobachtungsverfahren in Form einer aufgabenorientierten Interaktion angewandt (vgl. Kapitel 3.5.4), mit der die zweite Hypothese bezüglich der Beachtung sozialer Hinweise überprüft wurde. Für das Beobachtungsverfahren wurde das interaktive Verhalten der Experimentatorin und des Versuchsteilnehmers während der dargebotenen Wortlernsituation auf Video aufgezeichnet und anschließend kodiert. Zum anderen wurde das Eye-Tracking-Verfahren genutzt, wobei das Interactive IPLP (vgl. Kapitel 3.3.4) modifiziert wurde, um die erste und dritte Hypothese zu überprüfen.

¹⁸ Das SprachSpielLabor ist eine Einrichtung der Universität Paderborn, die zur Arbeitsgruppe Psycholinguistik (unter der Leitung von Prof. Dr. Katharina J. Rohlfing) gehört.

4.1 Die Eye-Tracking-Methode

Säuglinge können von Geburt an ihre Umwelt visuell wahrnehmen. Dies ist eine der ersten Fähigkeiten, mit der Neugeborene ausgestattet sind (Gredebäck et al., 2009), und die sie nutzen, auch wenn die Sehfähigkeit noch nicht komplett ausgebildet ist. Die Erfassung des Blickverhaltens ist beispielsweise mit der Aufzeichnung auf Videokameras möglich. Allerdings ist dies mit viel Zeit und Arbeitsaufwand verbunden, da die Videoaufnahmen manuell kodiert werden müssen. Eine Lösung hierfür bieten Eye-Tracking-Systeme. Sie stellen sich in dieser Hinsicht als besonders nützlich dar, da keine Kodierung notwendig ist und daher weniger zeit- und arbeitsaufwendig sind (Brône & Oben, 2018, S. 8–9). Da das Blickverhalten nicht nur der okulomotorischen, sondern auch der kommunikativen Funktion dient (S. 5), gibt die Erfassung von Augenbewegungen nicht nur Aufschluss darüber, wann wohin und wie lange Versuchsteilnehmer auf etwas schauen, sondern ermöglicht Rückschlüsse über die visuellen Präferenzen, Entstehung von Joint Attention, Strukturierung einer Interaktionsstruktur, die Referentenauswahl und das Wortverständnis zu ziehen (S. 1).

Im weiteren Verlauf soll die Eye-Tracking-Methode genauer vorgestellt werden. Zunächst werden allgemeine Informationen über das menschliche Sehsystem geschildert. Darauf aufbauend wird die für das Eye-Tracking-Verfahren gängige Cornea-Reflex-Methode erklärt. Hiernach wird der klassische Ablauf einer Eye-Tracking-Studie beschrieben. Dabei wird das Augenmerk auf einige Funktionen gelegt, die speziell in der vorliegenden Arbeit genutzt wurden. Im Anschluss werden die Vor- und Nachteile des Eye-Tracking-Verfahrens, insbesondere in Studien mit Kindern, vorgestellt.

An dieser Stelle ist zu betonen, dass sich die folgenden Angaben speziell auf den Eye Tracker von Tobii Studio (Tobii Pro X3-120) beziehen. Aus diesem Grund sind Abweichungen, beispielsweise in der Benutzeroberfläche oder in den Funktionen, bei anderen Eye-Tracker-Systemen oder -Versionen möglich.

4.1.1 Messung der Fixationen und Sakkaden

Zum Verständnis der Eye-Tracking-Methode soll die Funktionsweise des menschlichen Auges beschrieben werden. Die wichtigste Voraussetzung für die visuelle Wahrnehmung unserer Umwelt ist das Licht, welches durch zwei *Transformationen* im Auge zu einem Bild umgewandelt wird (Goldstein, 2015, S. 20). Zunächst treffen Lichtstrahlen auf Gegenstände, die das menschliche Auge reflektiert. Das Auge fängt das einfallende Licht auf der Hornhaut bzw. Cornea ein und bündelt es. Die Cornea ist die glasklare, frontale Schicht des Augapfels.

Dieses Licht tritt weiter durch zur Iris (Regenbogenhaut), in dessen Zentrum die Pupille liegt. Die Iris kontrolliert, wieviel Licht durch die Pupillenöffnung einfällt, indem sie die Pupillenöffnung entweder weitet oder verengt. Bei Dunkelheit weitet sich die Öffnung, so dass mehr Licht durchströmen kann, wohingegen bei Helligkeit die Pupille enger wird. Das durchströmende Licht wird von der dahinterliegenden Linse noch einmal gebündelt und schließlich auf der Netzhaut bzw. Retina abgebildet. Die Retina befindet sich in der Innenseite des Auges und ist mit visuellen *Rezeptoren*, bestehend aus *Stäbchen* (für das Hell-Dunkel-Sehen) und *Zapfen* (für das Farb-Sehen) ausgestattet (Holmqvist & Nyström, 2011, S. 21). Die erste Transformation ist die des reflektierenden Lichts in ein Bild, das auf der Netzhaut bzw. Retina verkehrt herum erzeugt werden. Danach erfolgt eine zweite Transformation. Die Rezeptoren transformieren das Licht in Nervenimpulse, wodurch die dabei entstehenden elektrischen Signale über den Sehnerv an das Gehirn weitergeleitet werden können, um dort weiterverarbeitet zu werden. Erst im Gehirn wird das Bild „richtig“ gedreht.

Für die Eye-Tracking-Methode ist wichtig zu wissen, dass das menschliche Auge nicht in der Lage ist, das komplette Bild in seiner vollen Schärfe abzubilden, sondern nur einen kleinen Bereich davon (Holmqvist & Nyström, 2011, S. 21). Dieser sogenannte *foveale* Bereich liegt in der Retina und hat die Besonderheit, dass er jede Menge Zapfen in Relation zu Stäbchen enthält. Während in der Fovea eine reichhaltige Zapfenanzahl vorliegt, sind in dem restlichen Bereich der Retina mehr Stäbchen als Zapfen vorhanden. Diese ungleiche Verteilung der Rezeptoren führt zu dem, dass das Objekt, welches fokussiert wird, zwar in der Fovea scharf wahrgenommen wird, aber hin zum *parafovealen* und *peripheralen* Bereich die Sehschärfe abnimmt. Das was nicht auf der Fovea abgebildet wird, wird daher auch vom Auge registriert, allerdings nicht in der vollen Schärfe und Farbe.

Durch diese visuelle Einschränkung entsteht die Notwendigkeit, Augenbewegungen bzw. Sakkaden durchzuführen. Ein Bild muss Punkt für Punkt fixiert werden, um eine Szene visuell komplett erfassen zu können. Wie bei einem Mosaikbild entsteht aus den einzelnen fixierten Bildern eine vollständige Szene der Umgebung. Dieses Blickverhalten, welches durch Fixationen und Sakkaden realisiert wird, ist auch als „visuelles Abtasten“ bekannt (Goldstein 2015, S. 128).

In den meisten Eye-Tracking-Studien, so wie auch in der Studie der vorliegenden Arbeit, wird über Fixationen berichtet, da lediglich in solchen Momenten, in denen das Auge etwas fixiert, visuelle Informationen verarbeitet werden (Holmqvist & Nyström, 2011, S. 22). Darunter versteht man Punkte, an denen das Auge für einige Millisekunden verweilt. Fixationen variieren in ihrer Dauer und sind üblicherweise etwa 50 bis 600 Millisekunden

lang (Holmqvist & Nyström, 2011, S. 23). Allerdings steht auch während einer Fixation das Auge nicht komplett still. Währenddessen spielen sich Mikrosakkaden ab, die jedoch nicht von allen Eye Trackern erfasst werden können. Solche Mikrobewegungen der Augen dienen hauptsächlich dem Ausgleich der Rezeptorenermüdung. Im Gegensatz zu Fixationen ist während den Sakkaden, die etwa 30 bis 80 Millisekunden dauern (S. 23), die Aufnahme sowie die Verarbeitung von Informationen nicht möglich. Daher wird auf Sakkaden im folgenden Verlauf dieser Arbeit nicht näher eingegangen.

In Eye-Tracking-Verfahren wird die *Cornea-Reflex-Methode* (engl.: pupil centre corneal reflection – PCCR) genutzt, um die Fixationen und Augenbewegungen zu erfassen und diese auf dem Monitor sichtbar zu machen (Holmqvist & Nyström, 2011, S. 24–29; Gredebäck et al., 2009). Damit ein Fixationspunkt errechnet werden kann, wird das für einen Menschen nicht sichtbare Infrarotlicht aus kurzer Entfernung vom Eye Tracker ausgestrahlt und somit das Auge ausgeleuchtet. Dadurch entsteht ein kleiner Reflexionspunkt der Lichtquelle auf der Cornea, der als *Glint-* oder *Purkinje-Bild* bezeichnet wird.

Zur Bestimmung des Fixationspunktes ist das Purkinje-Bild und die Pupillenposition notwendig. Das Eye-Tracking-System berechnet anhand eines Algorithmus aus dem Abstand zwischen der Pupillenposition und dem Purkinje-Bild den genauen Fixationspunkt. Der Abstand zwischen den beiden Punkten kann variieren und muss deshalb im nächsten Schritt durch eine *Kalibrierung* (vgl. Kapitel 4.1.2) ausgeglichen werden, woraus schließlich die exakte Augenposition rekonstruiert werden kann (Gredebäck et al., 2009; Holmqvist & Nyström, 2011, S. 25).

4.1.2 Ablauf eines Eye-Tracking-Experiments

Im Folgenden soll der Ablauf eines Eye-Tracking-Experiments exemplarisch beschrieben werden. Neben den üblichen Schritten, die für die meisten Eye-Tracker-Systeme zutreffen, werden auch einige gesonderte Aspekte beschrieben, die speziell in der Studie der vorliegenden Arbeit durchgeführt wurden.

Schritt 1: Areas of Interest anlegen

Eine zentrale Einstellung, die noch vor der Blickdatenerfassung vorgenommen werden sollte, da sie eine wichtige Rolle für die spätere Auswertung spielt, ist die Festlegung der AOIs (vgl. Abbildung 1, S. 79) (Holmqvist & Nyström, 2011, S. 187–230). Dabei handelt es sich um die Markierung eines oder mehrerer Bereiche des visuellen Stimulus für den die Fixationen gesammelt und ausgewertet werden sollen.

Für die Größe der AOIs gibt es zwar keine konkreten Vorgaben, aber es sollte beispielsweise beachtet werden, welche Zielgruppe untersucht wird. Im Vergleich zu erwachsenen Versuchsteilnehmern ist bei Kindern die Datenqualität meist schlechter. Daraus kann es zu Abweichungen zwischen dem tatsächlichen Fixationspunkt und dem gemessenen Fixationspunkt kommen, weswegen in Experimenten mit Kindern die AOIs als große Flächen angelegt werden (z. B. Bergelson & Swingley, 2012). Auf der einen Seite sollte bedacht werden, dass mit zunehmender Größe der AOIs mehr Fixationspunkte in die Analysen einfließen als vielleicht notwendig. Auf der anderen Seite besteht das Risiko, wichtige Blickdaten zu verlieren, wenn die AOIs zu klein sind.

Abschließend ist zu erwähnen, dass bei der Erstellung von AOIs Gruppenzuordnungen vorgenommen werden können, um beispielsweise unterschiedliche Stimuli als Gruppe zu kategorisieren und miteinander zu vergleichen. Da in der vorliegenden Studie unterschiedliche Stimuli verwendet wurden, wurden zwei AOI-Gruppen (vgl. Abbildung 1) angelegt (Zielobjekt, Ablenker).

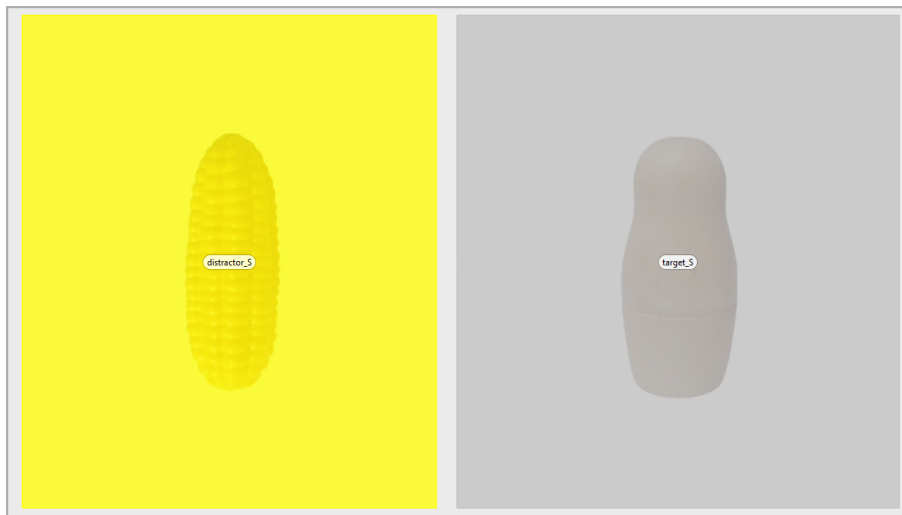


Abbildung 1. Zwei Areas of Interest aus der Studie der vorliegenden Arbeit.

Schritt 2: Kalibrierung

Die Kalibrierung wird mit jedem einzelnen Probanden vor dem Experiment durchgeführt, damit das Eye-Tracking-System aus dem Abstand zwischen der Pupillenposition und dem Purkinje-Bild den genauen Fixationspunkt des Probanden berechnen kann. Der Abstand zwischen den beiden Punkten kann aufgrund der Blickrichtung variieren und wird deshalb

durch eine Kalibrierung ausgeglichen (Gredebäck et al., 2009; Holmqvist & Nyström, 2011, S. 25).

Hierfür positioniert sich der Versuchsteilnehmer etwa 50–80 cm vor dem Bildschirm (optimal sind 65 cm), an dem der Eye Tracker angebracht ist. Während der Kalibrierung soll der Teilnehmer mindestens zwei bis maximal neun Punkte auf dem Monitor folgen, die sich über den Bildschirm bewegen. Für Kinder wird anstelle von Punkten eine Animation genutzt, die in allen Ecken des Bildschirms auftaucht und saliente Geräusche von sich gibt, um den visuellen Fokus des Kindes in diese Richtung zu lenken (vgl. Abbildung 2).

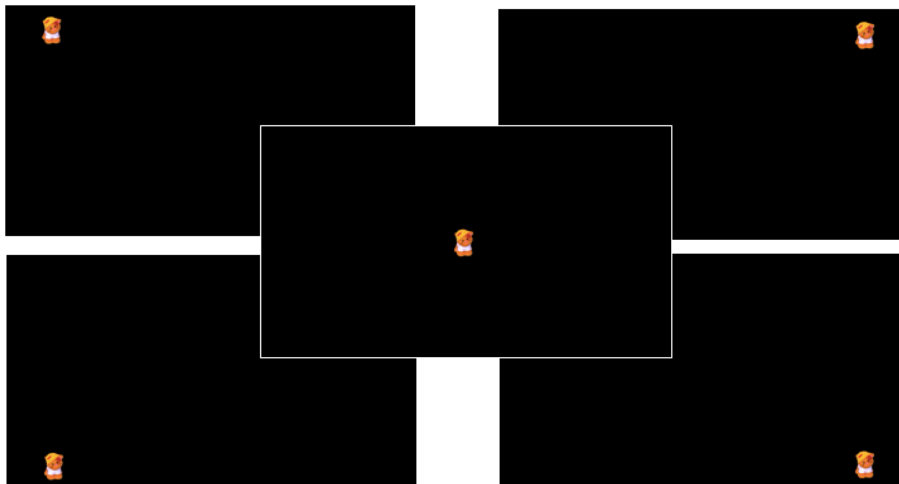


Abbildung 2. Präsentierte Animation während der Fünf-Punkt-Kalibrierung.

Der Punkt bzw. die Animation kann entweder automatisch ablaufen oder manuell gesteuert werden. In der Studie der vorliegenden Arbeit wurde eine manuelle Fünf-Punkt-Kalibrierung auf einem schwarzen Hintergrund in der Lautstärke 10 dB durchgeführt. Dabei wurden die Kalibrierungspunkte mit einer Bluetooth-Tastatur (KeySonic KSK-5220BT mit Touchpad) gesteuert. Eine manuelle Steuerung hat den Vorteil, dass die Animation per Touchpad in die nächste Ecke versetzen kann, sobald man sieht, dass das Kind den Punkt fixiert hat.

Direkt im Anschluss zeigt das System das Ergebnis über die Qualität der Kalibrierung an. In der Studie dieser Arbeit wurde als erstes auf die Anzahl der kalibrierten Punkte geachtet. Bei einem fehlenden Punkt wurde in der Regel neu kalibriert bis alle fünf Punkte erfasst waren. Als zweites wurde die Qualität der kalibrierten Daten anhand der *Offsets* (grüne Linien) bewertet (sehr gut, gut, befriedigend, ausreichend, keine). Wenn beispielsweise alle fünf Punkte erfasst, jedoch die Qualität als befriedigend eingestuft wurde, weil die Offsets zu lang waren, dann haben die Experimentatorinnen anhand der Stimmung des

Kindes individuell entschieden, ob eine Re-Kalibrierung durchgeführt werden soll. Diese Beobachtungen wurden in einem Studienprotokoll festgehalten, welches im Anhang der vorliegenden Arbeit (vgl. Anhang E, Studienprotokoll) zu finden ist.

Schritt 3: Aufnahme starten

Nach einer erfolgreichen Kalibrierung wird erst der eigentliche Test gestartet. Abhängig vom Studiendesign bietet das Eye-Tracking-System die Möglichkeit, Bilder, Videos oder PDF-Dateien zu präsentieren. In der Studie der vorliegenden Arbeit wurden statische Videos im AVI-Format hochgeladen, die Audioaufnahmen enthielten.

Eine nützliche Funktion, die in Studien mit Kindern eingesetzt werden kann, sind die Aufmerksamkeitslenker (engl.: attention getter). Dabei handelt es sich um zentrierte Animationen, die wenige Sekunden lang sind und die durch ihre visuelle und akustische Salienz die Aufmerksamkeit des Probanden auf den Bildschirm ziehen sollen. Der Einsatz von Aufmerksamkeitslenkern ist vor allem bei jungen Kindern sinnvoll. Zum einen, um ihre Aufmerksamkeit immer wieder zum Ort des Geschehens – den Bildschirm – zu holen. Zum anderen, um in Preferential-Looking-Paradigmen ihren Blick zwischen den Bild- bzw. Videosequenzen in die Mitte des Bildschirms zu lenken. Dabei kann der Experimentator selber entscheiden, ob die Aufmerksamkeitslenker automatisch an bestimmten Stellen auftauchen oder ob sie manuell gesteuert werden. In der vorliegenden Studie lief sowohl das Videomaterial als auch die Aufmerksamkeitslenker automatisch ab.

Bevor die Aufnahme gestartet wurde, wurden die Bezugspersonen darauf hingewiesen, dass sie ihre eingenommene Sitzposition sowie die des Kindes nicht enorm verändern und nicht auf den Bildschirm zeigen sollten, da dadurch unter Umständen die Augen des Probanden nicht erfasst werden können.

Schritt 4: Weighted Gaze Samples prüfen

Nach der Aufnahme kann im nächsten Schritt die Datenqualität eingesehen werden, die maximal 100 % betragen kann. Diese wird im System als Weighted Gaze Samples (WGS) angezeigt und wird im Tobii Handbuch (Tobii AB, 2016, S. 40) als die prozentual erfasste Blickanzahl definiert. Das heißt, dass ein Wert von 100 % dann erreicht werden kann, wenn beide Augen während der gesamten Aufnahme vom Eye Tracker uneingeschränkt erfasst wurden. Ein Wert von 50 % bedeutet also, dass entweder nur ein Auge während der gesamten Aufnahme oder beide Augen lediglich die Hälfte der Aufnahme erfasst wurden.

Ein niedriger WGS-Prozentwert kann ein Indiz für unterschiedliche Probleme sein. Es kann beispielsweise an unvorteilhaften Lichtverhältnissen liegen oder daran, dass der Eye Tracker Schwierigkeiten hatte, die Augen der Versuchsteilnehmer zu erkennen. Dies kann wiederum mit weiteren, unterschiedlichen Gründen verbunden sein. In der Studie der vorliegenden Arbeit war der häufigste Grund für einen niedrigen WGS-Wert bei Kindern ihre kurze Aufmerksamkeitsspanne, die bewirkte, dass sie sich nach kurzer Zeit vom Bildschirm abwendeten.

Da die WGS mit der Datenqualität in Verbindung gesehen werden, ist es sinnvoll, einen Richtwert festzulegen, der als Ausschlusskriterium (vgl. Kapitel 4.2.1.3) gelten soll. Liegt der Grenzwert beispielsweise bei 50 %, dann werden alle Probanden, deren Werte darunterliegen, für weitere Datenauswertungen nicht berücksichtigt.

Schritt 5: Szenen erstellen

Ein weiterer Schritt, der erst nach der Datenaufzeichnung erfolgen kann, ist die Erstellung von Szenen (Tobii AB, 2016, S. 47–49). Diese Funktion dient primär dazu, um bestimmte Ausschnitte eines Videos zu erstellen, die für Visualisierungen oder statistische Analysen benötigt werden. Auch für die Studie dieser Arbeit wurden Szenen festgelegt, um die Videoclips in Onset- und Offset-Phasen zu teilen (vgl. Kapitel 4.2.2.5). Dadurch konnten die Fixationszeiten nicht nur für das gesamte Video, sondern auch gesondert für den Videoausschnitt vor oder nach dem Zielwort ausgegeben werden.

Schritt 6: Blickdaten exportieren

Der letzte Schritt besteht aus dem Export der Daten für weitere Analysen in Programmen wie Excel, SPSS, R, Python oder Matlab. Die im ersten Schritt angelegten AOIs liefern nun für den jeweiligen Bereich erfassten Blickdaten, die in einer Tabelle oder als Diagramm eingesehen werden können. Zudem können die Fixationszeiten für den gesamten Videoclip ausgegeben werden und/oder für einzelne Szenen, die vorab erstellt wurden.

Durch die verschiedenen Parameter, zum Beispiel *Time to First Fixation*, *Fixations before*, uvm., die im Tobii System angeboten werden, können Blickdaten hinsichtlich unterschiedlicher Fragestellungen ausgewertet werden. Für die statistischen Testverfahren der vorliegenden Arbeit wurde die Gesamtblickdauer (engl.: total visit duration) und die Dauer bis zur ersten Fixation (engl. time to first fixation) zu den einzelnen AOIs verwendet. Für eine ausführliche Erläuterung aller Parameter wird empfohlen, das Handbuch von Tobii Eye Tracking heranzuziehen (Tobii AB, 2016, S. 100–120).

4.1.3 Möglichkeiten und Grenzen der Eye-Tracking-Methode

Die Eye-Tracking-Methode bringt Vorteile, aber auch gewisse Nachteile mit sich, die im Folgenden aufgeführt werden sollen, wobei das Augenmerk insbesondere auf die Kleinkind- bzw. Säuglingsforschung gelegt wird. Ein wichtiger Aspekt, der in der Forschung mit Kindern geschätzt wird, ist die Kopfbewegungsfreiheit der Kinder. Neben den *Head-Mounted-Eye-Trackern*, bspw. Eye-Tracking-Brillen, die am Kopf angebracht und deshalb meistens in der Erwachsenenforschung verwendet werden, gibt es auch monitorbasierte Geräte. Monitorbasierte Eye Tracker ermöglichen es, dass Kinder mit keinerlei Kabel oder anderen elektronischen Anbringungen in Berührung kommen. Sie können auf dem Schoß des Elternteils sitzen und verfügen dabei über eine relativ große Kopfbewegungsfreiheit. Auch in der Studie der vorliegenden Arbeit wurde ein monitorbasierter Eye Tracker (engl.: screen based; remote eye tracker) genutzt.

Ein weiterer Punkt, der für das Eye-Tracking-Verfahren in Studien mit Säuglingen und Kleinkindern spricht, lautet, dass die Probanden keine Instruktionen erhalten und keine Aufgabe erfüllen müssen. Da es sich um ein implizites Verfahren (Bjorklund & Causey, 2018, S. 95) handelt, erfordert es keine direkte oder bewusste Kommunikation des Wissens (vgl. Kapitel 3.1.). Allein durch Messung des Blickverhaltens auf sprachliche oder visuelle Stimuli können Rückschlüsse über Präferenzen, Differenzierungen und das Wortwissen gezogen werden.

Hinsichtlich der Bewegungsfreiheit hat die Technik allerdings auch ihre Grenzen. Eine solcher Grenzen wird kann an den WGS-Werten beobachtet werden, weil diese bei Erwachsenen meist höher sind als bei Kindern. Die Qualität der Daten kann zwar an den Laborbedingungen, zum Beispiel aufgrund von schlechten Lichtverhältnissen leiden, wird aber oft durch individuelle Faktoren beeinflusst. Die in der Studie der vorliegenden Arbeit beobachteten Probleme, aus denen eine schlechte Datenqualität hervorgeht, wurden im Studienprotokoll (vgl. Anhang E) festgehalten und nun beschrieben:

Der größte Eye-Tracking-Datenverlust entstand aufgrund einer erheblichen Veränderung der Sitzposition, weil Kinder von dem Schoß der Bezugsperson entweder runterrutschen oder versuchten sich hinzustellen. Des Weiteren waren einige Kinder von anderen Dingen aus ihrem visuellen Umfeld zwischenzeitig abgelenkt oder verloren das Interesse an den Stimuli ganz, sodass sie ihren Blick vom Blickschirm abwendeten. Aber auch Zeigegesten der Kinder, die auf die Stimuli auf dem Monitor gerichtet waren, haben die Strahlung des Infrarotlichts durchbrochen, sodass für diesen Moment keine Blickdaten erfasst werden

konnten. Im Tobii Handbuch (Tobii AB, 2016, S. 120–122) werden weitere individuelle Aspekte, wie zum Beispiel Schlupflieder oder nach unten zeigende Wimpern beschrieben, die ebenfalls auftreten können. Eine genauere Ausführung sowie weitere potenzielle Schwierigkeiten im Eye-Tracking-Verfahren sind in Holmqvist und Nyström (2011) nachzulesen.

Nichtsdestotrotz können einige Vorkehrungen getroffen werden, um eine gelungene Aufnahme des Blickverhaltens von Kindern zu erhalten. Zum einen sollten die Erziehungsberechtigten vor dem Experiment ausreichend informiert und über mögliche Schwierigkeiten mit dem Eye Tracker aufgeklärt werden. Dadurch können sich Erwachsene darauf einstellen und ihre Kinder in die richtige Sitzposition bringen oder wieder zum Bildschirm lenken. Da sich Kinder leicht ablenken lassen, kann man sich für solche Fälle mit Aufmerksamkeitslenkern behelfen (vgl. Kapitel 4.1.2), um deren Fokus wieder zum Monitor zu lenken. Da Kleinkinder eine kürzere Aufmerksamkeitsspanne haben im Vergleich zu Erwachsenen, sollte außerdem darauf geachtet werden, dass die präsentierten Videos nicht allzu lang sind, da ansonsten das Risiko besteht, dass sie schnell das Interesse verlieren und wegsehen. Letzten Endes bedarf es vor allem Übung im Ablauf einer Eye-Tracking-Studie und mit der Technik, um sich sowohl für das kindliche Verhalten zu sensibilisieren als auch mit technischen Schwierigkeiten umgehen zu können.

Eye-Tracking-Verfahren sind im Bereich der Spracherwerbsforschung nicht mehr wegzudenken. Der Trend geht dahin, dass monitorbasierte Eye Tracker von Head-Mounted-Eye-Trackern abgelöst werden, da Interaktionen immer mehr in den Vordergrund der Forschung treten und diese deshalb losgelöst vom Bildschirm gemessen werden sollen. Aber auch die Kombination von Eye Trackern mit anderen Technologien, zum Beispiel mit LENA (z. B. Odean et al., 2015) oder mit EEG (z. B. Kushnerenko et al., 2013), gewinnen in der aktuellen Forschung zunehmend an Bedeutung.

4.2 Mixed-Methods-Forschungsdesign

Für die vorliegende Studie wurden zwei technische Apparaturen genutzt – ein Eye Tracker und zwei Videokameras. Im Folgenden wird auf die Besonderheiten der Nutzung dieser Messinstrumente sowie auf die Laborbedingungen eingegangen.

Das Eye-Tracking-Verfahren wurde mittels eines monitorbasierten Eye Trackers (Tobii Pro Studio X3-120, Version 3.4.6, 120 Hz) durchgeführt. Der Eye Tracker wurde an die untere Leiste eines 24-Zoll-Monitors (mit Schwenkarm) eines Computers (Dell XPS

8900 Intel Core i7-6700K) mithilfe eines selbstklebenden Magnetstreifens angebracht. Während in dem Eye-Tracking-Bereich (vgl. Abbildung 4, S. 86) die Explorations-, Salienz- und Testphase durchgeführt wurde, fand in dem Interaktionsbereich die Trainingsphase statt.

Die Trainingsphase (vgl. Kapitel 4.2.2.4) wurde mit zwei Handycam-Videokameras (SONY, HDR-CX700, 12.3 Mega Pixels) aufgezeichnet, die an zwei Dreibein-Stativen (Sony VCT-80AV und Hama, Profi 76) angebracht und in dem Interaktionsbereich (vgl. Abbildung 3, S. 86) aufgestellt wurden. Die Videokamera bzw. -kameras, die in höhenverstellbaren Stativen befestigt waren, standen von Anfang bis Ende des Experiments in dem Raum. Sie wurden links oder/und rechts von dem Interaktionstisch so ausgerichtet, dass die Augen des Kindes, die Augen der Experimentatorin sowie beide Objekte auf dem Tisch eindeutig zu erkennen waren. Neben den Kameras befand sich in der Interaktionsecke ein Tisch, auf dem die Objekte lagen, zwei Stühle und eine Spieldecke, die auf dem Boden in der Ecke des Raumes platziert war.

Die Interaktionsphase wurde entweder mit einer und mit zwei Videokameras aufgezeichnet. In den Randomisierungsgruppen 1, 2, 5 und 6 war eine Videokamera auf der Seite des Zielobjekts ausreichend. Zwei Kameras waren lediglich dann zwingend notwendig, wenn die Seite der Zielobjekte auf dem Interaktionstisch im ersten und zweiten Versuch variierte. Um die Blickdauer zum Zielobjekt sowie den Blick zur Experimentatorin genau kodieren zu können, wurden zwei Kameras, sowohl von der rechten als auch von der linken Seite des Tisches, aufgestellt. Dies trifft auf die Randomisierungsgruppen 3, 4, 7 und 8 zu (vgl. Anhang A, Tabelle 24). Zum Beispiel lag in der Randomisierungsgruppe 3 das Zielobjekt während des ersten Versuchs auf der rechten und im zweiten Versuch auf der linken Seite des Tisches (aus der Perspektive der Experimentatorin).

Da Licht eine entscheidende Rolle für die Qualität der Blickdaten spielt, wurde das Labor durch Außenrollos und Jalousien abgedunkelt und somit für konstante Lichtverhältnisse gesorgt. Der Interaktionsbereich wurde von einer Deckenleuchte über dem Tisch erhellt. Der Eye-Tracking-Bereich war hingegen von dem restlichen Raum durch einen dunkelroten Paravent abgetrennt. Eine dimmbare LED-Leuchte (My BEAUTY LIGHT, 30 x 1.2 x 1.2 cm, 6W, 6500 Kelvin) wurde an den schwenkbaren Arm des Monitors mit zwei Saugnäpfen befestigt, sodass der Monitor samt des Eye Trackers von hinten beleuchtet wurde. Zur Reduktion von Umgebungsreizen wurde die Studie in einem Raum mit geschlossenen Fenstern und Türen durchgeführt. Des Weiteren haben beide Experimentatorinnen darauf geachtet, dass sie während des Experiments weder Arm- oder Ohrschmuck tragen, um zu verhindern, dass die Kinder sich davon ablenken lassen.

Abschließend sind die Stoppuhr und die Kappe, die die Experimentatorin trug, zu erwähnen, die für die Durchführung der Studie notwendig waren. Während der Trainingsphase stoppte eine Experimentatorin die Zeit für 30 Sekunden mithilfe einer Stoppuhr, damit alle Probanden die gleiche Zeit zur Verfügung hatten. Des Weiteren hatte die Experimentatorin, die mit dem Kind interagierte, während der gesamten Studie eine schwarze Kappe mit Schirm auf dem Kopf, deren Nutzen in Kapitel 4.2.2.4 expliziert wird.



Abbildung 3. Der Interaktionsbereich im SprachSpielLabor.



Abbildung 4. Der Eye-Tracking-Bereich im SprachSpielLabor.

4.2.1 Studienteilnehmer

Im Folgenden stehen die Angaben zur Stichprobe des ersten Experiments im Fokus. Zunächst soll beschrieben werden, wie die Versuchsteilnehmer rekrutiert wurden und wie sich die Stichprobe letztlich zusammensetzte. Außerdem sollen die Ausschlusskriterien benannt werden. Schließlich wird das Randomisierungsverfahren der Probanden vorgestellt werden.

4.2.1.1 Rekrutierung

Für die Studien der vorliegenden Arbeit wurden im Kreis Paderborn über PEKiP¹⁹-, DELFI²⁰-, Rückbildungs- und Babymassagekurse, offene Mutter-Kind-Treffen, auf Kinderflohmärkten sowie über Social-Media ungefähr 250 Kinder rekrutiert. Den Erziehungsberechtigten wurden in den Kursen Informationsblätter ausgeteilt (vgl. Anhang B, Informationsblatt), damit sie sich bei Interesse an der Studie bei den Ansprechpersonen melden konnten. Alternativ hatten sie die Möglichkeit sich in eine unverbindliche Kontaktliste eintragen, falls Interesse an einer Studienteilnahme bestand. Damit gaben sie der Autorin der vorliegenden Arbeit die Erlaubnis, dass sie telefonisch oder per Mail kontaktieren werden durften. Die Erziehungsberechtigten wurden kontaktiert, sobald ihre Kinder ungefähr neun Monate alt waren, um den Erhebungstermin zu vereinbaren.

4.2.1.2 Stichprobenzusammensetzung

Vor Beginn der Erhebung wurde die Stichprobengröße von mindestens 60 Versuchsteilnehmern festgelegt, damit jeder Gruppe jeweils 30 Kinder zugeteilt werden können. An dem ersten Experiment nahmen insgesamt 83 Versuchsteilnehmer freiwillig teil. Da 17 Probanden ausgeschlossen werden mussten, worauf in Kapitel 4.2.1.3 eingegangen wird, umfasste die endgültige Stichprobe 66 Probanden ($M_{Alter} = 305.4$ Tage, $SD = 3.6$; $N_{weiblich} = 33$, $N_{männlich} = 33$), von denen 33 Kinder an der Joint-Engagement- ($M_{Alter} = 304.8$ Tage, $SD = 3.4$; $N_{weiblich} = 16$, $N_{männlich} = 17$) und 33 Kinder an der Self-Exploring-Gruppe ($M_{Alter} = 305.4$ Tage, $SD = 3.7$; $N_{weiblich} = 17$, $N_{männlich} = 16$) partizipierten.

Zur Erhebung wurden die Familien in das SprachSpielLabor der Universität Paderborn eingeladen. Vorab wurden sie informiert, ungefähr eine Stunde für den Besuch einzuplanen, damit sich das Kind an die neue Umgebung gewöhnen kann. Für gewöhnlich dauerten die Termine ungefähr eine halbe Stunde. Als Anreiz zur Teilnahme bestand die Erfahrung bei

¹⁹ Das Prager Eltern-Kind-Programm (PEKiP) ist ein Kursangebot, das Eltern und ihre Kinder mit Spiel- und Bewegungsanregungen durch das erste Lebensjahr begleitet.

²⁰ DELFI (Denken, Entwickeln, Lieben, Fühlen, Individuell) ist ein Kursangebot für Eltern und ihre Kinder im ersten Lebensjahr zur Stärkung der Eltern-Kind-Beziehung.

einem Eye-Tracking-Experiment mitzumachen und ein Geschenk für das Kind zu erhalten. Nach der Studie erhielten die Probanden als Aufwandsentschädigung ein Buch oder ein Spielzeug sowie eine Urkunde. Bei einer Anfahrtstrecke von über 10 Kilometer, wurde die Bezugsperson mit einer Geldsumme von 10 EUR entschädigt.

Im Anschluss an das Experiment wurden die Erziehungsberechtigten gebeten, einen Fragebogen auszufüllen (vgl. Anhang D, Demographische Daten), der Angaben zur Entwicklung des Kindes sowie den Bildungsstand der Erziehungsberechtigten enthielt. Alle Erziehungsberechtigten gaben an, dass ihre Kinder ein normal ausgeprägtes Seh- und Hörvermögen haben. Hätte eine Erkrankung des Seh- oder Hörorgans vorgelegen, hätten die Daten nicht für die Analysezwecke genutzt werden können, da nicht ausgeschlossen werden könnte, ob die Probanden die, auf dem Bildschirm präsentierten visuellen und auditiven Stimuli wahrgenommen haben. Des Weiteren wurde abgefragt, ob die Versuchsteilnehmer vor Vollendung der 40. Schwangerschaftswoche (SSW)²¹ geboren wurden und falls ja, welches Geburtsgewicht das Kind hatte.

Zusätzlich wurde das Bildungsniveau der Erziehungsberechtigten mittels des demographischen Fragebogens erfasst, indem die Elternteile jeweils ihren höchst erreichten Bildungsabschluss angaben. Bezüglich des Bildungsstandes enthält die Stichprobe einen recht großen Anteil an Familien mit einem hohen Bildungsabschluss im Vergleich zu den anderen Abschlüssen. Aus den angegebenen Daten konnte des Weiteren entnommen werden, dass die Mütter signifikant mehr Zeit ($M_{Stunden} = 22.75$, $SD = 4.54$) mit ihren Kindern täglich verbringen als die Väter ($M_{Stunden} = 4.66$, $SD = 3.72$), $t(65) = 21.57$, $p = .000$, $r = .94$. Daher wurde für die Datenauswertungen (vgl. Kapitel 4.3) der Bildungsabschluss der Mütter als potenzieller Prädiktor für das Verhalten der Kinder in Erwägung gezogen.

²¹ In Kapitel 4.2.1.3 werden die Gründe für diese Abfrage genannt.

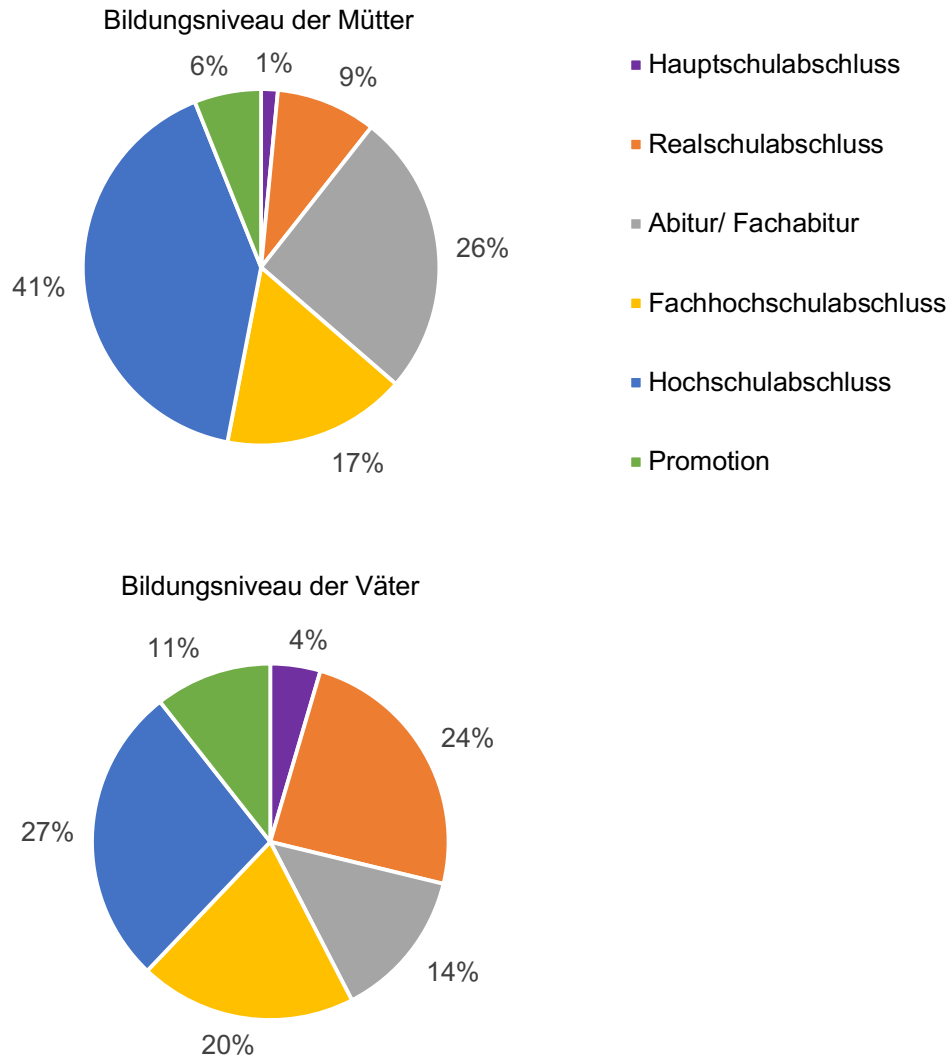


Abbildung 5. Höchster erreichter Bildungsabschluss der Erziehungsberechtigten (Experiment 1)

4.2.1.3 Ausschlusskriterien

Es wurden insgesamt 17 Versuchsteilnehmer von weiteren Analysen ausgeschlossen. Hierbei erfolgte der Ausschluss nach fünf Kriterien: Als erstes wurden diejenigen Versuchsteilnehmer ausgeschlossen ($N = 6$), die der Pilotierung des Studiendesigns dienten. Als zweites wurde ein Versuchsteilnehmer ausgeschlossen, da es während der Erhebung zu technischen Problemen kam, sodass das Videomaterial von der Trainingsphase fehlte. Das dritte Ausschlusskriterium betrifft die Eye-Tracking-Datenqualität (gemessen anhand der WGS), wonach vier weitere Versuchsteilnehmer ausgeschlossen wurden.

Bei den WGS (vgl. Kapitel 4.1.2) handelt es sich um Prozentwerte, die das System für jeden Eye-Tracking-Test anzeigt, um die Qualität der aufgezeichneten Daten zu beurteilen. Es ist jedoch kritisch zu erwähnen, dass in Studien die WGS-Werte nicht berichtet werden und man an diese Information nur durch Nachfragen kommt. In einigen Studien gilt beispielsweise ein Wert unter 50 % als Ausschlusskriterium der Versuchsteilnehmer, wohingegen in anderen Studien keiner der Probanden wegen eines niedrigen WGS ausgeschlossen wird. Beide Vorgehensweisen haben ihre Vor- und Nachteile, auf die kurz eingegangen werden soll.

Die Vorgehensweise die WGS-Grenze auf 50 % zu legen, hat den Vorteil, dass in solch einem Datensatz weniger fehlende Werte auftauchen. Allerdings gestaltet es sich bei jungen Kindern schwierig, hohe Prozentwerte aufgrund ihrer kurzen Aufmerksamkeitsspanne zu erreichen. Die andere Vorgehensweise, trotz niedriger WGS keinen der Probanden auszuschließen, ist einerseits praktisch, da dadurch die Anzahl der Nacherhebungen geringer wird. Allerdings ist dann fragwürdig, wie die Daten anschließend ausgewertet werden, wenn ein Großteil der Blickdaten fehlt. Werden die Versuchsteilnehmer mit den niedrigen WGS nicht ausgeschlossen, dann ist des Weiteren davon auszugehen, dass auch die Gründe für die niedrigen Werte nicht hinterfragt werden.

Für die Studie der vorliegenden Arbeit wurde zur Bestimmung des WGS-Wertes, der als Maßstab für den Ausschluss von Versuchsteilnehmern gelten sollte, zunächst der Durchschnittswert (in Prozent) der WGS aller Eye-Tracking-Tests für 77 Probanden (sechs Pilotkinder wurden hierfür nicht berücksichtigt) berechnet ($N = 77$, $M = 65.69$, $SD = 13.36$). Dabei lag die Spannweite der WGS-Werte bei 59.25 % ($Min = 31.25$, $Max = 90.5$). Da der durchschnittliche WGS-Wert bei ungefähr 66 % lag, wäre bei einem Grenzwert von 50 % 12 Probanden von den Auswertungen ausgeschlossen werden müssen. Keine der Versuchsteilnehmer aufgrund der niedrigen Blickdatenqualität auszuschließen, wäre aber auch nicht vertretbar gewesen, da diese zumeist mit anderen triftigen Gründen zusammenhängen, die wiederum als Ausschlusskriterium gelten würden.

Es wurde im Rahmen der Studie der vorliegenden Arbeit der Entschluss gefasst, dass man bei Kindern nicht von einem WGS-Höchstwert von 100 % ausgehen kann, zumal der Mittelwert mit circa 66 % deutlich darunterliegt. Daher wurde der durchschnittliche Wert von 66 % als Höchstmaß und 33 % als Mindestwert, der in den vier Eye-Tracking-Tests zu erreichen gilt, festgelegt. Sofern ein Wert in einem der Tests unter die 33-Prozent-Grenze fiel, wurde dieser Wert auf eine kurze Unaufmerksamkeit oder Ablenkung zurückgeführt. Lagen jedoch zwei oder mehr der insgesamt vier WGS-Werte unter 33 %, dann wurden die

Daten dieses Versuchsteilnehmers von weiteren Analysen ausgeschlossen. Zwei Werte unter dem Mindestprozentwert deuteten meistens auf weitere kritische Aspekte hin (Krankheit, schlechte Stimmung, schreckhafter Versuchsteilnehmer), sodass dies zum Ausschlusskriterium wurde.

Während aller Eye-Tracking-Phasen wurden die Probanden von den Experimentatorinnen, die hinter der Trennwand standen, beobachtet, und das Verhalten der Kinder protokolliert haben. Aus den Protokollen sind unterschiedliche Gründe, die womöglich für die geringen Prozentwerte verantwortlich waren, zu entnehmen. Es wurde beispielsweise beobachtet, dass einige Kinder zum Ende des Experiments das Interesse an den Videos verloren haben, deshalb unruhig wurden und weniger auf den Bildschirm schauten, wodurch die Prozentwerte gesunken sind. Außerdem kam es vor, dass die Probanden sich durchgehend viel bewegt haben, sodass die Augen von dem Eye Tracker schlecht erfasst werden konnten. Es gab auch Fälle, in denen sich das Kind von anderen Personen abgelenkt war und sich zum Beispiel zu der Bezugsperson umgedreht hat oder versucht hat, hinter die Trennwand zu den Experimentatorinnen zu schauen. Auch die Veränderung der Sitzposition war eines der Gründe, der zu niedrigen WGS-Werten führte. Die Kinder rutschten von dem Schoß des Elternteils herunter oder stellten sich auf, wodurch der Eye Tracker die Augen des Probanden nicht durchgehend erfassen konnte.

In das vierte Ausschlusskriterium fielen Probanden, die vor der Vollendung der 37. SSW, ausgegangen von einer Schwangerschaft von 40 SWS, geboren wurden und unter 2500 Gramm wogen, um den Einfluss von möglichen Störungen, die in Zusammenhang mit einer kurzen Schwangerschaftsdauer und einem niedrigen Geburtsgewicht stehen (Heubrock & Petermann, 2000, S. 88), ausschließen zu können. Das Gewicht wurde berücksichtigt, da erwiesen ist, dass ein geringes Geburtsgewicht aufgrund einer Frühgeburt mit Entwicklungsstörungen einhergehen gehen kann (S. 88–90). Dieses Ausschlusskriterium traf auf vier Versuchsteilnehmer zu, die 5 bis 8.5 Wochen vor dem errechneten Geburtstermin geboren wurden und bei der Geburt zwischen 1400 und 2100 Gramm wogen.

Bei dem fünften und somit letzten Ausschlusskriterium handelt es sich um das Wohlbefinden der Probanden. Zeigte ein Kind bereits zu Beginn des Experiments deutliche Ermüdungserscheinungen oder eine schlechte Stimmung, dann wurden dessen Blickdaten ausgeschlossen. Mögliche Gründe hierfür, wie zum Beispiel Krankheiten, schlechter Schlaf oder das „Zahnen“, wurden in einem Protokoll festgehalten. Auf Grundlage dieses Kriteriums wurden insgesamt drei Versuchsteilnehmer ausgeschlossen.

4.2.1.4 Randomisierung der Variablen Geschlecht und Bedingung

Das Randomisieren ist eine Technik, bei der die Versuchsteilnehmer nach dem Zufallsprinzip verschiedenen Bedingungen zugeordnet werden. Diese Technik wird bei Versuchsplänen angewandt, damit potenzielle Störvariablen nach dem Zufallsprinzip auf verschiedene Bedingungen verteilt werden. Mit dieser Art der Kontrolltechnik geht einher, dass die Reihenfolge der Stimuli oder die Probandenanzahl in den jeweiligen Gruppen komplett dem Zufall überlassen ist, wodurch es aber zu Problemen in der Gleichverteilung kommen kann. Es besteht das Risiko, dass beispielsweise der Einfluss der Variable *Geschlecht* unbeabsichtigt gemessen wird. Aufgrund dessen wurde in der vorliegenden Arbeit die Verteilung der Variablen systematisch randomisiert. Die Technik der kontrollierten Randomisierung wird auch *Ausbalancieren* bezeichnet (Meindl, 2011, S. 39–40).

Zunächst wurde festgelegt, dass 33 Kinder je Gruppe, d. h. insgesamt 66 Personen getestet werden sollen, damit die Daten parametrischen Verfahren unterzogen werden können (über den zentralen Grundwertsatz vgl. Kapitel 4.3.2). Für die Randomisierung wurde die Regel festgelegt, dass nicht die Objekte, sondern die Objekt-Sets zufällig verteilt werden, sodass sich daraus acht Möglichkeiten und damit acht Randomisierungsgruppen ergaben (vgl. Anhang A, Tabelle 24). Zudem wurden auch die Namen den Objekten randomisiert zugeordnet. Die 66 Probanden wurden auf die acht Gruppen kontrolliert aufgeteilt, wobei der ersten Randomisierungsgruppe zehn und den restlichen Gruppen jeweils acht Probanden zugewiesen wurden. Es wurde außerdem auf eine ausgeglichene Verteilung der Geschlechter geachtet. Des Weiteren wurde von jedem Geschlecht die Hälfte der Probanden der Self-Exploring- und die andere Hälfte der Joint-Engagement-Gruppe zugeordnet (vgl. Anhang A, Tabelle 24).

4.2.2 Versuchsablauf

Im Folgenden werden allgemeine Informationen zur Studie aufgeführt, die aus insgesamt drei Experimenten besteht. In der Tabelle 2 (S. 94) ist ein Überblick über den genauen Versuchsablauf dargestellt, der in allen drei Experimenten gleich ist. Unterschiede, die die Methode betreffen, werden in der Methodenbeschreibung des jeweiligen Kontroll-experiments explizit beschrieben.

Während des gesamten Experiments waren stets zwei Experimentatorinnen anwesend. Bei der Experimentatorin 1 handelte es sich um eine wissenschaftliche Hilfskraft, die für die technischen und formellen Angelegenheiten zuständig war. Sie kümmerte sich um die Besprechung der Einwilligungserklärung, das Eye-Tracking-Verfahren, die Messung der Zeit

für die Trainingsphase sowie die Bedienung der Videokameras. Experimentatorin 2 und zugleich Autorin der vorliegenden Arbeit nahm die Probanden in Empfang, spielte während der Aufwärmphase mit den Probanden, positionierte sie vor dem Eye Tracker und interagierte mit ihnen während der Trainingsphase.

Nach dem Empfang erhielten die jeweiligen Erziehungsberechtigten der Probanden eine Einführung in den Ablauf und Zweck der Untersuchung. Sowohl diese Einführung als auch weitere (Vor-)Absprachen fanden in Einzelsettings (sprich: Experimentatorin 1 und Erziehungsberechtigte/r) statt. Danach wurde die Einverständniserklärung besprochen (vgl. Anhang C, Einverständniserklärung), woraufhin die jeweiligen Erziehungsberechtigten die Möglichkeit hatten, Fragen zu stellen. Erst nach der freiwilligen Unterzeichnung der Einwilligungserklärung wurden die Videoaufzeichnungen und somit auch das Experiment gestartet. Nach jeder Datenerhebung wurden die Rohdaten der Videoaufnahmen umgehend auf einem geschützten Server der Universität Paderborn gesichert und der Speicherort auf den Kameras formatiert, um den Datenschutz der Versuchsteilnehmer zu gewährleisten.

Ein wesentlicher Bestandteil der Studie war außerdem die Protokollführung (vgl. Anhang E, Studienprotokoll), die lediglich unter Angabe der Versuchsteilnehmer-Nummer anonymisiert waren. Darin wurde für jeden Versuchsteilnehmer die Datenqualität der Kalibrierung, das Verhalten während der Eye-Tracking-Phasen sowie über andere Auffälligkeiten (z. B. Ermüdung) berichtet. Anhand der Notizen konnten potenzielle Gründe für eine schlechte Eye-Tracking-Datenqualität geschlussfolgert werden. Aber auch Fehler, die während der Erhebungen auftraten, wurden protokolliert, um die nächsten Erhebungen zu optimieren.

Das Mixed-Methods-Forschungsdesign, das für diese Studie herangezogen wurde, besteht aus dem Eye Tracking und dem Beobachtungsverfahren. Der Ablauf der damit verbundenen Datenerhebungen wurde zum größten Teil aus dem Interactive IPLP von Pruden et al. (2006) übernommen. Somit besteht die Studie der vorliegenden Arbeit aus vier Phasen: Der Explorations-, Salienz-, Trainings- und Testphase. Die Explorationsphase wurde nicht für Auswertungen genutzt. Bei den Salienz- und Testphasen wurde das Eye-Tracking-Verfahren verwendet, wohingegen die Trainingsphase auf Video aufgezeichnet wurde. Im Folgenden soll der Ablauf der Experimente schrittweise beschrieben werden, wobei die Punkte, in denen sich die vorliegende Studie von der Studie von Pruden et al. (2006) unterscheiden, explizit erklärt werden.

Tabelle 2

Ablauf der Experimente 1–3 der vorliegenden Studie

Phase	Erklärung	Annahme
Aufwärmphase	Freie Spielsituation zur Gewöhnung an die Umgebung: Die Experimentatorin und das Kind spielen auf einer Decke im Interaktionsbereich.	
Explorationsphase	Die Probanden explorieren im Eye-Tracking-Bereich ein salientes und ein weniger salientes Objekt für jeweils 26 Sekunden.	
Salienzphase	Auf dem Bildschirm im Eye-Tracking-Bereich erscheint zweimal nacheinander das zuvor explorierte Objekt-Set für jeweils neun Sekunden.	Alle Probanden haben eine visuelle Präferenz für das saliente Objekt.
Trainingsphase	Benennungsphase: Das bekannte Objekt-Set liegt auf dem Tisch im Interaktionsbereich. Die Experimentatorin benennt das weniger saliente Objekt fünfmal und hebt es durch ostensive Signale als Zielobjekt hervor, während das saliente Objekt unbenannt bleibt.	Alle Probanden folgen den ostensiven Signalen der Experimentatorin zum Zielobjekt.
	Spielphase: Die Probanden werden zwei Gruppen zugeordnet. Die Self-Exploring-Gruppe exploriert das Zielobjekt alleine. Die Joint-Engagement-Gruppe führt die Zielfunktion mit der Unterstützung der Experimentatorin aus.	Die Joint-Engagement- und die Self-Exploring-Gruppen zeigen ein unterschiedliches Blickverhalten, da sie an verschiedenen Lernformaten teilgenommen haben.
Testphase	Original Label: Die Probanden sehen auf dem Bildschirm wieder dasselbe Objekt-Set. Dabei hören sie das zuvor trainierte Zielwort (Modi, Nebo, Guse, Doma).	Wenn die Probanden das Zielobjekt mit dem Wort während der Trainingsphase verknüpft haben, schauen sie im Original-Label-Test das Zielobjekt länger an im Vergleich zum Wahrscheinlichkeitswert.
	New Label: Die Probanden sehen dasselbe Objekt-Set auf dem Bildschirm, aber hören ein neues Wort.	Laut des Prinzips der Ausschließbarkeit sollten die Probanden nun zum namenlosen, salienten Objekt länger schauen. Dies setzt voraus, dass sie die während der Trainingsphase die richtige Wort-Objekt-Verknüpfung gebildet haben.
	Recovery: Die Probanden hören wieder das Zielwort, welches in der Trainingsphase genannt wurde.	Wenn die Probanden das Zielobjekt mit dem Wort während der Trainingsphase richtig verknüpft haben, schauen sie im Recovery-Test länger zum Zielobjekt.

4.2.2.1 Aufwärmphase

Den Phasen (*Exploration Phase, Salience Phase, Training Phase, Testing Phase*), die alle aus der Studie von Pruden und Kollegen (2006) für die vorliegende Studie übernommen wurden, wurde eine Aufwärmphase vorangestellt. Nach der Begrüßung und dem Empfang im SprachSpielLabor hatten die eingeladenen Probanden die Möglichkeit, sich an die räumliche Umgebung sowie an die Experimentatorinnen zu gewöhnen. Da die Probanden sich in der darauffolgenden Trainingsphase in einem sozialen Lernformat befinden, sollten sie vor allem mit der Person, mit der sie interagieren werden (hier: Experimentatorin 2), schon vorher durch gemeinsames Spielen in Kontakt treten. Diese erste Phase hatte kein festes Format, sondern gestaltete sich individuell nach den Bedürfnissen des Kindes.

Die Aufwärmphase fand hauptsächlich auf einer Spieldecke in dem Interaktionsbereich statt (vgl. Abbildung 3, S. 86) und dauerte ungefähr 15 Minuten. Spielsachen, die den Stimuli optisch glichen oder ähnliche Funktionen hatten, wurden im Vornherein aussortiert. Von Anfang an hatte die Experimentatorin eine Kappe auf dem Kopf (für die Funktion der Kappe vgl. Kapitel 4.2.2.4), damit das Kind sich auch daran gewöhnen konnte. Sobald beide Seiten bereit waren, wurde der jeweilige Erziehungsberechtigte mit dem Kind hinter die Trennwand in die Eye-Tracking-Ecke begleitet und gebeten, für die anschließende Explorations- und Salienzphase auf einem Stuhl Platz zu nehmen. In der Zwischenzeit wurden die Spielsachen von der Experimentatorin 1 weggeräumt, damit das Kind in der auf die Explorations- und Salienzphase folgenden Trainingsphase, die in der Interaktionsecke stattfand, nicht abgelenkt wird. Danach schaltete die Experimentatorin 1 die Kamera/s an, um einen fließenden Ablauf von der Salienzphase zur Trainingsphase zu gewährleisten.

4.2.2.2 Explorationsphase

Die Explorationsphase fand hinter der Trennwand in der Eye-Tracking-Ecke statt (vgl. Abbildung 4, S. 86), währenddessen der Bildschirm jedoch ausgeschaltet war, damit die Versuchsperson nicht abgelenkt wird. Dabei saß das Kind auf dem Schoß des jeweiligen Erziehungsberechtigten.

Das Ziel der Explorationsphase bestand darin, die Kinder die Stimuli auf ihre Weise erforschen zu lassen, sei es haptisch, oral oder visuell, bevor sie diese auf dem Bildschirm sehen. Wie in der Studie von Pruden et al. (2006) beschrieben, hatten auch die Probanden der vorliegenden Studie für jedes der beiden Objekte 26 Sekunden Explorationszeit. Im Gegensatz zu der Originalstudie, in der die Reihenfolge der Objekte variierte, wurde für die vorliegende Arbeit beschlossen, dass zuerst das langweilige Objekt und danach das saliente

Objekt den Kindern gereicht werden soll. Diese Reihenfolge wurde festgelegt, da bei den Pilotkindern festgestellt worden ist, dass sie kaum an dem langweiligen Objekt interessiert waren, nachdem sie das saliente Objekt als erstes exploriert hatten.

Für die Explorationsphase wurden duplizierte Modelle der Zielobjekte (vgl. Kapitel 4.2.3.1) verwendet. Diese sahen im Vergleich zu den Zielobjekten identisch aus, waren jedoch zugeklebt, damit die Probanden die Zielfunktion des Objektes nicht vor der Trainingsphase entdecken. Zunächst erhielten die Probanden das langweilige Objekt des ersten Stimuli-Sets mit der Äußerung „Hey, schau mal, was ich hier für dich habe!“ Während der 26 Sekunden, in denen sich das Kind mit dem Objekt beschäftigte, sagte die Experimentatorin 2 nichts weiter und ihr Blick war auf das Objekt gerichtet. Nach Ablauf der Zeit wurde das saliente Spielzeug mit dem Satz „Schau, was ist das denn?“ hervorgeholt und gegen das Holzobjekt ausgetauscht. Nun hatten die Kinder wieder 26 Sekunden Zeit, das saliente Objekt zu explorieren. Nachdem die Zeit verstrichen war, wurde das saliente Objekt aus dem Sichtfeld der Kinder genommen und die Bezugsperson gebeten eine bequeme Position für das darauffolgende Eye-Tracking-Verfahren einzunehmen.

Zwischen der Explorations- und der Salienzphase fand die Kalibrierung des Eye Trackers statt. Beide Experimentatorinnen standen während des Kalibrierungsprozesses hinter der Trennwand und beobachteten das Verhalten des Kindes. Während der Kalibrierung sowie der darauffolgenden Präsentation der Stimuli wurden die Erziehungsberechtigten noch einmal explizit darauf hingewiesen, dass sie ihre Augen schließen müssen, damit sichergestellt werden kann, dass der Eye Tracker die Augen des Kindes und nicht die des Erwachsenen erfasst.

Die Nutzung eines Eye Trackers stellt eine wesentliche Modifizierung der Studie von Pruden et al. (2006) dar, in der die Stimuli in allen Phasen nebeneinander auf einem schwarzen Brett, das auf einem Tisch stand, befestigt wurden. Während in der Studie von Pruden et al. (2006) das Blickverhalten zu den Stimuli von einer Videokamera erfasst und anschließend manuell kodiert wurde, ist das Blickverhalten der Probanden während der Salienz- und Testphasen in den folgenden Experimenten der vorliegenden Arbeit mithilfe eines Eye Trackers erfasst worden. Ein weiterer Unterschied zu der Studie von Pruden und Kollegen (2006) besteht darin, dass in der vorliegenden Studie der auditive Input aus den Lautsprechern ertönte und somit gleichzeitig mit der visuellen Stimuli-Präsentation abgespielt wurde.

4.2.2.3 Salienzphase

Direkt im Anschluss an die Kalibrierung begann die Salienzphase. Diese hatte zum Ziel, die Präferenz für eines der Objekte zu erfassen. Der Salienztest besteht aus zwei aufeinanderfolgenden statischen Videoclips, die mithilfe der Software *iMovie* (Version 10.1.2) auf dem MacBook Air (OS X El Capitan, Version 10.11.6) erstellt wurden.

Zu Beginn der Aufnahme tauchte auf einem weißen Hintergrund ein Aufmerksamkeitslenker auf. Daraufhin startete der erste Videoclip der ersten Salienzphase (im Eye-Tracking-System als S1.1 benannt) mit einem Bild beider Stimuli nebeneinander auf einem weißen Hintergrund. Es ertönte ein 0,5 Sekunden langer Signalton und daraufhin folgte ein auditiver Stimulus „Hey! Was siehst du?“. Nach dem sprachlichen Input war das Bild des Objekt-Sets noch weitere 6 Sekunden auf dem Bildschirm zu sehen. Nach Ablauf der Zeit erschien ein weiterer aufmerksamkeitslenkender Stimulus. Hiernach wurde der zweite Videoclip mit den beiden Objekten gezeigt. Der zweite Clip begann wieder ein kurzer Signalton, auf den randomisiert entweder die Phrase „Oh! Schau doch mal!“ oder „Wow! Schau doch mal!“ als auditiver Stimulus folgte. Nach weiteren sechs Sekunden war die Salienzphase beendet.

Alle Audioaufnahmen wurden von einer Sprecherin mit Deutsch als Muttersprache zuerst auf einem Diktiergerät (Zoom H4n Pro 4-Kanal-Handyrekorder) aufgenommen und danach für die Videosequenzen geschnitten. Die Lautstärke der Audio-Stimuli wurde auf 18 dB reguliert. In der Abbildung 6 sind die beiden jeweils 9-sekündigen Videoclips der Salienzphase visuell dargestellt.

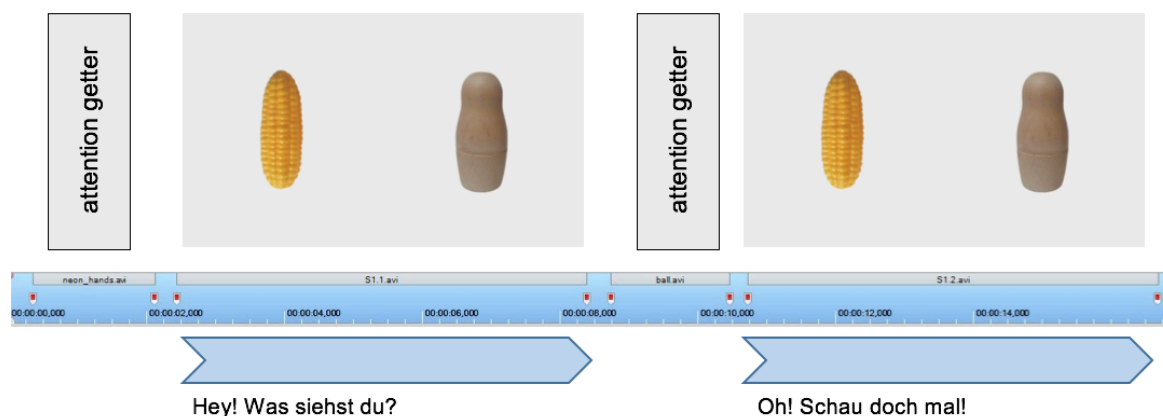


Abbildung 6. Zeitachse der Salienzphase am Beispiel der Randomisierungsgruppe 1.

4.2.2.4 Trainingsphase

Im Anschluss an die Salienzphase erfolgte die erste Trainingsphase. Das Ziel dieser Phase bestand darin, dass die 10 Monate alten Probanden das Wort mit dem Zielobjekt verknüpfen. Hierfür wurde die Bezugsperson gebeten, mit dem Kind am Tisch in der Interaktionsecke Platz zu nehmen. Währenddessen stand Experimentatorin 1 hinter dem Paravent und legte die Stoppuhr bereit. Experimentatorin 2 nahm dem Kind gegenüber Platz, welches auf dem Schoß der Bezugsperson am Tisch saß. Das Objekt-Set, welches die Versuchsperson soeben in der Explorations- und Salienzphasen gesehen hat, lag bereits auf dem Tisch und die Experimentatorin saß mittig zwischen den beiden Objekten.

In dem ersten Teil der Trainingsphase, der Benennungsphase, sollten die Enkodierungsstrategien des assoziativen und des sozio-pragmatischen Ansatzes einander gegenübergestellt werden. Die Benennungsphase startete mit einer *ostensiven Benennung* (engl.: *ostensive naming*) (vgl. Kapitel 3.4.3), in der das neue Wort insgesamt fünf Mal genannt wurde. Diese Phase begann mit einer namentlichen Adressierung an das Kind „Hallo, [*Name des Kindes*]!“, bis ein gegenseitiger Blickkontakt sichergestellt wurde. Danach drehte die Experimentatorin ihren Kopf in Richtung des Zielobjektes, schaute dieses an und benannte es ostensiv: „Schau! Ein *X*!“ Mit den folgenden drei Benennungsphrasen – „Guck mal, was man mit dem *X* machen kann!“, „Wow, was für ein schönes *X*!“ und „Schau dir das *X* mal an!“ gingen gleichzeitig drei Ausführungen der Zielfunktion einher. Währenddessen war der Blick der Experimentatorin stets auf das Zielobjekt gerichtet, die Bewegungen waren langsam und die Sätze wurden laut und deutlich gesprochen. Danach stellte die Experimentatorin das Zielobjekt in die Ausgangsposition zurück, schaute das Kind an und fragte „Kannst du das auch machen?“. Daraufhin wurde dem Kind mit der Äußerung „Hier, ich geb’ dir das *X*!“ das Zielobjekt gereicht.

In der Studie der vorliegenden Arbeit schaute die Experimentatorin das Zielobjekt nicht nur an und benannte es, sondern präsentierte zusätzlich mehrmals die Zielfunktion des langweiligen Objektes (vgl. konträr zu Pruden et al., 2006). Das heißt, dass das langweilige Objekt stets als Zielobjekt fungierte, wohingegen das saliente Objekt immer als Ablenker galt, welcher ebenfalls im visuellen Blickfeld des Kindes lag. Im Gegensatz dazu wurde in der Originalstudie von Pruden et al. (2006) unterschieden zwischen der Coincidental- und Conflict-Bedingung, da in der einen Bedingung das saliente Objekt (Coincidental) und in der anderen Bedingung das langweilige Objekt (Conflict) benannt wurde.

Die Benennungsphase bestand aus einer Reihe von ostensiven Signalen, die nach den Kriterien von Sperber und Wilson (1995) verwendet wurden (vgl. Kapitel 3.4.3). Sie definieren ostensive Signale als wechselseitigen Blickkontakt, kindgerichtete Sprache und an das Kind gerichtetes Handeln. Ein wechselseitiger Blickkontakt sei für beide Kommunikationspartner ein Hinweis dafür, dass der jeweils andere aufnahmebereit ist für neue Informationen bzw. etwas vermitteln möchte (Csibra, 2010, S. 145). Dies wurde mit dem direkten Anschauen und der namentlichen Adressierung des Kindes durchgeführt.

Eine kindgerichtete Sprache (Csibra, 2010, S. 147–148; Sperber & Wilson, 1995) ist das zweite Kriterium der ostensiven Signale, womit die Anpassung von Erwachsenen in ihrer Prosodie gemeint ist, wenn sie mit Kindern reden. Damit lenken sie nicht nur die Aufmerksamkeit auf sich, sondern zeigen dem Kind auch, dass es der Adressat ist (Csibra 2010, S. 147). Während der ostensiven Benennung in der Studie der vorliegenden Arbeit hat die Experimentatorin ebenfalls ihre Prosodie verändert, indem sie langsamer und deutlicher mit dem Kind sprach, als wenn sie zu einem Erwachsenen sprechen würde. Außerdem hat eine alternierende Nutzung von Imperativ- und Interrogativsätzen zu einer wechselnden Intonation beigetragen.

Neben der an das Kind gerichteten Sprache bzw. *Motherese* wurde auch *Motionese* in die Trainingsphase eingebaut. Darunter versteht man, ähnlich wie Motherese, an das Kind gerichtete Bewegungen, Handlungen und Gesten, weil diese in ihrer Komplexität einfacher sind, häufiger wiederholt und mit mehr Enthusiasmus präsentiert werden als einem Erwachsenen (Brand, Baldwin & Ashburn, 2002; vgl. Rohlfsing, 2019, S. 291). In dem Fall der vorliegenden Studie allerdings nicht von der Mutter ausgehend, sondern von der Experimentatorin. In der Studie dieser Arbeit wurden die Zielobjekte in der Trainingsphase mit Rücksichtnahme auf Motionese demonstriert, indem die Objektfunktion durch einfache, wiederholende Bewegungen enthusiastisch vorgeführt wurde.

Neben dem Vergleich der perzeptuellen Salienz und den sozialen Hinweisen hinsichtlich der Aufmerksamkeitslenkung, sollten auch die zwei sozialen Wortlernformate – der sozio-pragmatische und der interaktionistische Ansatz – einander gegenübergestellt werden. Aus diesem Grund folgte im Training auf die Benennungsphase eine Spielphase. Während der Ablauf im ersten Teil der Trainingsphase für alle Probanden gleich war, wurden die Probanden im zweiten Teil der Trainingsphase zwei Gruppen zugeordnet – der Joint-Engagement- und der Self-Exploring-Gruppe. In der Self-Exploring-Gruppe zog die Experimentatorin ihre Kappe herunter und unterbrach somit den Blickkontakt zum Kind. Damit signalisierte die Experimentatorin, dass sie sich nicht aktiv an der Interaktion

beteiligen wird. Die Probanden waren also frei in ihren Objekthandlungen und hatten die Möglichkeit, die präsentierte Zielfunktion zu imitieren oder aber das Objekt auf eine ganz eigene Art und Weise zu manipulieren.

Die Kappe wurde deshalb als Utensil verwendet, um nicht nur den Gesichtsausdruck zu verbergen, sondern auch den Blickkontakt zum Kind zu unterbrechen, um den Probanden deutlich zu machen, dass sich die Experimentatorin aus dem weiteren Handeln herauszieht. Eine Variante ohne die Kappe würde bedeuten, dass die Experimentatorin einen neutralen Ausdruck aufsetzen muss, denn ein positiver Gesichtsausdruck (z. B. Lächeln) könnte die Bereitschaft für eine Interaktion signalisieren. Allerdings könnte auch ein neutraler Gesichtsausdruck negative Emotionen bei dem Kind hervorrufen. Dies wird vor allem in einem Still-Face-Experiment (Weinberg & Tronick, 1996) deutlich, in dem die Forscher zeigten, wie sozial Kinder bereits im Alter von sechs Monaten sind und sogar mit Tränen auf das Zurückhalten der Reaktion bzw. einen starren Blick seitens des Elternteils reagieren können. Es wurden auch alternative Möglichkeiten durchdacht, zum Beispiel, dass die Experimentatorin den Raum verlässt oder einer anderen Tätigkeit, bspw. Lesen nachgeht, während die Versuchsperson das Objekt manipuliert. Allerdings würde das Verlassen des Raumes mit einer enormen Veränderung der Situation einhergehen. Es ist unbekannt, wie sich eine neue Situation auf das kindliche Verhalten oder gar das Lernen auswirken kann. Bei der zweiten Option besteht das Risiko, dass sich die Aufmerksamkeit der Probanden von der eigentlichen Aufgabe auf die neue Tätigkeit der Experimentatorin, zum Beispiel auf das Lesen oder das Schreiben, verschiebt.

Im Gegensatz zur Self-Exploring-Gruppe wurde die Joint-Engagement-Gruppe von der Experimentatorin in eine gemeinsame Interaktion eingebunden, in der sie anstrebte, die Zielfunktion mit dem Kind mehrmals zu erreichen. In der Joint-Engagement-Gruppe hielt die Experimentatorin das Zielobjekt so fest, dass das Kind beispielsweise an dem oberen Teil der Schachtelpuppe ziehen konnte. Die Joint-Engagement-Gruppe wurde dabei nicht nur bei der Ausführung der Funktion unterstützt, sondern ihr wurde auch jedes Mal, wenn die Zielfunktion erreicht wurde, ein positives Feedback in Form von Lob gegeben und gebeten, die Handlung zu wiederholen. Obwohl sich die Experimentatorin aus der Interaktion mit der Self-Exploring-Gruppe herausgezogen hat, wurde auch diese Gruppe nach Erfüllen der Zielfunktion gelobt, um das positive Feedback in beiden Gruppen vergleichbar zu halten.

Nach Ablauf der 30-sekündigen Spielphase signalisierte die Experimentatorin 1 durch ein leises Räuspern, dass die Trainingsphase beendet ist. Das Zielobjekt wurde wieder zurück in die Ausgangsposition gebracht und die Bezugsperson wurde gebeten, mit dem Kind wieder vor dem Bildschirm in dem Eye-Tracking-Bereich Platz zu nehmen.

4.2.2.5 Testphase

Nach der Trainingsphase wurde der Eye-Tracking-Test durchgeführt. Die Bezugsperson platzierte sich genauso wie in der Salienzphase vor dem Bildschirm, während das Kind auf dem Schoß saß. Die Testphase sollte überprüfen, ob die Kinder nach dem Training die enkodierte Wort-Objekt-Verknüpfung rezeptiv abrufen können. Zur Überprüfung des Wortverständnisses wurden den Versuchsteilnehmern insgesamt drei Tests (2x *Original Label*, *New Label*, *Recovery*), die in vier Videoclips umgesetzt wurden, präsentiert. In der Abbildung 7 (S. 102) ist die Abfolge der Videoclips dargestellt.

Vor dem ersten Original-Label-Test wurde als erstes ein aufmerksamkeitslenkender Stimulus eingeblendet, der die kindliche Aufmerksamkeit zum Bildschirm lenken sollte. Dann ertönte ein Signalton, woraufhin das Stimuli-Set auf einem weißen Bildschirmhintergrund präsentiert wurde. In dem ersten 9-sekündigen Videoclip wurden die Probanden mit den Worten „Schau! Da ist das *X*!“ aufgefordert, das zuvor in der Trainingsphase benannte Zielobjekt anzuschauen. An den ersten Videoclip schloss sich der zweite Original-Label-Test mit Frage „Siehst du das *X*?“ an. Der New-Label-Test bezieht sich auf den dritten Videoclip, in dem die Probanden einen neuen Objektnamen hörten („Wo ist das *Y*?“). Nach dem New-Label-Test folgt der Recovery-Test, in dem die Probanden aufgefordert werden wieder zum Zielobjekt zu schauen („Schau! Das *X*!“). Es wird angenommen, dass die Probanden ihren Blick wieder auf das Zielobjekt richten, sobald sie das Zielwort hören.

Pruden et al. (2006) sowie Hollich (2000) argumentieren, dass in bisherigen Wortlernstudien lediglich der Original-Label-Test angewandt wurde. Die Annahme des New-Label-Tests ist, dass Kinder, sofern sie die Wort-Objekt-Verbindung gelernt haben, ihren Blick vom Zielobjekt abwenden werden, um nach dem Referenten für das neue Wort zu suchen. Somit müsste die Blickzeit im New-Label-Test zum salienten Objekt steigen. Diese Annahme greift das Prinzip der Ausschließbarkeit auf, nach welchem Kinder neue Wörter mit neuen Objekten in Zusammenhang bringen.

Die Videoclips der Testphase wurden in Onset- (vor und während der Wortäußerung) und Offset-Phasen (nach der Wortäußerung) aufgeteilt, um zu untersuchen, wo die Probanden zu diesen Zeitpunkten hinschauen. Aufgrund des jungen Alters der Kinder wurde

das Wortverständnis erst nach der Wortäußerung gemessen, da sie sprachlichen Input noch nicht so schnell verarbeiten wie ältere Kinder (Zangl & Fernald, 2007). Da in der Studie von Pruden und Kollegen (2006) nicht explizit zwischen dem Wortverständnis vor und nach der Wortäußerung unterschieden wurde, ist davon auszugehen, dass sie die Gesamtblickdauer in den jeweiligen Tests ausgewertet haben.

Im Vergleich zu der Studie von Pruden und Kollegen (2006) wurden in der vorliegenden Studie die Präsentationsseiten der Stimuli in der vorliegenden Arbeit kontrollierter und häufiger randomisiert. Dies war in der Originalstudie nicht möglich, da die Objekte auf einem Brett befestigt waren, während der Versuchsleiter hinter einer Trennwand den sprachlichen Input gab. In dem zweiten Experiment modifizierten die Forscher zwar das Verfahren, indem sie das Zielobjekt während der Salienzphase beispielsweise auf der linken Seite und während der Testphase auf der rechten Seite befestigten. Allerdings behielt es während der Trainingsphase dieselbe Seite wie in der Salienzphase. In der vorliegenden Studie wurden die Seiten nicht nur zwischen der Salienz-, Trainings- und Testphase, sondern auch innerhalb der einzelnen Testphasen variiert (vgl. Abbildung 7).

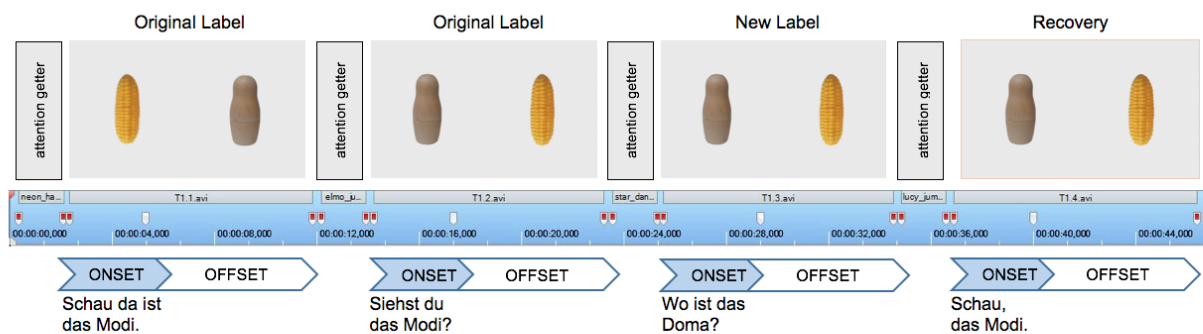


Abbildung 7. Zeitachse der Testphase am Beispiel der Randomisierungsgruppe 1.

Abschließend ist zu erwähnen, dass genauso wie in der Studie von Pruden et al. (2006) auch in der vorliegenden Studie jedes Experiment aus zwei Versuchen bestand. Das bedeutet, dass im Anschluss an den soeben beschriebenen Versuchsablauf mit dem ersten Objekt-Set, die Explorationsphase mit dem zweiten Objekt-Set (vgl. Kapitel 4.2.3.1) folgte.

4.2.3 Stimulusmaterial

Im Folgenden soll darauf eingegangen werden, nach welchen Kriterien die Objektauswahl getroffen wurde und wie diese Objekte benannt wurden.

4.2.3.1 Objekte

Den Kindern wurden zwei Stimuli-Sets präsentiert, die jeweils aus zwei Objekten, einem salienten und einem „langweiligen“, bestanden (vgl. Abbildung 8, S. 104). Als Kriterium für die Stimuli-Auswahl wurde festgelegt, dass die beiden Objekte innerhalb der Sets in ihrer Form und Größe vergleichbar sein sollen. Für die salienten Objekte wurde zudem beschlossen, dass diese eine ähnliche Komplexität haben und perzeptuell auffällig sein sollen. Die perzeptuelle Salienz wurde durch eine grelle Farbe und eine interessante Form des Objekts erreicht. Zu den salienten Objekten zählte ein gelber Maiskolben (9.5 cm x 3.2 cm) und ein blaues Schiff (5.5 cm x 10.5 cm), deren materielle Beschaffenheit Plastik war.

Für die Auswahl der Stimuli, die als weniger-saliente Zielobjekte dienen sollten, stand eine dezente, optisch unauffällige Beschaffenheit im Vordergrund. Die Objekte sollten eine unauffällige Farbe und eine glatte, matte Oberfläche haben. Als eine sinnvolle Lösung haben sich Holzspielzeuge angeboten. Das zweite Kriterium war, dass die Objekte eine Funktion haben sollten, die 10 Monate alte Kinder alleine, aber auch in einer gemeinsamen Handlung mit der Experimentatorin ausführen können. Daraus wurde geschlussfolgert, dass die langweiligen Objekte idealerweise aus zwei Teilen bestehen sollten und die Objektbestandteile von der Größe her für Kinder greifbar sind. Hierfür wurde ein Holzklotz (4 cm x 6 cm) in der *Objektwerkstatt* in Paderborn²² in zweifacher Ausführung hergestellt. Als zweites „langweiliges“ Objekt wurden zwei hölzerne Schachtelpuppen (8.3 cm x 4 cm) verwendet. Die Zielfunktion der Schachtelpuppe bestand darin, sie auf und zu zumachen. In dem Klotz steckte außerdem noch ein Rundholz (5 cm x 3.3 cm), welches man hineinstecken und herausziehen konnte.

Der Maiskolben und die hölzerne Schachtelpuppe sowie der Klotz und das Schiff wurden als ein Objekt-Set zusammengeschlossen. Damit die Kinder die Funktionen der langweiligen Objekte nicht schon in der Explorationsphase herausfinden, sondern erst während des Trainings erfahren, wurden die zweiten Anfertigungen der beiden Holzobjekte zugeklebt, sodass sich das Rundholz nicht aus dem Klotz ziehen ließ und die Schachtelpuppe nicht mehr aufging. Anschließend wurden die Holzobjekte mit einem Klackklack speziell für Kinderspielzeuge beschichtet.

²² Eine Werkstatt in der Möbelstücke, dekorative Objekte und Kleinkunst aus heimischen Hölzern gefertigt werden (siehe: www.objektwerkstatt.de).

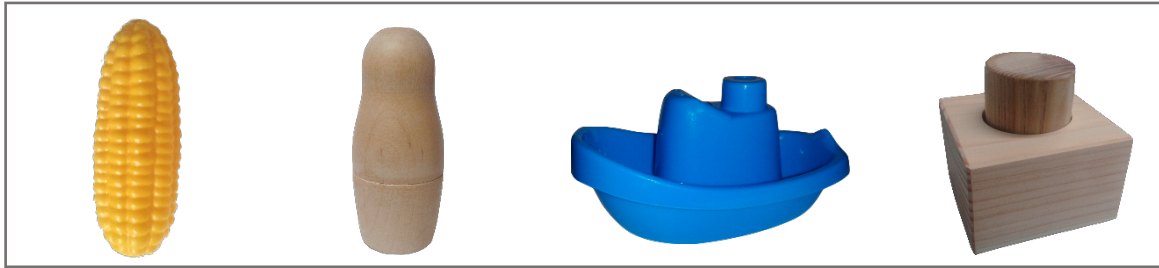


Abbildung 8. Verwendete Stimuli: (1) Maiskolben, (2) Schachtelpuppe, (3) Schiff, (4) Klotz

Als Nächstes soll auf die kontrollierte Randomisierung der vier Stimuli eingegangen werden (vgl. Anhang A, Tabelle 24). Dabei handelt es sich zum einen um die Bestimmung der Reihenfolge der präsentierten Stimuli-Sets, die Festlegung der Seitenposition der Objekte in den jeweiligen Phasen und um die Zuordnung der Wörter zu den Objekten. Da es sich insgesamt um vier verschiedene Objekte handelt, die in derselben Zweier-Kombination in zwei Durchgängen präsentiert werden sollen, gibt es für das erste Objekt vier (n) Auswahlmöglichkeiten: Es kann in dem ersten Versuch des Experiments entweder links oder rechts auf dem Monitor erscheinen und es wird in dem zweiten Durchgang entweder links oder rechts eingefügt. Somit bleiben für das zweite Objekt drei Positionen übrig ($n - 1$), für das dritte bleiben zwei Möglichkeiten ($n - 2$) und für das vierte Objekt nur noch eine Möglichkeit ($n - 3$). Wenn y die Objektanzahl ist und n für die Positionsmöglichkeiten eines Objekts steht, dann ergeben sich daraus die gesamten Randomisierungsmöglichkeiten x . Der folgenden Formel nach gibt es 24 Möglichkeiten der Stimulusanordnung:

$$y \times (n - 1) \times (n - 2) \times (n - 3) = x$$

$$4 \times (4 - 1) \times (4 - 2) \times (4 - 3) = 24$$

Um die 24 Möglichkeiten einzugrenzen, wurde die Regel festgelegt, dass der Maiskolben mit der Schachtelpuppe sowie das Schiff und der Klotz nur gemeinsam in der Konstellation auftauchen. Daraus ergaben sich acht Möglichkeiten mit der Stimulusanordnung und somit acht Randomisierungsgruppen (vgl. Anhang A, Tabelle 24).

4.2.3.2 Pseudowörter

In der Studie von Pruden et al. (2006) wurden die Objektamen *modi*, *glorp*, *dawnoo* und *blicket* verwendet (S. 269). Unter Beachtung der phonotaktischen Regeln des Deutschen wurden die Namen in der Studie der vorliegenden Arbeit verändert. Zum einen sollten die Wörter zwar einen deutschen Wortlaut haben, jedoch nicht Teil des deutschen Wortschatzes

sein. Durch die Verwendung solcher Pseudowörter soll kontrolliert werden, dass keines der Kinder die Wörter kennt und die Lernvoraussetzungen somit gleich sind. Des Weiteren sollten die Wörter vergleichbar sein, damit ausgeschlossen werden kann, dass ein Wort im Gedächtnis bleibt, weil es phonotaktisch besonders auffällig klingt im Vergleich zu den anderen Wörtern. Aus diesem Grund wurde festgelegt, dass alle Wörter zweisilbig sind und aus stimmhaften Konsonanten und Vokalen zusammengesetzt sein sollen. Daraus ergaben sich die Kunstwörter *Modi*, *Doma*, *Guse* und *Nebo*.

Nachdem die Reihenfolge der Stimuli sowie die Objektnamen festgelegt wurde, konnten die Objektnamen den Stimuli kontrolliert zugeordnet werden. Da es insgesamt acht Randomisierungsgruppen mit je vier Objekten sind, wurde bei der Zuordnung jedem Objekt zweimal derselbe Name zugewiesen. Die im Anhang der vorliegenden Arbeit eingefügte Tabelle 24 (vgl. Anhang A) stellt die Anordnung und Benennung der Stimuli dar.

4.2.4 Kodierschema

Das während der Trainingsphase erhobene Videomaterial wurde mit der ELAN-Version²³ 5.2 annotiert (Sloetjes & Wittenburg, 2008). Bei ELAN handelt es sich um eine Annotationssoftware, entwickelt im Max-Planck-Institut für Psycholinguistik in Nijmegen, mit der sich Audio- sowie Videoaufnahmen einlesen lassen, um diese beispielsweise hinsichtlich der Sprache oder der Gestik zu transkribieren. Anschließend können die angelegten Annotationen zum Zwecke statistischer Auswertungen exportiert werden.

Das Kodierschema der vorliegenden Arbeit umfasst insgesamt sechs unterschiedliche Kategorien, die im Hinblick auf die wissenschaftliche Fragestellung und Hypothesen der vorliegenden wissenschaftlichen Abhandlung entworfen wurden. Dazu zählen fünf Kategorien, die das kodierte Verhalten der Kinder enthalten: *Blickrichtung* (in ELAN als *Gaze 1*, *Gaze 2*, *Gaze 3*, *Gaze 4* kodiert) und *Handlungsstrategien* (als *Strategies* kodiert). In der Kategorie *Soziale Hinweise* (als *Social Cues* kodiert) wurde das soziale Verhalten der Experimentatorin kodiert. In den folgenden Unterkapiteln werden zunächst die Kategorien vorgestellt, die sich auf das Verhalten der Kinder beziehen, und anschließend die Kategorie Soziale Hinweise beschrieben.

Zum allgemeinen Verständnis des Kodierschemas ist es wichtig, zwei Besonderheiten der Trainingsphase noch einmal zu betonen: Erstens sind insgesamt zwei Trainingsphasen

²³ Das Akronym ELAN steht für *EUDICO Linguistic Annotator*, ein Annotationsprogramm für Video- und Audio-Dateien (siehe http://www.mpi.nl/tools/elan/ELAN_Manual.pdf für nähere Informationen).

vorhanden, da es zwei Objekt-Sets gibt. Zweitens ist jede Trainingsphase in zwei Abschnitte gegliedert (vgl. Kapitel 4.2.2.4). Während im ersten Abschnitt (Benennungsphase) die Experimentatorin das neue Wort sowie die Funktion des Zielobjekts präsentiert und das Kind dieses Geschehen beobachtet, darf das Kind im zweiten Abschnitt selber mit dem Objekt handeln (Spielphase).

4.2.4.1 Blickrichtung

Eine Kategorie des Kodierschemas ist die Blickrichtung der Probanden. Diese Kategorie wurde aufgrund der Abfolge der beiden Objekt-Sets und der Aufteilung in eine Benennungs- und Spielphase in vier einzelnen Zeilen kodiert (Gaze 1, Gaze 2, Gaze 3, Gaze 4) (vgl. Tabelle 3, S. 107). Für jede Zeile wurden die fünf Annotationen *Zielobjekt*, *Ablenker*, *Experimentatorin*, *Bezugsperson* und *Sonstiges* (in ELAN kodiert als *Target object*, *Distractor*, *Experimentator*, *Caregiver* und *Other*) als mögliche Blickrichtungen der Probanden angelegt. In den Transkriptionen folgen diese Annotationen direkt aufeinander, da Sakkaden und Augenblinzeln nicht gesondert kodiert wurden. Als Startpunkt einer Annotation galt bereits die Augenbewegung in Richtung des Fixationsziels und sie endete, sobald dieser Bereich nicht mehr fixiert wurde.

Die Zeilen Gaze 1 und Gaze 3 stehen für das Blickverhalten der Kinder zum ersten bzw. zweiten Objekt-Set während der Benennungsphase. Die Kodierung beginnt mit der Äußerung der Experimentatorin „Hallo *Name des Kindes*“ und der endet mit „Hier, ich geb’ dir das *X!*“. Das vorrangige Ziel dieser beiden Zeilen der Gaze-Kategorie ist die Erfassung des kindlichen Blickverhaltens während der Benennungsphase, um zu untersuchen, ob 10-Monatige den ostensiven Signalen der Experimentatorin zu dem Zielobjekt folgen oder ob ihre Aufmerksamkeit auf den salienten Distraktor gerichtet ist.

Die Zeilen Gaze 2 und Gaze 4 folgen direkt im Anschluss auf Gaze 1 bzw. Gaze 3 und stehen für das Blickverhalten der Kinder während der ungefähr 30-sekündigen Spielphase. Abhängig davon, in welcher Gruppe sich die Versuchsteilnehmer befanden, manipulierten sie in diesem Zeitraum das Objekt entweder alleine oder mit der Unterstützung der Experimentatorin.

Tabelle 3

Die Kategorie Blickrichtung mit den jeweiligen Annotationen

Blickrichtung	Beschreibung
Zielobjekt	Das Kind schaut zum nicht-salienten Objekt (Holzpuppe oder Holzklötz). Ist das Objekt auf dem Kamerabild nicht zu sehen, da es auf den Boden fällt, dann wird die Blickrichtung als Sonstiges kodiert.
Ablenker	Das Kind richtet seinen Blick zum salienten Ablenker (Maiskolben oder Schiff). Ist der Ablenker auf dem Kamerabild nicht zu sehen, da es auf den Boden fällt, dann wird die Blickrichtung als Sonstiges kodiert.
Experimentatorin	Die Blickrichtung wird zur Experimentatorin kodiert, sofern sie auf dem Video zu sehen ist. Wenn sie aufsteht, um ein Objekt aufzuheben, und dadurch nicht im Kamerabild zu sehen ist, wird die Stelle als Sonstiges kodiert.
Bezugsperson	Die Blickrichtung zur Bezugsperson wird kodiert, wenn sich das Kind nach hinten zu ihr umdreht.
Sonstiges	Eine Blickrichtung wird jeweils dann als Sonstiges kodiert, wenn das Blickverhalten keiner der anderen vier Annotationen zugeordnet werden kann.

4.2.4.2 Handlungsstrategien

Schließlich ist noch die Kategorie der Handlungsstrategien (Strategies) aufzuführen (vgl. Tabelle 4, S. 108). Darunter wurden jegliche Manipulationen des Zielobjektes erfasst, die bei den Kindern während der Spielphase beobachtet und als Annotationen *Öffnen/Schließen*, *Rausziehen/Einsetzen*, *Sonstige*, *Verweigerung* (in ELAN kodiert als *Open/Close*, *Pull/Put*, *Other* und *Refuse*) angelegt wurden.

Zweck dieser Kategorie war es, die Strategietypen sowie die Anzahl der erreichten Zielfunktionen festzuhalten. Zur Annotation Öffnen/Schließen, Rausziehen/Einsetzen zählt die Ausführung der Zielfunktion mit den Holzobjekten²⁴, worunter das Auf- und Zumachen der Holzpuppe sowie das Einsetzen und Rausziehen des Rundholzes aus dem Klotz fällt. Alle anderen Strategien, die die Kinder ausführten, aber zur Zielfunktion zählten, wurden als Sonstige kodiert. Eine weitere Annotation ist die Verweigerung der Kinder, das Zielobjekt zu manipulieren.

²⁴ In den Transkripten sind die Strategien der Joint-Engagement-Gruppe mit dem Suffix *_joint* und die Strategien der Self-Exploring-Gruppe mit *_self* markiert, um zu erkennen, welcher Gruppen die Versuchsteilnehmer zugehören. In der vorliegenden Dissertationsschrift wird das Suffix in den Bezeichnungen jedoch ausgelassen.

Tabelle 4
Die Kategorie Handlungsstrategien mit den jeweiligen Annotationen

Handlungsstrategien	Beschreibung
Öffnen/Schließen und Rausziehen/Einsetzen	Öffnen und Schließen beziehen sich auf die Zielfunktionen der Holzpuppe. Rausziehen und Einsetzen auf den Holzklotz. Wurde das Ziel erreicht, unabhängig davon, ob es zufällig war oder nicht, wird die entsprechende Annotation gesetzt.
Sonstige	Hiermit wurden alle Strategien annotiert, z. B. mit dem Objekt auf den Tisch klopfen, über den Tisch streichen, auf den Boden werfen, die nicht zu den Zielfunktionen zählen. Wurde mit dem Objekt zuerst auf den Tisch geklopft und danach über den Tisch gestrichen, dann wurden die Handlungen als zwei verschiedene Annotationen vermerkt. Wenn das Kind eine Handlung ausführt, stoppt und dann dieselbe Handlung fortsetzt, wurde es als eine Annotation kodiert.
Verweigerung	Diese Annotation steht für die Verweigerung des Kindes mit dem Objekt zu handeln oder mit der Experimentatorin in die Interaktion zu treten. Beispiel: Die Experimentatorin reicht dem Kind das Zielobjekt, aber das Kind zeigt Desinteresse, indem es sich zur Bezugsperson umdreht.

4.2.4.3 Soziale Hinweise

In der Social-Cues-Zeile in ELAN sind die drei Annotationen *Ostension*, *Positives Feedback* und *Aufforderung*²⁵ (in ELAN kodiert als *Ostension*, *Positive Feedback*, *Request*) enthalten, die das soziale Verhalten der Experimentatorin während der Trainingsphase beschreiben (vgl. Tabelle 5, S. 109). Die Erstellung dieser Kategorien war dadurch motiviert, dass neben den festgelegten Äußerungen der Experimentatorin während der Benennungsphase auch eine teilweise freie Interaktion in der Spielphase stattgefunden hat, deren Merkmale ebenfalls berücksichtigt werden sollten.

Als *Ostension* wird eine direkte Adressierung des Kindes definiert, mit der dahinterstehenden Intention ein bestimmtes Wissen an das Kind weiterzugeben. Die Bezeichnung für diese Annotation wurde als sinnvoll empfunden, da es in der Fachliteratur als Voraussetzung für eine Informationsvermittlung genannt wird (vgl. Csibra & Gergely, 2006). In den Transkripten der Videodaten wurden Annotationen als *Ostension* bezeichnet, wenn die Voraussetzung für einen Lernzustand nicht erfüllt war (vgl. Csibra & Gergely, 2006). Damit sind Situationen sowohl während der Benennungs- als auch während der Spielphase (hier

²⁵ Neben den vier aufgezählten Annotationen wurden außerdem *ostensive breathing* (ein lautes Einatmen der Experimentatorin) und *onomatopoeia* (Wörter, die lautmalerisch ein Geräusch imitieren oder sich auch auf eine Handlung beziehen) nachträglich kodiert, um das Verhalten der Experimentatorin in den Transkripten festzuhalten. Da diese Annotationen für die Analysen der vorliegenden Arbeit nicht hinzugezogen wurden, wird auf sie nicht weiter eingegangen. Für künftige Studien ist aber denkbar, diese Kategorien auf weitere Fragestellungen zu untersuchen.

aber nur bei der Joint-Engagement-Gruppe) gemeint, in denen das Kind beispielsweise unruhig wurde und sich umdrehte, während die Experimentatorin versuchte, dessen Aufmerksamkeit wieder zur Interaktion zu lenken. In solchen Fällen adressierte die Experimentatorin das Kind meistens mit dessen Namen und versuchte, einen gegenseitigen Blickkontakt herzustellen.

Im Anschluss an eine erfolgreiche Ausführung einer Zielfunktion folgte die Annotation Positives Feedback in Form von Lob (z. B. „Das hast Du super gemacht!“, „Wow, toll gemacht!“). Abschließend wurde der Versuchsteilnehmer durch Aufforderungen motiviert, die Zielfunktion noch einmal zu wiederholen („Mach das nochmal!“).

Tabelle 5

Die Kategorie Soziale Hinweise mit den jeweiligen Annotationen

Soziale Hinweise	Beschreibung
Ostension	Das Kind wird namentlich adressiert, um seine Aufmerksamkeit zum Ort des Geschehens zu lenken.
Positives Feedback	Die Experimentatorin lobt das Kind für die Durchführung einer bestimmten Zielfunktion, z. B. „Super gemacht!“, „Das machst du toll!“.
Aufforderung	Die Experimentatorin fordert das Kind auf, die Zielfunktion noch einmal auszuführen, z. B. „Mach das nochmal!“; folgt meistens nach dem positiven Feedback.

4.2.4.4 Reliabilität der Kodierungen

Damit gewährleistet werden konnte, dass die Kodierungen der Videos stets nach dem vorgeschriebenen Kodierschema durchgeführt wurden, wurden sie von zwei studentischen Hilfskräften unabhängig voneinander kodiert und die Kodierungen anschließend auf Übereinstimmung überprüft. Bei der Überprüfung von mindestens 20 % des Videomaterials gilt es, das höchste Maß an Reliabilität zu erreichen. Während eine Hilfskraft (A) alle 66 Videos kodiert hat, hat eine zweite studentische Hilfskraft (B) 19 von insgesamt 66 Videos und damit 28.8 % statt der empfohlenen 20 % kodiert, die sie stichprobenartig ausgewählt hat.

Zur Reliabilitätsprüfung der 19 Videos wurde Krippendorffs Alpha (α) als statistisches Maß hinzugezogen (KALPHA). Der Vorteil des KALPHA-Tests besteht darin, dass er für unterschiedliche Skalenniveaus sowie für kleine Stichproben geeignet ist. Außerdem setzt KALPHA keine bestimmte Anzahl an Kodierern voraus und kann in der Berechnung mit fehlenden Werten umgehen (De Swert, 2012; Hayes & Krippendorff, 2007; Krippendorff, 2004, 2007). In der Berechnung der Reliabilität werden mögliche, zufällig auftretende Über-

einstimmungen herausgerechnet, die die tatsächliche Übereinstimmung überschätzen können. Laut Krippendorff (2004, S. 241) liegt der α -Richtwert für eine zuverlässige Reliabilität der Kodierungen bei $\alpha > .800$. Werte zwischen $\alpha = .667$ und $\alpha = .800$ sollten allerdings mit Vorsicht betrachtet werden (De Swert, 2012, S. 5). In solchen Fällen wird empfohlen, sich die Kodierungen genauer anzuschauen, um den Grund für die geringe Reliabilität zu identifizieren.

Die Prüfung auf Reliabilität der kodierten Kategorie Gaze sowie der Anzahl der ausgeführten Strategien wurde mithilfe des KALPHA-Befehls in der Statistik- und Analysesoftware SPSS nach De Swert (2012) durchgeführt. Dabei wurde ein vorgeschriebener Basis-Befehl (De Swert, 2012, S. 4) an das Skalenniveau sowie an die Variablennamen der eigenen Daten angepasst und anschließend in der SPSS Syntax ausgeführt:

Basisbefehl:

```
KALPHA judges = judgelist/level = lev/detail = det/boot = z.
```

Beispiel für die Variable Öffnen, ausgeführt von der Self-Exploring-Gruppe:

```
KALPHA judges = open_self_P2_count_A open_self_P2_count_B/  
level = 4/detail = 0/boot = 10000.
```

Für die in Kursiv gesetzten Abschnitte wurde eine Anpassung an die eigenen Daten durchgeführt. *Judgelist* wurde durch die zu untersuchenden Variablen, beispielsweise *open_self_P2_count_A* und *open_self_P2_count_B* ersetzt, wobei das Suffix *_A* und *_B* die beiden Kodiererinnen repräsentiert, die in diesem Beispiel die Handlungsstrategie Öffnen kodiert haben. Weiterhin ist statt *lev* eine 4 eingesetzt worden, die für die rationalskalierte Merkmalsausprägung der Werte steht. Die Wert 0 anstelle von *det* bestimmt, dass nur der KALPHA Wert angezeigt werden soll. Ersetzt man ihn durch die Zahl 1, so erhält man detaillierte Informationen der Auswertung. Der Befehl *Boot* ist die Abkürzung für Bootstrapping und somit auch verantwortlich dafür, wie viel Bootstrapping durchgeführt werden soll. Als Richtwert wird 10 000 vorgegeben, der als solcher übernommen werden konnte. Die Berechnung des α -Werts ergab, dass die Kodierungen aller Variablen übereinstimmen (alle α -Werte > 0.84) und somit weitere statistische Auswertungen vorgenommen werden können. Bei einigen Variablen ergab sich sogar ein $\alpha = 1.00$, da die Kodierungen beider Kodiererinnen vollkommen übereinstimmten.

4.3 Ergebnisse

Nachdem die Reliabilitätsprüfung eine einheitliche Kodierung bestätigte, wurden die erhobenen Verhaltens- sowie Eye-Tracking-Daten in der Statistik- und Analyse-Software SPSS ausgewertet, um die aufgestellten Hypothesen (vgl. Kapitel 3.7) zu untersuchen.

4.3.1 Aufbereitung der Verhaltens- und Eye-Tracking-Daten

Bevor die erhobenen Verhaltensdaten der Trainingsphase in SPSS ausgewertet wurden, wurden die kodierten Blickdaten (in Sekunden) zu Durchschnittswerte sowie in anteilige Blickzeiten umgerechnet. Dies hatte zum einen den Grund, dass die Objekt-Sets in den beiden aufeinanderfolgenden Versuchen randomisiert wurden. Zum anderen war die Dauer der Benennungs- und Spielphasen nicht exakt dieselbe, wodurch sich auch die Blickdauer der Probanden um wenige Sekunden unterscheiden konnte. So wurde beispielsweise nicht die Blickdauer zum Maiskolben und zum Schiff in Sekunden für die weiteren Auswertungen verwendet, sondern die durchschnittliche Blickdauer, die sich aus der Fixationsdauer zu den beiden salienten Ablenkern ergibt, als eine Variable (z. B. *gaze_distractor*) angelegt. Dieser Durchschnittswert wurde in Verhältnis zu den anderen Blickdaten gesetzt und dadurch der anteilige Prozentwert ermittelt. Die anteilige Blickdauer der Probanden zur Experimentatorin berechnet sich anhand der folgenden Formel:

$$\frac{\text{Ablenker}}{\text{Zielobjekt} + \text{Experimentatorin} + \text{Bezugsperson} + \text{Sonstiges}}$$

Ebenso wie für die Verhaltensdaten wurde auch für die Eye-Tracking-Daten eine vorbereitende Datenanalyse durchgeführt. Es sollten drei unterschiedliche Messgrößen – *Total Visit Duration*, *Time to First Fixation* und *Weighted Gaze Samples* – für die statistischen Auswertungen hinzugezogen werden. Die *Total Visit Duration*, mit der die Gesamtblickdauer zu den AOIs im Eye-Tracking-Verfahren gemeint ist, gilt als wesentliche Messgröße dieser Arbeit. Mithilfe der Gesamtblickdauer zu den Objekt-Sets sollen Erkenntnisse über die visuelle Präferenz in der Salienzphase und über die Referentenauswahl in den Testphasen gewonnen werden. Die Gesamtblickdauer ist die Zeit jeder einzelnen Fixation innerhalb einer AOI, die von dem Tobii Pro Studio System automatisch als summierte Gesamtblickdauer für jede Versuchsperson in Sekunden herausgegeben wird (Tobii AB, 2016, S. 110). Für die folgenden statistischen Auswertungen wurde aus den gegebenen Blickwerten die anteilige

Blickzeit zum Ablenker bzw. zum Zielobjekt aus der Gesamtblickzeit (Ablenker + Zielobjekt) berechnet.

Neben der Auswertung der durchschnittlichen prozentualen Blickdauer wurde außerdem die Messgröße Time to First Fixation zur Auswertung des Blickverhaltens hinzugezogen. Dabei handelt sich um die Zeit, bis ein Versuchsteilnehmer zum ersten Mal eine AOI fixiert. Indem lediglich die erste Fixation nach der Wortäußerung näher betrachtet wird, sollen genauere Aussagen über das Wortverständnis während der Testphasen getroffen werden. Auch für diesen Messwert wurden aus zusammengehörenden Time-to-First-Fixations-Variablen die Durchschnittswerte berechnet. Allerdings wurden diese Daten nicht als Prozentwerte, sondern als Sekundenwerte belassen.

Schließlich ist die dritte Messgröße der Weighted Gaze Samples aufzuführen (vgl. Kapitel 4.1.2), die als Prozentwert im Tobii System angezeigt wird. Anhand dieser Messgröße sollen Aussagen über die Qualität der erfassten Blickdaten sowie über den Aufmerksamkeitsverlauf getroffen werden. Diese Messgröße wurde einbezogen, um die Aufmerksamkeit der Probanden während der vier Eye-Tracking-Phasen (Salienzphase 1, Testphase 1, Salienzphase 2, Testphase 2) zu kontrollieren. Es ist anzumerken, dass diese Messgröße in der bisherigen Forschungsliteratur, soweit bekannt, nicht in Eye-Tracking-Studien mit Kindern berichtet worden ist.

4.3.2 Auswahl der statistischen Testverfahren

Neben der Aufbereitung der Verhaltens- und Eye-Tracking-Daten ist außerdem noch die Entscheidung bezüglich der Überprüfung der Normalverteilung und zu dem Umgang mit fehlenden Daten zu erwähnen, die vor den statistischen Analysen getroffen wurde.

Die Stichprobengröße für die Studie der vorliegenden Arbeit wurde unter Berücksichtigung des zentralen Grundwertsatzes (engl.: central limit theorem) festgelegt. Demnach ist es erlaubt, auch dann parametrische Verfahren zu nutzen, wenn keine Normalverteilung vorliegt, sofern die Stichprobengröße über 30 liegt, da sich in solch einem Fall die Stichprobenverteilung der Normalverteilung annähert (Field, 2009, S. 42; Luhmann, 2015, S. 176). Da der Stichprobenumfang der ersten Experiments 33 Versuchsteilnehmer in jeder der beiden Gruppen umfasst, ist damit die Voraussetzung > 30 erfüllt und es kann von einem normalverteilten Datensatz ausgegangen werden. Zudem ist darauf hinzuweisen, dass individuelle Unterschiede entwicklungsbedingt hoch sind (vgl. Rohlfing, 2019, S. 260) und eine Ausreißeranalyse mit Säuglingsdaten in der vorliegenden Arbeit daher nicht als sinnvoll er-

achtet wurde. Hinzu kommt, dass mögliche, abweichende bzw. extreme Werte nicht das Resultat eines Fehlers (der Technik oder der Kodierer) waren, sodass diese nicht ausgeschlossen wurden. Daher wurden die Eye-Tracking- sowie die Verhaltensdaten mit parametrischen Verfahren analysiert. Dennoch wurde der Shapiro-Wilk-Test durchgeführt, um einen Überblick über die Datenverteilung zu erhalten. In Bezug auf die Blickdauer der Probanden während der Trainingsphase stellte sich heraus, dass die Daten einiger Variablen in beiden Gruppen nicht normalverteilt waren. Auch einige der Eye-Tracking-Daten verletzen die Voraussetzung auf Normalverteilung.

Zur Bestimmung der Effektstärke wurde der Pearson-Korrelationskoeffizient r aus der Formel, $r = \sqrt{\frac{t^2}{t^2 + df}}$ berechnet (Field, 2009, S. 332). Neben dem Cohen's d (Cohen, 1969) ist dies ein gängiges Maß, welches ebenfalls häufig verwendet wird. Liegt der Pearson-Korrelationskoeffizient r zwischen 0.1 und 0.3, so handelt es sich um einen kleinen Effekt. Ein Wert zwischen 0.3 und 0.5 zählt als mittlerer Effekt, wohingegen ein Wert, der größer ist als 0.5 für eine große Effektstärke steht (Field, 2009, S. 57). Somit liegt eine maximale Effektstärke bei dem Wert 1 vor und die niedrigste bei 0.

Abschließend soll auf den Umgang mit fehlenden Werten in den Eye-Tracking-Daten eingegangen werden. Fehlende Werte ergaben sich beispielsweise dadurch, dass Probanden vom Bildschirm weggeschaut haben und deren Augen dadurch nicht erfasst werden konnten. Da in den meisten Fällen lediglich einer Eye-Tracking-Tests von fehlenden Werten betroffen war, wäre es nicht folgerichtig gewesen, diesen Versuchsteilnehmer komplett aus den Auswertungen auszuschließen (listenweiser Ausschluss). Ein paarweiser Ausschluss, bei dem die einzelnen fehlenden Werte ausgeschlossen werden, hätte allerdings dazu geführt, dass unterschiedliche Stichprobengrößen in dem Original-, New-Label- und Recovery-Test zustande gekommen wären. Aus den erwähnten Gründen und mit dem Ziel, den kompletten Datensatz zu erhalten, wurden fehlende Werte mithilfe einer einfachen Imputation ersetzt.

Für die vorliegende Arbeit ist eine einfache Mittelwertimputation ausgewählt worden (Little & Rubin, 1987, 2014; Pérez, Dennis, Gil, Rondón & López, 2002; Saunders et al., 2006). Dabei werden die fehlenden Werte jeder Variable mit einem plausiblen Wert ersetzt, der sich aus dem Stichprobenmittelwert der jeweiligen Gruppe ergibt. Zum Beispiel schaut die Joint-Engagement-Gruppe während des ersten New-Label-Tests im Durchschnitt 0.69 Sekunden auf das saliente Objekt. Dieser Wert ersetzte fortan zwei fehlende Werte in der Variable der Joint-Engagement-Gruppe. Allerdings geht mit diesem Verfahren auch einher, dass dabei die Varianz der Datenverteilung unterschätzt wird (Little & Rubin, 1987, S. 62).

4.3.3 Blickdauer während der Eye-Tracking-Salienzphase (Hypothese 1)

Als erste Hypothese wurde formuliert, dass Kinder im Alter von 10 Monaten eine visuelle Präferenz für saliente Objekte haben, weil sie die kindliche Aufmerksamkeit stärker auf sich lenken als weniger saliente Objekte. Somit besteht die Annahme, dass die Probanden das saliente Objekt während der Salienzphase länger anschauen werden als das langweilige Objekt. Für die Überprüfung dieser Hypothese werden die Daten aller Probanden gruppenunabhängig ausgewertet, da die Probanden zu diesem Zeitpunkt noch keine unterschiedlichen Lernformate erfahren haben.

Um die erste Hypothese zu testen, wurde ein Einstichproben *t*-Test (einfacher *t*-Test; engl.: one-sample *t*-test) durchgeführt, bei dem die anteilige Blickdauer (für Erläuterungen vgl. Kapitel 4.3.1) zum Distraktor sowie zum langweiligen Objekt gegen den Wahrscheinlichkeitswert (engl.: chance level) von 50 getestet wird, da die Wahrscheinlichkeit auf eines der beiden Objekte auf dem Bildschirm zu schauen bei jeweils 50 % liegt: $\frac{100}{2} = 50$. Als Signifikanzniveau für den *t*-Test wurde $\alpha = 0.05$ festgelegt. Ist der *p*-Wert kleiner als der angegebene α -Wert, dann bedeutet dies, dass die Irrtumswahrscheinlichkeit unter fünf Prozent liegt und das Ergebnis deshalb signifikant ist. Der Test ergab, dass die prozentuale Blickzeit aller Probanden ($N = 66$) zum salienten Objekt (Schiff oder Maiskolben) ($M = 55.42$, $SD = 17.66$) signifikant über dem Wahrscheinlichkeitswert von 50 % liegt, $t(65) = 2.5$, $p = .02$, $r = .3$. Dementsprechend liegt die Blickzeit zum weniger salienten Objekt (Holzpuppe oder Klotz) ($M = 44.58$, $SD = 17.66$) signifikant unter dem Wahrscheinlichkeitswert, $t(65) = -2.5$, $p = .02$, $r = .3$. Die Bedeutsamkeit des vorliegenden Ergebnisses unterstreicht die mittlere Effektstärke.

Abschließend wurde eine Regressionsanalyse durchgeführt, die eingesetzt wird, um einen Zusammenhang zwischen zwei Variablen zu erklären. Darüber hinaus, kann anhand einer linearen Regressionsanalyse die abhängige Variable aus der unabhängigen vorhergesagt werden (Field, 2009, S. 198). Damit sollte getestet werden, ob der Bildungsabschluss der Mütter (vgl. Kapitel 4.2.1.2) in Zusammenhang steht mit den Blickdaten in der Salienzphase. Es stellte sich heraus, dass der Bildungsabschluss der Mütter kein signifikanter Prädiktor für die Blickdauer der Kinder zum salienten Objekt war, $R^2 = .02$, $F(1, 64) = 1.33$, $p = .253$.

Diese Auswertung untermauert nicht nur die Ergebnisse der Salienzphase aus der Studie von Pruden und Kollegen (2006), sondern bestätigt auch den assoziativen Ansatz und somit auch die Hypothese, dass perzeptuell saliente Objekte aufmerksamkeitslenkend für

Kinder im Alter von 10 Monaten sind. Dieses Phänomen, dass die Fixationszeit zu salienten Stimuli länger andauert im Vergleich zu weniger salienten Stimuli, wird im folgenden Verlauf als *Salienz-Effekt* bezeichnet. An dieser Stelle ist außerdem hinzuzufügen, dass in der Salienzphase noch keine sozialen Hinweise präsentiert wurden. Es wurde allein die visuelle Präferenz für eines der beiden Objekte, die sich in ihrer perzeptuellen Salienz unterschieden, getestet.

4.3.4 Blickdauer während der Benennungsphase (Hypothese 2)

Die zweite Hypothese besagt, dass Kinder im Alter von 10 Monaten soziale Hinweise als Informationsquelle zur Referentenauswahl nutzen. Wenn die Versuchsteilnehmer die meiste Zeit in der Benennungsphase das Zielobjekt anschauen werden, welches von der Experimentatorin ostensiv benannt wird, dann kann die zweite Hypothese bestätigt werden. Damit ginge einher, dass die Probanden den sozialen Hinweisen selbst dann folgen würden, wenn zeitgleich ein salientes Objekt im visuellen Blickfeld der Kinder läge. An dieser Stelle ist zu betonen, dass sich die Probanden in der Benennungsphase noch nicht in ihrem Lernformat unterscheiden.

Zur Überprüfung der zweiten Hypothese wurde zunächst die Blickdauer, abhängig von der Gruppenzuordnung, während der Benennungsphasen zum Zielobjekt ausgewertet. Dafür wurde als erstes mithilfe des Einstichproben-*t*-Tests analysiert, ob sich die Blickdauer zum Zielobjekt, Ablenker, zur Experimentatorin, Bezugsperson und Sonstigem von dem Wahrscheinlichkeitswert von 20 % unterscheidet. Der Wahrscheinlichkeitswert ergab sich aus dem Prozentwert 100 (Gesamtblickzeit), dividiert durch die Anzahl der möglichen Blickrichtungen: $\frac{100}{5} = 20$.

Die Ergebnisse zeigen (vgl. Tabelle 6, S. 116), dass bei der Joint-Engagement-Gruppe ($N = 33$) die anteilige Blickzeit zum Zielobjekt ($M = 64.82$, $SD = 10.53$) während der Benennungsphase signifikant über dem Wahrscheinlichkeitswert liegt, $t(32) = 24.44$, $p < .001$, $r = .97$. Auch bei der Self-Exploring-Gruppe ($N = 33$) liegt die prozentuale anteilige Blickzeit zum Zielobjekt ($M = 60.69$, $SD = 11.83$) signifikant über dem Wahrscheinlichkeitswert von 20 %, $t(32) = 19.76$, $p < .001$, $r = .96$. Dabei ist auch auf die hohen Effektstärken hinzuweisen.

Somit zeigen diese Ergebnisse, dass beide Gruppen das weniger saliente Objekt während der Benennungsphase als Referenten identifiziert haben. Da die Blickdauer zum Ablenker, zur Experimentatorin, Bezugsperson und zu Sonstigem bei beiden Gruppen ebenfalls

entweder signifikant über oder unter dem Wahrscheinlichkeitswert von 20 % liegt, kann daraus geschlossen werden, dass keine der Fixationszeiten auf eine Zufallswahrscheinlichkeit zurückzuführen ist.

Tabelle 6

Prozentuale Blickdauer der beiden Gruppen während der Benennungsphase im Vergleich zum Wahrscheinlichkeitswert von 20 %

Blickrichtung	Blickdauer		<i>t</i> (32)	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>			
Joint Engagement					
Zielobjekt	64.82	10.53	24.44	<.001***	.97
Ablenker	5.78	3.67	−22.24	<.001***	.97
Experimentatorin	24.03	10.06	2.3	.028*	.38
Bezugsperson	0.39	1.29	−87.42	<.001***	.99
Sonstiges	4.98	5.42	−15.92	<.001***	.94
Self Exploring					
Zielobjekt	60.69	11.83	19.76	<.001***	.96
Ablenker	8.36	6.06	−11.04	<.001***	.89
Experimentatorin	24.50	10.58	2.44	.02*	.4
Bezugsperson	0.53	1.77	−63.22	<.001***	.99
Sonstiges	5.93	6.26	−12.92	<.001***	.92

Da die beiden Gruppen zu dem Zeitpunkt der Benennungsphase denselben Input erhalten haben und die Versuchsdurchführung dieselbe war, sollten im nächsten Schritt die Blickdaten der Joint-Engagement- und der Self-Exploring-Gruppe zusammengelegt werden, sofern sie sich nicht signifikant voneinander unterscheiden. Aus diesem Grund wurden die Blickzeiten der beiden Gruppen mithilfe des unabhängigen *t*-Test miteinander verglichen. Ein unabhängiger *t*-Test (engl.: independent *t*-test) wird dafür verwendet, um beispielsweise zwei verschiedene Gruppen hinsichtlich einer Variable zu vergleichen.

Die Auswertung ergab, dass die Self-Exploring-Gruppe lediglich den Ablenker ($M = 8.36$, $SD = 6.06$) signifikant länger anschaut im Vergleich zur Joint-Engagement-Gruppe ($M = 5.78$, $SD = 3.67$), $t(52,71) = -2.09$, $p = .042$, $r = .29$ (vgl. Tabelle 7, S. 117). Da es die einzige Variable ist, in der sich die Gruppen in ihrer Fixationsdauer unterscheiden und diese nur einen schwachen Effekt ($r = .29$) aufzeigte ($0.1-0.3$ = niedriger Korrelationskoeffizient), wurden die Blickdaten beider Gruppen in der Benennungsphase für die weiteren Auswertungen zusammengefügt.

Tabelle 7

Prozentuale Blickdauer während der Benennungsphase in Abhängigkeit von der Gruppenzuordnung

Blickrichtung	Self Exploring		Joint Engagement		t(64)	p	r
	M	SD	M	SD			
Zielobjekt	60.69	11.83	64.82	10.53	1.50	.139	.18
Ablenker	8.36	6.06	5.78	3.67	– 2.09	.042*	.29
Experimentatorin	24.50	10.58	24.03	10.06	– 0.19	.854	.02
Bezugsperson	0.53	17.70	0.39	1.29	0.35	.725	.04
Sonstiges	5.93	6.26	4.98	5.42	– 0.66	.511	.08

Der anschließende Einstichproben-*t*-Test zeigte in der Auswertung aller Versuchsteilnehmer ($N = 66$), dass deren anteilige Blickdauer zur Experimentatorin ($M = 24.26$, $SD = 10.25$), $t(65) = 3.38$, $p < .001$, $r = .39$, und zum Zielobjekt ($M = 62.75$, $SD = 11.31$), $t(65) = 30.72$, $p < .001$, $r = .97$, signifikant über dem Wahrscheinlichkeitslevel von 20 Prozent liegt (vgl. Tabelle 8). Die Fixationsdauer zur Experimentatorin, zum Ablenker, zur Bezugsperson und zu Sonstigem liegt signifikant unter diesem Level. Dabei ist insbesondere auffällig, dass die Probanden dem salienten Objekt (Maiskolben oder Schiff), welches stets im Sichtfeld des Kindes auf dem Tisch lag, kaum Beachtung schenkten ($M = 0.46$, $SD = 1.53$). Eine daran anknüpfende Regressionsanalyse zeigte, dass der Bildungsabschluss der Mütter nicht in Zusammenhang stand mit der Blickdauer aller Probanden zum Zielobjekt während der Benennungsphase, $R^2 = .025$, $F(1,62) = 1.624$, $p = .207$.

Tabelle 8

Prozentuale Blickdauer aller Probanden während der Benennungsphase im Vergleich zum Wahrscheinlichkeitswert von 20 %

Blickrichtung	Blickdauer		t(65)	p	r
	M	SD			
Zielobjekt	62.75	11.31	30.72	<.001***	.97
Ablenker	7.07	5.14	– 20.44	<.001***	.93
Experimentatorin	24.26	10.25	3.38	<.001***	.39
Bezugsperson	0.46	1.53	– 103.26	<.001***	.99
Sonstiges	5.46	5.83	– 20.28	<.001***	.93

Da sich in den bisherigen Auswertungen herausstellte, dass die Blickzeiten zum Zielobjekt ($M = 62.75$, $SD = 11.31$) und zur Experimentatorin ($M = 24.26$, $SD = 10.25$) am längsten waren, wurde zusätzlich ein *t*-Tests für verbundenen Stichproben (abhängiger *t*-Test,

engl.: dependent *t*-test) durchgeführt, indem die beiden Werte verglichen wurde. Ein solcher *t*-Test wird verwendet, wenn dieselben Versuchsteilnehmer zweimal im Datensatz auftauchen, beispielsweise bei zwei Messzeitpunkten oder wenn zwei Variablen miteinander zusammenhängen, z. B. zwei Stimuli nebeneinander präsentiert werden. Die Auswertung ergab, dass die Probanden ($N = 66$) während der Benennungsphase eine signifikant längere Blickzeit zum Zielobjekt ($M = 62.75$, $SD = 11.31$) im Vergleich zur Experimentatorin ($M = 24.26$, $SD = 10.25$) haben, $t(65) = -15.79$, $p < .001$, $r = .89$.

Aus den Ergebnissen wird ersichtlich, dass die Blickzeiten zum Zielobjekt und zur Experimentatorin im Vergleich zu den anderen Blickrichtungen dominieren. Daraus kann die zweite Hypothese, dass 10 Monate alte Kinder auf den Kommunikationspartner achten und den sozialen Hinweisen zum Referenten folgen, bestätigt werden. Dieser Befund spricht gegen den assoziativen Ansatz von Pruden et al. (2006), dass junge Kinder soziale Hinweise ignorieren. Der sozio-pragmatische Ansatz, dass sie solche Signale zur Referentenauswahl nutzen, wird hingegen unterstützt.

Außerdem stehen diese Ergebnisse in Kontrast zu den Ergebnissen der Studie von Yurovsky und Frank (2017), laut der es jungen Kindern schwerfällt, sich von Gesichtern zu lösen, um einem referentiellen Hinweis zu folgen. Im Gegensatz dazu zeigen die vorliegenden Ergebnissen dieses Experiments, dass die längste Blickdauer nicht der Experimentatorin, sondern dem Zielobjekt gilt. Vermutlich ist die längere Blickzeit zum Gesicht der Experimentatorin anstatt zum Zielobjekt auf die methodische Durchführung der Studie von Frank und Yurovsky (2017) zurückzuführen, die im Zwischenfazit der vorliegenden Arbeit ausführlich diskutiert wird (vgl. Kapitel 4.4).

4.3.5 Zusammenhang der Enkodierungsstrategien und der Retention von Wort-Objekt-Verknüpfungen (Hypothese 3)

Die dritte Hypothese besagt, dass ein Zusammenhang zwischen den Enkodierungsstrategien, ergebend aus den beiden sozialen Lernformaten, und der Retention der Wort-Objekt-Verknüpfung in der Testphase vorliegt. Hinsichtlich der Retention wird angenommen, dass Unterschiede im Blickverhalten der Self-Exploring- und der Joint-Engagement-Gruppe vorzufinden sein werden. Dabei stellt sich die Frage, ob bzw. welche der beiden Gruppen die Wort-Objektverknüpfung beibehalten kann. Zur Untersuchung der dritten Hypothese sind drei Schritte in der Auswertung notwendig: Als erstes gilt zu überprüfen, ob sich die beiden Gruppen während der Spielphase in ihren Blickzeiten voneinander unterscheiden. Im zweiten Schritt soll ausgewertet werden, ob Unterschiede in den Eye-Tracking-Daten zu finden

sind. Wenn sowohl in der Trainings- als auch in der Testphase gruppenabhängige Unterschiede vorzufinden sind, kann der Zusammenhang überprüft werden.

4.3.5.1 Gruppenabhängiges Blickverhalten während der Spielphase

Als erstes wurde das Blickverhalten beider Gruppen während der Spielphase untersucht. Es liegt die Vermutung nahe, dass sich das Blickverhalten der beiden Gruppen von der Benennungsphase bis hin zur Spielphase verändert hat. Hierfür wurde ein abhängiger *t*-Test durchgeführt (vgl. Tabelle 9, S. 120).

Die Auswertung ergab, dass die Blickdauer der Joint-Engagement-Gruppe zur Experimentatorin von der Benennungsphase ($M = 24.03$, $SD = 10.06$) bis hin zur Spielphase ($M = 12.69$, $SD = 12.07$) signifikant abnimmt, $t(32) = 6.11$, $p < .001$, $r = .73$. Im Gegenzug nimmt die Blickzeit zum Zielobjekt von der Benennungsphase ($M = 64.82$, $SD = 10.53$) zur Spielphase ($M = 80.23$, $SD = 14.41$) signifikant zu, $t(32) = -6.07$, $p < .001$, $r = .73$. Vergleicht man die Blickzeit zum salienten Ablenker während der Benennungsphase ($M = 5.78$, $SD = 3.67$) und der Spielphase ($M = 2.04$, $SD = 2.55$), so ist diese in der letzteren Phase signifikant niedriger, $t(32) = 5.15$, $p < .001$, $r = .67$. Die Blickdauer zur Bezugsperson und in sonstige Richtungen hat sich hingegen nicht signifikant verändert.

Die Self-Exploring-Gruppe zeigt signifikante Veränderungen in ihrer Blickzeit zur Experimentatorin, zum Zielobjekt, zum salienten Objekt sowie in sonstige Blickrichtungen. Im Gegensatz zur Joint-Engagement-Gruppe nimmt die Blickdauer der Self-Exploring-Gruppe zur Experimentatorin von der Benennungsphase ($M = 24.5$, $SD = 10.58$) bis hin zur Spielphase ($M = 30.35$, $SD = 16.33$) signifikant zu, $t(32) = -2.11$, $p = .043$, $r = .35$. Die Blickzeit zum Zielobjekt nimmt allerdings von der Benennungsphase ($M = 60.69$, $SD = 11.83$) zur Spielphase ($M = 50.81$, $SD = 18.25$) signifikant ab, $t(32) = 3.11$, $p < .001$, $r = .48$. Die Blickdauer zum Ablenker während der Benennungsphase ($M = 8.36$, $SD = 6.06$) ist signifikant länger als in der darauffolgenden Spielphase ($M = 5.6$, $SD = 6.55$), $t(32) = 2.53$, $p = .017$, $r = .41$. Schließlich ist noch aufzuführen, dass die Self-Exploring-Gruppe in der Spielphase signifikant länger in sonstige Richtungen ($M = 12.29$, $SD = 10.37$) schaut als in der Benennungsphase ($M = 5.93$, $SD = 6.26$), $t(32) = -3.56$, $p < .001$, $r = .53$. Hinsichtlich der Blickdauer zur Bezugsperson wurden keine signifikanten Unterschiede zwischen den beiden Phasen gefunden.

Tabelle 9

Vergleich der prozentualen Blickdauer beider Gruppen zwischen der Benennungs- und Spielphase

Blickrichtung	Blickdauer				t(32)	p	r
	Benennungsphase		Spielphase				
	M	SD	M	SD			
Joint Engagement							
Zielobjekt	64.82	10.53	80.23	14.41	− 6.07	<.001***	.73
Ablenker	5.78	3.67	2.04	2.55	5.15	<.001***	.67
Experimentatorin	24.03	10.06	12.69	12.07	6.11	<.001***	.73
Bezugsperson	0.39	1.29	0.51	2.25	− 0.25	.804	.04
Sonstiges	4.98	5.42	4.54	6.24	0.29	.777	.05
Self Exploring							
Zielobjekt	60.69	11.83	50.81	18.25	3.11	.004**	.48
Ablenker	8.36	6.06	5.6	6.55	2.53	.017*	.41
Experimentatorin	24.50	10.58	30.35	16.33	− 2.11	.043*	.35
Bezugsperson	0.53	1.77	0.95	2.77	− 0.71	.485	.12
Sonstiges	5.93	6.26	12.29	10.37	− 3.56	<.001***	.53

Als nächstes wird die Frage, ob sich beiden Gruppen hinsichtlich ihrer Blickdauer zum Zielobjekt, salienten Objekt, zur Experimentatorin, Bezugsperson und zu Sonstigem in der Spielphase unterscheiden, mithilfe eines unabhängigen *t*-Tests untersucht (vgl. Tabelle 10, S. 121). Die Auswertungen ergaben, dass die Blickdauer der Self-Exploring-Gruppe zur Experimentatorin ($M = 30.35$, $SD = 16.33$) signifikant länger ist als die der Joint-Engagement-Gruppe ($M = 12.69$, $SD = 12.07$), $t(58.93) = -4.99$, $p < .001$, $r = .55$. Die Joint-Engagement-Gruppe schaut hingegen länger zum Zielobjekt ($M = 80.23$, $SD = 14.41$) als die Self-Exploring-Gruppe ($M = 50.81$, $SD = 18.25$), $t(64) = 7.27$, $p < .001$, $r = .67$. Das saliente Objekt fixiert die Self-Exploring-Gruppe signifikant länger ($M = 5.6$, $SD = 6.55$) im Vergleich zur Joint-Engagement-Gruppe ($M = 2.04$, $SD = 2.55$), $t(41.49) = -2.91$, $p = .006$, $r = .41$. Ebenso gibt es einen signifikanten Unterschied in der Blickdauer hinsichtlich sonstiger Blickrichtungen zwischen der Self-Exploring- ($M = 12.29$, $SD = 10.37$) und der Joint-Engagement-Gruppe ($M = 4.54$, $SD = 6.24$), $t(52.48) = -3.681$, $p < .001$, $r = .45$.

Eine anschließende Regressionsanalyse zeigte, dass der Bildungsabschluss der Mütter weder in Zusammenhang stand mit der Blickdauer der Joint-Engagement-Gruppe, $R^2 = .001$, $F(1, 31) = .04$, $p = .85$, noch mit der Blickdauer der Self-Exploring-Gruppe zum Zielobjekt während der Spielphase, $R^2 = .02$, $F(1, 31) = .6$, $p = .45$.

Tabelle 10

Prozentuale Blickdauer während der Spielphasen in Abhängigkeit von der Gruppenzuordnung

Blickrichtung	Joint Engagement		Self Exploring		t(64)	p	r
	M	SD	M	SD			
Zielobjekt	80.23	14.41	50.81	18.25	7.27	<.001***	.67
Ablenker	2.04	2.55	5.6	6.55	– 2.91	.006**	.41
Experimentatorin	12.69	12.07	30.35	16.33	– 4.99	<.001***	.53
Bezugsperson	0.51	2.25	0.95	2.77	– 0.71	.48	.09
Sonstiges	4.54	6.24	12.29	10.37	– 3.68	<.001***	.45

Zusammenfassend kann festgehalten werden, dass sich die Joint-Engagement- und die Self-Exploring-Gruppe in ihrer Blickdauer während der Benennungsphase kaum unterscheiden (vgl. Kapitel 4.3.4), wohingegen während der Spielphase Unterschiede in der Blickdauer, insbesondere zum Zielobjekt und zur Experimentatorin zu finden sind. Des Weiteren wurde festgestellt, dass sich die Blickzeiten beider Gruppen in der Benennungs- und Spielphase voneinander unterscheiden.

4.3.5.2 Gruppenabhängiges Blickverhalten während der Testphasen

Im nächsten Schritt wird auf das Blickverhalten während der Eye-Tracking-Testphasen eingegangen. Die Original-Label, New-Label- und Recovery-Tests wurden in jeweils zwei Abschnitte geteilt, um das Blickverhalten vor (Onset) und nach dem Wortende (Offset) auszuwerten (vgl. Kapitel 4.2.2.5). Hinsichtlich des Wort-Onset-Abschnitts besteht die Annahme, dass eine visuelle Präferenz für das saliente Objekt (Ablenker) im Vergleich zum langweiligen Objekt (Zielobjekt) in beiden Gruppen zu beobachten ist. Da das Zielwort erst zum Ende der Onset-Phase fällt und ab diesem Zeitpunkt die Referentenauswahl erfolgen soll, liegt die Vermutung nahe, dass sich die Gruppen erst in den Offset-Phasen, aber nicht in den Onset-Phasen, in ihren Fixationszeiten voneinander unterscheiden werden.

Neben der Retention der während der Trainingsphase enkodierten Wort-Objekt-Verbindung wird außerdem eine weitere Enkodierungsstrategie überprüft. Es handelt sich um das Ausschließbarkeitsprinzip, welches auf einem vorhandenen Wortwissen basiert. Es stellt sich die Frage, ob die jungen Versuchsteilnehmer das saliente Objekt mit einem neuen Wort assoziieren können, indem sie auf das zuvor trainierte Wortwissen zurückgreifen. Damit geht einher, dass die zuvor enkodierte Wort-Objekt-Verbindung bis zum Test beibehalten werden muss, um das Wortwissen im New-Label-Test als Ausschlusskriterium nutzen zu können.

4.3.5.2.1 Blickdauer zu den AOIs in den Onset-Phasen

Die ersten Auswertungen der Testphase betreffen die Onset-Phasen der Original-Label-, New-Label- und Recovery-Tests. Anhand eines Einstichproben- t -Tests soll zunächst ausgewertet werden, ob die Probanden während der Onset-Phasen eines der beiden Objekte länger anschauen im Vergleich zum Wahrscheinlichkeitslevel von 50 %. Da erst zum Ende der Onset-Phasen das trainierte Zielwort fällt, wird vermutet, dass eine längere Blickdauer zum salienten Objekt zu beobachten sein wird.

Zunächst wurden die Blickdaten beider Gruppen gesondert voneinander ausgewertet, indem ein Einstichproben t -Test durchgeführt wurde (vgl. Tabelle 12, S. 124). Die Auswertung ergab, dass in den drei Sekunden des Original-Label-Onset-Tests die Blickdauer der Joint-Engagement-Gruppe zum Ablenker ($M = 60.03$, $SD = 14.38$) signifikant über dem Wahrscheinlichkeitswert von 50 % liegt, $t(32) = 4.01$, $p < .001$, $r = .58$. Die Blickdauer zum Zielobjekt ($M = 39.97$, $SD = 14.38$) liegt dementsprechend signifikant unter dem 50-Prozent-Level, $t(32) = -4.01$, $p < .001$, $r = .58$. Im New-Label und Recovery-Test ist die Blickdauer zum Ablenker ebenfalls höher als der Wahrscheinlichkeitswert, aber nicht signifikant. Von der Self-Exploring-Gruppe liegt die Fixationsdauer zum Ablenker ($M = 59.89$, $SD = 12.67$) während des Original-Label-Tests ebenfalls signifikant über dem Wahrscheinlichkeitswert, $t(32) = 4.48$, $p < .001$, $r = .62$, wohingegen die Fixationsdauer des Zielobjektes ($M = 40.11$, $SD = 12.67$) signifikant unter dem Wahrscheinlichkeitswert liegt, $t(32) = -4.48$, $p < .001$, $r = .62$. Auch in dieser Gruppe ist die Blickdauer zum Ablenker in den ersten drei Sekunden des New-Label- und Recovery-Tests über dem Wahrscheinlichkeitswert, unterscheidet sich aber nicht signifikant von diesem.

Ein anschließender unabhängiger t -Test ergab, dass sich die beiden Gruppen in ihren Blickzeiten zu den AOIs während der Onset-Phasen nicht signifikant unterscheiden. Daher wurden die Daten beider Gruppen zusammengelegt und in der Tabelle 11 (vgl. S. 123) veranschaulicht. In dieser Tabelle wird deutlich, dass alle Probanden während der Original-Label-Onset-Phase eine visuelle Präferenz für den Ablenker haben ($M = 59.96$, $SD = 13.45$), $t(65) = 6.02$, $p < .001$, $r = .60$. In den weiteren Onset-Tests zeigen die Probanden keine signifikanten Unterschiede in ihren Blickzeiten. Somit ist festzuhalten, dass nicht nur während der Salienzphase, sondern auch während der Original-Label-Onset-Testphase ein Salienz-Effekt bei den Probanden festzustellen ist.

Tabelle 11

Prozentuale Blickdauer aller Probanden zum Ablenker während der Eye-Tracking-Onset-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 %

Test	Blickdauer zum Ablenker		<i>t</i> (65)	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>			
Original-Label-Onset	59.96	13.45	6.02	<.001***	.60
New-Label-Onset	54.25	22.48	1.54	.13	.19
Recovery-Onset	54.41	24.03	1.49	.14	.18

4.3.5.2.2 Blickdauer zu den AOIs in den sechs-sekündigen Offset-Phasen

Nachdem festgestellt wurde, dass alle Probanden während der Original-Label-Onset-Phase das salientere der beiden Objekte präferieren, wird nun das Blickverhalten beider Gruppen in den Offset-Phasen ausgewertet. Jede Offset-Phase umfasst ein sechs-sekündiges Zeitfenster, welches direkt nach dem Zielwort startet und bis zum Ende des Videoclips andauert. Es wird angenommen, dass die Joint-Engagement-Gruppe ein anderes Blickverhalten während der Original-Label-, New-Label-, und Recovery-Offset-Phasen zeigt als die Self-Exploring-Gruppe, da die Probanden an unterschiedlichen, sozialen Lernbedingungen teilgenommen haben. Die Vermutung lautet, dass die Self-Exploring-Gruppe die Wort-Objekt-Verbindung nicht beibehält, wodurch weiterhin ein Salienz-Effekt in den Offset-Phasen zu beobachten sein wird. Im Gegensatz dazu wird vermutet, dass die Joint-Engagement-Gruppe die Wort-Objekt-Verbindung beibehält und das rezeptive Wortwissen anhand einer längeren Blickdauer zum Zielobjekt demonstriert. Zusätzlich soll anhand des New-Label-Offset-Tests das Ausschließbarkeitsprinzip überprüft werden, ob die Versuchsteilnehmer eine längere Blickzeit zum Ablenker aufweisen.

Die Offset-Blickzeiten zum salienten Objekt und zum Zielobjekt wurden in einem Einstichproben-*t*-Test gegen den Wahrscheinlichkeitswert von 50 % getestet (vgl. Tabelle 12, S. 124). Dem Einstichproben *t*-Test nach sind in der Joint-Engagement-Gruppe weder während des Original-Label-, noch während des New-Label- oder des Recovery-Tests Unterschiede zum Wahrscheinlichkeitswert zu finden. In der Self-Exploring-Gruppe sind hingegen ein marginaler (Rasch, Frieze, Hofmann & Naumann, 2006, S. 58) und zwei signifikante Unterschiede vorhanden: In dem Original-Label-Test liegt die Blickdauer zum Ablenker ($M = 54.7$, $SD = 14.72$) marginal über dem Wahrscheinlichkeits-Level von 50 %, $t(32) = 1.83$, $p = .076$, $r = .31$, und die Blickdauer zum Zielobjekt ($M = 45.3$, $SD = 14.72$) daher marginal darunter, $t(32) = 1.83$, $p = .76$, $r = .31$. In dem New-Label-Test fixiert die Self-Exploring-Gruppe den Ablenker durchschnittlich 57.35 % ($SD = 20.21$) der

Gesamtzeit, was signifikant über dem Wahrscheinlichkeits-Level liegt, $t(32) = 2.09$, $p = .045$, $r = .35$. Dementsprechend ist die prozentuale Blickdauer zum Zielobjekt ($M = 42.65$, $SD = 20.21$) während des New-Label-Tests signifikant unter dem Wahrscheinlichkeitswert, $t(32) = -2.09$, $p = .045$, $r = .35$. Ebenso zeigt sich ein signifikanter Unterschied in der Blickdauer zum Ablenker ($M = 57.03$, $SD = 19.61$) und dem Wahrscheinlichkeitswert während des Recovery-Tests, $t(32) = 2.06$, $p = .048$, $r = .34$. Auch hier liegt die Blickdauer zum Zielobjekt ($M = 42.97$, $SD = 19.61$) daher signifikant unter 50%, $t(32) = -2.06$, $p = .048$, $r = .34$.

Die beschriebenen Ergebnisse bestätigen die aufgestellte Annahme insofern, als die Self-Exploring-Gruppe in allen Offset-Testphasen eine visuelle Präferenz für den salienten Ablenker zeigt und daher die enkodierte Wort-Objekt-Verbindung nicht beibehalten hat. Bei der Joint-Engagement-Gruppe konnte keine visuelle Präferenz für das Zielobjekt im Eye-Tracking-Test festgestellt werden.

Des Weiteren kann hinsichtlich des Ausschließbarkeitsprinzips geschlussfolgert werden, dass zwar eine signifikant längere Blickzeit zum Ablenker in der New-Label-Phase der Self-Exploring-Gruppe zu finden ist, dieses Ergebnis aber – in Anbetracht der Blickzeiten während des Original-Label- und Recovery-Tests – für eine andauernde visuelle Präferenz für den Ablenker spricht, anstatt für eine bewusste Referentenauswahl.

Tabelle 12

Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye-Tracking-Onset- und Offset-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 %

Test	Blickdauer zum Ablenker									
	Joint Engagement					Self Exploring				
	<i>M</i>	<i>SD</i>	<i>t</i> (32)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (32)	<i>p</i>	<i>r</i>
Onset-Phase (3 Sekunden)										
Original Label	60.03	14.38	4.01	<.001***	.58	59.89	12.68	4.48	<.001***	.29
New Label	54.55	20.85	1.25	.219	.22	53.95	24.33	0.93	.359	.16
Recovery	53.51	20.38	0.99	.33	.17	55.31	27.49	1.11	.275	.19
Offset-Phase (6 Sekunden)										
Original Label	51.82	15.07	0.69	.493	.12	54.7	14.72	1.83	.076†	.31
New Label	54.7	22.05	1.22	.23	.21	57.35	20.21	2.09	.045*	.35
Recovery	48.03	19.69	-0.58	.57	.10	57.03	19.61	2.06	.048*	.34

4.3.5.2.3 Blickdauer zu den AOIs in den drei-sekündigen Offset-Phasen

Wie zuvor zusammengefasst, konnte die Annahme in Bezug auf die Referentenauswahl der Joint-Engagement-Gruppe während der Offset-Tests nicht bestätigt werden. Die Ergebnisse lassen vermuten, dass die Blickdaten aufgrund der sechs-sekündigen Offset-Phase „verzerrt“ sein könnten. Daher lautet die Annahme in Bezug auf die folgende Auswertung, dass in den ersten Sekunden, nachdem das Zielwort fällt, die Probanden eine Referentenauswahl treffen und sich in der restlichen Zeit visuell anders orientieren. Das heißt, dass sie in der restlichen Zeit vermutlich den Blick zwischen den beiden Stimuli hin und her wechseln.

Da auch aus den Protokolleinträgen der vorliegenden Studie hervorging, dass die Probanden zum Ende der Videoclips unruhiger wurden als am Anfang, werden im nächsten Schritt die Eye-Tracking-Testphasen nochmal überprüft, indem dieses Mal lediglich das Blickverhalten während der ersten drei Sekunden des insgesamt sechs Sekunden langen Offset-Fensters ausgewertet wird. Wie schon in den Auswertungen zuvor, wurde mithilfe des Einstichproben-*t*-Tests die Fixationszeit zum Ablenker gegen den Wahrscheinlichkeitswert von 50 % getestet (vgl. Tabelle 13, S. 126).

Doch auch in dieser Auswertung konnte bei der Joint-Engagement-Gruppe in keinem der Eye-Tracking-Tests Unterschiede zwischen den Fixationszeiten und dem Wahrscheinlichkeitswert verzeichnet werden. In der Self-Exploring-Gruppe ist hingegen ein signifikanter Unterschied in dem Original-Label-Test und ein marginaler Unterschied in dem New-Label-Test zu finden: Die Self-Exploring-Gruppe zeigt eine signifikant längere Blickdauer zum Ablenker ($M = 56.85$, $SD = 13.05$) im Vergleich zum Wahrscheinlichkeitswert, $t(32) = 3.014$, $p = .005$, $r = .47$. Dementsprechend liegt die Blickdauer zum Zielobjekt ($M = 43.15$, $SD = 13.05$) signifikant unter dem Wahrscheinlichkeitswert, $t(32) = -3.014$, $p = .005$, $r = .48$. Die Blickpräferenz für den Ablenker ($M = 56.62$, $SD = 18.89$) bleibt auch noch im New-Label-Test erhalten, $t(32) = 2.01$, $p = .047$, $r = .34$. Auch hier bedeutet es, dass die Fixationsdauer zum Zielobjekt ($M = 43.38$, $SD = 18.89$) unter dem Wahrscheinlichkeitswert liegt, $t(32) = -2.01$, $p = .047$, $r = .34$. Trotz des marginalen Effekts kann man hier von einem mittelstarken Effekt sprechen, da der Wert über 0.3 liegt. Im Vergleich zum vorherigen Recovery-Test findet sich in der Auswertung des drei-sekündigen Ausschnitts kein signifikanter Unterschied. Mithilfe einer anschließenden Regressionsanalyse konnte außerdem gezeigt werden, dass der Bildungsabschluss der Mütter nicht in Zusammenhang steht mit der Fixationsdauer der Joint-Engagement-Gruppe, $R^2 = .007$, $F(1, 31) = 0.213$, $p = .648$, und der Self-Exploring-Gruppe zum Zielobjekt im Original-Label-Offset-Test, $R^2 = .003$, $F(1, 31) = 0.088$, $p = .768$.

Die Auswertung der Blickdaten aus einem kürzeren Zeitfenster untermauert die vorherigen Ergebnisse insofern, als die Self-Exploring-Gruppe während des Original-Label- und des New-Label-Tests den Ablenker länger fixiert als das Zielobjekt, wodurch wieder der Salienz-Effekt ersichtlich wird. Bei der Joint-Engagement-Gruppe zeigt sich auch im kürzeren Zeitfenster zu keinem der beiden Objekte eine visuelle Präferenz. In Bezug auf das Ausschließbarkeitsprinzip kann geschlussfolgert werden, dass eine längere Blickzeit zum Ablenker in der New-Label-Phase der Self-Exploring-Gruppe zu finden ist. Diese spricht allerdings – wie auch schon in der vorherigen Auswertung (vgl. Kapitel 4.3.5.2.2) – für eine fortlaufende, visuelle Präferenz für das saliente Objekt, anstatt für eine explizite Referentenauswahl.

Tabelle 13

Prozentuale Blickdauer beider Gruppen zum Ablenker während der drei-sekündigen Eye-Tracking-Offset-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 %

Test	Blickdauer zum Ablenker									
	Joint Engagement					Self Exploring				
	<i>M</i>	<i>SD</i>	<i>t</i> (32)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (32)	<i>p</i>	<i>r</i>
Original Label	53.32	16.67	1.15	.26	.20	56.85	13.05	3.01	.005**	.47
New Label	53.33	21.34	0.89	.38	.16	56.62	18.89	2.01	.047†	.34
Recovery	52.91	26.21	0.64	.53	.11	49.69	26.59	– 0.07	.947	.01

4.3.5.2.4 Time to First Fixation

Da die Auswertung der anteiligen Blickdauer während der Offset-Testphasen die Annahmen bezüglich des Salienz-Effekt in der Self-Exploring-Gruppe untermauern, aber keine Rückschlüsse über das Wortverständnis in der Joint-Engagement-Gruppe gezogen werden können, wird im nächsten Schritt die Zeit bis zur ersten Fixation nach der Äußerung des Zielwortes untersucht. Indem die Dauer der visuellen Orientierung zu einem der beiden Objekte für die Auswertung hinzugezogen wird, sollen dadurch nähere Erkenntnisse über die Referentenauswahl der Joint-Engagement-Gruppe gewonnen werden. Da in den bisherigen Auswertungen die Joint-Engagement-Gruppe keine längere Blickdauer zum Zielobjekt zeigte, sollte untersucht werden, ob die Referentenauswahl noch früher, also direkt nach der Wortäußerung, erfolgt. Somit wird angenommen, dass sich die Joint-Engagement-Gruppe im Original-Label-Test schneller zum Zielobjekt als zum Ablenker orientiert, wenn sie das Wort mit dem Zielobjekt verknüpft hat und beibehalten konnte.

Mithilfe eines abhängigen t -Tests wurde ausgewertet, ob es einen Unterschied zwischen der Zeit bis zur ersten Fixation des Zielobjektes und des Ablenkers gibt (vgl. Tabelle 14, S. 127). Bei der Joint-Engagement-Gruppe konnten keine signifikanten Unterschiede in den Time-to-First-Fixation-Zeiten während der Offset-Tests festgestellt werden. Im Gegensatz zu der Joint-Engagement-Gruppe, ist die Self-Exploring-Gruppe während des Original-Label-Tests signifikant schneller in der visuellen Orientierung zum Ablenker ($M = 0.83$, $SD = 0.7$) als zum Zielobjekt ($M = 1.2$, $SD = 0.7$), $t(32) = -3.68$, $p < .001$, $r = .55$.

Die aufgestellte Annahme, dass sich die Joint-Engagement-Gruppe schneller zum Zielobjekt als zum salienten Objekt orientiert, sobald sie das Zielwort hört, konnte nicht bestätigt werden. Doch die Ergebnisse zeigen, dass die Self-Exploring-Gruppe sich in dem Original-Label-Test schneller zum salienten Objekt als zu dem weniger salienten Zielobjekt schaut. Damit verstärkt dieses Ergebnis den Befund des Salienz-Effekts in der Self-Exploring-Gruppe.

Tabelle 14

Die Zeitdauer in Sekunden bis zur ersten Fixation eines der beiden Objekte während der Eye-Tracking-Offset-Testphasen

Test	Dauer bis zur ersten Fixation				<i>t</i> (32)	<i>p</i>	<i>r</i>
	Ablenker		Zielobjekt				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Joint Engagement							
Original Label	1.03	0.67	0.97	0.63	0.44	.663	.08
New Label	1.02	0.78	0.94	0.79	0.42	.676	.07
Recovery	1.28	1.05	1.26	0.80	0.09	.929	.02
Self Exploring							
Original Label	0.83	0.70	1.20	0.70	− 3.68	<.001***	.55
New Label	0.92	0.88	1.16	0.94	− 0.94	.354	.16
Recovery	1.22	1.06	1.17	0.62	0.22	.829	.04

4.3.5.3 Zusammenhang der Blickzeiten während der Spiel- und Testphase

In den zuvor beschriebenen Auswertungen stellte sich nicht nur ein unterschiedliches Blickverhalten zwischen den beiden Gruppen während der Spielphase (vgl. Kapitel 4.3.5.1), sondern auch während der Eye-Tracking-Tests heraus (vgl. Kapitel 4.3.5.2): Die Teilnehmer der Joint-Engagement-Gruppe fixierten das Zielobjekt während der Spielphase länger als die Self-Exploring-Gruppe, wohingegen die Self-Exploring-Gruppe die Experimentatorin

länger anschaute im Vergleich zur Joint-Engagement-Gruppe. Im Eye-Tracking-Test stellte sich außerdem heraus, dass die Joint-Engagement-Gruppe keines der beiden Objekte länger anschaute, bei der Self-Exploring-Gruppe hingegen ein Salienz-Effekt im Test auftrat. Als dritte Hypothese wurde aufgestellt, dass es einen Zusammenhang zwischen den sozialen Lernformaten der Spielphase und den Eye-Tracking-Daten im Test gibt. Daher gilt es nun zu untersuchen, welches Verhalten in dem Lernformat der Joint-Engagement- und der Self-Exploring-Gruppe in Zusammenhang mit den unterschiedlichen Eye-Tracking-Daten steht.

Als mögliche Erklärung der unterschiedlichen Eye-Tracking-Daten wird angenommen, dass gemeinsame Interaktionen die visuelle Aufmerksamkeit von Kindern zum Zielobjekt länger aufrechterhalten als alleinige Spielphasen. Bei der Self-Exploring-Gruppe ist während der alleinigen Spielphase die Blickdauer zum Zielobjekt gesunken, wohingegen sie zu anderen Dingen im visuellen Umfeld gestiegen ist (vgl. Kapitel 4.3.5.1). Im Vergleich zur Self-Exploring-Gruppe hat die Blickdauer der Joint-Engagement-Gruppe zum Zielobjekt während der Spielphase zugenommen. Diese Annahme wird zudem von der aktuellen Forschungsliteratur (Deák et al., 2018, Yu & Smith, 2016) gestützt. In Bezug auf diese vorliegenden Ergebnisse besteht die Annahme, dass die gemeinsamen Handlungen in der Joint-Engagement-Bedingung die visuelle Aufmerksamkeitsdauer der Kinder zum Zielobjekt erhöht und damit die Enkodierung der Wort-Objekt-Verknüpfung verstärkt haben. Wenn diese Annahme bestätigt wird, dann kann daraus geschlussfolgert werden, dass die Joint-Engagement-Gruppe aufgrund einer längeren Blickdauer zum Zielobjekt während der Spielphase, die durch gemeinsame, zielführende Objekthandlungen erreicht wurde, dem Salienz-Effekt im Eye-Tracking-Test entgegenwirken konnte.

In den folgenden Auswertungen werden die Blickzeiten der Probanden während der Spielphase mit den Blickzeiten während des Eye-Tracking-Tests in Zusammenhang gebracht. Zur Auswertung wurde die prozentuale Blickdauer zum Zielobjekt bzw. zum Ablenker während der Spielphase sowie die prozentuale Fixationszeit des Zielobjekts und des Ablenkers während des drei-sekündigen Original-Label-Offset-Tests ausgewählt. Es wurde das kürzere Zeitfenster ausgewählt, weil im drei-sekündigen Original-Label-Offset-Test ein signifikanter Unterschied zwischen der Blickdauer zum salienten Objekt und dem Wahrscheinlichkeitswert in der Self-Exploring-Gruppe vorliegt (vgl. Kapitel 4.3.5.2.3), wohingegen in dem sechs-sekündigen Zeitfenster der Unterschied marginal war.

Die beschriebenen Variablen wurden mithilfe der bivariaten Korrelation nach Pearson für alle Probanden ($N = 66$) korreliert, weil von einer Normalverteilung ausgegangen wird und die Daten intervallskaliert sind (Field, 2009, S. 186). Allerdings bestand keine

signifikante Korrelation zwischen der Blickdauer zum Zielobjekt während der Spielphase und dem Original-Label-Offset-Test, $r = .18$, $p > .073$. Die Blickdauer zum Ablenker während der Spielphase korrelierte ebenfalls nicht mit der Blickdauer zum Ablenker während des Original-Label-Offset-Tests, $r = .19$, $p > .066$. Auch der Korrelationskoeffizient, der zwischen -1 und 1 liegen kann, zeigt, dass es sich um einen kleinen Effekt handelt, da r kleiner als $.30$ ist. Aus den Ergebnissen kann geschlussfolgert werden, dass kein Zusammenhang zwischen der Fixationsdauer der Objekte während der Spielphase und der Fixationsdauer dieser Objekte im Original-Label-Test besteht.

4.3.5.4 Gruppenabhängige Handlungsstrategien während der Spielphase

Da die Blickdauer zu den beiden Objekten während der Spielphase nicht in Zusammenhang zu der Blickdauer im drei-sekündigen Original-Label-Offset-Test steht, werden alternative Erklärungen in Erwägung gezogen, um das unterschiedliche Blickverhalten der beiden Gruppen im Eye-Tracking-Test zu erklären. Im Folgenden werden die Handlungsstrategien der Joint-Engagement- und der Self-Exploring-Gruppe miteinander verglichen, um zu erfahren, ob sich die beiden Gruppen in den Objekthandlungen voneinander unterscheiden.

Es wurde ein unabhängiger t -Test durchgeführt, um die Anzahl der Handlungsstrategien beider Gruppen miteinander zu vergleichen (vgl. Tabelle 15 S. 130). Dabei stellte sich heraus, dass es hochsignifikante Unterschiede zwischen den beiden Gruppen in der Durchführung der Zielstrategien gibt. Die Joint-Engagement-Gruppe führte alle Zielfunktionen signifikant häufiger aus als die Self-Exploring-Gruppe. Die Self-Exploring-Gruppe ($M = 9.45$, $SD = 3.21$) praktizierte hingegen signifikant häufiger sonstige Objekthandlungen (z. B. orale Exploration, auf den Boden werfen, mit dem Objekt auf den Tisch klopfen) als die Joint-Engagement-Gruppe ($M = 6.36$, $SD = 2.12$), $t(64) = -4.61$, $p < .001$, $r = .49$. Die sonstigen Strategien wurden allerdings in dem Kodierschema der vorliegenden Arbeit sowie in der Tabelle 15 nicht weiter unterteilt. Außerdem besteht ein signifikanter Unterschied der beiden Gruppen in der Verweigerung, $t(64) = 2.87$, $p < .001$, $r = .34$. Dieser Unterschied wird damit erklärt, dass der Self-Exploring-Gruppe ($M = 0.06$, $SD = 0.35$) die Verweigerung der Interaktion nicht in demselben Maße möglich war wie der Joint-Engagement-Gruppe ($M = 0.55$, $SD = 0.91$), weil sich die Experimentatorin in der Self-Exploring-Bedingung zurückgezogen hat. Daher wird dieses Ergebnis für die weitere Diskussion der Daten nicht berücksichtigt.

Abschließend ist zu erwähnen, dass eine Regressionsanalyse zeigte, dass der Bildungsabschluss der Mütter weder in Zusammenhang steht mit der Anzahl der

Handlungen der Joint-Engagement-Gruppe mit dem Zielobjekt, $R^2 = .009$, $F(1,31) = 0.269$, $p = .607$, noch mit der Anzahl der Handlungen der Self-Exploring-Gruppe, $R^2 = .052$, $F(1,31) = 1.689$, $p = .203$.

Tabelle 15

Durchschnittliche Anzahl der angewandten Handlungsstrategien während der Spielphase in Abhängigkeit von der Gruppenzuordnung

Strategie	Joint Engagement		Self Exploring		t(64)	p	r
	M	SD	M	SD			
Öffnen	3.24	0.97	0.30	0.47	15.70	<.001***	.89
Schließen	0.61	0.93	0.00	0.00	3.73	<.001***	.42
Rausziehen	3.42	1.12	0.82	0.53	12.10	<.001***	.83
Einsetzen	1.21	1.32	0.03	1.74	5.11	<.001***	.54
Sonstige	6.36	2.12	9.45	3.21	- 4.61	<.001***	.49
Verweigerung	0.55	0.91	0.06	0.35	2.87	<.001**	.34

4.3.5.5 Zusammenhang der Handlungsstrategien und der Eye-Tracking-Daten

Da festgestellt wurde, dass sich die beiden Gruppen in der Anzahl der in der Spielphase ausgeführten Zielhandlungen voneinander unterscheiden (vgl. Tabelle 15), besteht die Annahme, dass eine häufigere Ausführung der Zielfunktionen mit einer längeren Blickdauer zum Zielobjekt im drei-sekündigen Original-Label-Offset-Test in Zusammenhang steht. Das wiederum bedeutet, dass eine geringe Anzahl oder gar keine Zielhandlungen mit einer kurzen Blickdauer zum Zielobjekt im Test zusammenhängen. Diese Annahme wird damit begründet, dass die Anzahl der zielführenden Objekthandlungen mit der Tiefe der Wort-Objekt-Enkodierung zusammenhängt.

Um diesen Zusammenhang mithilfe der Pearson-Korrelation zu überprüfen, wurde zunächst der Durchschnittswert aus allen Zielstrategien (Öffnen, Schließen, Rausziehen, Einsetzen) berechnet. Durch die Ermittlung einer neuen Variable, sollte auch der Einfluss der Schwierigkeitsgrade der Objektfunktionen als Störvariable ausgeschlossen werden. Alle Probanden ($N = 66$) haben im Durchschnitt 1.2 ($SD = 1.11$) Zielfunktionen erreicht. Die darauffolgende Korrelation nach Pearson ergab, dass die Anzahl der Funktionshandlungen während der Spielphase weder mit der Blickdauer zum Ablenker, $r = -.07$, $p > .295$, noch mit der Blickdauer zum Zielobjekt, $r = .07$, $p > .295$, während des drei-sekündigen Original-Label-Offset-Test korrelierte. Somit konnte kein Zusammenhang zwischen den Handlungsstrategien und dem Blickverhalten zu den beiden Objekten festgestellt werden.

Eine weitere Annahme lautet, dass sich die Kinder nicht nur zwischen den Gruppen, sondern auch innerhalb der jeweiligen Gruppen hinsichtlich der Anzahl der erreichten Zielfunktionen unterscheiden. Im Folgenden werden die Versuchsteilnehmer in „schwache“ und „starke“ Teilnehmer unterteilt, um zu untersuchen, ob sie ein unterschiedliches Blickverhalten im Test haben (vgl. Tabelle 16, S. 132). Es wird angenommen, dass diejenigen Probanden der Joint-Engagement-Gruppe, die in den Handlungsstrategien weniger erfolgreich waren, den Ablenker im Test länger anschauen, als diejenigen, die die Handlungsstrategien erfolgreich gemeistert haben. Für diese Annahme spricht das Ergebnis vorheriger Auswertungen, dass in der Joint-Engagement-Gruppe keine Unterschiede in der Blickdauer zu den beiden Objekten während der Eye-Tracking-Testphase zu beobachten waren (vgl. Kapitel 4.3.5.2). Dies könnte damit zusammenhängen, dass die Probanden unterschiedlich erfolgreich in den Ausführungen der Zielfunktionen waren und daher nicht alle Probanden das Wortwissen im Test abrufen konnten.

Um diese Annahme zu prüfen, wurden die Versuchsteilnehmer der Self-Exploring- und der Joint-Engagement-Gruppe in Subgruppen aufgeteilt. Probanden, die in den Handlungsstrategien erfolgreich waren, wurden als *Achiever* bezeichnet. Diejenigen, die weniger erfolgreich waren, wurden der Subgruppe *Low Achiever* zugeordnet. Als Maß der Zuordnung zu der Achiever bzw. der Low-Achiever Gruppe gilt die Ausführung der Zielfunktionen mit den Holzobjekten: Das Öffnen der Holzfigur und das Rausziehen des Rundholzes aus dem Klotz zählen zur Funktion 1, wohingegen das Schließen der Holzfigur und das Einsetzen des Rundholzes in den Klotz zur Funktion 2 gehören.

Damit die Versuchsteilnehmer der Joint-Engagement-Gruppe den Achievern ($N = 20$) zugeordnet werden, sollten sie zwei Punkte erfüllt haben. Zum einen sollten sie eine der beiden Objekthandlungen der Funktion 1 – d. h., das Öffnen der Holzfigur oder das Herausziehen des Klotzes – mindestens drei Mal erreicht haben. Diese Angabe beruht auf der durchschnittlichen Anzahl dieser Funktion in der Joint-Engagement-Gruppe ($M_{\text{Öffnen}} = 3.24$, $SD = 0.95$; $M_{\text{Rausziehen}} = 3.42$, $SD = 1.1$). Zusätzlich sollte die Achiever-Gruppe mindestens einmal eine der beiden Objekthandlungen der Funktion 2 durchgeführt haben, da die Funktion des Schließens der Holzfigur im Durchschnitt 0.61 Mal ($SD = 0.92$) und die Funktion des Einsetzens des Rundholzes durchschnittlich 1.2 Mal ($SD = 1.3$) erreicht wurde. Versuchsteilnehmer, die die genannten Kriterien *nicht* erfüllten, wurden der Gruppe der Low-Achiever ($N = 13$) zugewiesen.

In der Self-Exploring-Gruppe galten etwas andere Voraussetzungen: Die Achiever der Self-Exploring-Gruppe ($N = 8$) unterscheiden sich von den Low Achiever ($N = 25$) insofern,

als sie die Funktion 1 mindestens einmal erreicht haben sollten. Auch dieser Wert ist auf den Durchschnittswert der Funktion des Öffnens ($M = 0.3$, $SD = 0.45$) und der des Herausziehens zurückzuführen ($M = 0.82$, $SD = 0.52$).

Nachdem die Untergruppen festgelegt waren, wurde ein unabhängiger t -Test durchgeführt, mithilfe dessen die Blickzeiten zum Distraktor zwischen den Achievern und Low Achievern der Joint-Engagement- und Self-Exploring-Gruppe verglichen wurden (vgl. Tabelle 16). Allerdings hat sich in keiner der Testphasen das Blickverhalten zwischen den jeweiligen Subgruppen zu den Stimuli unterschieden.

Tabelle 16

Prozentuale Blickdauer zum Ablenker während der Eye-Tracking-Testphasen in Abhängigkeit von der Subgruppen-Zuordnung

Test	Blickdauer zum salienten Objekt				<i>t</i> (31)	<i>p</i>	<i>r</i>
	Achiever		Low Achiever				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Joint Engagement							
Original Label	57.98	13.45	55.30	12.82	0.58	.568	.10
New Label	56.64	19.45	56.59	18.81	0.01	.995	.00
Recovery	53.59	26.57	44.39	26.67	0.98	.334	.17
Self Exploring							
Original Label	53.75	15.47	52.34	20.02	0.22	.827	.04
New Label	52.42	22.12	55.43	20.41	−0.37	.716	.07
Recovery	51.49	28.63	56.16	20.56	−0.46	.646	.08

Anschließend wurde die Blickzeit der Joint-Engagement-Achiever-Gruppe zum Ablenker ($N = 23$) mit der Blickzeit der Self-Exploring-Low-Achiever ($N = 10$) untersucht, um die „stärkste“ Gruppe mit der „schwächsten“ Gruppe zu vergleichen. Dies hat zum Ziel zu erfahren, ob sich die Blickzeiten der beiden „Extrem“-Gruppen unterscheiden. Aber auch hier zeigte sich in keinem der Eye-Tracking-Tests ein signifikanter Unterschied zwischen den Fixationszeiten. Unter Berücksichtigung der vorliegenden Ergebnisse, können auch die durchgeführten Handlungsstrategien während des Trainings die unterschiedlichen Blickzeiten der beiden Gruppen im Eye-Tracking-Test nicht erklären. In der Zwischendiskussion (vgl. Kapitel 4.4) werden die Ergebnisse hinsichtlich der aufgestellten Hypothesen ausführlich diskutiert. Im anschließenden Kapitel werden weiterführende Analysen beschrieben, die unabhängig von den aufgestellten Hypothesen durchgeführt wurden.

4.3.6 Weiterführende Analysen

Aus den bisherigen Befunden ergaben sich weitere Fragen in Bezug auf die Video- und Eye-Tracking-Daten, die in weiterführenden Analysen untersucht wurden. In den Auswertungen der Objektmanipulation (vgl. Kapitel 4.3.5.4) hat sich herausgestellt, dass die Joint-Engagement-Gruppe signifikant mehr Zielhandlungen ausgeführt hat im Vergleich zu der Self-Exploring-Gruppe. Daraus resultierte die Annahme, dass die Joint-Engagement-Gruppe mehr positive Rückmeldungen und Ansporn in Form von Aufforderungen erhalten hat als die Self-Exploring-Gruppe. Des Weiteren wurden die Eye-Tracking-Daten hinsichtlich der Weighted Gaze Samples zur Analyse hinzugezogen, um den Aufmerksamkeitsverlauf der Probanden zu kontrollieren.

4.3.6.1 Soziale Hinweise während der Trainingsphase

Zu der Kategorie der sozialen Hinweise zählt neben der Ostension auch die positive Rückmeldung und Aufforderung (vgl. Kapitel 4.2.4.3). Die Menge der sozialen Hinweise seitens der Experimentatorin wurde gruppenabhängig in einem unabhängiger t -Test untersucht. Die Auswertung zeigt, dass sich die beiden Gruppen tatsächlich hinsichtlich des sozialen Inputs signifikant unterscheiden (vgl. Tabelle 17). Die Joint-Engagement-Gruppe bekam nicht nur signifikant mehr positives Feedback ($M = 4.26$, $SD = 1.08$) im Vergleich zu der Self-Exploring-Gruppe ($M = 0.65$, $SD = 0.48$), $t(64) = 17.51$, $p < .001$, $r = .91$, sondern auch mehr Aufforderungen eine Zielfunktion zu wiederholen, $t(64) = 16.03$, $p < .001$, $r = .89$. Auch in der Kategorie Ostension liegt ein signifikanter Unterschied zwischen der Joint-Engagement-Gruppe ($M = 0.08$, $SD = 0.18$) und der Self-Exploring-Gruppe vor ($M = 0.00$, $SD = 0.00$) vor, $t(64) = 2.31$, $p = .02$, $r = .28$, da sich die Experimentatorin in der Spielphase der Self-Exploring-Bedingung aus der Interaktion herausgezogen hat und keine ostensiven Signale gab. Daher wird diese Kategorie für die Interpretation der Ergebnisse nicht berücksichtigt.

Tabelle 17

Durchschnittliche Anzahl der sozialen Hinweise während der Spielphase in Abhängigkeit von der Gruppenzuordnung

Soziale Hinweise	Joint Engagement		Self Exploring		$t(64)$	p	r
	M	SD	M	SD			
Ostension	0.08	0.18	0.00	0.00	2.31	.02*	.28
Positives Feedback	4.26	1.08	0.65	0.48	17.51	<.001***	.91
Aufforderung	3.31	0.94	0.55	0.32	16.03	<.001***	.89

Anschließend wurden die sozialen Hinweise mit den Handlungsstrategien korreliert, um den Zusammenhang dieser Variablen zu untersuchen. Das Ziel der Anwendung dieses statistischen Testverfahrens bestand darin zu erfahren, wie das soziale Feedback seitens der Experimentatorin und die Handlungsstrategien der Kinder zusammenhängen. Hierfür wurde zunächst die Anzahl der zielgerichteten Funktionen (Öffnen, Schließen, Rausziehen, Einsetzen) als eine übergreifende Variable *objectfunction* zusammengefasst und die Variablen Sonstige und Verweigerung als solche beibehalten. Im nächsten Schritt wurde die bivariate Korrelation nach Pearson angewandt.

In der Joint-Engagement-Gruppe ist der signifikante Zusammenhang zwischen den Variablen Verweigerung und Ostension aufzuführen, $r = .41, p = .02$. Während bei der Joint-Engagement-Gruppe keine weiteren Korrelationen gefunden wurden, korreliert bei der Self-Exploring-Gruppe die durchschnittlich erreichte Anzahl der Zielfunktionen positiv mit der Anzahl der positiven Feedbacks, $r = .79, p < .001$, sowie mit der Anzahl der Aufforderungen, $r = .76, p < .001$. Auf die Interpretation dieser Ergebnisse wird in der Zwischendiskussion (vgl. Kapitel 4.4) eingegangen.

4.3.6.2 Aufmerksamkeitsverlauf – Weighted Gaze Samples

Aufgrund der Notizen in den Studienprotokollen bestand die Annahme, dass die WGS zum Ende des Eye-Tracking-Tests abnehmen. Da alle technischen Voraussetzungen für das Eye-Tracking-Verfahren erfüllt waren, könnte eine Reduktion der WGS auf eine abfallende Aufmerksamkeitsspanne hindeuten. Um der Annahme nachzugehen, wurden in der Datenauswertung die WGS hinzugezogen.

Für die Auswertung ist eine Mixed-Design-ANOVA in SPSS durchgeführt worden (Field, 2009, S. 506 ff.). Dabei handelt es sich um eine Form der Varianzanalyse, die sowohl Innersubjektfaktoren, als auch Zwischensubjektfaktoren erfasst. In dem vorliegenden Fall handelt es sich bei dem Zwischensubjektfaktor um die *Gruppenzugehörigkeit* (Joint Engagement vs. Self Exploring). Die Innersubjektfaktoren sind vier Prozentwerte der WGS, die für die Datenqualität der jeweiligen Eye-Tracking-Tests (Salienzphase 1, Salienzphase 2, Testphase 1, Testphase 2) stehen (vgl. Tabelle 18, S. 135).

Da die Sphärizität der Daten gegeben war ($p > .05$) sowie die Homogenität der Fehlervarianzen gemäß des Levene-Tests für alle Variablen erfüllt waren ($p > .05$), wurde der Interaktionseffekt bestimmt, bei dem es sich um eine Abhängigkeit zweier Variablen handelt. Eine Interaktion dieser Variablen würde bedeuten, dass sich die Variable Weighted Gaze Samples in Abhängigkeit von der Variable Gruppenzugehörigkeit unterscheidet.

Allerdings ist keine statistisch signifikante Interaktion zwischen den WGS und den Untersuchungsgruppen zu finden, $F(3, 192) = 0.6, p = .62, \eta^2 = .01$. Aber es ergibt sich ein signifikanter Haupteffekt der Weighted Gaze Samples auf, wenn dieser unabhängig von der Gruppenzugehörigkeit der Versuchsteilnehmer betrachtet wird, $F(3, 192) = 5.83, p < .001, \eta^2 = .08$.

Da kein signifikanter Interaktionseffekt, sondern ein einfacher, gruppenunabhängiger Haupteffekt vorliegt, wurden zusätzlich paarweise Vergleiche durchgeführt, um zu bestimmen, wo die signifikanten Unterschiede zwischen den einzelnen WGS auftauchten. Die paarweisen Auswertungen zeigen, dass sich die WGS zwischen der ersten Testphase (T1) und der zweiten Testphase (T2) sowie zwischen des zweiten Salienztests (S2) und der zweiten Testphase (T2) signifikant unterscheiden. Ein anschließender abhängiger t -Test zeigte, dass sich der Prozentwert der WGS in T1 ($M = 70.76, SD = 14.33$) hochsignifikant von dem Wert in T2 ($M = 62.79, SD = 16.52$) unterscheidet, $t(65) = 3.94, p < .001, r = .44$. Aber auch in S2 ($M = 70.59, SD = 14.26$) ist der Prozentwert signifikant höher als in T2 ($M = 62.79, SD = 16.52$), $t(65) = 3.65, p < .001, r = .41$.

Zusammenfassend lässt sich festhalten, dass der WGS-Prozentwert in dem letzten Eye-Tracking-Test (T2) niedriger ist im Vergleich zu den vorherigen Tests (vgl. Tabelle 18). Dieser Nebenfund zeigt, dass die visuelle Aufmerksamkeit zu den Stimuli, die auf dem Bildschirm präsentiert sind, in T2 am niedrigsten ist. Die möglichen Gründe hierfür werden in der Zwischendiskussion (vgl. Kapitel 4.4) aufgeführt.

Tabelle 18

Weighted Gaze Samples aller Probanden ($N = 66$) in den Eye-Tracking-Tests

Test	Weighted Gaze Samples	
	<i>M</i>	<i>SD</i>
Salienzphase 1	68.27	17.46
Testphase 1	70.76	14.33
Salienzphase 2	70.59	14.26
Testphase 2	62.79	16.52

4.4 Zwischendiskussion

In dem theoretischen Teil der vorliegenden Arbeit wurden vier Wortlerntheorien vorgestellt, die sich hinsichtlich der frühkindlichen Enkodierungsstrategien unterscheiden. Das Ziel des ersten Experiments bestand zunächst darin, diese Enkodierungsstrategien in einer neuen Studie, die alle vier Theorien zusammenbringt, zu überprüfen. Dabei sollte untersucht werden, welche dieser Enkodierungsstrategien 10 Monate alte Kinder im Fast-Mapping-Prozess tatsächlich anwenden und ob die im Fast-Mapping-Prozess enkodierten Wort-Objekt-Verbindungen im Test erfolgreich abgerufen werden können. Zur Untersuchung dieser Forschungsfragen wurden die Probanden zwei unterschiedlichen Gruppen zugeordnet – der Self-Exploring- und der Joint-Engagement-Gruppe – und mit ihnen zwei neue Wörter trainiert.

Als Forschungsmethode wurde das Beobachtungsverfahren angewandt, indem die soziale Lernsituation in den beiden Gruppen auf Video aufgezeichnet und mithilfe von ELAN kodiert wurde. Diese Methode wurde mit dem Eye-Tracking-Verfahren kombiniert, um anhand der Blickzeitmessung zu überprüfen, ob die enkodierte Wort-Objekt-Verknüpfung von den Kindern abgerufen werden kann. Im folgenden Verlauf dieses Kapitels werden zunächst die Ergebnisse des Experiments zusammengefasst und die im Theorieteil formulierten Hypothesen beantwortet. Daran anknüpfend wird auf die Bedeutung dieser Ergebnisse für die Befunde der bisherigen Forschung eingegangen. Abschließend werden noch offene Fragen und Kritikpunkte, die sich aus dem ersten Experiment dieser Arbeit ergeben, thematisiert.

Die erste Hypothese besagt, dass die perzeptuelle Salienz für 10 Monate alte Kinder aufmerksamkeitslenkend ist. Hinsichtlich dieser Hypothese bestand das Ziel darin, die visuelle Präferenz der Versuchsteilnehmer anhand der Blickzeiten während der Salienzphase im Eye-Tracking-Verfahren gruppenunabhängig zu analysieren (vgl. Kapitel 4.3.3). Aufgrund der längeren Blickdauer aller Probanden zum salienten Stimulus konnte aus den Ergebnissen geschlussfolgert werden, dass 10-Monatige in der Tat eine Präferenz für perzeptuell saliente im Vergleich zu weniger salienten Objekten haben. Dieses Blickverhalten wurde in der vorliegenden Arbeit unter den Begriff *Salienz-Effekt* gefasst.

Dieser Befund spricht für die aufmerksamkeitslenkende Wirkung der perzeptuellen Salienz, die im assoziativen Ansatz von Hollich et al. (2000) sowie Pruden et al. (2006) beschrieben wird. Doch die Ergebnisse der vorliegenden Arbeit gehen über deren Sichtweise, dass Kinder am Anfang des Spracherwerbs von der perzeptuellen Salienz geleitet sind, hinaus. Der Salienz-Effekt in dem Experiment der vorliegenden Arbeit trat bei 10-

Monatigen lediglich in der Salienzphase auf, als das perzeptuell saliente Objekte nicht mit den sozialen Hinweisen gleichzeitig präsentiert wurde. In der Trainingsphase, als auch die sozialen Hinweise präsent waren, wurde das saliente Objekt von den Probanden kaum noch beachtet. Daraus kann geschlussfolgert werden, dass die Enkodierungsstrategie der perzeptuellen Objektsalienz nicht altersabhängig, sondern kontextabhängig ist.

Als zweite Hypothese wurde aufgestellt, dass soziale Hinweise 10 Monate alten Kindern als Informationsquelle zur Referentenauswahl dienen. Um diese Hypothese zu überprüfen, wurde das auf Video aufgezeichnete Blickverhalten während der Trainingsphasen ausgewertet. Die Auswertungen ergaben, dass die Versuchsteilnehmer in der Benennungsphase das Zielobjekt, welches durch soziale Hinweise der Experimentatorin hervorgehoben wurde, am längsten angeschauten. Auch in der Spielphase war die Blickdauer beider Gruppen zum Zielobjekt weiterhin am längsten. Damit zeigt das Experiment der vorliegenden Arbeit, dass die Probanden die meiste Zeit auf das langweilige Zielobjekt achteten, welches ostensiv hervorgehoben wurde, obwohl das saliente Spielzeug auf dem Tisch lag. Damit unterstützen die beschriebenen Ergebnisse die zweite Hypothese und damit die sozio-pragmatische Sichtweise, dass Kinder soziale Signale beachten und diese als Informationsquelle zur Referentenauswahl nutzen.

Die Ergebnisse der Benennungsphase sprechen daher gegen die bisherigen Forschungsbefunde von Pruden et al. (2006), dass Kinder vor ihrem ersten Lebensjahr soziale Hinweise ignorieren, und damit auch gegen die assoziative Wortlerntheorie, dass lediglich allgemeine, kognitive Mechanismen für das frühe Wortlernen notwendig sind. Allerdings sind auch Unterschiede zu den Befunden des sozio-pragmatischen Ansatzes zu finden, da beispielsweise die Ergebnisse der Studie von Yurovsky und Frank (2017) hinsichtlich der beschriebenen Schwierigkeit bei jungen Kindern, sich von Gesichtern loszulösen, nicht unterstützt werden. Die Auswertung des Blickverhaltens während der Benennungsphase demonstriert, dass sich die Probanden der vorliegenden Studie von dem Gesicht der Experimentatorin ohne Schwierigkeiten lösen konnten. Darüber hinaus schauten sie sogar die meiste Zeit auf das Zielobjekt.

Bezüglich der Studie von Yurovsky & Frank (2017) sollen zwei Kritikpunkte erwähnt werden, die einen möglichen Einfluss auf deren Ergebnisse gehabt haben könnten. Als erster Punkt ist aufzuführen, dass die Entfernung des Gesichts der Experimentatorin zu den Objekten relativ gering war, wodurch die festgelegten AOIs nah beieinanderlagen. Da sich eine Kalibrierung bei Kindern manchmal schwierig gestalten kann, kann es dazu kommen, dass der vom System berechnete Fixationspunkt versetzt ist. Aus diesem Grund sollten die AOIs

in Studien mit Kindern groß angelegt werden, sich nicht überschneiden und optimaler Weise weit genug auseinanderliegen, damit ein versetzter Fixationspunkt nicht in die falsche AOI gelangt. Aufgrund dessen, dass die AOIs nah beieinander liegen, besteht des Weiteren das Risiko, dass eines der Stimuli nicht gesondert fixiert werden muss, da es vermutlich auch peripher wahrgenommen werden kann. Eine Modifizierung dieser beiden Punkte könnte dazu beitragen, genauere Aussagen über die Bedeutung von sozialen Hinweisen in Wortlernstudien zu treffen.

Als dritte Hypothese der vorliegenden Arbeit wurde formuliert, dass es einen Zusammenhang zwischen den angewandten Enkodierungsstrategien und der Retention der Wort-Objekt-Verknüpfung während der Testphase gibt. Für die Überprüfung dieser Hypothese waren drei Schritte notwendig:

Zunächst wurde überprüft, ob es ein gruppenabhängiges Blickverhalten während der Spielphase gab, da die Probanden ein unterschiedliches Lernformat erhalten haben (vgl. Kapitel 4.3.5.1). Dabei wurde festgestellt, dass sich die beiden Gruppen in der Blickdauer zu nahezu allen AOIs unterscheiden. Während die Joint-Engagement-Gruppe vor allem das Zielobjekt länger anschaute als die Self-Exploring-Gruppe, fixierte die Self-Exploring-Gruppe die Experimentatorin, den salienten Ablenker und alle andere Richtungen länger im Vergleich zur Joint-Engagement-Gruppe. Die Ergebnisse der Spielphase stützen die Befunde von Yu & Smith (2016) und Wass, Clackson et al. (2018), dass die Aufmerksamkeitsdauer zu einem Zielobjekt durch gemeinsame Handlungen expandiert werden kann.

Als nächstes wurde das Blickverhalten während der Eye-Tracking-Testphasen ausgewertet, um auch dieses auf gruppenabhängige Unterschiede zu überprüfen (vgl. Kapitel 4.3.5.2). Es stellte sich heraus, dass während der Original-Label-Onset-Phase in beiden Gruppen eine visuelle Präferenz für das saliente Objekt zu beobachten war. Die anschließende Auswertung der sechs-sekündigen Offset-Phasen ergab, dass die Joint-Engagement-Gruppe in keinem der Tests eine Präferenz für das saliente Objekt zeigte, wohingegen die Self-Exploring-Gruppe das saliente Objekt präferierte.

Da die Joint-Engagement-Gruppe weder das saliente Objekt noch das Zielobjekt länger anschaute, bestand die Vermutung, dass die Versuchsteilnehmer bereits in den ersten Sekunden eine Referentenauswahl getroffen haben und sich danach visuell umorientieren. Aus diesem Grund wurden die Offset-Zeitfenster gekürzt und die Blickdaten noch einmal ausgewertet. Nachdem statt der sechs, lediglich die ersten drei Sekunden der Offset-Phasen ausgewertet wurden, zeigte sich in der Joint-Engagement-Gruppe weiterhin keine längere Blickdauer zu einem der Objekte im Vergleich zum Wahrscheinlichkeitswert. In der Self-

Exploring-Gruppe war, wie schon in der Auswertung davor, eine visuelle Präferenz für das saliente Objekt festzustellen. Die längere Blickdauer der Self-Exploring-Gruppe zum salienten Objekt wurde als Salienz-Effekt und nicht als Referentenauswahl gedeutet, da während des Trainings das Zielobjekt am längsten angeschaut und der saliente Ablenker ausgeblendet wurde. Somit kann ein „mismatching“ (Pruden, 2006, S. 274), also die Verknüpfung des Zielwortes mit dem falschen Objekt, ausgeschlossen werden.

Eine weitere Messgröße, die in der Eye-Tracking-Datenauswertung genutzt wurde, war die Zeit bis zu der ersten Fixation während den Offset-Testphasen. Aber auch hier konnten keine signifikanten Unterschiede in der Schnelligkeit der visuellen Orientierung zu einem der beiden Objekte in der Joint-Engagement-Gruppe festgestellt werden. Die Self-Exploring-Gruppe orientierte sich zum Ablenker schneller als zum Zielobjekt.

Die Ergebnisse der Joint-Engagement-Gruppe werden dahingehend interpretiert, dass die Probanden aufgrund des interaktiven Lernformats dem Salienz-Effekt entgegenwirken konnten, da sie weder im sechs-sekündigen, noch im drei-sekündigen Fenster oder in der First-Fixation-Auswertung eine visuelle Präferenz für das saliente Objekt zeigten. Angesichts der Eye-Tracking-Ergebnisse der Self-Exploring-Gruppe kann geschlussfolgert werden, dass sie von dem sozio-pragmatischen Lernformat, in dem das Zielobjekt mehrmals benannt und ostensiv hervorgehoben wurde, profitiert hat, da sie die Wort-Objekt-Verbindung nicht im Eye-Tracking-Test rezeptiv abrufen konnte und stattdessen das saliente Objekt länger fixierte. Schlussfolgernd wird die sozio-pragmatische Wortlerntheorie für 10 Monate alte Kinder nicht unterstützt.

Abhängig davon, ob die Versuchsteilnehmer die zuvor enkodierte Wort-Objekt-Verknüpfung beibehalten haben, sollte anhand des New-Label-Tests untersucht werden, ob 10 Monate alte Kinder vorhandenes Wortwissen als Prinzip der Ausschließbarkeit als Enkodierungsstrategie nutzen können. Wie zuvor beschrieben, konnte weder für die sechs-, noch für die drei-sekündigen Offset-Testphasen eine im Vergleich zum Wahrscheinlichkeitswert längere Blickdauer zum Zielobjekt in der Joint-Engagement-Gruppe festgestellt werden. Angesichts dieser Ergebnisse ist unklar, ob die Joint-Engagement-Gruppe die Wort-Objekt-Verknüpfung beibehalten hat, sodass auch keine Rückschlüsse über das Prinzip der Ausschließbarkeit gezogen werden können. In der Self-Exploring-Gruppe war eine visuelle Präferenz im New-Label-Test zu verzeichnen, die allerdings, wie zuvor argumentiert, auf den Salienz-Effekt und nicht auf eine bewusste Referentenauswahl zurückgeführt wird.

Nachdem Unterschiede im Blickverhalten sowohl während der Spielphase als auch während der Eye-Tracking-Testphasen festgestellt wurden, wurde die dritte Hypothese

hinsichtlich des Zusammenhangs dieser Variablen überprüft. Zunächst wurde die Dauer der visuellen Aufmerksamkeit zum Zielobjekt als Erklärung für die unterschiedlichen Eye-Tracking-Daten der beiden Gruppen untersucht. Doch die Auswertungen ergaben, dass die Länge der Blickzeit zum Zielobjekt bzw. zum salienten Objekt während der Spielphase nicht mit der Blickdauer zum Zielobjekt bzw. salienten Objekt während des Original-Label-Offset-Tests zusammenhängt.

Als nächstes wurden die Handlungsstrategien der Probanden als alternative Erklärung der Eye-Tracking-Daten hinzugezogen. Es hat sich herausgestellt, dass die Joint-Engagement-Gruppe in der Interaktion deutlich mehr Zielhandlungen mit den Objekten ausführt als die Self-Exploring-Gruppe, wohingegen die Self-Exploring-Gruppe häufiger andere Objekthandlungen als die Zielhandlung verwendet. Allerdings stand die Anzahl der Ausführungen der Zielfunktion nicht in Zusammenhang mit der Blickdauer zum Zielobjekt bzw. zum salienten Objekt im Original-Label-Offset-Test.

Als weitere Erklärung der unterschiedlichen Eye-Tracking-Daten wurde die von den Probanden durchgeführte Anzahl der Zielfunktionen betrachtet. Da die Probanden der Joint-Engagement- und der Self-Exploring-Gruppe innerhalb ihrer Gruppen unterschiedlich erfolgreich in der Ausführung der Zielfunktionen waren, wurden sie in die Subgruppen der Achiever und Low-Achiever aufgeteilt. Bei einem Vergleich der Subgruppen hinsichtlich der Blickdauer im Eye-Tracking-Test, zeigten sich jedoch keine Unterschiede.

Aus den Ergebnissen ist zu schlussfolgern, dass die dritte Hypothese nicht bestätigt werden konnte, da das unterschiedliche Blickverhalten im Test weder durch die Blickdauer während des Trainings, noch durch die Handlungsstrategien erklärt werden konnte. Dies ist ein wesentlicher Punkt, der die interaktionistische Sichtweise verdeutlicht, dass die visuelle Aufmerksamkeit und die Aufmerksamkeitsdauer zwar zur Enkodierung notwendig, aber nicht ausreichend sind, um dem Salienz-Effekt im Test entgegenzuwirken oder die enkodierte Information sogar abzurufen. Erst durch das interaktive Verhalten zwischen der Experimentatorin und dem Kind, indem sie gemeinsam eine zielführende Handlung ausführten, erfahren die Kinder die Relevanz des Zielobjektes und messen der Objektsalienz weniger Bedeutung bei. In Bezug auf die Gedächtnisprozesse würde es bedeuten, dass die geteilte Aufmerksamkeit, eine schwächere Gedächtnisspur erzeugt im Vergleich zu den geteilten Handlungen.

Abschließend wurden weiterführende Analysen durchgeführt, die die aufgestellten Hypothesen der vorliegenden Arbeit nicht betrafen. Es wurde ein Zusammenhang zwischen

einigen Formen des sozialen Verhaltens der Experimentatorin und dem Verhalten der Probanden während der Spielphase festgestellt (vgl. Kapitel 4.3.6.1). Die Korrelation der Verweigerung der Probanden und der ostensiven Signale kann dahingehend interpretiert werden, dass auf ein verweigerndes Verhalten in der Joint-Engagement-Gruppe ein ostensives Signal von der Experimentatorin eingesetzt wurde, um die Aufmerksamkeit des Probanden wieder auf das Geschehen zu lenken. Die Korrelation der Anzahl des positiven Feedbacks und den Aufforderungen, die seitens der Experimentatorin kamen, konnte lediglich in der Self-Exploring-Gruppe festgestellt werden. Dies wird damit erklärt, dass aufgrund des schnellen Tempos in der Ausführung der Zielfunktionen, die Experimentatorin in der Joint-Engagement-Gruppe nicht immer mit einem positiven Feedback und einer Aufforderung reagieren konnte. Die Self-Exploring-Gruppe erreichte hingegen eine Zielfunktion höchstens einmal, sodass diese Handlung nicht nur gelobt, sondern auch mit einer Aufforderung kommentiert wurde.

Außerdem wurde der Aufmerksamkeitsverlauf der Probanden in den Eye-Tracking-Tests kontrolliert und dabei festgestellt, dass die Aufmerksamkeitsspanne zum Ende hin abnimmt (vgl. Kapitel 4.3.6.2). Als ein Nebebefund basierend auf den Eye-Tracking-Daten stellte sich heraus, dass die Datenqualität, gemessen anhand der Weighted Gaze Samples, zum Ende der Eye-Tracking-Tests abnahm. Da alle notwendigen technischen Vorkehrungen getroffen waren, werden die niedrigen WGS mit der sinkenden Aufmerksamkeit der Kinder erklärt.

Nach der Zusammenfassung und Interpretation der bisherigen Ergebnisse soll der Erkenntnisgewinn aus dem ersten Experiment der vorliegenden Arbeit für die bestehende Forschung dargelegt werden. Die vorgestellten Ergebnisse haben gezeigt, dass die perzeptuelle Objektsalienz für Kinder aufmerksamkeitslenkend ist, aber soziale Hinweise der Objektsalienz unter bestimmten Bedingungen vorgezogen werden, auch dann, wenn diese gleichzeitig präsent sind. Hierfür müssen die sozialen Signale ostensiv hervorgehoben werden, damit sie von Kindern als Enkodierungsstrategie genutzt werden. Daher ist an dem ECM (Hollich et al., 2000) zu kritisieren, dass es kein altersabhängiges Stufenmodell ist, laut dem Kinder zunächst perzeptionsgeleitet lernen und später soziale Hinweise als Informationsquelle berücksichtigen. Abhängig von dem gegebenen Kontext können Kinder vor dem ersten Lebensjahr sowohl die perzeptuelle Salienz als auch sozialen Hinweise als Enkodierungsstrategie nutzen.

Aus der Blickdatenauswertung der Testphase ging jedoch hervor, dass die Enkodierung der sozio-pragmatischen Hinweise, in Form einer ostensiven Benennungsphase und

einer eigenen Objektexploration, den Kindern nicht ausreichte, um die Wort-Objekt-Verknüpfung beizubehalten und im Eye-Tracking-Test erfolgreich abzurufen. Die Experimentator-Kind-Interaktion erzielte hingegen, dass die Probanden dem Salienz-Effekt im Eye-Tracking-Test entgegensteuern konnten. Dies wird dahingehend erklärt, dass in gemeinsamen, zielführenden Handlungen die Kinder die Relevanz eines Objekts erfahren und dadurch eine tiefere Gedächtnisspur im Gedächtnis entsteht als bei alleinigen Objekthandlungen. Eine weiterführende und ausführlichere Diskussion hinsichtlich der Enkodierungsstrategien (vgl. Kapitel 7.1) und der Behaltensleistung der Wort-Objekt-Verknüpfung (vgl. Kapitel 7.2) findet in der Gesamtdiskussion statt.

Abschließend ist als Kritikpunkt am ersten Experiment die Länge der Eye-Tracking-Videoclips zu erwähnen, die möglicherweise Einfluss auf die Datenqualität (WGS) gehabt haben, da sich in den weiterführenden Analysen herausstellte, dass die WGS hin zum letzten Eye-Tracking-Test abgenommen haben (vgl. Kapitel 4.3.6.2). Daraus folgen die Bedenken, ob sich die langen Videoclips negativ auf die Aufmerksamkeitsspanne der Eye-Tracking-Daten ausgewirkt haben und sich womöglich ein anderes Blickverhalten in den Tests ergeben hätte, wenn die Videoclips kürzer gewesen wären. Dies wirft die Frage auf, ob der Joint-Engagement-Gruppe der Abruf der Wort-Objekt-Verknüpfung doch gelungen wäre, wenn sich ihre Aufmerksamkeitsspanne aufgrund der langen Videoclips nicht reduziert hätte.

4.5 Ableitung der Hypothese für ein Kontrollexperiment

In den Auswertungen des ersten Experiments der Studie wurde beobachtet, dass die Aufmerksamkeitsspanne der Probanden zu den Videos zum Ende des Experiments gesunken ist (vgl. Kapitel 4.3.6.2). Dies kann damit verbunden sein, dass die präsentierten Videoclips zu lang waren und die Kinder sich schnell gelangweilt fühlten. Daher besteht die Vermutung, dass die Ergebnisse der Gesamtblickdauer aufgrund der sinkenden Aufmerksamkeit zum Bildschirm im Hinblick auf den Abruf des rezeptiven Wortwissens nicht aussagekräftig sind.

Aus diesem Grund besteht der Bedarf ein Kontrollexperiment durchzuführen, um die Aussagekraft der Befunde des ersten Experiments zu überprüfen. Für das Kontrollexperiment wird die Hypothese aufgestellt, dass nach einer Kürzung der Videos die Aufmerksamkeit der Kinder konstant bleiben wird. Des Weiteren wird der Frage nachgegangen, ob dadurch auch ein anderes Blickverhalten während der Eye-Tracking-Tests zu beobachten sein wird. Hierfür wird die Blickdauer der beiden Gruppen zu den Objekten nach denselben statistischen Verfahren, wie im ersten Experiment, noch einmal ausgewertet.

5 Stabilität der Aufmerksamkeitsspanne – Kontrollexperiment 1

In den folgenden Unterkapiteln wird zunächst auf das methodische Vorgehen des Kontrollexperiments eingegangen (vgl. Kapitel 5.1), wobei die Unterschiede zu dem ersten Experiment explizit beschrieben werden. Da in dem Ablauf der Trainingsphase nichts verändert wurde, wurden für das Kontrollexperiment lediglich die Eye-Tracking-Daten ausgewertet.

5.1 Methode

Die Methode des Kontrollexperiments unterscheidet sich hinsichtlich der Stichprobe sowie der Testphase zum ersten Experiment der Studie der vorliegenden Arbeit. Für das Kontrollexperiment 1 wurden neue Versuchsteilnehmer rekrutiert, die zuvor nicht an der Studie teilgenommen haben. Vor Beginn der Erhebung wurde die Stichprobengröße auf mindestens 12 Versuchsteilnehmer festgelegt, damit jeder Gruppe jeweils sechs Kinder zugeteilt werden können. Von insgesamt getesteten 14 Versuchsteilnehmern wurden zwei ausgeschlossen (für eine ausführliche Beschreibung der Ausschlusskriterien vgl. Kapitel 4.2.1.3), sodass die Stichprobe schlussendlich 12 Probanden umfasste ($M_{Alter} = 308.33$ Tage, $SD = 4.42$; $N_{weiblich} = 9$, $N_{männlich} = 3$), von denen sechs in der Joint-Engagement-Gruppe ($M_{Alter} = 310$ Tage, $SD = 5.48$; $N_{weiblich} = 4$, $N_{männlich} = 2$) und sechs in der Self-Exploring-Gruppe ($M_{Alter} = 306.67$ Tage, $SD = 2.5$; $N_{weiblich} = 5$, $N_{männlich} = 1$) partizipierten. Ein Versuchsteilnehmer wurde ausgeschlossen, weil zwei WGS-Werte unter dem Mindestwert von 33 % lagen. Die Daten des zweiten ausgeschlossenen Versuchsteilnehmers wurden nicht für weitere Analysen hinzugezogen, da aufgrund des unruhigen Verhaltens weder die Trainingsphase noch die Eye-Tracking-Phase nach Plan durchgeführt werden konnte.

Laut der Fragebögen (vgl. Anhang D), die die Erziehungsberechtigten ausgefüllt haben, haben alle Probanden ein normal ausgeprägtes Seh- und Hörverständnis und keines der getesteten Kinder kam zu früh auf die Welt. Die Mehrheit der Erziehungsberechtigten der Probanden haben als ihren höchsten Bildungsabschluss den Hochschulabschluss angegeben (vgl. Abbildung 9, S. 144).

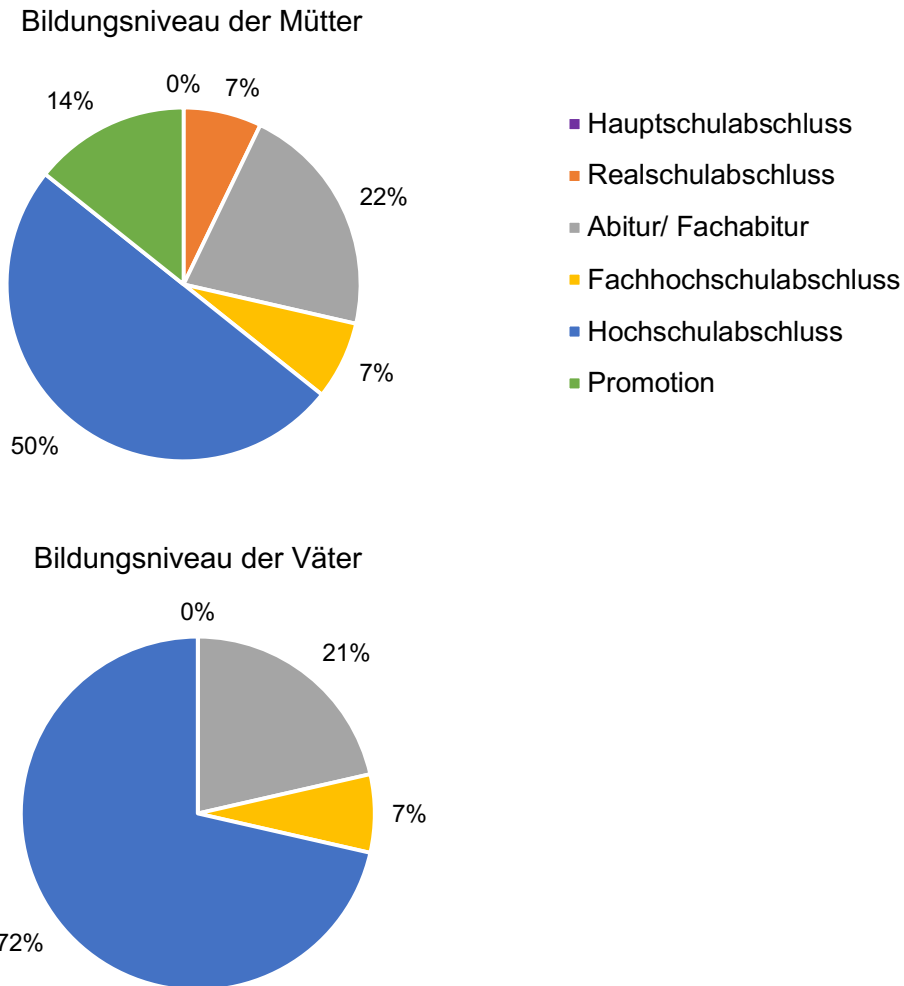


Abbildung 9. Höchster erreichter Bildungsabschluss der Erziehungsberechtigten (Kontroll-experiment 1).

Der Ablauf des Kontroll-experiments war derselbe wie in dem ersten Experiment der vorliegenden Arbeit. Der wesentliche Unterschied bestand in der Modifizierung des Eye-Tracking-Videos, indem die Gesamtlänge der Videoclips sowohl der Salienzphasen als auch der Testphasen von neun auf sechs Sekunden gekürzt wurde. In Experiment 1 war die Onset-Testphase drei Sekunden und die Offset-Phasen etwa sechs Sekunden lang. Aufgrund der Befunde hinsichtlich der abnehmenden WGS-Werte in dem ersten Experiment, wurden die Offset-Zeitfenster für das Kontroll-experiment von sechs Sekunden auf drei Sekunden gekürzt.

5.2 Ergebnisse der Eye-Tracking-Daten

Die Eye-Tracking-Daten des Kontrollexperiments wurden denselben vorbereitenden Datenanalysen unterzogen wie die Daten des ersten Experiments. Das bedeutet, dass die fehlenden Werte mithilfe der Mittelwertimputation ergänzt wurden, die Blickdaten in den einzelnen Phasen gemittelt und in anteilige Prozentwerte umgerechnet wurden. Da in dem Kontrollexperiment die Stichprobengröße < 30 ist, wurde stets der Shapiro-Wilk-Test auf Normalverteilung vor den weiteren statistischen Testverfahren durchgeführt.

5.2.1 Aufmerksamkeitsverlauf – Weighted Gaze Samples

Zunächst wurde die Hypothese, dass die Kürzung der Videolänge der Aufmerksamkeitsreduktion entgegenwirkt, anhand der WGS überprüft (vgl. Tabelle 19). Der Shapiro-Wilk-Test auf Normalverteilung der WGS-Werte ergab, dass bis auf die Werte in der zweiten Salienzphase, die Prozentwerte der restlichen drei Phasen (S1, T1 und T2) normalverteilt waren.

Aufgrund der Mehrheit an normalverteilten Daten wurde das parametrische Testverfahren, in Form einer ANOVA-Messwiederholung durchgeführt. Zunächst wurde überprüft, dass die Sphärizität der Daten gegeben war ($p > .05$) und dass die Homogenität der Fehlervarianzen gemäß des Levene-Tests für die Variablen erfüllt war ($p > .05$). Da beides zutraf, wurden im nächsten Schritt mögliche Interaktionseffekte untersucht. Eine Interaktion würde bedeuten, dass sich die Variable *WGS* in Abhängigkeit von der Variable *Gruppenzugehörigkeit* (Joint-Engagement- vs. Self-Exploring-Gruppe) unterscheidet. Allerdings zeigt die Auswertung, dass es weder einen Interaktionseffekt zwischen den WGS und den beiden Gruppen, $F(3, 24.6) = 1.43, p = .26, \eta^2 = .13$, noch einen einfachen Haupteffekt der WGS gibt, $F(3, 24.6) = 1.1, p = .37, \eta^2 = .1$. Das heißt, dass die WGS nicht abgenommen haben, wodurch die Hypothese bestätigt werden konnte.

Tabelle 19

Weighted Gaze Samples aller Versuchsteilnehmer ($N = 12$) in den Eye-Tracking-Tests

Test	Weighted Gaze Samples	
	<i>M</i>	<i>SD</i>
Salienzphase 1	77.75	17.34
Testphase 1	74.83	13.74
Salienzphase 2	72.83	19.06
Testphase 2	71.17	13.42

5.2.2 Blickverhalten während der Salienzphase

Da dem Shapiro-Wilk Test nach die Blickdaten der Probanden ($N = 12$) während der Salienzphase normalverteilt waren, wurde wie in dem ersten Experiment ein Einstichproben- t -Test durchgeführt, um die visuelle Präferenz der Probanden zu messen. Die relativen Blickzeiten zum salienten und zum Zielobjekt wurden gegen den Wahrscheinlichkeitswert von 50 Prozent getestet. Die Auswertung der Salienzphase ergab, dass die Blickzeit aller Probanden ($N = 12$) zum Ablenker ($M = 53.02$, $SD = 16.69$) sich nicht signifikant zum Wahrscheinlichkeitswert von 50 % unterscheidet, $t(11) = .63$, $p = .54$, $r = .19$. Im Gegensatz zu dem ersten Experiment ist im ersten Kontrollexperiment keine visuelle Präferenz zum salienten Objekt (Ablenker) bei den Probanden zu beobachten. Auf diesen Punkt soll in der Zwischendiskussion (vgl. Kapitel 5.3) eingegangen werden.

5.2.3 Blickverhalten während der Testphasen

Da sich die Aufmerksamkeitsspanne der Probanden während der Eye-Tracking-Tests nicht signifikant reduziert hat, wurde ihr Blickverhalten weiterführend untersucht. Vorab wurde der Shapiro-Wilk-Test auf Normalverteilung durchgeführt, der eine Normalverteilung der Blickdaten während der Onset-Phasen der Joint-Engagement-Gruppe zeigte. Bei den Werten der Self-Exploring-Gruppe lag, bis auf die Recovery-Onset-Phase, ebenfalls eine Normalverteilung vor. Die Blickdaten der Offset-Testphasen waren in beiden Gruppen normalverteilt. Da die Mehrheit der Blickdaten normalverteilt war, wurden im weiteren Verfahren parametrische Tests durchgeführt (vgl. Tabelle 20, S. 147). Im Folgenden werden zunächst die Ergebnisse der Onset-Testphasen und anschließend die der Offset-Testphasen vorgestellt.

Der Einstichproben- t -Test ergab, dass die Joint-Engagement-Gruppe während der Original-Label-Onset Phase eine visuelle Präferenz für den Ablenker hat ($M = 61.79$, $SD = 4.83$), $t(5) = 5.98$, $p = .002$, $r = .94$. Für den New-Label- und Recovery-Test war dies nicht der Fall. Die Self-Exploring-Gruppe schaute den Ablenker während des New-Label-tests signifikant länger an ($M = 68.13$, $SD = 14.32$), $t(5) = 3.1$, $p = .027$, $r = .81$. Während der Recovery-Onset-Phase ist ein marginaler Effekt hinsichtlich einer längeren Blickdauer zum Ablenker zu beobachten ($M = 76.22$, $SD = 25.78$), $t(5) = 2.49$, $p = .06$, $r = .74$. Hier lässt sich jedoch eine hohe Effektstärke feststellen. Da sich die beiden Gruppen in ihrer Blickdauer zum Ablenker während der New-Label-Onset-Phase unterscheiden, $t(5) = -2.72$, $p = .022$, $r = .77$, wurden die Werte nicht zusammengelegt, sondern gruppenabhängig in der Tabelle 20 (vgl. S. 147) dargestellt.

Tabelle 20

Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye-Tracking-Onset-Testphasen im Vergleich zum Wahrscheinlichkeitswert von 50 %

Test	Blickdauer zum Ablenker									
	Joint Engagement					Self Exploring				
	<i>M</i>	<i>SD</i>	<i>t</i> (5)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (5)	<i>p</i>	<i>r</i>
Original Label	61.79	4.83	5.98	.002 **	.94	.55.44	12.12	1.1	.322	.44
New Label	47.94	11.18	- 0.45	.67	.19	68.13	14.32	3.1	.027 *	.81
Recovery	60.79	13.31	1.99	.104	.66	76.22	25.78	2.49	.055 †	.74

In der Analyse der prozentualen Gesamtblickdauer der Eye-Tracking-Offset-Testphasen (vgl. Tabelle 21) konnte festgestellt werden, dass die Joint-Engagement-Gruppe während der New-Label-Offset-Phase eine signifikant längere Blickzeit zum Ablenker ($M = 45.57$, $SD = 16.46$) im Vergleich zum Wahrscheinlichkeitswert zeigte, $t(5) = 3.1$, $p < .001$, $r = .81$. Die Self-Exploring-Gruppe schaute den Ablenker ($M = 56.58$, $SD = 7.66$) während der Original-Label-Offset-Phase marginal länger an, $t(5) = 2.1$, $p = .089$, $r = .69$. Trotz des marginalen Effekts, ist hier auf die hohe Effektstärke hinzuweisen. In der New-Label-Offset-Phase wurde eine signifikant längere Blickdauer zum Ablenker ($M = 75.24$, $SD = 21.68$) festgestellt, $t(5) = 2.85$, $p < .036$, $r = .79$. Auch in der Recovery-Phase ist eine signifikant höhere Blickdauer im Vergleich zum Wahrscheinlichkeitswert vorzufinden, $t(5) = 6.78$, $p < .001$, $r = .95$.

Tabelle 21

Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye-Tracking-Offset-Testphasen im Vergleich zum Wahrscheinlichkeitswert von 50 %

Test	Blickdauer zum Ablenker				
	<i>M</i>	<i>SD</i>	<i>t</i> (5)	<i>p</i>	<i>r</i>
Joint Engagement					
Original Label	45.57	16.46	1.1	.539	.44
New Label	74.98	7.71	3.1	<.001***	.81
Recovery	62.35	26.53	2.49	.306	.74
Self Exploring					
Original Label	56.58	7.66	2.10	.089†	.69
New Label	75.24	21.68	2.85	.036*	.79
Recovery	85.59	12.86	6.782	<.001***	.95

5.3 Zwischendiskussion

Abschließend sollen die Ergebnisse des Kontrollperiments zusammengefasst und mit den Ergebnissen des ersten Experiments verglichen werden. Das Ziel des Kontrollversuchs bestand in erster Linie darin, die Aufmerksamkeitsspanne der 10 Monate alten Kinder in der Eye-Tracking-Untersuchung aufrechtzuhalten. Es wurde die Hypothese aufgestellt, dass dies durch eine Kürzung der Videoclips erreicht wird und konnte angesichts der beschriebenen Ergebnisse (vgl. Kapitel 4.3.6.2) bestätigt werden.

Während im ersten Experiment festgestellt wurde, dass die Aufmerksamkeit der Probanden abnahm, konnte durch eine Kürzung der Videos in dem zweiten Experiment eine Stabilität in der Aufmerksamkeitsspanne erreicht werden. Dieser Befund verdeutlicht, dass die Länge der Videoclips im ersten Experiment dazu führte, dass Aufmerksamkeitsspanne der jungen Probanden gesunken ist. Um der Fragestellung nachzugehen, ob sich durch die Aufrechterhaltung der Aufmerksamkeit zum Bildschirm, das Blickverhalten zu den Stimuli verändert hat, wurde das Blickverhalten der Kinder in den Eye-Tracking-Phasen des zweiten Experiments nochmal überprüft.

Im Vergleich zum ersten Experiment stellte sich in den Auswertungen des Kontrollperiments heraus, dass die Probanden in der Salienzphase keine visuelle Präferenz für das saliente Objekt zeigten. Doch die Blickdaten beider Experimente wiesen Ähnlichkeiten in den Onset- und Offset-Testphasen auf. Wie schon im ersten Experiment, zeigten auch die Probanden des zweiten Experiments eine längere Blickdauer zum salienten Objekt in einigen der Onset-Phasen. Hinsichtlich der Offset-Phasen fiel auf, dass die Self-Exploring-Gruppe des ersten und zweiten Experiments in allen drei Testphasen eine deutliche Präferenz für das saliente Objekt zeigt. In der Joint-Engagement-Gruppe fanden sich im ersten Experiment keinerlei Unterschiede in der Blickdauer, aber im Kontrollexperiment ergab sich eine signifikant längere Fixationsdauer zum Ablenker im Vergleich zum Wahrscheinlichkeitswert in der New-Label-Offset-Phase.

Als Erstes soll auf die Interpretation des Blickverhaltens während der Salienzphase eingegangen werden, da in dem Kontrollexperiment kein Salienz-Effekt festgestellt werden konnte. Dieses Ergebnis wird auf die kleine Stichprobengröße zurückgeführt, da im ersten Experiment bei einer Stichprobengröße von 66 Probanden eine visuelle Präferenz für die Objektsalienz bestand. Der Salienz-Effekt zeigte sich jedoch bei beiden Gruppen in den Onset-Testphasen des Kontrollperiments.

Aus den Ergebnissen der Fixationszeiten in den Offset-Testphasen kann wie bereits im ersten Experiment auch für das Kontrollexperiment geschlussfolgert werden, dass die Self-Exploring-Gruppe die enkodierte Wort-Objekt-Verknüpfung nicht beibehalten konnte und ihre Aufmerksamkeit wieder von der perzeptuellen Objektsalienz bottom-up-gesteuert war. Für die Joint-Engagement-Gruppe kann zusammengefasst werden, dass sie in der Original-Label- und Recovery-Phase dem Salienz-Effekt entgegenwirkte und für keines der beiden Objekte eine visuelle Präferenz zeigte. Damit unterstützen diese Ergebnisse die Befunde des ersten Experiments. Im Kontrollexperiment ergab sich allerdings eine längere Blickdauer der Joint-Engagement-Gruppe zum salienten Objekt während der New-Label-Phase. Auf der einen Seite könnte dies für die Enkodierungsstrategie des Ausschließbarkeitsprinzips sprechen, welches auf bestehendem Wortwissen beruht. Auf der anderen Seite sollte diese Interpretation mit Vorsicht betrachtet werden, da die Probanden weder in dem Original-Label- noch in dem Recovery-Test ein Wortverständnis durch eine längere Blickdauer zum Zielobjekt gezeigt haben.

Als Kritikpunkt der Kontrollstudie ist die kleine Stichprobengröße aufzuführen, weswegen die Vermutung besteht, dass in der Salienzphase kein signifikanter Salienz-Effekt zu beobachten war. Des Weiteren wurde in den Ergebnissen der Offset-Testphasen wieder der Salienz-Effekt in der Self-Exploring-Gruppe ersichtlich. Bezüglich des Blickverhaltens der Joint-Engagement-Gruppe blieben jedoch weiterhin einige Fragen offen, die sich bereits aus den Ergebnissen des ersten Experiments der vorliegenden Arbeit ergaben: Zeigt die Joint-Engagement-Gruppe während der Original-Label-Offset-Phasen keine visuelle Präferenz für das Zielobjekt, weil sie die Wort-Objekt-Verknüpfung nach dem Training bereits vergessen hat? Doch wenn die enkodierte Verknüpfung vergessen wurde, warum tritt bei ihnen kein Salienz-Effekt wie in der Self-Exploring-Gruppe auf?

Angesichts der offenen Fragen ergab sich der Bedarf im Anschluss an das erste Kontrollexperiment ein weiteres Kontrollexperiment durchzuführen. Um Rückschlüsse über den Abruf der enkodierten Wort-Objekt-Verknüpfung der Joint-Engagement-Gruppe ziehen zu können, wurde im folgenden Kontrollexperiment die Konsolidierung als weiterer Gedächtnisprozess zur Erklärung der Blickdaten hinzugezogen.

6 Konsolidierung der Wort-Objekt-Verknüpfung – Kontrollexperiment 2

Aus den Ergebnissen der vorliegenden Arbeit ergab sich der Bedarf ein weiteres Kontroll-experiment durchzuführen, um das Blickverhalten anhand von Gedächtnisprozessen erklären zu können. Das erste Experiment fokussierte die Enkodierungsstrategien, die von Kindern im Alter von 10 Monaten angewandt werden. Dabei stellte sich heraus, dass sowohl die perzeptuelle Salienz als auch soziale Hinweise die kindliche Aufmerksamkeit auf einen Referenten lenken. Allerdings hatten sozio-pragmatische Hinweise im Vergleich zur Objektsalienz mehr Gewichtung, wenn beide gleichzeitig präsentiert wurden. Aus weiteren Analysen ging hervor, dass mithilfe der sozialen Hinweise zwar die Wort-Objekt-Verknüpfung enkodiert werden konnte, allerdings war die dabei entstandene Gedächtnisspur zu schwach, um die Information im anschließenden Test rezeptiv abzurufen. Während die Self-Exploring-Gruppe die enkodierte Wort-Objekt Verknüpfung nicht beibehalten konnte und eine visuelle Präferenz für das saliente Objekt zeigte, war bei der Joint-Engagement-Gruppe während des Tests kein Salienz-Effekt zu beobachten. Daraus wurde geschlussfolgert, dass aufgrund der gemeinsamen Interaktion in der Joint-Engagement-Gruppe eine tiefere Gedächtnisspur entstanden ist, wodurch die Probanden sich nicht von der Objektsalienz ablenken ließen. Auch ein Kontrollexperiment, bei dem die Videos gekürzt wurden, um die Aufmerksamkeit der Kinder länger aufrecht zu erhalten, unterstütze den beschriebenen Befund.

Hinsichtlich der Ergebnisse der Experimente besteht die Annahme, dass die Joint-Engagement-Gruppe dem Salienz-Effekt zwar entgegenwirken, aber die enkodierte Wort-Objekt-Verknüpfung rezeptiv nicht abrufen konnte, weil der enkodierte Gedächtniseintrag zu diesem Zeitpunkt noch nicht konsolidiert war. Neu enkodierte Informationen, die zunächst im Hippocampus gespeichert sind, werden nach der Konsolidierung in den Neocortex übertragen, wodurch das Wissen verfestigt wird. Für die Konsolidierung spielt vor allem Schlaf eine wesentliche Rolle, da „aufgeschnappte“ Informationen, in Sinne des Fast Mappings, erst nach dem Schlaf erfolgreich abgerufen werden (Axelsson et al., 2018; Henderson et al., 2012, S. 20). Da das bedeutet, dass die Probanden Schlaf benötigen, damit die Wort-Objekt-Verknüpfung im Gedächtnis verfestigt wird, wurden für das zweite Kontrollexperiment zwei Erhebungszeitpunkte festgelegt.

Anlehnend an die Ergebnisse des ersten Experiments und des Kontrollexperiments wurden folgende Annahmen in Bezug auf den ersten Erhebungstermin aufgestellt: In der Salienzphase wird ein Salienz-Effekt zu beobachten sein. Auch für die Onset-Phasen wird angenommen, dass beide Gruppen eine visuelle Präferenz für das saliente Objekt zeigen werden. Zudem wird erwartet, dass die Probanden der Self-Exploring-Gruppe, wie auch schon in den Experimenten zuvor, in den Offset-Testphasen einen Salienz-Effekt zeigen werden, wohingegen die Joint-Engagement-Gruppe keines der beiden Objekten länger anschauen wird.

Für den zweiten Erhebungstermin lautet eine Annahme, dass sich in den Onset-Testphasen wieder ein Salienz-Effekt in beiden Gruppen zeigen wird. Es ist außerdem davon auszugehen, dass die Aufmerksamkeit der Self-Exploring-Gruppe auch in den Offset-Testphasen des zweiten Tages bottom-up-gesteuert sein wird und sie deshalb länger auf das saliente Objekt schauen wird. Für die Joint-Engagement-Gruppe wird die Annahme aufgestellt, dass sie den Gedächtniseintrag während des Schlafs konsolidiert hat und am zweiten Tag das Wortwissen erfolgreich abrufen wird, indem sie eine längere Blickdauer im Eye-Tracking-Test zu den jeweiligen Zielobjekten zeigt.

6.1 Methode

Für das zweite Kontrollexperiment wurden neue Probanden rekrutiert und zwei Erhebungszeitpunkte festgelegt, um deren Wortverständnis an zwei aufeinanderfolgenden Tagen zu überprüfen. Vor Beginn der Erhebung wurde die Stichprobengröße auf mindestens 12 Versuchsteilnehmer festgelegt, damit jeder Gruppe jeweils sechs Kinder zugeteilt werden können. Von den insgesamt 14 getesteten Versuchsteilnehmern wurden zwei Kinder ausgeschlossen, wodurch sich die Anzahl auf 12 Probanden reduzierte ($M_{Alter} = 311.92$, $SD = 16.5$; $N_{weiblich} = 8$, $N_{männlich} = 4$). Von den 12 Probanden wurden sechs der Joint-Engagement-Gruppe ($M_{Alter} = 308.33$, $SD = 6.41$; $N_{weiblich} = 2$, $N_{männlich} = 4$) und sechs der Self-Exploring-Gruppe ($M_{Alter} = 315.5$, $SD = 22.96$; $N_{weiblich} = 6$, $N_{männlich} = 0$) zugewiesen. Ein Versuchsteilnehmer wurde ausgeschlossen, weil dieser zum Zeitpunkt der Erhebung noch keine 10 Monate alt war. Die Daten der zweiten, ausgeschlossenen Person konnten für weitere Analysen nicht verwendet werden, da der zweite Erhebungstermin krankheitsbedingt ausgefallen ist.

Laut der Fragebögen (vgl. Anhang D), die die Erziehungsberechtigten der 12 neuen Probanden ausfüllten, hatten alle Kinder zum Zeitpunkt der Erhebung ein normal ausgeprägtes Seh- und Hörverständnis und keines der getesteten Kinder wurde vor der Vollendung der 37. SSW geboren. Die Mehrheit der Erziehungsberechtigten der Probanden haben einen Hochschulabschluss (vgl. Abbildung 10).

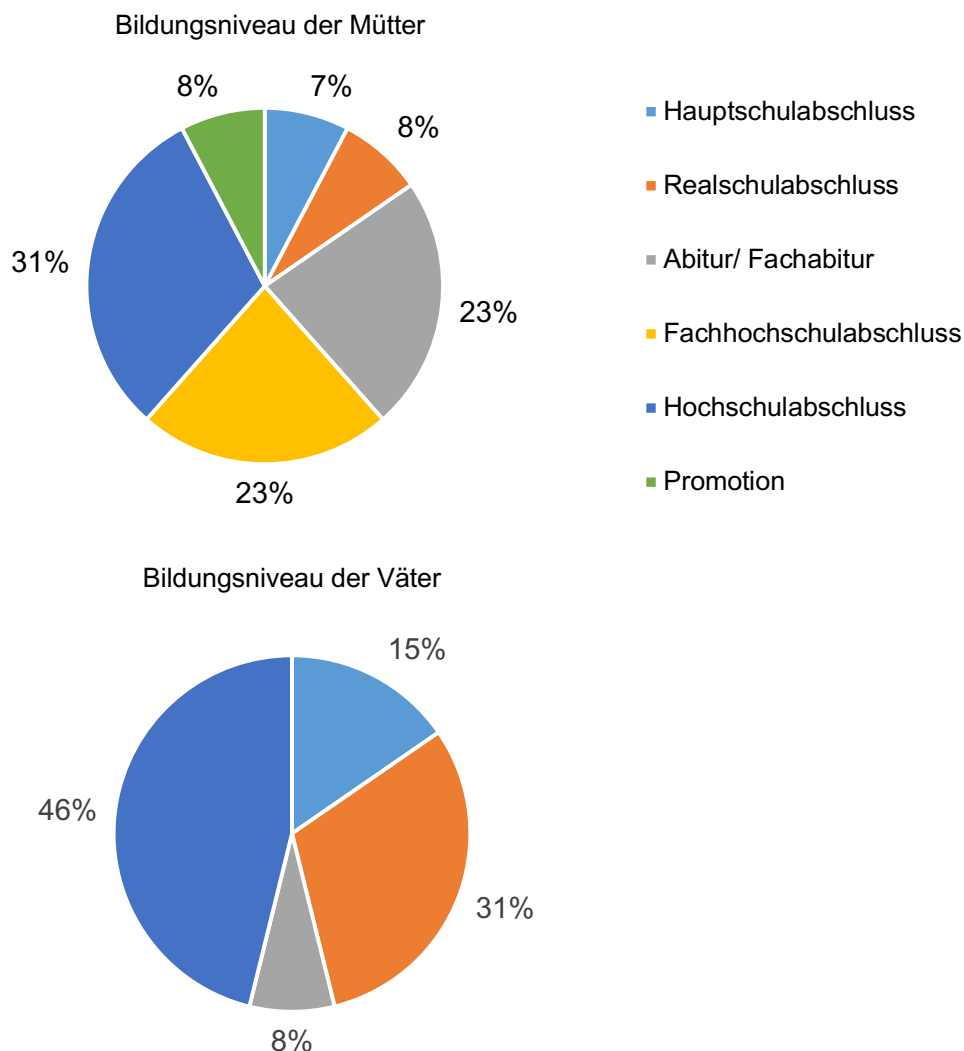


Abbildung 10. Höchster erreichter Bildungsabschluss der Erziehungsberechtigten (Kontrollexperiment 2).

Die Erziehungsberechtigten außerdem wurden gebeten, die Schlafzeiten ihres Kindes bis zum darauffolgenden Erhebungstermin zu protokollieren. Der Termin am nächsten Tag hat, soweit es möglich war, zur selben Uhrzeit wie am ersten Tag stattgefunden. In dieser Zeit schliefen die getesteten 12 Probanden im Durchschnitt 12.96 Stunden ($SD = 1.54$). Am zweiten Erhebungstermin wurden die beiden Testphasen, die das Kind bereits am Tag zuvor

gesehen hat, noch einmal nacheinander präsentiert, um das Wortverständnis erneut zu untersuchen.

Am ersten Erhebungstermin lief das zweite Kontrollexperiment genauso ab wie in den ersten beiden Experimenten, wobei die kürzeren Videoclips des ersten Kontrollperiments genutzt wurden. Daher konnten die Daten des ersten Kontrollperiments und die am ersten Erhebungstermin des zweiten Kontrollperiments erhobenen Daten zusammengeführt werden, um eine größere Stichprobe zu bilden. Aus der Stichprobe des ersten Kontrollperiments ($N = 12$) und der des zweiten Kontrollperiments ($N = 12$) ergab sich somit eine Stichprobengröße von 24 Kindern ($M_{Alter} = 310.13$, $SD = 11.7$; $N_{weiblich} = 17$, $N_{männlich} = 7$), von denen 12 Versuchsteilnehmer in der Joint-Engagement-Gruppe ($M_{Alter} = 309.17$, $SD = 5.51$; $N_{weiblich} = 6$, $N_{männlich} = 6$) und 12 in der Self-Exploring-Gruppe ($M_{Alter} = 311.08$, $SD = 15.55$; $N_{weiblich} = 6$, $N_{männlich} = 1$) teilgenommen haben. Die größere Stichprobe war allerdings nur für die Datenauswertung des ersten Erhebungstermins möglich. Für die Auswertung der Blickdaten, die am zweiten Tag erhoben wurden, lagen nur die Daten der 12 Probanden des zweiten Kontrollperiments vor.

6.2 Ergebnisse der Eye-Tracking-Daten

Auch für diesen Kontrollversuch wurden, wie in den Experimenten davor, dieselben vorbereitenden Datenanalysen sowie statistischen Verfahren angewendet. Das bedeutet, dass die fehlenden Werte mithilfe der Mittelwertimputation ersetzt wurden, die Blickdaten gemittelt und in anteilige Prozentwerte umgerechnet wurden. Im Folgenden werden die Ergebnisse der beiden Tage chronologisch vorgestellt.

6.2.1 Blickverhalten während der Salienzphase – Tag 1

Dem Shapiro-Wilk-Test nach waren Fixationswerte aller Probanden ($N = 24$) normalverteilt. Im Anschluss wurde ein Einstichproben t -Test vorgenommen, bei dem die Blickzeiten gegen den Wahrscheinlichkeitswert von 50 % verglichen wurden. Der Test ergab, dass während der Salienzphase die Probanden den Ablenker ($M = 58.05$, $SD = 16.99$) signifikant länger anschauen im Vergleich zum Wahrscheinlichkeitslevel, $t(23) = 2.32$, $p = .03$, $r = .44$. Aufgrund des signifikanten Unterschieds sowie der mittleren Effektstärke, können die Ergebnisse des ersten Experiments, dass bei den 10 Monate alten Kinder während der Salienzphase ein Salienz-Effekt zu beobachten ist, unterstützt werden.

6.2.2 Blickverhalten während der Testphase – Tag 1

Bei den Daten der einzelnen Tests während der Onset-Phasen lag sowohl für die Joint-Engagement-Gruppe als auch für die Self-Exploring-Gruppe eine Normalverteilung vor. Daher wurde weiterhin der Einstichproben t -Test für die Analysen genutzt. Wie erwartet ergaben sich für beide Gruppen Unterschiede in den Onset-Testphasen (vgl. Tabelle 22, S. 155). Die Joint-Engagement-Gruppe schaute den Ablenker während des Original-Label-Onset-Tests ($M = 58.01$, $SD = 8.66$), $t(11) = 3.21$, $p = .008$, $r = .70$, sowie während des Recovery-Onset-Tests ($M = 64.65$, $SD = 14.09$), $t(11) = 3.6$, $p = .004$, $r = .74$, signifikant länger an im Vergleich zum Wahrscheinlichkeitswert von 50 %. Während des New-Label-Onset-Tests zeigte sich ein marginaler Unterschied in der Blickdauer zum Ablenker ($M = 59.74$, $SD = 16.83$) und dem Wahrscheinlichkeitswert von 50 %, $t(11) = 2.0$, $p = .07$, $r = .52$. Bei der Self-Exploring-Gruppe ergaben sich ebenfalls signifikante Unterschiede zwischen dem Blickverhalten und dem Wahrscheinlichkeitswert: Im New-Label-Onset-Test ($M = 65.74$, $SD = 13.75$), $t(11) = 3.96$, $p = .002$, $r = .77$, sowie in dem Recovery-Onset-Test ($M = 73.65$, $SD = 22.05$), $t(11) = 3.72$, $p = .003$, $r = .75$, lag die Blickdauer zum Ablenker signifikant über dem Level von 50 %.

Während für alle Offset-Testphasen der Self-Exploring-Gruppe eine Normalverteilung vorlag, waren die Blickdaten der Joint-Engagement-Gruppe während der New-Label-Offset-Testphase nicht normalverteilt. Da aber die Mehrheit der Daten normalverteilt war, wurde dennoch für die weitere Datenauswertung der parametrische Einstichproben t -Test durchgeführt. Die Joint-Engagement-Gruppe zeigte während des Original-Label-Offset-Tests keinen Unterschied in der Blickdauer zum Wahrscheinlichkeitswert von 50 %, wohingegen im New-Label-Offset-Test eine signifikant längere Fixationszeit zum salienten Objekt ($M = 68.32$, $SD = 23.53$) vorlag, $t(11) = 2.7$, $p = .021$, $r = .63$. Außerdem konnte eine Tendenz im Recovery-Test festgestellt werden, da der Wert zum Ablenker ($M = 62.99$, $SD = 21.89$) höher war als der Wahrscheinlichkeitswert, $t(11) = 2.06$, $p = .064$, $r = .53$. Anhand einer längeren Blickdauer zum Ablenker ($M = 59.7$, $SD = 9.52$) im Original-Label-Offset-Test, $t(11) = 3.53$, $p = .005$, $r = .73$, einer marginal längeren Fixationszeit des Ablenkens ($M = 64.95$, $SD = 24.25$) im New-Label-Offset-Test, $t(11) = 2.15$, $p = .055$, $r = .54$, sowie auch einer visuellen Präferenz des Ablenkens im Recovery-Test ($M = 74.35$, $SD = 22.81$), $t(11) = 3.7$, $p = .004$, $r = .74$, wurde in der Self-Exploring-Gruppe, wie erwartet, ein Signifikanz-Effekt festgestellt.

Tabelle 22

Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye-Tracking-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 % (1. Erhebungstermin)

Test	Blickdauer zum Ablenker									
	Joint Engagement					Self Exploring				
	<i>M</i>	<i>SD</i>	<i>t</i> (11)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (11)	<i>p</i>	<i>r</i>
Onset-Phase (3 Sekunden)										
Original Label	58.01	8.66	3.21	.008**	.70	54.80	9.02	1.85	.092†	.49
New Label	59.74	16.83	2.00	.07†	.52	65.74	13.75	3.96	.002**	.77
Recovery	64.65	14.09	3.60	.004**	.74	73.65	22.05	3.72	.003**	.75
Offset-Phase (3 Sekunden)										
Original Label	53.29	17.95	0.64	.539	.19	59.70	9.52	3.53	.005**	.73
New Label	68.32	23.53	2.7	.021*	.63	64.95	24.15	2.15	.055†	.54
Recovery	62.99	21.89	2.06	.064†	.53	74.35	22.81	3.7	.004**	.74

6.2.3 Blickverhalten während der Testphase – Tag 2

Im Folgenden werden die Ergebnisse der erhobenen Blickdaten des zweiten Erhebungstermins beschrieben. Dabei wird zunächst auf die Ergebnisse der Onset-Testphasen und anschließend auf die Offset-Testphasen eingegangen.

Als Erstes ist zu erwähnen, dass bei der Joint-Engagement-Gruppe lediglich in der New-Label-Onset-Phase die Voraussetzung auf Normalverteilung verletzt war. Da die Mehrheit der Daten normalverteilt war, wurde weiterhin mit dem parametrischen Einstichproben *t*-Test gerechnet. Den Ergebnissen nach sind keine signifikanten Unterschiede in der Blickdauer zu dem Ablenker und dem Wahrscheinlichkeitswert von 50 % in den Original-Label- und New-Label-Onset-Phasen der Joint-Engagement-Gruppe zu finden (vgl. Tabelle 23, S. 156). Allerdings ist eine marginal längere Blickdauer zum salienten Objekt ($M = 70.67$, $SD = 22.09$) im Vergleich zum Wahrscheinlichkeitswert vorhanden, $t(5) = 2.29$, $p = .071$, $r = .72$. Für die Daten der Self-Exploring-Gruppe lag in allen Onset-Tests eine Normalverteilung der Blickdaten vor. Den Auswertungen nach war in der Self-Exploring-Gruppe während des New-Label-Tests eine signifikant längere Blickdauer zum Distraktor ($M = 74.36$, $SD = 14.55$) festzustellen, $t(5) = 3.71$, $p = .014$, $r = .87$.

In beiden Gruppen bestand in jedem der Offset-Tests eine Normalverteilung der Daten. Weitere Analysen mit dem Einstichproben *t*-Test ergaben, dass die Joint-Engagement-Gruppe eine längere Blickdauer zum Distraktor ($M = 61.68$, $SD = 13.46$) im Vergleich zum Wahrscheinlichkeitswert während des Original-Label-Offset-Tests zeigt,

$t(5) = 2.13$, $p = .087$, $r = .69$, wobei die Differenz zum Wahrscheinlichkeitslevel marginal ist. Bei der Self-Exploring-Gruppe konnte wieder ein Salienz-Effekt aufgrund der signifikant längeren Fixationszeit des salienten Objektes im Vergleich zum Wahrscheinlichkeitswert während des New-Label-Offset-Tests ($M = 76.29$, $SD = 23.66$), $t(5) = 2.72$, $p = .042$, $r = .77$, und des Recovery-Offset-Tests ($M = 72.67$, $SD = 7.37$), $t(5) = 7.53$, $p < .001$, $r = .96$, festgestellt werden.

Tabelle 23

Prozentuale Blickdauer beider Gruppen zum Ablenker während der Eye-Tracking-Tests im Vergleich zum Wahrscheinlichkeitswert von 50 % (2. Erhebungstermin)

Test	Blickdauer zum Ablenker									
	Joint Engagement (N = 6)					Self Exploring (N = 6)				
	M	SD	t(5)	p	r	M	SD	t(5)	p	r
Onset-Phase (3 Sekunden)										
Original Label	51.38	15.25	0.22	.834	.10	60.96	17.89	1.5	.194	.56
New Label	71.78	35.94	1.49	.198	.55	74.36	16.09	3.71	.014*	.87
Recovery	70.67	22.09	2.29	.071†	.72	61.76	14.55	1.98	.105	.66
Offset-Phase (3 Sekunden)										
Original Label	61.68	13.46	2.13	.087†	.69	57.19	10.69	1.65	.161	.59
New Label	41.26	20.89	-1.03	.352	.42	76.29	23.66	2.72	.042*	.77
Recovery	50.96	14.86	0.16	.881	.07	72.67	7.37	7.53	.001***	.96

6.3 Zwischendiskussion

Das Ziel des zweiten Kontrollexperiments bestand darin, den Versuchsteilnehmern die Möglichkeit zu bieten, den Gedächtniseintrag zu konsolidieren und sie am darauffolgenden Tag hinsichtlich des Wortwissens nochmal zu testen. Indem die Daten der Probanden des ersten und zweiten Kontrollexperiments zusammengelegt wurden, konnte die Stichprobengröße für den ersten Erhebungstermin erweitert werden.

Die Ergebnisse des zweiten Kontrollexperiments (1. Erhebungstermin) lassen sich dahingehend zusammenfassen, dass alle Probanden in der Salienzphase das saliente Objekt präferierten und die visuelle Präferenz in den Onset-Testphasen weiterhin zu beobachten war (vgl. Tabelle 22, S. 155). Damit bestätigt die durchgeführte Kontrollstudie den Befund des Salienz-Effekts aus dem ersten Experiment der vorliegenden Studie. Dies bedeutet, dass 10 Monate alte Kinder eine visuelle Präferenz für saliente Objekte zeigen, aber nur dann, wenn die Objekte nicht zeitgleich mit sozialen Hinweisen auftauchen.

Während die Aufmerksamkeit der Self-Exploring-Gruppe im zweiten Kontrollexperiment (1. Erhebungstermin) in allen Testphasen (vgl. Tabelle 22, S. 155) weiterhin von der Objektsalienz gelenkt wird, kann die Joint-Engagement-Gruppe dem Salienz-Effekt im Original-Label-Test entgegenwirken. Die längere Blickdauer der Joint-Engagement-Gruppe zum salienten Objekt im New-Label-Offset-Test wird als visuelle Präferenz für die perzeptuelle Objektsalienz interpretiert. Sie wird nicht auf das Prinzip der Ausschließbarkeit zurückgeführt, da im Original-Label-Test kein Wortverständnis festgestellt werden konnte. Für diese Erklärung steht auch die im Recovery-Offset-Test darauffolgende schwache Präferenz für das saliente Objekt. Aus den Ergebnissen ist zu schlussfolgern, dass in den Testphasen dieses Kontrollexperiments (1. Erhebungstermin) bei der Joint-Engagement-Gruppe ein schwacher Salienz-Effekt festgestellt wurde, der in den vorherigen Experimenten nicht vorhanden war.

Die Auswertung der Blickdaten des zweiten Erhebungstermins betraf die Frage nach der Konsolidierung (vgl. Tabelle 23, S. 156). Zunächst wurde festgestellt, dass in einigen der Onset-Phasen ein Salienz-Effekt in beiden Gruppen auftauchte. In den Offset-Phasen zeigte die Self-Exploring-Gruppe, wie erwartet, eine eindeutige Präferenz für das saliente Objekt. Bei der Joint-Engagement-Gruppe ergab sich in der Original-Label-Offset-Testphase ein schwacher Salienz-Effekt. Die zuvor aufgestellte Annahme hinsichtlich der längeren Blickdauer zum Zielobjekt in der Joint-Engagement-Gruppe konnte daher nicht bestätigt werden. Somit wird in Bezug auf den Konsolidierungsprozess geschlussfolgert, dass sich die Wort-Objekt-Verknüpfung nach dem Schlaf bei der Joint-Engagement-Gruppe nicht verfestigt hat, da am zweiten Erhebungstermin immer noch ein Salienz-Effekt bestand.

Im Folgenden werden mögliche Erklärungen bezüglich des Blickverhaltens der Joint-Engagement-Gruppe diskutiert, da die Frage, warum die Joint-Engagement-Gruppe keine visuelle Präferenz für das Zielobjekt während der Testphasen zeigt, ungeklärt blieb. Eine Erklärung lautet, dass der Enkodierungsprozess während des Trainings nicht ausreichend von der Experimentatorin unterstützt wurde, sodass die entstandene Gedächtnisspur nicht tief genug war, um im Test abgerufen zu werden.

In einer Studie von Craik und Tulvin (1975) mit Erwachsenen wurde festgestellt, dass neue Wörter besser enkodiert und damit auch erfolgreicher abgerufen werden, wenn die Probanden nicht nur oberflächliche, sondern tiefer gehende Fragen in Bezug auf die Wörter beantworten. Die Autoren zogen daraus den Rückschluss, dass die Verarbeitungstiefe einer neuen Information einen Einfluss auf den Abruf hat. Zu ähnlichen Erkenntnissen sind auch Vlach und Sandhofer (2012) in einer Studie mit Dreijährigen gekommen. Sie fanden heraus,

dass unterschiedliche Gedächtnisstützen während der Trainingsphase den Enkodierungsprozess der Wort-Objekt-Verknüpfung unterstützen und eine tiefere Gedächtnisspur hinterlassen als wenige oder gar keine Gedächtnisstützen. Vlach und Sandhofer (2012) fanden außerdem heraus, dass der Vergessensprozess schwach enkodierter Wörter direkt nach dem Training einsetzt und innerhalb der ersten Woche exponentiell abnimmt (vgl. Searleman & Herrmann, 1994, S. 99–100).

Vlach und Sandhofer (2012) gelang es, dem Vergessensprozess bei dreijährigen Kindern entgegenzuwirken, indem die Probanden mit insgesamt drei Gedächtnishilfen während der Trainingsphase unterstützt wurden. Neben der sechsmaligen Wiederholung des Wortes wurde die Objektsalienz erhöht, indem den Versuchsteilnehmern gesagt wurde, dass es ein ganz besonders Objekt sei. Darüber hinaus sollten die Dreijährigen das Zielwort nachsprechen. Die Mehrheit der Probanden, die die drei Gedächtnishilfen erfahren haben, konnten das Wort auch noch nach einen Monat rezeptiv abrufen. Bei denjenigen, die aber lediglich zwei oder nur eine Gedächtnishilfe erhalten haben, setzte der Vergessensprozess unmittelbar nach dem Training ein und nahm stetig ab. Nach einem Monat konnte kaum einer der Probanden das Wort dem Zielobjekt richtig zuordnen. Diese Studie verdeutlicht, dass encodierte Informationen vergessen werden, wenn diese nicht während des Enkodierungsprozesses durch Gedächtnisstützen im Gedächtnis verankert werden.

Anknüpfend an diese Befunde besteht die Annahme, dass in der Studie der vorliegenden Arbeit den Kindern während des Trainings nicht genug Gedächtnisstützen geboten wurden, sodass die Wort-Objekt-Verknüpfung bereits unmittelbar nach dem Training nicht abgerufen werden konnte und am zweiten Erhebungstermin sogar der Vergessensprozess einsetzte. Der Unterschied im Blickverhalten zwischen der Self-Exploring- und der Joint-Engagement-Gruppe besteht darin, dass die Self-Exploring-Gruppe in den Testphasen einen Salienz-Effekt zeigte, dem die Joint-Engagement-Gruppe hingegen entgegenwirken konnte. Aufgrund dieses Befunds ist davon auszugehen, dass beide Gruppen die Wort-Objekt-Verknüpfung nicht gelernt haben, aber die Joint-Engagement-Gruppe die Wort-Objekt-Verknüpfung in der Spielphase durch gemeinsame zielführende Handlungen tiefer encodiert hat als die Self-Exploring-Gruppe.

Im Folgenden werden einige Möglichkeiten vorgestellt, wie eine tiefere Gedächtnisspur bei Kindern unter einem Jahr erzeugt werden kann. Da sich die Studie von Vlach und Sandhofer (2012) auf dreijährige Kinder bezieht und die angewandten Gedächtnisstützen für jüngere Kinder nicht geeignet sind, werden die Gedächtnisstützen aus den Experimenten von Hollich et al. (2000) mit 12 Monate alten Probanden aufgeführt. Allerdings wurden die

Stimuli aus der Studie von Hollich et al. (2000) nicht in ihrer Salienz unterschieden, sodass soziale Hinweise mit der perzeptuellen Salienz nicht gleichzeitig präsent waren. In der Experiment-Reihe von Hollich et al. (2000, Experiment 8) zeigte sich beispielsweise, dass eine 10-malige Benennung des Referenten sich positiv auf die anschließende Referentenauswahl im Test auswirkt. Eine längere Benennungszeit des Objektes ging ebenfalls mit einem erfolgreichen rezeptiven Abruf der Wort-Objekt-Verbindung einher (Experiment 9). Am Beispiel dieser Experimente wird deutlich, dass Hollich und Kollegen (2000) der Benennung des Zielobjekts eine wichtige Rolle für das Wortlernen beimessen.

Vergleicht man das methodische Vorgehen dieser Experimente (Hollich et al., 2000) mit der Studie der vorliegenden Arbeit, wird ersichtlich, dass den 10 Monate alten Probanden der vorliegenden Studie mehr Möglichkeiten zur Objekt-Enkodierung (Beobachtung und Ausführung der Objekt-Zielfunktion) als zur Wort-Enkodierung (fünfmaliges Benennen während der Benennungsphase) geboten wurden. Aus diesem Grund besteht die Annahme, dass die Probanden der vorliegenden Studie mehr Gedächtnisstützen in Bezug auf die Wortform gebraucht hätten, zum Beispiel durch eine häufigere Benennung oder durch eine längere Trainingsphase bzw. Benennungszeit, um nicht nur das Objekt, sondern auch die Wortform erfolgreich enkodieren zu können.

Eine alternative Erklärung des Blickverhaltens im Kontrollexperiment (2. Erhebungstermin) lautet, dass die Probanden einen wiederkehrenden, pragmatischen Rahmen benötigt hätten, um die Wort-Objekt-Verknüpfung vom Vortag abrufen zu können. Im Vergleich zum ersten Erhebungstermin fehlte den Probanden am zweiten Erhebungstermin das soziale Format vor dem Beginn der Testphase. Im Vergleich zum ersten Tag, wurden sie am zweiten Tag in die Testsituation direkt hineingebracht, ohne das soziale Miteinander vorher zu erfahren. Es ist daher möglich, dass die Probanden aufgrund der ausbleibenden sozialen Interaktion keine Referentenauswahl im Test getroffen haben. Auch dieser Punkt könnte in weiterführenden Studien aufgegriffen werden, indem der zweite Erhebungstermin um eine Spielphase mit den Objekt-Sets, die vor dem Eye-Tracking-Test stattfindet, erweitert wird.

7 Gesamtdiskussion

Das Forschungsvorhaben der vorliegenden Dissertationsschrift wurde durch bisherige Forschungsarbeiten motiviert, die im Folgenden zusammengefasst werden sollen. Die bestehenden Theorien bezüglich der frühen Enkodierungsstrategien im Fast-Mapping-Prozess bei Kindern vor dem ersten Lebensjahr sind in der Studie der vorliegenden Arbeit zusammengebracht worden (vgl. Kapitel 3). In den folgenden Unterkapiteln (vgl. Kapitel 7.1 und 7.2) werden die Ergebnisse der Studie der vorliegenden Arbeit auf die zuvor im Theorieteil formulierten Hypothesen (vgl. Kapitel 3.7) bezogen und diskutiert, was die neu gewonnenen Befunde über die Enkodierungsstrategien und die Behaltensleistung von Wort-Objekt-Verknüpfungen für den bisherigen Forschungsstand im Bereich des Spracherwerbs bedeuten. Abschließend werden in dem Unterkapitel 7.3 die Forschungsdesiderate der vorliegenden Studie benannt und ein Ausblick für künftige Studien gegeben.

Im Fokus bisheriger empirischer Studien zum Spracherwerb (z. B. Bergelson & Swingley, 2012, 2015) stand vor allem die Frage, ob Kinder bereits vor ihrem ersten Lebensjahr alltägliche Wörter verstehen. Mithilfe der Blickmessungsmethode wurde aufgedeckt, dass bereits ab dem 6. Lebensmonat ein rudimentäres Wortverständnis vorhanden ist und Wörter nicht nur mit Personen (Tincoff & Jusczyk 1999, 2012), sondern auch mit Objekten (Bergelson & Swingley, 2012, 2015) und Handlungen (Nomikou et al., 2018) assoziiert werden (vgl. Kapitel 2.3). Im Zusammenhang mit diesen Befunden zum frühen Wortverständnis stellt sich die Frage, *wie* Kinder ihre ersten Referenzen erwerben. Damit verbunden sind auch die Fragen, welche Reize Kinder zur Enkodierung einer Wort-Referent-Verknüpfung im Fast-Mapping-Prozess nutzen und ob die durch die entsprechende Enkodierungsstrategie entstandene Verknüpfung beibehalten wird. Für den Bereich der Spracherwerbsforschung ist die Beantwortung dieser Fragen insofern wichtig, um zu verstehen, welche Enkodierungsstrategien Kinder am Anfang im Wortlernprozess nutzen. Diese Erkenntnisse könnten dazu genutzt werden, um Bezugspersonen dafür zu sensibilisieren, auf welche Reize junge Kinder achten und wie sie die Enkodierungsstrategien in der Interaktion mit Kindern verstärken können.

In Bezug auf die Frage, wie erste Referenzen entstehen, liefern eine Reihe an Wortlerntheorien unterschiedliche Erklärungsansätze: Auf der einen Seite steht die Theorie der Hypothesen einschränkenden Prinzipien (vgl. Kapitel 3.2), laut der, z. B. das Prinzip der Ausschließbarkeit die Referentenauswahl, vor der Kinder in einer neuen Wortlernsituation stehen, einschränkt (Halberda, 2003; Mather & Plunkett, 2010, 2012; Markman, 1989,

1990). Dieser theoretische Ansatz hat mit der assoziativen Wortlernlernertheorie (vgl. Kapitel 3.3) gemeinsam, dass das frühe Wortlernen als ein mechanistischer Prozess beschrieben wird, für den einfache, kognitive Prozesse, so wie die selektive Aufmerksamkeit (vgl. Kapitel 2.1.1), verantwortlich sind. Laut dem assoziativen Ansatz hat vor allem die perzeptuelle Objektsalienz (vgl. Kapitel 2.1.2) eine große Bedeutung im Bottom-up gesteuerten Wortlernprozess (vgl. Kapitel 3.3.2), um einen Referenten für ein neues Wort auszuwählen (Hollich et al., 2000; Pruden et al., 2006).

Auf der anderen Seite stehen zwei Theorien, die für das Wortlernen in einer sozialen Interaktion plädieren, aber unterschiedliche Sichtweisen auf die Rolle der Bezugsperson in der Interaktion haben. Laut der sozio-pragmatischen Wortlerntheorie (z. B. Tomasello, 2000) spielt insbesondere die geteilte Aufmerksamkeit (vgl. Kapitel 2.1.1) eine wichtige Rolle. Im Gegensatz zur einfachen, selektiven Aufmerksamkeit, wird die geteilte Aufmerksamkeit in einer Triade koordiniert. Das heißt, dass in der sozial koordinierten Interaktion Kinder nicht nur lernen, auf einen gemeinsamen Referenten zu achten, sondern auch, *über* diesen Referenten *mit* einem Gesprächspartner zu kommunizieren und sich über die Wortbedeutung dieses Referenten auszutauschen.

Die in der vorliegenden Arbeit vertretene interaktionistische Sichtweise (vgl. Kapitel 3.5) geht über den sozio-pragmatischen Ansatz hinaus, denn sie besagt, dass eine Wortbedeutung nicht durch die Interaktion im geteilten Aufmerksamkeitsfokus, sondern durch gemeinsame, zielführende Handlungen in Koordination mit verbalen Handlungen, entsteht (Rohlfing et al., 2016; Wildt et al., 2019). Diese Annahme wird gestützt von bisherigen wahrnehmungspsychologischen Studien (vgl. Kapitel 3.5.1 und 3.5.2), aus denen hervorgeht, dass gemeinsame Handlungen, im Vergleich zu alleinigen Objekthandlungen, mit einer längeren Aufmerksamkeitsspanne in Zusammenhang stehen (z. B. Wass et al., 2018). Obwohl ein Zusammenhang zwischen gemeinsamen Handlungen und der Aufmerksamkeitslenkung und -spanne bei Kindern gezeigt wurde, blieb in der bisherigen Forschung die Rolle der gemeinsamen Objekthandlungen als Enkodierungsstrategie im Kontext des Wortlernens weitestgehend berücksichtigt. Des Weiteren wurde in den Studien auch nicht untersucht, um welche Art von gemeinsamen Handlungen es sich handelte, die die Aufmerksamkeitsspanne der Kinder verlängerte.

Angesichts der unterschiedlichen Theorien bezüglich der Enkodierungsstrategien im Fast-Mapping-Prozess und der aufgeführten Forschungslücke im interaktionistischen Ansatz, wurden die beschriebenen Enkodierungsstrategien sowie die Behaltensleistung von Wort-Objekt-Verknüpfungen in der Studie der vorliegenden Arbeit überprüft. Die aus den

bisherigen Theorien abgeleitete Annahme für die vorliegende Arbeit lautete, dass die Objektsalienz, im Sinne der assoziativen Sichtweise, für Kinder aufmerksamkeitslenkend ist, aber auch soziale Informationen, im Sinne der sozio-pragmatischen Sichtweise, von Kindern zur Wort-Referent-Enkodierung hinzugezogen werden. Allerdings könne eine Wort-Objekt-Verbindung nur dann rezeptiv abgerufen werden, wenn sie zuvor durch gemeinsame, zielführende Handlungen beider Interaktanten enkodiert wurde (vgl. Wildt et al., 2019). Somit wird in der vorliegenden Arbeit der interaktionistische Ansatz vertreten, der in Bezug auf die Ergebnisse dieser Studie noch näher erläutert wird.

7.1 Enkodierungsstrategien im frühen Alter

Das Ziel dieses Unterkapitels besteht darin, die Ergebnisse der Studie der vorliegenden Arbeit zusammenzufassen und in den bestehenden Forschungsstand einzuordnen. Hierfür wird auf die Kritikpunkte der Wortlerntheorien und der jeweiligen Studien hingewiesen und erläutert, wie sie in der vorliegenden Arbeit aufgegriffen wurden. Anschließend wird beschrieben, ob die Ergebnisse der vorliegenden Arbeit die zuvor aufgestellten Hypothesen unterstützen und inwiefern die Befunde die bestehende Forschungsliteratur untermauern oder widerlegen. Zuletzt wird geschlussfolgert, inwieweit die neu gewonnenen Erkenntnisse hinsichtlich der Enkodierungsstrategien von Kindern im frühen Fast-Mapping-Prozess mit den bisherigen Befunden verbunden bzw. ergänzt werden können.

Als Beispiel für die assoziative Wortlerntheorie wurde die Studie von Pruden und Kollegen (2006) ausgewählt. Diese ist repräsentativ für den assoziativen Ansatz, da hier die Rolle der Objektsalienz im Wortlernprozess besonders hervorgehoben wird. In der Studie wurde die Wirkung von sozialen im Vergleich zu perzeptuellen Stimuli auf die Aufmerksamkeit und das Fast Mapping von 10 Monate alten Kindern verglichen. Dabei stellten die Autoren fest, dass in dieser Altersgruppe die Objektsalienz mehr gewichtet wird als soziale Hinweise. Aufgrund dieser visuellen Präferenz leiteten Pruden et al. (2006) her, dass das Wortlernen perzeptionsgeleitet abläuft und sozialen Signalen in der ersten Entwicklungsphase noch keine Bedeutung zugeschrieben wird.

Bezüglich dieser Schlussfolgerung wurde in der Studie der vorliegenden Arbeit überprüft, ob 10-Monatige die perzeptuelle Salienz priorisieren, und ob soziale Hinweise im Wortlernkontext nicht beachtet werden. Die **erste Hypothese** des ersten Experiments der vorliegenden Arbeit betraf die Eye-Tracking-Daten der Salienzphase. Es wurde angenommen, dass sich in der Salienzphase, wie auch in der Studie von Pruden et al. (2006), eine

visuelle Präferenz für das saliente Objekt zeigen wird. Die Ergebnisse der Salienzphasen der vorliegenden Arbeit ergaben, dass Kinder in der Tat saliente Objekte präferieren, die durch ihre Form und Farbe aus dem Kontext hervorstechen. Damit bekräftigen sie die Befunde aus der Salienzphase der Studie von Pruden et al. (2006), dass die selektive Aufmerksamkeit von 10 Monate alten Kindern von der Objektsalienz gelenkt werden kann.

Im Allgemeinen wird an dem assoziativen Ansatz kritisiert, dass den sozialen Signalen, die von einer Bezugsperson zur Aufschlüsselung der Referenz im Wortlernprozess ausgehen, zu wenig Bedeutung beigemessen wird, obwohl Studien belegen, dass Kinder von Anfang an sensibel sind für soziale Signale (z. B. Gredebäck et al., 2018; für eine ausführliche Diskussion der assoziativen Wortlerntheorie vgl. Kapitel 3.6). An der Studie von Pruden und Kollegen (2006) ist kritisch anzumerken, dass in der auf die Salienzphase folgende Trainingsphase, die gebotenen sozialen Hinweise lediglich in Form der Blickrichtung operationalisiert wurden, obwohl aus früheren Studien (Hollich et al., 2000) hervorging, dass die Blickrichtung als referentieller Hinweis nicht ausreicht, um die Aufmerksamkeit von jungen Kindern zu lenken. Dieser Kritikpunkt wurde in der Studie der vorliegenden Arbeit aufgegriffen, indem die Blickrichtung der Experimentatorin während der Benennungsphase, die den ersten Teil der Trainingsphase ausmachte, mit weiteren ostensiven Signalen kombiniert wurde, um die Aufmerksamkeit der Probanden zum Referenten zu lenken (vgl. Kapitel 4.2.2.4).

Ein weiterer Kritikpunkt an der Studie von Pruden und Kollegen (2006) lautet, dass die Autoren lediglich von der Blickdauer zu den Objekten, aber nicht zum Experimentator berichteten, sodass unklar ist, ob die Probanden die wechselnde Blickrichtung überhaupt wahrgenommen haben. Daher wurde in der vorliegenden Arbeit die Blickrichtung der Probanden nicht nur zu den Objekten, sondern auch zur Experimentatorin und zu anderen Richtungen kodiert (vgl. Kapitel 4.2.4.1) und ausgewertet.

Unter Berücksichtigung der Modifizierung der zuvor benannten Punkte betraf die **zweite Hypothese** dieser Arbeit die Trainingsphase und besagte, dass 10-Monatige empfänglich sind für soziale Signale und diesen daher zum Referenten folgen werden, auch dann, wenn gleichzeitig das saliente Objekt als Ablenker präsent ist. Somit bestand die Herausforderung für die Kinder darin, sich während der Benennungsphase im Training nicht vom salienten Objekt ablenken zu lassen und den referentiellen Hinweisen zum Zielobjekt zu folgen. Aus den Auswertungen ging hervor, dass die Probanden während der Benennungsphase die meiste Zeit das Zielobjekt anschauten und somit die zweite Hypothese bestätigt wurde. Die Tatsache, dass das Zielobjekt und die Experimentatorin für den Großteil

der Zeit fixiert und das saliente Objekt, ihre Bezugsperson und andere Dinge im Umfeld kaum beachtet wurden, widerlegt die Aussage von Pruden et al. (2006), dass soziale Hinweise ignoriert werden (S. 274). Die in der Studie von Pruden und Kollegen (2006) festgestellte visuelle Präferenz für die salienten Objekten in den Testphasen ist nicht darauf zurückzuführen, dass soziale Hinweise nicht berücksichtigt werden, sondern dass die Blickrichtung des Experimentators als referenzieller Hinweis nicht ausreichend war, damit die Wort-Referent-Verknüpfung beibehalten wird. Dies hatte zur Folge, dass alle Probanden im Test stets das saliente Objekt länger anschauten. Die Ergebnisse hinsichtlich der Blickdauer während der Benennungsphase (vgl. Kapitel 4.3.4) unterstützt die zweite Hypothese der vorliegenden Arbeit und somit die Befunde sozio-pragmatischer Studien, dass Kinder bereits im Alter von 10 Monaten soziale Hinweise zur Referentenauswahl nutzen (Baldwin et al., 1991; Yurovsky & Frank, 2017).

Im Zusammenhang mit den Ergebnissen aus der Salienz- und der Trainingsphase des ersten Experiments der vorliegenden Studie können die Befunde des assoziativen und des sozio-pragmatischen Ansatzes verbunden und die Schlussfolgerung gezogen werden, dass Kinder hauptsächlich dann bottom-up-gesteuert handeln und die perzeptuelle Objektsalienz als Enkodierungsstrategie anwenden, wenn keine weiteren Informationsquellen zur Verfügung stehen. Sobald aber soziale Hinweise in demselben Kontext auftauchen, nutzen Kinder diese zur Enkodierung der Wort-Objekt-Verknüpfung, auch dann, wenn sie in Konkurrenz mit der Objektsalienz stehen. Damit Kinder den sozialen Hinweisen „Vorrang gewähren“, müssen die Signale ostensiv sein. Zu ostensiven Signalen zählt beispielweise ein wechselseitiger Blickkontakt, kindgerichtete Sprache und ein an das Kind gerichtetes Handeln. Damit lenken Erwachsene nicht nur die Aufmerksamkeit auf sich, sondern zeigen dem Kind, dass es der Adressat ist und ihm eine neue Information vermittelt werden soll (Csibra 2010, S. 147).

Aufgrund der aus der Salienz- und Trainingsphase dieser Arbeit gewonnene Erkenntnis, dass 10 Monate alte Kinder sowohl auf die perzeptuelle Objektsalienz, als auch auf soziale Hinweise achten, sollte das Emergentist Coalition Model nach Hollich und Kollegen (2000) überdacht werden. Bei dem ECM handelt es sich um ein hybrides Modell, demnach Kinder zu unterschiedlichen Zeitpunkten in ihrer Entwicklung perzeptuelle, soziale und linguistische Signale unterschiedlich stark im Wortlernprozess gewichten (vgl. Kapitel 3.3.1, für eine kritische Reflexion des ECM vgl. Kapitel 3.6). Der oben beschriebene Befund der vorliegenden Arbeit bedeutet für das ECM, dass die Beachtung der perzeptuellen Salienz und der sozialen Hinweise kein altersabhängiges Stufenmodell, sondern vielmehr

kontextabhängig ist (vgl. Kapitel 4.4). Die Kontextabhängigkeit der Enkodierungsstrategien könnte mit unterschiedlichen Altersgruppen weiterführend untersucht werden (vgl. Yurovsky & Frank, 2017). Sollte sich in solch einer Untersuchung ergeben, dass ältere Kinder ein ähnliches Blickverhalten wie die 10-Monatigen zeigen, indem sie in der Salienzphase perzeptionsgeleitet handeln, aber in der Trainingsphase den sozialen Hinweisen folgen, dann würde es ein Umdenken in der Sichtweise auf die Gewichtung unterschiedlicher Faktoren im Wortlernprozess erfordern.

Im Gegensatz zur assoziativen Wortlerntheorie spielen im sozio-pragmatischen Ansatz soziale Achtungshinweise eine wesentliche Rolle für die Vermittlung bzw. den Erwerb neuer Referenzen. Zum einen werden soziale Hinweise, beispielsweise in Form ostensiver Signale von Erwachsenen für die Herstellung einer geteilten Aufmerksamkeit (vgl. Kapitel 2.1.1) genutzt, in der die Referenz hergestellt werden soll (Csibra, 2010, S. 145). Zum anderen achten Kindern auf soziale Signale, um nicht nur den Referenten für ein neues Wort zu bestimmen, sondern auch, um falsche Referenzbildungen zu vermeiden (Baldwin, 1991).

Bisherige Studien zeigen eine frühe Sensibilität von Kindern für Augen und Gesichter bereits vor dem ersten Lebensjahr (z. B. D’Entremont et al., 1997; Michel et al., 2019) und sogar einen Zusammenhang mit deren späteren Sprachentwicklung (z. B. Morales et al., 1998). Allerdings wurde die Wirkung von sozialen Hinweisen als Enkodierungsstrategie in einer Wortlernsituation erst bei Kindern ab dem 12. Lebensmonat untersucht (z. B. Yurovsky & Frank, 2017; Booth et al., 2008; Brandone et al., 2007; Hollich et al., 2000). Daher blieb in der Forschung weitestgehend unberücksichtigt, ob Kinder auch schon vor dem ersten Lebensjahr sozio-pragmatische Hinweise (z. B. Blickrichtung oder Gesten), die von einer Bezugsperson ausgehen, nutzen, um im gemeinsamen Aufmerksamkeitsfokus neue Wörter mit dessen Referenten zu verknüpfen und beizubehalten.

Diese Forschungslücke wurde in der Studie der vorliegenden Arbeit aufgegriffen, indem 10 Monate alten Probanden in der Benennungsphase die Möglichkeit zur Enkodierung einer neuen Wort-Objekt-Verbindung mithilfe von sozialen Hinweisen geboten wurde. Wie bereits zuvor erwähnt, ergaben die Auswertungen der Blickdaten während der Trainingsphase, dass bereits im Alter von 10 Monaten soziale Hinweise als Informationsquelle zur Enkodierung genutzt werden. Ob nach der Enkodierung, auch der Abruf der Wort-Referent-Verknüpfung gelungen ist, wird in dem folgenden Kapitel 7.2 diskutiert.

An dem sozio-pragmatischen Ansatz ist kritisch anzumerken, dass es sich um eine passive Enkodierungsstrategie handelt, für die hauptsächlich der geteilte Aufmerksamkeitsfokus eine zentrale Rolle spielt. Das heißt, dass Kinder das enkodieren, was in der sozialen

Interaktion von der Bezugsperson vorgeführt wird, sie aber nicht eigenständig mit dem Referenten handeln. Kritisch gesehen, spiegelt dies die natürliche Situation nicht wieder, die Kinder im Alltag erleben, wenn sie mit ihren Mitmenschen interagieren.

Aus diesem Grund wurde den Probanden in der Trainingsphase der vorliegenden Studie dieser Arbeit neben der Benennungsphase, in der die Kinder in einem geteilten Aufmerksamkeitsfokus involviert sind, eine anschließende Spielphase geboten, in der sie entweder mit dem Objekt alleine (Self-Exploring-Gruppe) oder zusammen mit der Experimentatorin (Joint-Engagement-Gruppe) spielten (vgl. Kapitel 4.2.2.4). Durch das Spielformat in der Self-Exploring-Gruppe sollte der sozio-pragmatische Ansatz an eine alltagsnahe Situation angenähert werden, in der Kinder zwar alleine, aber selber aktiv sind. Dies wurde erreicht, indem die Experimentatorin aus der Interaktion ausgetreten ist und die Probanden mit dem Zielobjekt eigenständig handeln konnten. Im Gegensatz zur Self-Exploring-Gruppe, wurde die Joint-Engagement-Gruppe von der Experimentatorin in eine gemeinsame Interaktion eingebunden, in der sie anstrebte die Zielfunktion mit dem Kind mehrmals zu erreichen. Durch die Spielphase in der Joint-Engagement-Gruppe, sollte der interaktionistische Ansatz aufgegriffen und mit dem sozio-pragmatischen Ansatz kontrastiert werden.

Der interaktionistischen Sichtweise nach entsteht eine Wortbedeutung nicht nur durch die Interaktion in einem geteilten Aufmerksamkeitsfokus, sondern vor allem durch gemeinsame, zielführende Handlungen in Koordination mit verbalen Handlungen (Rohlfing et al., 2016; Wildt et al., 2019). Das heißt, dass eine Referenz erst in einem interaktiven Prozess (Heller & Rohlfing, 2017) geformt wird, in der die Bezugsperson nicht nur der Aufmerksamkeitskoordination dient, sondern mit dem Kind gemeinsam auf ein bestimmtes Ziel hinarbeitet (Rohlfing, 2019, S. 217).

Die Auswertung der während der Spielphase erfassten Blickdauer (vgl. Kapitel 4.3.5.1) ergab, dass die Joint-Engagement-Gruppe, die im Sinne des interaktionistischen Ansatzes in einer gemeinsamen, zielführenden Interaktion eingebunden war, die meiste Zeit der Spielphase das Zielobjekt fixierte. Die Self-Exploring-Gruppe, die mit dem Zielobjekt alleine spielte, schaute zum Zielobjekt zwar weiterhin am längsten, aber deren Blickdauer nahm auch zu anderen Richtungen (z. B. Ablenker, Experimentatorin) zu. Diese Ergebnisse untermauern die Befunde vorheriger Studien bezüglich gemeinsamer Handlungen (z. B. Yu & Smith, 2016; Wass et al., 2018) insofern, als die Aufmerksamkeitsdauer und somit auch die Dauer der Enkodierung in der Joint-Engagement-Gruppe, im Gegensatz zur Self-Exploring-Gruppe, durch die Koordination verbaler und nonverbaler Handlungen expandiert werden konnte. Ein weiterer Befund der Spielphase war die Durchführung unterschiedlicher

Objekthandlungen der beiden Gruppen (vgl. Kapitel 4.3.5.4). Während die Joint-Engagement-Gruppe durch die Unterstützung der Experimentatorin die Zielhandlung des Objektes häufiger ausführte als die Self-Exploring-Gruppe, hat die Self-Exploring-Gruppe häufiger andere Objekthandlungen ausprobiert. Daraus lässt sich schlussfolgern, dass diese beschriebenen Befunde den bisherigen Forschungsstand (z. B. Yu & Smith, 2016; Wass et al., 2018), welcher für das Voranbringen des interaktionistischen Ansatzes relevant ist, insofern ergänzen, als 10 Monate alte Kinder sowohl alleinige als auch gemeinsame Objekthandlungen als Enkodierungsstrategie einsetzen. Das Letztere hält die visuelle Aufmerksamkeit zum Zielobjekt während des Trainings jedoch länger aufrecht. Außerdem wurde deutlich, dass in alleinigen Objekthandlungen die Zielfunktion von jungen Kindern kaum ausgeführt wird, was vermutlich damit zu erklären ist, dass sie mehr Unterstützung brauchen, um sie zu durchschauen und durchführen zu können.

Neben den drei beschriebenen Enkodierungsstrategien – der Objektsalienz, den sozialen Hinweisen und den Objekthandlungen – wurde außerdem das Ausschließbarkeitsprinzip als weitere Enkodierungsstrategie überprüft. In der bestehenden Forschungsliteratur (z. B. Mather & Plunkett 2010, 2012) wurde das Prinzip der Ausschließbarkeit, beruhend auf der Objektneuheit, bereits mit 10 Monate alten Kindern getestet (vgl. Kapitel 3.2.1). Diese Studien plädieren dafür, dass Kinder im frühen Alter einen Referenten für ein neues Wort auf Basis der Objektneuheit auswählen. Doch an der Studie von Mather und Plunkett (2010) ist kritisch anzumerken, dass die Probanden die Namen für die familiären Stimuli nicht kannten und die Objektneuheit daher die einzige Enkodierungsstrategie war, auf die sie zurückgreifen konnten. Die weiterführende Studie (Mather & Plunkett 2012), allerdings schon mit älteren Kindern, zeigte, dass die Probanden in der Referentenauswahl langsamer waren, als sich die Stimuli lediglich in ihrem Neuheitsgrad unterschieden. Sofern aber familiäre Objekte mit dabei waren, deren Namen die Kinder kannten, resultierte es in einer schnelleren Referentenauswahl (vgl. Kapitel 3.2.2).

Die beiden Studien warfen die Frage auf, welchen Referenten Kinder in der Studie der vorliegenden Arbeit auswählen, wenn sich die präsentierten Stimuli nicht im Neuheitsgrad, sondern nur in dem Wortwissen unterscheiden (familiäres-benanntes vs. familiäres-unbenanntes Objekt). Da Studien zeigen (Tincoff & Jusczyk, 2012; Bergelson & Swingley, 2012, 2015; Nomikou et al., 2018), dass Kinder bereits vor dem ersten Lebensjahr über ein Wortwissen verfügen (vgl. Kapitel 2.3), sollte in der vorliegenden Arbeit untersucht werden, ob

Kinder bereits im Alter von 10 Monaten die Referentenauswahl in einer ambigen Wortlernsituation eingrenzen können, indem sie auf das gelernte Wortwissen anstatt auf die Objektneuheit zurückgreifen.

Das Ausschließbarkeitsprinzip, beruhend auf dem Wortwissen, wurde als Enkodierungsstrategie anhand des New-Label-Tests in der Eye-Tracking-Testphase überprüft, indem das in der Trainingsphase zuvor benannte Objekt mit dem salienten, unbenannten Objekt präsentiert wurde und ein neues Wort fiel (vgl. Abbildung 7, S. 102). Es wurde angenommen, dass die Probanden länger zum salienten, familiären Objekt schauen, wenn sie ein neues Wort hören, da das weniger-saliente, ebenfalls familiäre Objekt bereits einen Namen hat. Allerdings konnten keine aussagekräftigen Rückschlüsse über das Prinzip der Ausschließbarkeit gezogen werden. Dies hat zum einen den Grund, dass in den drei Experimenten der Studie dieser Arbeit ein unterschiedliches Blickverhalten bei der Joint-Engagement-Gruppe vorlag (zur Veranschaulichung siehe Anhang A, Tabelle 25). Im ersten Experiment bestand während des New-Label-Tests keine visuelle Präferenz, wohingegen sich in den Kontrollexperimenten eine längere Blickdauer zum salienten Objekt zeigte. Zum anderen kann das Prinzip der Ausschließbarkeit nicht bekräftigt werden, da es auf dem Wortwissen von Objektnamen beruht, aber die während des Trainings enkodierte Wort-Objekt-Verknüpfung nicht durch eine längere Blickdauer zum Zielobjekt im Test rezeptiv abgerufen werden konnte.

In weiterführenden Studien, die das Ausschließbarkeitsprinzip nach demselben Versuchsdesign wie in der vorliegenden Arbeit (siehe auch Pruden et al., 2006) weiterführend untersuchen, sollte zunächst sichergestellt werden, dass die zuvor enkodierte Wort-Objekt-Verknüpfung rezeptiv abgerufen werden kann. Um das zu erreichen, sollte der Enkodierungsprozess der Wort-Objekt-Verknüpfung während des Trainings verstärkt werden (für eine ausführliche Diskussion vgl. Kapitel 6.3). Erst dann, wenn eine korrekte Referentenauswahl während der Original-Label-Phase anhand einer längeren Blickdauer zum Zielobjekt erfolgt, kann das Ausschließbarkeitsprinzip in der New-Label-Phase überprüft werden. Des Weiteren sollte darauf geachtet werden, dass beide Stimuli sich nicht in ihrem Neuheitsgrad unterscheiden. Außerdem wäre denkbar, eine stärkere Vergleichbarkeit zu schaffen, indem die Probanden auch mit dem salienten Objekt in der Spielphase handeln dürfen.

Abschließend wird nun der Erkenntnisfortschritt für die Wortlerntheorien und den Gedächtnisprozess der Enkodierung zusammengefasst. Der Gewinn aus der Studie der vorliegenden Arbeit für die Spracherwerbsforschung besteht vor allem darin, dass unterschiedli-

che Ansätze hinsichtlich des frühen Wortlernprozesses durch die Kombination zweier Methoden in einer Studie überprüft und die daraus gewonnen Erkenntnisse zusammengebracht werden konnten. Es stellte sich heraus, dass sich der assoziative und sozio-pragmatische Theorie in Bezug auf die Enkodierungsstrategien nicht gegenseitig ausschließen, sondern diese vielmehr in unterschiedlichen Kontexten genutzt werden. Das heißt, dass Kinder hauptsächlich dann perzeptionsgeleitet handeln, wenn die Objektsalienz nicht mit ostensiven Hinweisen in Kontrast steht. Für die Wort-Objekt-Enkodierung bedeuten diese Erkenntnisse, dass Kinder auch von der Salienz profitieren können, wenn Erwachsene saliente Objekte, auf die Kinder in dem Moment achten, benennen und zusätzlich ostensiv hervorheben, anstatt für einen „Wettstreit“ zwischen perzeptuellen und sozialen Reizen zu sorgen. Referieren Erwachsene jedoch auf weniger saliente Objekte, während ein salienter Ablenker präsent ist, dann müssen sie deutliche, ostensive Signale an das Kind senden, um einen geteilten Aufmerksamkeitsbezug zu erreichen. Der Erkenntnisgewinn für den interaktionistischen Ansatz besteht darin, dass alleinige und gemeinsame Objekthandlungen von Kindern zur Enkodierung eines Referenten genutzt werden. Allerdings unterscheidet sich die Probanden, die mit der Experimentatorin interagieren von den Probanden, die mit dem Zielobjekt alleine spielen, hinsichtlich ihres Blickverhaltens sowie der ausgeführten Objekthandlungen. Lediglich zum Prinzip der Ausschließbarkeit konnten keine Erkenntnisse gewonnen werden, da der rezeptive Abruf, auf den im Folgenden näher eingegangen werden soll, der zuvor enkodierten Wort-Objekt-Verknüpfung scheiterte

7.2 Retention und Konsolidierung der Wort-Objekt-Verknüpfung

Im Unterkapitel 7.1 wurden die Erkenntnisse zu den Enkodierungsstrategien, die Kinder im Alter von 10 Monaten nutzen, diskutiert. Die Frage, nach welcher Enkodierungsstrategie die Wort-Objekt-Verknüpfung abgerufen werden kann, wird in dem jetzigen Kapitel diskutiert, indem zunächst beschrieben werden soll, ob die Ergebnisse die dritte Hypothese unterstützen und inwieweit sie den bisherigen Forschungsstand des interaktionistischen Ansatzes ergänzen.

Da die Probanden die Objektsalienz als Enkodierungsstrategie im Training ausschlossen haben und stattdessen den sozialen Hinweisen während der Benennungsphase folgten, wurde die Behaltensleistung von Wort-Objekt-Verknüpfungen überprüft, die entweder in gemeinsamen Handlungen oder in alleinigen Objekthandlungen enkodiert wurden. In Anbetracht der aktuellen Forschungsliteratur (z. B. Wass et al., 2018), laut der, gemeinsame

Handlungen im Vergleich zu alleinigen Objekthandlungen, mit einer längeren Aufmerksamkeitsspanne zum Zielobjekt in Zusammenhang stehen, bestand die Vermutung, dass sie auch für den Abruf der Wort-Objekt-Verknüpfung wirkungsvoller sein könnten. Da Objekthandlungen bisher noch nicht als Enkodierungsstrategie im Kontext des Wortlernens untersucht wurden, wurde diese Forschungslücke in der vorliegenden Arbeit aufgegriffen und hinsichtlich des interaktionistischen Ansatzes in Form zielführender Handlungen operationalisiert.

In Bezug auf die Spielphase der vorliegenden Arbeit wurde als **dritte Hypothese** aufgestellt, dass ein Zusammenhang zwischen den in der Spielphase genutzten Enkodierungsstrategien und der Behaltensleistung (Retention) der enkodierten Wort-Objekt-Verknüpfung besteht. Zur Überprüfung dieser Annahme wurde die Enkodierung zunächst anhand der Blickdauer zu den Stimuli während der Spielphase und anschließend anhand der ausgeführten Zielhandlungen operationalisiert. Sowohl die Blickdauer als auch die Anzahl der Objekthandlungen wurden im Zusammenhang mit der Retention der Wort-Objekt-Verknüpfung, die anhand des Blickverhaltens in der Eye-Tracking-Testphase gemessen wurde, in beiden Gruppen untersucht.

Wie bereits in Kapitel 7.1 erwähnt, zeigten die beiden Gruppen in der Spielphase nicht nur ein unterschiedliches Blickverhalten zu den Objekten, sondern auch eine unterschiedliche Anzahl an Objekthandlungen. Weitere Auswertungen der Eye-Tracking-Daten ergaben, dass bereits nach wenigen Minuten keine der beiden Gruppen den enkodierten Wort-Objekt-Link rezeptiv abrufen konnte (vgl. Kapitel 4.3.5.2). Außerdem stellte sich heraus, dass sich bei der Self-Exploring-Gruppe Offset-Testphasen ein Salienz-Effekt zeigte, wohingegen die Joint-Engagement-Gruppe keines der beiden Objekte präferierte. Des Weiteren konnte weder ein Zusammenhang zwischen den Fixationszeiten (vgl. Kapitel 4.3.5.3) noch zwischen der Anzahl der erreichten Zielhandlungen im Training und der Fixationsdauer des Zielobjekts/des salienten Objekts im Test (vgl. Kapitel 4.3.5.5) festgestellt werden. Daher konnte die dritte Hypothese hinsichtlich des Zusammenhangs der Enkodierungsstrategien und der Beibehaltensleistung der enkodierten Wort-Objekt-Verknüpfung nicht unterstützt werden. Im Folgenden soll zum einen die Frage diskutiert werden, weshalb der rezeptive Abruf gescheitert ist und zum anderen, wie das unterschiedliche Blickverhalten der Self-Exploring- und Joint-Engagement-Gruppe erklärt werden kann.

Eine Vermutung bezüglich der Frage, warum der rezeptive Abruf der enkodierten Wort-Objekt-Verknüpfung gescheitert ist, lautet, dass die Trainingsphase nicht genügend Gedächtnishilfen zur Unterstützung des Enkodierungsprozesses der Wort-Objekt-Verknüpfung bot, sodass die Gedächtnisspur zu schwach war, um aktiviert werden zu können (für

eine ausführliche Diskussion vgl. Kapitel 6.3). Daher hätte insbesondere die Enkodierung des Zielwortes von der Experimentatorin stärker unterstützt werden sollen. Diese Annahmen werden auch von den Ergebnissen des zweiten Kontrollexperiments untermauert. Da die korrekte Referentenauswahl auch am darauffolgenden Erhebungstermin nicht gelungen ist, ist davon auszugehen, dass die enkodierte Information nicht in die Retentionsphase gelangt ist und daher auch nicht konsolidiert werden konnte. Stattdessen setzte nach der Enkodierung der Vergessensprozess ein. In Kapitel 6.3 wurden Ideen für künftige Studien vorgeschlagen, um den Enkodierungsprozess der Wort-Objekt-Verbindung zu verstärken, damit diese beibehalten werden kann.

Da das unterschiedliche Blickverhalten der Probanden im Test weder mit der Fixationsdauer zu den Stimuli, noch mit der Anzahl der Objekthandlungen während der Spielphase erklärt werden konnte, wird geschlussfolgert, dass es aufgrund der verschiedenen sozialen Spielformate zustande gekommen ist. Das heißt, dass durch die zielgerichteten Handlungen im gemeinsamen Spielformat die Relevanz des Referenten hervorgehoben und gleichzeitig der Salienz-Effekt des Ablenkers verhindert werden konnte. Die Relevanz wurde in der vorliegenden Studie bei der Joint-Engagement-Gruppe durch gemeinsame, zielgerichtete Handlungen geschaffen, die zusätzlich sprachlich untermauert wurden. Zu solchen sprachlichen Handlungen flossen beispielsweise ostensive Signale, positives Feedback und Aufforderungen (vgl. Kapitel 4.2.4.3) mit ein, die die Kinder der Self-Exploring-Gruppe kaum oder gar nicht erfahren haben, da sie mit dem Objekt alleine spielten.

Die vorliegende Arbeit leistet somit einen wesentlichen Beitrag zur Spracherwerbsforschung, indem sie demonstriert, dass 10 Monate alte Kinder eine Wort-Objekt-Verknüpfung als relevant auffassen müssen, um im Abrufttest die Salienz anderer Objekte auszublenden zu können. In Bezug auf den Gedächtnisprozess der Enkodierung bedeutet es, dass gemeinsame Zielhandlungen eine tiefere Gedächtnisspur der Wort-Objekt-Verknüpfung hinterlassen als alleinige Objekthandlungen, wodurch der Salienz-Effekt in der Joint-Engagement-Gruppe verhindert wurde.

Angesichts dieser Erkenntnis unterstützt die Studie der vorliegenden Arbeit das interaktionistische Wortlernformat. Dieses sollte in der künftigen Forschung noch weiter ausgebaut werden, indem beispielsweise die Enkodierung, insbesondere die der Wortform, verstärkt werden sollte (vgl. Kapitel 6.3). Außerdem sind weitere Studien notwendig, die den Begriff *Relevanz* messbar machen, um zu erfahren, was 10–12 Monate alte Kinder neben gemeinsamen, zielführenden Objekthandlungen noch als relevant auffassen. Des Weiteren

könnten die von Bezugspersonen ausgeführten, sozialen Hinweise (z. B. Feedback, Aufforderungen, ostensive Signale) in natürlichen, gemeinsamen Interaktionen näher untersucht werden, um zu erfahren, welche sozialen Signale für das Erreichen eines Ziels besonders wirksam sind. Dies kann man anhand der Art und Menge des sprachlichen Inputs und der Anzahl der durchgeführten Zielhandlungen messbar machen, während die Dyaden in einer freien Spielsituation auf ein gemeinsames Funktionsziel hinarbeiten. Wenn sich aus solchen Untersuchungen ergibt, welche sozialen Signale besonders wirksam in gemeinsamen Interaktionen mit Kindern sind, dann können Bezugspersonen dafür sensibilisiert werden, diese in Interaktionen gezielt einzusetzen.

Auf der methodologischen Ebene können die benannten Forschungsfragen und Vorschläge in weiterführenden Studien beispielsweise durch die Messung von Hirnaktivitäten während gemeinsamer Spielphasen untersucht werden. Aus der bisherigen Forschung ist bereits bekannt, dass durch die Präsenz von Augen und direkten Augenkontakt bei Kindern nicht nur eine längere Blickdauer (z. B. Farroni et al., 2002), sondern auch eine erhöhte Hirnaktivität hervorgerufen wird (Höhl et al., 2014; Michel et al., 2019). Da aber Kinder in Interaktionen keine passiven Beobachter sind, sondern mit ihren Mitmenschen aktiv interagieren, besteht die Notwendigkeit neurologisches Verhalten in solchen Interaktionen sowohl von Kindern, als auch von deren Bezugspersonen zu erfassen. Wass und Kollegen (2018) bieten in ihrer Studie erste Einblicke in die Hirnaktivitäten von 12 Monate alten Kindern und ihrer Bezugspersonen während der Interaktion. Die Studie ergab, dass die Kinder während gemeinsamer Spielphasen aufmerksamer einen Referenten explorierten als in alleinigen Spielphasen. Zudem stand die Aufmerksamkeitsspanne zu dem Objekt, welches die Kinder fokussierten, in Zusammenhang mit einer erhöhten Hirnaktivität der Bezugsperson. Höhl und Markova (2018) regen dazu an, die Methode dieser Studie (Wass et al., 2018) in der weiterführenden Forschung anzuwenden und die Hirnaktivitäten während sozialer Interaktionen von Kindern und deren Bezugspersonen zu untersuchen, um das Verständnis über die reziproken, dynamischen Prozesse zu erweitern.

7.3 Forschungsdesiderate der Studie und Ausblick auf künftige Forschung

Hinsichtlich des methodischen Vorgehens weist die vorliegende Arbeit einige Einschränkungen auf, die im Folgenden aufgeführt sowie mögliche Verbesserungsvorschläge benannt werden. Als erstes ist kritisch anzumerken, dass die Eye-Tracking-Testphasen den Kindern

nicht nur abverlangten, den korrekten Referenten zu bestimmen, sondern auch, die Handlung des Suchens („Wo ist das Modi?“) zu erfüllen (Rohlfing, 2019, S. 243). Um den jungen Probanden die Tests zu vereinfachen, könnte die Suchaufgabe bereits in der Trainingsphase eingeführt werden („Wo ist das Modi? Schau, da ist das Modi!“). Dadurch wäre die Suchaufgabe in der Trainingsphase familiarisiert, wodurch es die Suche nach dem Referenten in der Testphase erleichtern würde. Alternativ könnte die Suchaufgabe in einem vorangehenden Eye-Tracking-Versuch mit familiären Objekten geübt werden. Solch ein Versuch hätte außerdem den Vorteil, dass die Kinder die neue Eye-Tracking-Situation kennenlernen. Ein weiterer Vorteil wäre, dass die erfassten Blickzeiten zu den familiären Objekten für Auswertungen hinzugezogen werden können, um das Blickverhaltensmuster im Antwortverhalten von Kindern näher zu untersuchen.

Ein weiterer Kritikpunkt betrifft die Trainingsphase der Self-Exploring-Gruppe, da während der Spielphase der geteilte Aufmerksamkeitsfokus unterbrochen wurde, indem die Experimentatorin ihre Augen mithilfe einer Kappe versteckte. Um die Bedingungen der Joint-Engagement- und Self-Exploring-Gruppe noch vergleichbarer zu gestalten, könnte die Kappe weggelassen und der geteilte Aufmerksamkeitsfokus in der Self-Exploring-Gruppe beibehalten werden. Dadurch wäre die Self-Exploring-Gruppe weiterhin in einem geteilten Aufmerksamkeitsfokus eingebunden, auch wenn sie das Objekt alleine exploriert.

Hinsichtlich der Datenaufbereitung ist kritisch aufzuführen, dass sich fehlende Werte, insbesondere in den Testphasen des ersten Experiments ergeben haben, die durch eine einfache Mittelwertimputation ersetzt wurden. Bei diesem Verfahren besteht das Risiko, dass die Varianz der Datenverteilung unterschätzt wird (Little & Rubin, 1987, S. 62). Erst durch die Kürzung der Videoclips konnte die Anzahl der fehlenden Werte reduziert werden. Für künftige Eye-Tracking-Studien wird deshalb vorgeschlagen, mehrere kurze Videos zu erstellen, um fehlende Werte zu verhindern, weil durch eine größere Videoclip-Anzahl die Wahrscheinlichkeit geringer ist, dass in gar keinem der Videos Blickdaten erfasst werden. Da Kinder eine kürzere Aufmerksamkeitsspanne haben als Erwachsene, sollten die einzelnen Videoclips nicht länger als sechs Sekunden sein.

Nach den beschriebenen Kritikpunkten sind abschließend einige Vorschläge für weitere Datenanalysen zu erwähnen. So könnte zum Beispiel die Auswertung des Blickverlaufs mithilfe der LWL-Methode (vgl. Tabelle 1, S. 27–28) mehr Aufschluss über die Sprachverarbeitung geben, weil durch sie erkennbar wird, ob und wie sich das Blickverhalten ändert, nachdem das Zielwort fällt.

Hinsichtlich der Auswertung ist außerdem anzumerken, dass es interessant wäre, die Daten des zweiten Kontrollexperiments mit einer ANOVA mit Messwiederholungen auszuwerten. Allerdings wird für diese Art der Analyse ein vollständiger Datensatz benötigt, da SPSS einen listenweisen Ausschluss der Probanden vornimmt, wenn Werte fehlen. Dies ist aufgrund der Zusammenfügung der Kontrollexperimente der Fall, da für den ersten Tag Daten von insgesamt 24 Versuchspersonen vorliegen und für den zweiten Tag Daten von 12 Personen. Daher werden die Versuchsteilnehmer, die nicht am zweiten Erhebungstermin teilgenommen haben, listenweise für solch eine Analyse ausgeschlossen.

Durch die in den letzten Jahren gewonnenen, inhaltlichen Erkenntnisse und aufgrund des technischen Fortschritts (vgl. Golinkoff et al., 2013) konnten bisherige Forschungsmethoden optimiert und an neue Fragestellungen angepasst werden (vgl. Tabelle 1, S. 27–28). Die Besonderheit der Methode in der Studie der vorliegenden Arbeit liegt im Mixed-Methods-Forschungsdesign, da bereits bestehende Verfahren kombiniert wurden. Neben der Beobachtungsmethode wurde das Eye-Tracking-Verfahren verwendet, wodurch nicht nur unterschiedliche Wortlerntheorien in einer Studie untersucht, sondern auch die Eye-Tracking-Daten mit den aus dem Beobachtungsverfahren kodierten Daten in Zusammenhang gebracht werden konnten.

Einige Vorschläge, wie das vorliegende Mixed-Methods-Design modifiziert werden könnte, wurden bereits erwähnt (vgl. S. 175). In weiterführenden Studien ist außerdem denkbar, andere als die in der vorliegenden Arbeit genutzten Verfahren miteinander zu verknüpfen. Um das Wortlernen in naturalistischen Interaktionen zwischen Kindern und ihren Bezugspersonen zu näher zu untersuchen, könnte beispielsweise ein mobiler Eye Tracker (wie eine Eye-Tracking-Brille) und LENA kombiniert werden. Abgesehen von der Technik, werden für künftige Studien Labore umgebaut, um eine heimische Atmosphäre für Bezugspersonen und ihre Kinder zu schaffen (Yu, Slone & Smith, 2018), die der alltäglichen Umgebung der Versuchsteilnehmer ähneln soll. Alltagsnahe Studiendesigns sowie die dafür entwickelte Technik und Labore beruhen somit auf dem stets wachsenden Forschungsinteresse für dynamische Interaktionen.

8 Literaturverzeichnis

- Adams, R. J. (1987). An evaluation of color preference in early infancy. *Infant Behavior and Development*, 10(2), 143–150. [https://doi.org/10.1016/0163-6383\(87\)90029-4](https://doi.org/10.1016/0163-6383(87)90029-4)
- Adams, R. J. & Courage, M. L. (1995). Development of chromatic discrimination in early infancy. *Behavioural Brain Research*, 67(1), 99–101. [https://doi.org/10.1016/0166-4328\(94\)00143-4](https://doi.org/10.1016/0166-4328(94)00143-4)
- Akhtar, N. & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff & K. A. Hirsh-Pasek (Hrsg.), *Becoming a word learner: A debate on lexical acquisition* (S. 115–135). Oxford: Oxford University Press.
- Anderson, D. R. & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, 48(5), 505–522. <https://doi.org/10.1177/0002764204271506>
- Anderson, J. R. (2007). *Kognitive Psychologie*. Spektrum Akademischer Verlag.
- Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6), 396–408. <https://doi.org/10.1111/j.1467-9280.1991.tb00174.x>
- Ansorge, U. & Leder, H. (2017). *Wahrnehmung und Aufmerksamkeit*. <https://doi.org/10.1007/978-3-658-12912-5>
- Aslin, R. N. & Smith, L. B. (1988). Perceptual development. *Annual Review of Psychology*, 39(1), 435–473. <https://doi.org/10.1146/annurev.ps.39.020188.002251>
- Atkinson, J. (2000). *The developing visual brain*. Oxford: Oxford University Press.
- Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, 2, 89–95. [https://doi.org/10.1016/S0079-7421\(08\)60422-3](https://doi.org/10.1016/S0079-7421(08)60422-3)
- Atkinson, R. C. & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, 225(2), 82–91.
- Axelsson, E. L., Churchley, K. & Horst, J. S. (2012). The right thing at the right time: Why ostensive naming facilitates word learning. *Frontiers in Psychology*, 3, 88. <https://doi.org/10.3389/fpsyg.2012.00088>

- Axelsson, E. L., Swinton, J., Winiger, A. I. & Horst, J. S. (2018). Napping and toddlers' memory for fast-mapped words. *First Language*, 38(6), 582–595. <https://doi.org/10.1177/0142723718785490>
- 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/10.1111/j.1468-0017.2009.01384.x>
- Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 62(5), 875–890. <https://doi.org/10.1111/j.1467-8624.1991.tb01577.x>
- Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, 29(5), 832.
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M. & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child development*, 67(6), 3135–3153. <https://doi.org/10.1111/j.1467-8624.1996.tb01906.x>
- Bergelson, E. & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253–3258. <https://doi.org/10.1073/pnas.1113380109>
- Bergelson, E. & Swingle, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, 11(4), 369–380. <https://doi.org/10.1080/15475441.2014.979387>
- Bion, R. A., Borovsky, A. & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, 126(1), 39–53. <https://doi.org/10.1016/j.cognition.2012.08.008>
- Bjorklund, D. F. & Causey, K. B. (2018). *Children's thinking: Cognitive development and individual differences*. Los Angeles: SAGE Publications.
- Bornstein, M. H. (1985). Habituation of attention as a measure of visual information processing in human infants: Summary, systematization, and synthesis. In G. Gottlieb & N. A. Krasnegor (Hrsg.), *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (S. 253–300). Westport, CT, US: Ablex Publishing.
- Bornstein, M. H. & Benasich, A. A. (1986). Infant habituation: Assessments of individual differences and short-term reliability at five months. *Child Development*, 87–99.

- Bloom, P. (2001). Précis of how children learn the meanings of words. *Behavioral and brain Sciences*, 24(6), 1095–1103. <https://doi.org/10.1017/S0140525X01000139>
- Bloom, L., Tinker, E. & Margulis, C. (1993). The words children learn: Evidence against a noun bias in early vocabularies. *Cognitive Development*, 8(4), 431–450. [https://doi.org/10.1016/S0885-2014\(05\)80003-6](https://doi.org/10.1016/S0885-2014(05)80003-6)
- Booth, A. E., McGregor, K. K. & Rohlfing, K. J. (2008). Socio-pragmatics and attention: Contributions to gesturally guided word learning in toddlers. *Language Learning and Development*, 4, 179–202. <https://doi.org/10.1080/15475440802143091>
- Brand, R. J., Baldwin, D. A. & Ashburn, L. A. (2002). Evidence for „motionese“: Modifications in mothers’ infant-directed action. *Developmental Science*, 5, 72–83. <https://doi.org/10.1111/1467-7687.00211>
- Brandone, A. C., Pence, K. L., Golinkoff, R. M. & Hirsh-Pasek, K. (2007). Action speaks louder than words: young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development*, 78(4), 1322–1342. <https://doi.org/10.1111/j.1467-8624.2007.01068.x>
- Bremner, J. G. & Slater, A. (2008). *Theories of infant development*. Oxford: Blackwell Publishing Ltd.
- Brône, G. & Oben, B. (2018). *Eye-tracking in interaction: Studies on the role of eye gaze in dialogue*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Bruner, J. S. (1983). *Child’s talk: Learning to use language*. New York: W W Norton & Co Inc.
- Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language*, 2(1), 1–19. <https://doi.org/10.1017/S0305000900000866>
- Capone, N. C. & McGregor, K. K. (2004). Gesture development: A review for clinical and research practices. *Journal of Speech, Language, and Hearing Research*, 47, 173–186. [https://doi.org/10.1044/1092-4388\(2004/015\)](https://doi.org/10.1044/1092-4388(2004/015))
- Capone, N. C. & McGregor, K. K. (2005). The effect of semantic representation on toddlers’ word retrieval. *Journal of Speech, Language, and Hearing Research*. [https://doi.org/10.1044/1092-4388\(2005/102\)](https://doi.org/10.1044/1092-4388(2005/102))
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan & G. A. Miller (Hrsg.), *Linguistic theory and psychological reality* (S. 264–293). Cambridge, Massachusetts: MIT Press.

- Carey, S. (2010). Beyond fast mapping. *Language Learning and Development*, 6(3), 184–205. <https://doi.org/10.1080/15475441.2010.484379>
- Carey, S. & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, 15, 17–29.
- Carpenter, M., Nagell, K. & Tomasello, M. (1998b) Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society of Research in Child Development*, 63(4), 1–143.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525. <https://doi.org/10.1016/j.visres.2011.04.012>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). New York: Lawrence Erlbaum Associates.
- Connor, C. E., Egeth, H. E. & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology*, 14(19), R850–R852. <https://doi.org/10.1016/j.cub.2004.09.041>
- Craik, F. I. M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology*, 104(3), 268–294. <https://doi.org/10.1037/0096-3445.104.3.268>
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25(2), 141–168. <https://doi.org/10.1111/j.1468-0017.2009.01384.x>
- Csibra, G. & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Hrsg.), *Processes of change in brain and cognitive development. Attention and performance*, Xxi (S. 249–274). Oxford: Oxford University Press.
- Csibra, G. & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148–153. <https://doi.org/10.1016/j.tics.2009.01.005>
- De Swert, K. (2012). Calculating inter-coder reliability in media content analysis using Krippendorff's Alpha. *Center for Politics and Communication*, 1–15.
- Deák, G. O., Krasno, A. M., Jasso, H. & Triesch, J. (2018). What leads to shared attention? Maternal cues and infant responses during object play. *Infancy*, 23(1), 4–28. <https://doi.org/10.1111/infa.12204>
- DeCasper, A. J. & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208(4448), 1174–1176. <https://doi.org/10.1126/science.7375928>

- DeCasper, Anthony J. & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9(2), 133–150. [https://doi.org/10.1016/0163-6383\(86\)90025-1](https://doi.org/10.1016/0163-6383(86)90025-1)
- Dudai, Y. (2012). The restless engram: consolidations never end. *Annual review of neuroscience*, 35, 227–247. <https://doi.org/10.1146/annurev-neuro-062111-150500>
- Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology*, 55(1), 51–86. <https://doi.org/10.1146/annurev.psych.55.090902.142050>
- Dunham, P. J., Dunham, F. & Curwin, A. (1993). Joint-attentional states and lexical acquisition at 18 months. *Developmental Psychology*, 29(5), 827–831.
- Esseily, R. & Fagard, J. (2013). Ostensive cues orient 10-month-olds' attention toward the task but delay learning. *Psychology*, 04(07), 20–25. <http://dx.doi.org/10.4236/psych.2013.47A003>
- Fantz, R. L. (1958). Pattern vision in young infants. *The psychological record*, 8(2), 43–47.
- Fantz, R. L. (1961). The origin of form perception. *Scientific American*, 204(5), 66–72. <https://doi.org/10.1038/scientificamerican0561-66>
- Fantz, R. L. (1963). Pattern vision in newborn infants. *Science*, 140(3564), 296–297. <https://doi.org/10.1126/science.140.3564.296>
- Fantz, R. L. (1964). Visual experience in infants: decreased attention to familiar patterns relative to novel ones. *Science*, 146(3644), 668–670. <https://doi.org/10.1126/science.146.3644.668>
- Farroni, T., Csibra, G., Simion, F. & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, 99(14), 9602–9605. <https://doi.org/10.1073/pnas.152159999>
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant behavior and development*, 8(2), 181–195.
- Fernald, A. E., Zangl, R., Portillo, A. L. & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández & H. Clahsen (Hrsg.), *Language Acquisition and Language Disorders* (Bd. 44, S. 97–135). <https://doi.org/10.1075/lald.44.06fer>

- Field, A. P. (2009). *Discovering statistics using SPSS: And sex, drugs and rock „n“ roll* (3. Aufl.). Los Angeles: SAGE Publications.
- Forbes, S. H. & Plunkett, K. (2019). Infants show early comprehension of basic color words. *Developmental psychology*, 55(2), 240. <http://dx.doi.org/10.1037/dev0000609>
- Franklin, A., Bevis, L., Ling, Y. & Hurlbert, A. (2010). Biological components of colour preference in infancy. *Developmental Science*, 13(2), 346–354. <https://doi.org/10.1111/j.1467-7687.2009.00884.x>
- Frick, J. E., Colombo, J. & Saxon, T. F. (1999). Individual and developmental differences in disengagement of fixation in early infancy. *Child Development*, 70(3), 537–548.
- Gampe, A., Brauer, J. & Daum, M. M. (2016). Imitation is beneficial for verb learning in toddlers. *European Journal of Developmental Psychology*, 13(5), 594–613. <https://doi.org/10.1080/17405629.2016.1139495>
- Geeraerts, D. (2006). *Cognitive linguistics: Basic readings*. Berlin; New York: Mouton de Gruyter.
- Gershkoff-Stowe, L. & Hahn, E. R. (2007). Fast mapping skills in the developing lexicon. *Journal of Speech, Language, and Hearing Research*, 50(3), 682–697. [https://doi.org/10.1044/1092-4388\(2007/048\)](https://doi.org/10.1044/1092-4388(2007/048))
- Gliga, T. & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, 20(3), 347–353. <https://doi.org/10.1111/j.1467-9280.2009.02295.x>
- Goldstein, E. B. & Gegenfurtner, K. R. (2015). *Wahrnehmungspsychologie: Der Grundkurs* (9. Aufl.). Berlin: Springer.
- Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M. & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, 28(1), 99–108. <https://doi.org/10.1037/0012-1649.28.1.99>
- Golinkoff, R. M., Can, D. D., Soderstrom, M. & Hirsh-Pasek, K. A. (2015). (Baby)Talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science*, 24(5), 339–344. <https://doi.org/10.1177/09637214155595345>

- Golinkoff, R. M. & Hirsh-Pasek, K. A. (2006). Baby wordsmith from associationist to social sophisticate. *Current Directions in Psychological Science*, 15(1), 30–33. <https://doi.org/10.1111/j.0963-7214.2006.00401.x>
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M. & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of child language*, 14(1), 23–45. <https://doi.org/10.1017/S030500090001271X>
- Golinkoff, R. M., Ma, W., Song, L. & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on Psychological Science*, 8(3), 316–339. <https://doi.org/10.1177/1745691613484936>
- Golinkoff, R. M., Mervis, C. B. & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, 21(1), 125–155. <https://doi.org/10.1017/S0305000900008692>
- Goodman, J. C., McDonough, L. & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development*, 69(5), 1330–1344. <https://doi.org/10.1111/j.1467-8624.1998.tb06215.x>
- Gredebäck, G., Astor, K. & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: A critical test of natural pedagogy. *Child Development*. <https://doi.org/10.1111/cdev.13026>
- Gredebäck, G., Johnson, S. & von Hofsten, C. (2009). Eye Tracking in infancy research. *Developmental Neuropsychology*, 35(1), 1–19. <https://doi.org/10.1080/87565640903325758>
- Grimminger, A. (2017). Gestische und sprachliche Kommunikation von 12–16 Monate alten Kindern und ihren Bezugspersonen in wiederkehrenden, semi-naturalistischen Interaktionen und individuelle Unterschiede in der späteren Sprachentwicklung (Dissertation). Universität Paderborn.
- Grimminger, A., Rohlfing, K. J. & Stenneken, P. (2010). Children’s lexical skills and task demands affect gestural behavior in mothers of late-talking children and children with typical language development. *Gesture*, 10(2–3), 251–278. <https://doi.org/10.1075/gest.10.2-3.07gri>
- Gruber, T. (2018). *Gedächtnis* (2. Aufl.). Berlin, Germany: Springer.
- Hagen, J. W. (1967). The effect of distraction on selective attention. *Child Development*, 38(3), 685–694.

- Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87(1), B23–B34. [https://doi.org/10.1016/S0010-0277\(02\)00186-5](https://doi.org/10.1016/S0010-0277(02)00186-5)
- Hallé, P. A. & Boysson-Bardies, B. de. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development*, 17(2), 119–129. [https://doi.org/10.1016/0163-6383\(94\)90047-7](https://doi.org/10.1016/0163-6383(94)90047-7)
- Hamilton, A., Plunkett, K. & Schafer, G. (2000). Infant vocabulary development assessed with a British Communicative Development Inventory. *Journal of Child Language*, 27, 689–705. <https://doi.org/10.1017/S0305000900004414>
- Hayes, A. F. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89. <https://doi.org/10.1080/19312450709336664>
- He, A. X. & Arunachalam, S. (2017). Word learning mechanisms: Word learning mechanisms. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(4), e1435. <https://doi.org/10.1002/wcs.1435>
- Heibeck, T. H. & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development*, 58(4), 1021. <https://doi.org/10.2307/1130543>
- Heller, V. & Rohlfing, K. J. (2017). Reference as an interactive achievement: Sequential and longitudinal analyses of labeling interactions in shared book reading and free play. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00139>
- Henderson, A. M. E., Sabbagh, M. A. & Woodward, A. L. (2013). Preschoolers' selective learning is guided by the principle of relevance. *Cognition*, 126(2), 246–257. <https://doi.org/10.1016/j.cognition.2012.10.006>
- Henderson, L. M., Weighall, A. R., Brown, H. & Gareth Gaskell, M. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, 15(5), 674–687. <https://doi.org/10.1111/j.1467-7687.2012.01172.x>
- Heubrock, D. & Petermann, F. (2000). *Lehrbuch der Klinischen Kinderneuropsychologie: Grundlagen, Syndrome, Diagnostik und Intervention*. Göttingen: Hogrefe Verlag.
- Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee & H. S. Cairns (Hrsg.), *Language, speech, and communication. Methods for assessing children's syntax* (S. 105-124). Cambridge, MA, US: The MIT Press.
- Hollich, G. J., Hirsh-Pasek, K. & Golinkoff, R. M. (1998). Introducing the 3-D intermodal preferential looking paradigm: A new method to answer an age-old

- question. *Advances in infancy research*, 12, 355–374.
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., ... Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the society for research in child development*, i–135.
- Hood, B. M., Willen, J. D. & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, 9(2), 131–134. <https://doi.org/10.1111/1467-9280.00024>
- Höhl, S. & Markova, G. (2018). Moving developmental social neuroscience toward a second-person approach. *PLoS biology*, 16(12), e3000055. <https://doi.org/10.1371/journal.pbio.3000055>
- Höhl, S., Michel, C., Reid, V. M., Parise, E. & Striano, T. (2014). Eye contact during live social interaction modulates infants' oscillatory brain activity. *Social neuroscience*, 9(3), 300–308. <https://doi.org/10.1080/17470919.2014.884982>
- Horst, J. S. & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128–157. <https://doi.org/10.1080/15250000701795598>
- Horst, J. S., Samuelson, L. K., Kucker, S. C. & McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition*, 118(2), 234–244. <https://doi.org/10.1016/j.cognition.2010.10.015>
- Horváth, K., Myers, K., Foster, R. & Plunkett, K. (2015). Napping facilitates word learning in early lexical development. *Journal of Sleep Research*, 24(5), 503–509. <https://doi.org/10.1111/jsr.12306>
- Houston-Price, C., Caloghris, Z. & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, 15(2), 125–150. <https://doi.org/10.1111/j.1532-7078.2009.00009.x>
- Houston-Price, C., Plunkett, K. & Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology*, 95(1), 27–55. <https://doi.org/10.1016/j.jecp.2006.03.006>
- Itti, L. & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194–203. <https://doi.org/10.1038/35058500>
- Itti, L., Koch, C. & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20(11), 1254–1259.

- James, W. (1890). *The principles of psychology*. New York: Dover Publications.
- Jusczyk, P. W., Hohne, E. A. & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & psychophysics*, 61(8), 1465–1476.
- Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze*. Berlin: Walter de Gruyter.
- Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A. & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development*, 18(1), 111–116. [https://doi.org/10.1016/0163-6383\(95\)90012-8](https://doi.org/10.1016/0163-6383(95)90012-8)
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2. Aufl.). Thousand Oaks, CA.: Sage.
- Krippendorff, K. (2007). Computing Krippendorff's alpha reliability. *Departmental Papers (ASC)*, 43, 1–10.
- Kucker, S. C. & Samuelson, L. K. (2012). The first slow step: Differential effects of object and word-form familiarization on retention of fast-mapped words. *Infancy*, 17(3), 295–323. <https://doi.org/10.1111/j.1532-7078.2011.00081.x>
- Kushnerenko, E., Tomalski, P., Ballieux, H., Ribeiro, H., Potton, A., Axelsson, E. L. et al. (2013). Brain responses to audiovisual speech mismatch in infants are associated with individual differences in looking behaviour. *European Journal of Neuroscience*, 38(9), 3363–3369. <https://doi.org/10.1111/ejn.12317>
- LENA: Building brains through early talk (o. D.). Abgerufen am 23.10.2019, von <https://lenafoundation.staging.wpengine.com>
- Lane, D. M. & Pearson, D. A. (1982). The development of selective attention. *Merrill-Palmer Quarterly*, 317–337.
- Lawson, K. R., Parrinello, R. & Ruff, H. A. (1992). Maternal behavior and infant attention. *Infant Behavior and Development*, 15(2), 209–229. [https://doi.org/10.1016/0163-6383\(92\)80024-O](https://doi.org/10.1016/0163-6383(92)80024-O)
- Lebovici, S. (1990). *Der Säugling, die Mutter und der Psychoanalytiker*. Stuttgart: Klett-Cotta.
- Leung, E. H. & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology*, 17(2), 215–220. <http://dx.doi.org/10.1037/0012-1649.17.2.215>

- Linke, A., Nussbaumer, M. & Portmann, P. R. (2004). *Studienbuch Linguistik* (5. Aufl.). Tübingen: Niemeyer.
- Liszkowski, U. (2015). Kommunikative und sozial-kognitive Voraussetzungen des Spracherwerbs. In S. Sachse (Hrsg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen* (S. 27–38). München: Elsevier.
- Little, R. J. A. & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons
- Little, R. J. A. & Rubin, D. B. (2014). *Statistical analysis with missing data*. New York: John Wiley & Sons.
- Luhmann, M. (2015). *R für Einsteiger: Einführung in die Statistiksoftware für die Sozialwissenschaften. Mit Online-Material* (4. Aufl.). Weinheim Basel: Beltz.
- Mampe, B., Friederici, A. D., Christophe, A. & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology*, 19(23), 1994–1997. <https://doi.org/10.1016/j.cub.2009.09.064>
- Marcus, G. F., Fernandes, K. J. & Johnson, S. P. (2012). The role of association in early word-learning. *Frontiers in Psychology*, 3, 283. <http://doi.org/10.3389/fpsyg.2012.00283>
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, Mass.: Bradford Books.
- Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14(1), 57–77. https://doi.org/10.1207/s15516709cog1401_4
- Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, 92, 199–227. [https://doi.org/10.1016/0024-3841\(94\)90342-5](https://doi.org/10.1016/0024-3841(94)90342-5)
- Markman, E. M. & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121–157. [https://doi.org/10.1016/0010-0285\(88\)90017-5](https://doi.org/10.1016/0010-0285(88)90017-5)
- Markman, E. M., Wasow, J. L. & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47(3), 241–275. [https://doi.org/10.1016/S0010-0285\(03\)00034-3](https://doi.org/10.1016/S0010-0285(03)00034-3)
- Markson, L. & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, 385(6619), 813–815. <https://doi.org/10.1038/385813a0>

- Mather, E. (2013). Novelty, attention, and challenges for developmental psychology. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00491>
- Mather, E. & Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology*, 105(3), 232–242. <https://doi.org/10.1016/j.jecp.2009.11.004>
- Mather, E. & Plunkett, K. (2012). The role of novelty in early word learning. *Cognitive Science*, 36(7), 1157–1177. <https://doi.org/10.1111/j.1551-6709.2012.01239.x>
- Maurer, D. (1985). Infants' perception of facedness. In T. M. Field & N. A. Fox (Hrsg.). *Social perception in infants*. Norwood, NJ, US: Ablex Pub.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178.
- Meindl, C. (2011). *Methodik für Linguisten: Eine Einführung in Statistik und Versuchsplanung*. Tübingen: Gunter Narr Verlag.
- Merriman, W. E., Bowman, L. L. & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development*, 54(3/4), i–129. <https://doi.org/10.2307/1166130>
- Merriman, W. E., Marazita, J. & Jarvis, L. (1995). Children's disposition to map new words onto new referents. In M. Tomasello & W. E. Merriman (Hrsg.) *Beyond names for things: Young children's acquisition of verbs* (S. 147–183). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Mervis, C. B. & Bertrand, J. (1994). Acquisition of the novel name-nameless category (N3C) Principle. *Child Development*, 65(6), 1646–1662. <https://doi.org/10.1111/j.1467-8624.1994.tb00840.x>
- Michel, C., Wronski, C., Pauen, S., Daum, M. M. & Hoehl, S. (2017). Infants' object processing is guided specifically by social cues. *Neuropsychologia*, 126, 54–61. <https://doi.org/10.1016/j.neuropsychologia.2017.05.022>
- Moon, C., Cooper, R. P. & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495–500. [https://doi.org/10.1016/0163-6383\(93\)80007-U](https://doi.org/10.1016/0163-6383(93)80007-U)
- Moore, C. & Dunham, P. J. (2014). *Joint attention: Its origins and role in development*. New York: Psychology Press.

- Moore, C., Angelopoulos, M. & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental psychology*, 35(1), 60.
- Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R. & Schwartz, H. K. (2000). Responding to joint attention across the 6- through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, 21(3), 283–298. [https://doi.org/10.1016/S0193-3973\(99\)00040-4](https://doi.org/10.1016/S0193-3973(99)00040-4)
- 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](https://doi.org/10.1016/S0163-6383(98)90014-5)
- Muir, D., Hains, S., Cao, Y. & D’Entremont, B. (1996). 3- to 6-month-olds’ sensitivity to adult intentionality: The role of adult contingency and eye direction in dyadic interactions. *Infant Behavior and Development, Supplement 1*(19), 199. [https://doi.org/10.1016/S0163-6383\(96\)90254-4](https://doi.org/10.1016/S0163-6383(96)90254-4)
- Müller, G. B., Pradeu, T. & Schäfer, K. (2004). *Evolution of communication systems: A comparative approach*. Cambridge, London: MIT Press.
- Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V. & 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>
- Munro, N., Baker, E., McGregor, K. & Arciuli, J. (2012). Why word learning is not fast. *Frontiers in Developmental Psychology*, 3, 41. <https://doi.org/10.3389/fpsyg.2012.00041>
- Nagata, Y. & Dannemiller, J. L. (1996). The selectivity of motion-driven visual attention in infants. *Child Development*, 67(6), 2608–2620. <https://doi.org/10.2307/1131742>
- Namy, L. L. (2012). Getting specific: Early general mechanisms give rise to domain-specific expertise in word learning. *Language Learning and Development*, 8(1), 47–60. <https://doi.org/10.1080/15475441.2011.617235>
- Niedźwiecka, A., Ramotowska, S. & Tomalski, P. (2017). Mutual gaze during early mother-infant interactions promotes attention control development. *Child Development*. <https://doi.org/10.1111/cdev.12830>
- Nomikou, I. & Rohlfing, K. J. (2011). Language does something: Body action and language in maternal input to three-month-olds. *IEEE Transactions on*

- Autonomous Mental Development*, 3(2), 113–128. <https://doi.org/10.1109/TAMD.2011.2140113>
- Nomikou, I., Leonardi, G., Radkowska, A., Rączaszek-Leonardi, J. & Rohlfing, K. J. (2017). Taking up an active role: Emerging participation in early mother–infant interaction during peekaboo routines. *Frontiers in Psychology*, 8, 1656. <https://doi.org/10.3389/fpsyg.2017.01656>
- Nomikou, I., Rohlfing, K. J., Cimiano, P. & Mandler, J. M. (2018). Evidence for early comprehension of action verbs. *Language Learning and Development*, 15(1), 64–74. <https://doi.org/10.1080/15475441.2018.1520639>
- Nomikou, I., Rohlfing, K. J. & Szufnarowska, J. (2013). Educating attention: Recruiting, maintaining, and framing eye contact in early natural mother–infant interactions. *Interaction Studies*, 14(2), 240–267. <https://doi.org/10.1075/is.14.2.05nom>
- Nothdurft, H.-C. (2000). Salience from feature contrast: Variations with texture density. *Vision Research*, 40(23), 3181–3200. [https://doi.org/10.1016/S0042-6989\(00\)00168-1](https://doi.org/10.1016/S0042-6989(00)00168-1)
- Odean, R., Nazareth, A. & Pruden, S. M. (2015). Novel methodology to examine cognitive and experiential factors in language development: Combining eye-tracking and LENA technology. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01266>
- Parish-Morris, J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M. & Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intention in word learning. *Child Development*, 78(4), 1265–1287. <https://doi.org/10.1111/j.1467-8624.2007.01065.x>
- Pereira, A. F., Smith, L. B. & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, 21(1), 178–185. <https://doi.org/10.3758/s13423-013-0466-4>
- Pérez, A., Dennis, R. J., Gil, J. F. A., Rondón, M. A. & López, A. (2002). Use of the mean, hot deck and multiple imputation techniques to predict outcome in intensive care unit patients in Colombia. *Statistics in Medicine*, 21(24), 3885–3896. <https://doi.org/10.1002/sim.1391>
- Plunkett, K. (1997). Theories of early language acquisition. *Trends in Cognitive Sciences*, 1(4), 146–153. [https://doi.org/10.1016/S1364-6613\(97\)01039-5](https://doi.org/10.1016/S1364-6613(97)01039-5)
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25. <https://doi.org/10.1080/00335558008248231>

- Posner, M. I., Snyder, C. R. & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109(2), 160
- Preissler, M. A. & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, 97(1), B13–B23. <https://doi.org/10.1016/j.cognition.2005.01.008>
- Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M. & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child development*, 77(2), 266–280. <https://doi.org/10.1111/j.1467-8624.2006.00869.x>
- Quine W. V. O. (1960) *Word and object*. Cambridge: MIT Press.
- Rasch, B. & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*, 93(2), 681–766. <https://doi.org/10.1152/physrev.00032.2012>
- Rasch, B., Frieze, M., Hofmann, W. & Naumann, E. (2006). *Quantitative Methoden: Einführung in die Statistik*, (Bd. 2, 2. Aufl.). Heidelberg: Springer Medizin Verlag.
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T. & Reissland, N. (2017). The human fetus preferentially engages with face-like visual stimuli. *Current Biology*, 27(12), 1825–1828.e3. <https://doi.org/10.1016/j.cub.2017.05.044>
- Rice, N. (1980). *Cognition to language*. Baltimore, MD: University Park Press.
- Richards, J. E. & Anderson, D. R. (2004). Attentional inertia in children's extended looking at television. In R. V. Kail (Hrsg.), *Advances in Child Development and Behavior* (Bd. 32, S. 163–212). [https://doi.org/10.1016/S0065-2407\(04\)80007-7](https://doi.org/10.1016/S0065-2407(04)80007-7)
- Rohlfing, K. J. (2013). *Frühkindliche Semantik*. Tübingen: Narr Francke Attempto Verlag.
- Rohlfing, K. J. (2019). *Frühe Sprachentwicklung*. Tübingen: Narr Francke Attempto Verlag.
- Rohlfing, K. J., Grimminger, A. & Lüke, C. (2017). An interactive view on the development of deictic pointing in infancy. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01319>
- Rohlfing, K. J., Wrede, B., Vollmer, A.-L. & Oudeyer, P. Y. (2016). An alternative to mapping a word onto a concept in language acquisition: Pragmatic frames. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00470>

- Rubin, D. C. & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103(4), 734–760. <https://doi.org/10.1037/0033-295X.103.4.734>
- Ruff, H. A. & Lawson, K. R. (1990). Development of sustained, focused attention in young children during free play. *Developmental Psychology*, 26(1), 85.
- Samuelson, L. K. & McMurray, B. (2016). What does it take to learn a word? *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1–2), n/a-n/a. <https://doi.org/10.1002/wcs.1421>
- Samuelson, L. K. & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, 69(1), 94–104. <https://doi.org/10.1111/j.1467-8624.1998.tb06136.x>
- Saunders, J. A., Morrow-Howell, N., Spitznagel, E., Dore, P., Proctor, E. K. & Pescarino, R. (2006). Imputing missing data: A comparison of methods for social work researchers. *Social Work Research*, 30(1), 19–31. <https://doi.org/10.1093/swr/30.1.19>
- Schafer, G. & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, 69(2), 309–320. <https://doi.org/10.2307/1132166>
- Searleman, A. & Herrmann, D. (1994). *Memory from a broader perspective*. New York: McGraw-Hill Inc., US.
- Senju, A. & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), 668–671. <https://doi.org/10.1016/j.cub.2008.03.059>
- Sloetjes, H. & Wittenburg, P. (2008). Annotation by Category: ELAN and ISO DCR. In *LREC Proceedings* (S. 816–820). Marrakesch, Marokko.
- Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In *Basic and applied perspectives on learning, cognition, and development. The Minnesota Symposia on Child Psychology* (Bd. 28, S. 1–32). Mahwah, NJ: Lawrence Erlbaum Associates.
- Smith, L. B., Colunga, E. & Yoshida, H. (2010). Knowledge as process: Contextually cued attention and early word learning. *Cognitive Science*, 34(7), 1287–1314. <https://doi.org/10.1111/j.1551-6709.2010.01130.x>
- Smith, L. B., Jones, S. S. & Landau, B. (1996). Naming in young children: A dumb

- attentional mechanism? *Cognition*, 60(2), 143–171. [https://doi.org/10.1016/0010-0277\(96\)00709-3](https://doi.org/10.1016/0010-0277(96)00709-3)
- Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology*, 15(6), 626–636. <https://doi.org/10.1037/0012-1649.15.6.626>
- Spencer, J. P., Perone, S., Smith, L. B. & Samuelson, L. K. (2011). Learning words in space and time: Probing the mechanisms behind the suspicious-coincidence effect. *Psychological Science*, 22(8), 1049–1057. <https://doi.org/10.1177/0956797611413934>
- Sperber, D. & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell Publishers.
- Spiegel, C. & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology*, 109(1), 132–140. <https://doi.org/10.1016/j.jecp.2010.10.013>
- Styles, S. J. & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, 1(1), 1–24. <https://doi.org/10.1515/LANGCOG.2009.001>
- Swingle, D. (2010). Fast mapping and slow mapping in children's word learning. *Language Learning and Development*, 6(3), 179–183. <https://doi.org/10.1080/15475441.2010.484412>
- Swingle, D. & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76(2), 147–166. [https://doi.org/10.1016/S0010-0277\(00\)00081-0](https://doi.org/10.1016/S0010-0277(00)00081-0)
- Swingle, D. & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13(5), 480–484. <https://doi.org/10.1111/1467-9280.00485>
- Szgun, G. (2013). *Sprachentwicklung beim Kind: Ein Lehrbuch* (5. Aufl.). Weinheim: Beltz.
- Szfnarowska, J., Rohlfing, K. J., Fawcett, C. & Gredebäck, G. (2015). Is ostension any more than attention? *Scientific Reports*, 4(1). <https://doi.org/10.1038/srep05304>
- Thom, E. E. & Sandhofer, C. M. (2009). More is more: The relationship between vocabulary size and word extension. *Journal of Experimental Child Psychology*, 104(4), 466–473. <https://doi.org/10.1016/j.jecp.2009.07.004>
- Tincoff, R. & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-

- month-olds. *Psychological Science*, 10(2), 172–175. <https://doi.org/10.1111/1467-9280.00127>
- Tincoff, R. & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, 17(4), 432–444. <https://doi.org/10.1111/j.1532-7078.2011.00084.x>
- Tobii AB (2016). Tobii Studio user’s manual. Version 3.4.5. Zugriff am 16.9.2019. Verfügbar unter <https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf>
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Hrsg.), *Joint attention: Its origins and role in development* (S. 103–130). Hillsdale, England: Lawrence Erlbaum Associates.
- Tomasello, M. (1999). The human adaptation for culture. *Annual Review of Anthropology*, 28(1), 509–529. <https://doi.org/10.1146/annurev.anthro.28.1.509>
- Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics*, 10(4), 401–413. <https://doi.org/10.1075/prag.10.4.01tom>
- Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. *Language acquisition and conceptual development*, 3, 132–158.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(05), 675–691. <https://doi.org/10.1017/S0140525X05000129>
- Tomasello, M. & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 57(6), 1454. <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>
- Trevarthen, C. (2011). What is it like to be a person who knows nothing? Defining the active intersubjective mind of a newborn human being. *Infant and Child Development*, 20(1), 119–135. <https://doi.org/10.1002/icd.689>
- Trevarthen, C. & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(1), 3–48.
- Troseth, G. L., Strouse, G. A., Verdine, B. N. & Saylor, M. M. (2018). Let’s chat: On-screen social responsiveness is not sufficient to support toddlers’ word learning

- from video. *Frontiers in Psychology*, 9, 2195. <https://doi.org/10.3389/fpsyg.2018.02195>
- Ulrich, W. (2002). *Wörterbuch: Linguistische Grundbegriffe* (5. Aufl.). Berlin, Stuttgart: Gebrüder Borntraeger.
- Vlach, H. A. & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00046>
- Vouloumanos, A., Hauser, M. D., Werker, J. F. & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development*, 81(2), 517–527. <https://doi.org/10.1111/j.1467-8624.2009.01412.x>
- Vouloumanos, A. & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science*, 7(3), 270–276. <https://doi.org/10.1111/j.1467-7687.2004.00345.x>
- Vouloumanos, A. & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10(2), 159–164. <https://doi.org/10.1111/j.1467-7687.2007.00549.x>
- Wass, S. V., Clackson, K., Georgieva, S. D., Brightman, L., Nutbrown, R. & Leong, V. (2018). Infants' visual sustained attention is higher during joint play than solo play: Is this due to increased endogenous attention control or exogenous stimulus capture? *Developmental Science*, 21(6), e12667. <https://doi.org/10.1111/desc.12667>
- Wass, S. V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R. et al. (2018). Parental neural responsivity to infants' visual attention: How mature brains influence immature brains during social interaction. *PLoS biology*, 16(12), e2006328. <https://doi.org/10.1371/journal.pbio.2006328>
- Waxman, S. R. & Booth, A. E. (2000). Principles that are invoked in the acquisition of words, but not facts. *Cognition*, 77(2), B33–B43. [https://doi.org/10.1016/S0010-0277\(00\)00103-7](https://doi.org/10.1016/S0010-0277(00)00103-7)
- Weinberg, M. K. & Tronick, E. Z. (1996). Infant affective reactions to the resumption of maternal interaction after the still-face. *Child development*, 67(3), 905–914.
- Werker, J. F. & Stager, C. L. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 388. 381–382.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M. & Stager, C. L. (1998). Acquisition

- of word-object associations by 14-month-old infants. *Developmental psychology*, 34(6), 1289.
- White, K. G. (2001). Forgetting functions. *Animal Learning & Behavior*, 29(3), 193–207. <https://doi.org/10.3758/BF03192887>
- Wilcox, B. M. & Clayton, F. L. (1968). Infant visual fixation on motion pictures of the human face. *Journal of Experimental Child Psychology*, 6(1), 22–32. [https://doi.org/10.1016/0022-0965\(68\)90068-4](https://doi.org/10.1016/0022-0965(68)90068-4)
- Wildt, E., Rohlfing, K. J. & Scharlau, I. (2019). The role of saliency in learning first words. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01150>
- Wildt, E. & Rohlfing, K. J. (2018, Juli). *What type of interactional presentation helps 10-month-olds to overcome the saliency-effect during referent selection?* Posterbeitrag auf dem International Congress of Infant Studies, Philadelphia.
- Wilson, D. & Sperber, D. (2012). *Meaning and Relevance*. Cambridge University Press.
- Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, 55(1), 235–269. <https://doi.org/10.1146/annurev.psych.55.090902.141555>
- Wojcik, E. H. (2013). Remembering new words: Integrating early memory development into word learning. *Frontiers in Developmental Psychology*, 4, 151. <https://doi.org/10.3389/fpsyg.2013.00151>
- Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, 6(3), 297–311. <https://doi.org/10.1111/1467-7687.00286>
- Woodward, A. L., Markman, E. M. & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, 30(4), 553–566. <https://doi.org/10.1037/0012-1649.30.4.553>
- Wu, R., Tummeltshammer, K. S., Gliga, T. & Kirkham, N. Z. (2014). Ostensive signals support learning from novel attention cues during infancy. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00251>
- Yont, K. M., Snow, C. E. & Vernon-Feagans, L. (2003). The role of context in mother-child interactions: An analysis of communicative intents expressed during toy play and book reading with 12-month-olds. *Journal of Pragmatics*, 35, 435–454. [https://doi.org/10.1016/S0378-2166\(02\)00144-3](https://doi.org/10.1016/S0378-2166(02)00144-3)

- Yoshida, H. & Burling, J. M. (2012). Highlighting: A mechanism relevant for word learning. *Frontiers in Developmental Psychology*, 3, 262. <https://doi.org/10.3389/fpsyg.2012.00262>
- Yu, C., Slone, L. & Smith, L. (2018, Juli). Quantity meets quality: Examining a statistical solution of referential uncertainty in parent-child naturalistic interaction. In *Everyday inputs and learning*. Symposium conducted at the International Congress of Infant Studies, Philadelphia.
- 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*, 26(9), 1235–1240. <https://doi.org/10.1016/j.cub.2016.03.026>
- Yurovsky, D. & Frank, M. C. (2017). Beyond naïve cue combination: Salience and social cues in early word learning. *Developmental Science*, 20(2), e12349. <https://doi.org/10.1111/desc.12349>
- Zangl, R. & Fernald, A. (2007). Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development*, 3(3), 199–231. <https://doi.org/10.1080/15475440701360564>
- Zukow-Goldring, P. (1996). Sensitive caregiving fosters the comprehension of speech: When gestures speak louder than words. *Early Development and Parenting*, 5(4), 195–211. [https://doi.org/10.1002/\(SICI\)1099-0917\(199612\)5:4<195::AID-EDP133>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1099-0917(199612)5:4<195::AID-EDP133>3.0.CO;2-H)

9 Anhang

Anhangsverzeichnis

Anhang A: Tabellen

Anhang B: Informationsblatt für die Erziehungsberechtigten

Anhang C: Einverständniserklärung

Anhang D: Demographische Daten

Anhang E: Studienprotokoll

Anhang F: Erklärung

Anhang A: Tabellen

Tabelle 24

Randomisierte Seite der Objekte [(1) Maiskolben und (2) Schachtelpuppe, (3) Schiff und (4) Klotz] im ersten und zweiten Versuch sowie die randomisierte Zuordnung der Objekt-namen (Doma, Modi, Guse, Nebo) zu den Objekten in den acht Randomisierungsgruppen (Experiment 1)

Randomisie- rungsgruppe	Alter und Geschlecht der Probanden	Versuch 1		Versuch 2	
1	$N = 10$, davon $N_{\text{weiblich}} = 5$ (davon $N_{JE} = 2$, $N_{SE} = 3$), $N_{\text{männlich}} = 5$ (davon $N_{JE} = 3$, $N_{SE} = 2$)	1 Doma	2 Modi	3 Guse	4 Nebo
2	$N = 8$, davon $N_{\text{weiblich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$), $N_{\text{männlich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$)	2 Doma	1 Modi	4 Guse	3 Nebo
3	$N = 8$, davon $N_{\text{weiblich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$), $N_{\text{männlich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$)	1 Doma	2 Modi	4 Guse	3 Nebo
4	$N = 8$, davon $N_{\text{weiblich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$), $N_{\text{männlich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$)	2 Doma	1 Modi	3 Guse	4 Nebo
5	$N = 8$, davon $N_{\text{weiblich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$), $N_{\text{männlich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$)	3 Doma	4 Modi	1 Guse	2 Nebo
6	$N = 8$, davon $N_{\text{weiblich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$), $N_{\text{männlich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$)	4 Doma	3 Modi	2 Guse	1 Nebo
7	$N = 8$, davon $N_{\text{weiblich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$), $N_{\text{männlich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$)	3 Doma	4 Modi	2 Guse	1 Nebo
8	$N = 8$, davon $N_{\text{weiblich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$), $N_{\text{männlich}} = 4$ (davon $N_{JE} = 2$, $N_{SE} = 2$)	4 Doma	3 Modi	1 Guse	2 Nebo

* SE = Self-Exploring-Gruppe, JE = Joint-Engagement-Gruppe

Tabelle 25

Prozentuale Blickdauer zum Ablenker während der Offset-Testphasen im Vergleich zum Wahrscheinlichkeitswert von 50 % (Experimente 1–3)

	Blickdauer zum Ablenker									
	Joint Engagement					Self Exploring				
Experiment 1	<i>M</i>	<i>SD</i>	<i>t</i> (32)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (32)	<i>p</i>	<i>r</i>
Original Label (6 Sek.)	51.82	15.07	0.69	.493	.12	54.7	14.72	1.83	.076†	.31
New Label	54.7	22.05	1.22	.23	.21	57.35	20.21	2.09	.045*	.35
Recovery	48.03	19.69	−0.58	.57	.10	57.03	19.61	2.06	.048*	.34
Original Label (3 Sek.)	53.32	16.67	1.15	.26	.20	56.85	13.05	3.01	.005**	.47
New Label	53.33	21.34	0.89	.38	.16	56.62	18.89	2.01	.047†	.34
Recovery	52.91	26.21	0.64	.53	.11	49.69	26.59	−0.07	.947	.01
Kontrollexperiment 1	<i>M</i>	<i>SD</i>	<i>t</i> (5)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (5)	<i>p</i>	<i>r</i>
Original Label	45.57	16.46	1.1	.539	.44	56.58	7.66	2.10	.089†	.69
New Label	74.98	7.712	3.1	.001***	.81	75.24	21.68	2.85	.036*	.79
Recovery	62.35	26.53	2.492	.306	.74	85.59	12.86	6.782	<.001***	.95
Kontrollexperiment 2 – Tag1	<i>M</i>	<i>SD</i>	<i>t</i> (11)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (11)	<i>p</i>	<i>r</i>
Original Label	53.29	17.95	0.635	.539	.19	59.70	9.52	3.532	.005**	.73
New Label	68.32	23.53	2.697	.021*	.63	64.95	24.15	2.145	.055†	.54
Recovery	62.99	21.89	2.055	.064†	.53	74.35	22.81	3.698	.004**	.74
Kontrollexperiment 2 – Tag 2	<i>M</i>	<i>SD</i>	<i>t</i> (5)	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i> (5)	<i>p</i>	<i>r</i>
Original Label	61.68	13.459	2.126	.087†	.69	57.19	10.692	1.646	.161	.59
New Label	41.26	20.893	−1.025	.352	.42	76.29	23.664	2.722	.042*	.77
Recovery	50.96	14.856	0.158	.881	.07	72.67	7.372	7.534	.001***	.96

Anhang B: Informationsblatt für die Erziehungsberechtigten (Vorderseite)



SPRACHSPIELLABOR



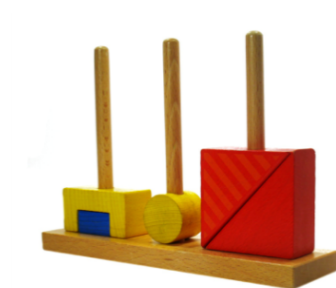
Liebe Eltern,

in unserem SprachSpielLabor untersuchen wir in kindgerechten Situationen die Ursprünge des Worterwerbs bei Kleinkindern.



Im Rahmen unserer aktuellen Studie suchen wir

- Eltern mit ihren **8-10 Monate** alten Kindern,
- die uns für ca. **30 Minuten**
- im **Technologiepark der Universität Paderborn** besuchen



Anhang B: Informationsblatt für die Erziehungsberechtigten (Rückseite)



Während des Experiments

- beobachten wir Ihr Kind in einer natürlichen Spielsituation
- messen wir die Augenbewegung mithilfe eines Geräts (ohne störende Kabel, Kopfstützen oder Brillen!)

Anschließend wartet ein Geschenk auf Ihr Kind!

Bei Fragen und Interesse melden Sie sich gerne telefonisch oder per E-Mail.




Wir freuen uns auf Ihren Besuch!

Kontakt:

Eugenia Wildt
M.A. Linguistin
05251 / 60-4309
(Mo-Fr 9-17 Uhr)

SprachSpielLabor
Technologiepark 21
33098 Paderborn

eugenia.wildt@uni-paderborn.de

 <https://www.facebook.com/Psycholinguistik/>



Anhang C: Einverständniserklärung



Universität Paderborn | Technologiepark 21 | 33098 Paderborn

Early Semantic Development (EASE)

SPRACHSPIELLABOR

Prof. Dr.
Katharina J. Rohlfing

Technologiepark 21
33098 Paderborn

Raum TP21.2.07
Tel. 05251 / 60-5717
E-Mail katharina.rohlfing@uni-paderborn.de

Einwilligungserklärung der Eltern zur Teilnahme an einem Forschungsprojekt:

Studie zur frühkindlichen Sprachentwicklung

Die Studie wird von Eugenia Wildt unter der Leitung von Prof. Dr. Katharina Rohlfing durchgeführt. Wir bitten Sie, dieses Formular zu lesen und mögliche Fragen zu stellen, bevor Sie sich einverstanden erklären, an der Studie teilzunehmen.

Ziel der Studie

Das Ziel dieser Studie ist es, zu untersuchen, worauf Kinder im Alter von 10 Monaten achten. Richtet sich die Aufmerksamkeit eines Kindes auf einen farblich auffälligen Gegenstand oder einen, der durch soziale Hinweise (z. B. Blickrichtung) hervorgehoben wird? Uns interessiert, welche dieser Strategien Kinder für das Wortlernen nutzen. Da 10 Monate alte Kinder noch keine Wörter selbst sprechen können, aber bereits ein großes Sprachverständnis haben, nutzen wir die Eye Tracking Methode. Ein Eye Tracker erfasst die Augenbewegung und ermöglicht uns, nachzuverfolgen, auf welches Bild Ihr Kind genau schaut, wenn es ein Wort hört. Dadurch können wir Aussagen über das Sprachverständnis treffen.

Beschreibung der Studie

Falls Sie sich entschließen, an der Studie teilzunehmen, laden wir Sie herzlich zu uns ins SprachSpielLabor der Universität Paderborn ein:

Als Erstes interessiert uns, auf welche Art von Objekte Kinder im Alter von 10 Monaten ihre Aufmerksamkeit richten. Dafür lassen wir sie mit zwei unterschiedlichen Objekten spielen. Im Anschluss daran laden wir Sie zu einer Eye Tracking Untersuchung ein: Während Ihr Kind auf Ihrem Schoß sitzt, werden ihm auf einem Bildschirm Bilder von den Objekten präsentiert. Wir untersuchen hierbei, wie lange Ihr Kind die präsentierten Bilder anschaut. Ihr Kind hat dabei keine Einschränkungen in seiner Kopfbewegung, da wir keine störenden Kabel, Kopfstützen oder Brillen verwenden.

Danach bitten wir Sie, sich mit Ihrem Kind auf dem Schoß an einen Tisch zu setzen. Eine Mitarbeiterin wird Ihrem Kind die beiden Gegenstände nochmal zeigen. Aber nur eines dieser Objekte wird durch eine Handlung hervorgehoben und benannt. Wir untersuchen, worauf Kinder in dieser Situation achten – auf das Objekt, das durch soziale Hinweise hervorgehoben wird, oder auf das bunte Spielzeug?

Schließlich wird das von Ihrem Kind erworbene Wissen erneut mithilfe eines Eye Trackers getestet.

Das gesamte Experiment dauert ungefähr 15 – 30 Minuten. Sie werden zu keiner Zeit Ihres Besuchs von Ihrem Kind getrennt. Falls Ihr Kind aus irgendeinem Grund nicht weitermachen möchte, beenden wir die Spielsituation oder die Untersuchung mit dem Eye Tracker. Sie können das Experiment ebenfalls jederzeit, aus jeglichem Grund beenden, ohne dadurch einen Nachteil zu erfahren.

Risiko und mögliche Unannehmlichkeiten

Das Experiment beinhaltet ein geringes Risiko: Um vergleichbares Lernen für alle Kinder zu gewährleisten, setzen wir unbekannte Wörter / Kunstwörter ein. Während des Besuchs wird Ihr Kind ein Kunstwort hören oder vielleicht sogar 'aufschnappen', das es in der Normalsprache nicht gibt. Unserer Erfahrung nach gefallen Kindern unsere Sprachspiele, und sie sind an den Objekten und Handlungen interessiert, die wir ihnen zeigen. Wir werden jeglichen Versuch unternehmen, die Situation für Sie und Ihr Kind so angenehm wie möglich zu gestalten. Falls Ihr Kind müde wird oder das Interesse verliert, können Sie die Teilnahme entweder vorübergehend oder dauerhaft abbrechen.

Nutzen der Studie

Die Teilnahme an dieser Studie verschafft Ihnen oder Ihrem Kind keinen direkten Nutzen, aber es hilft uns, mehr über den komplexen Prozess der Sprachentwicklung von Kindern zu lernen. Unserer Erfahrung nach machen die Spielsituationen den Kindern viel Spaß und sie interagieren gerne mit den MitarbeiterInnen. Auch die Eltern beobachten gern ihre Kinder mit anderen in der Interaktion und schauen sich hinterher die erfassten Blickdaten ihrer Kinder auf dem Bildschirm an. Eine geplante Veröffentlichung der Daten erfolgt ausschließlich in anonymisierter Form voraussichtlich im Jahr 2019.

Datenverarbeitung

Mit Fragebögen erheben wir personenbezogene Daten (z. B. das Geburtsdatum Ihres Kindes, Schulabschluss der Eltern, o.ä.). Diese nutzen wir in anonymisierter Form zur Beschreibung der Stichprobe und zu Analyse Zwecken.

Die Interaktion Ihres Kindes mit der Mitarbeiterin am Tisch wird aus zwei Perspektiven auf Video aufgezeichnet. Unter Umständen können Sie darauf durch den Winkel der Kamera auch zu sehen sein. Der Eye Tracker erfasst Augenbewegungen, die später als Punkte und Striche auf dem Bildschirm zu sehen sind.

Diese Daten nutzen wir für weitere Analysen. Die Videoaufnahmen werden benutzt, um das Blickverhalten und die Ausführung der Objektfunktion während der Interaktion festzuhalten. Die Videoaufnahmen werden auf den Rechnern in unseren Laborräumen mit der Annotationssoftware ELAN kodiert. Nach der Kodierung sind die Daten anonymisiert, sodass keine Rückschlüsse mehr auf die Person gezogen werden können. Aus Eye Tracking Daten gewinnen wir genaue Blickzeiten, die wir in Zusammenhang mit der Interaktion betrachten können.

Vertraulichkeit

Jegliche Informationen aus dieser Studie, die Sie oder Ihr Kind betreffen, werden streng vertraulich behandelt. Veröffentlichungen zu dieser Studie werden keinerlei Informationen beinhalten, die Sie oder Ihr Kind identifizieren könnten.

Die Aufzeichnungen werden auf einem **geschützten Server der Universität Paderborn** aufbewahrt und nur die MitarbeiterInnen aus unserer Forschungsgruppe haben darauf Zugriff. Die Video- und Audiodaten werden von den MitarbeiterInnen der Forschungsgruppe bearbeitet, die bei der Kodierung der Videos von ForschungsassistentInnen unterstützt werden. Falls die MitarbeiterInnen oder die ForschungsassistentInnen Sie oder Ihr Kind kennt, könnte derjenige Sie auf den Aufnahmen erkennen. Die AssistentInnen haben jedoch keinen Zugriff auf Ihre personenbezogenen Daten (Einverständniserklärung, demographische Daten, Emailadresse oder Adresse).

Forschungsdaten (z. B. Transkriptionen), personenbezogene Daten (z. B. Einverständniserklärung) und Kontaktdaten (z. B. Telefonnummer) werden **getrennt voneinander in verschlossenen Räumen** unseres Labors aufbewahrt. Die studienbezogenen Daten werden ausschließlich unter einer ID-Nummer geführt. Ihre Kontaktdaten speichern wir in einer **digitalen Datenbank**, auf die nur die MitarbeiterInnen Zugriff haben.

Nur für den Fall, dass Sie zustimmen, werden die Eye Tracking-, Audio- und Videoaufnahmen von Ihrem Kind zu Lehrzwecken genutzt, jedoch werden keine Informationen über die Identität Ihres Kindes weitergegeben. Bitte beachten Sie, dass jemand, der Ihr Kind zu dem Zeitpunkt der Aufnahme kennt, Ihr Kind aus Audio- und Videoaufnahmen wiedererkennen könnte. Sie können auch dann an der Studie teilnehmen, wenn Sie nicht wünschen, dass die Interaktion zwischen Ihrem Kind und den MitarbeiterInnen oder Eye Tracking Daten zu Lehrzwecken benutzt werden.

Wenn Sie sich bereit erklären, dass wir einzelne Videosequenzen / Bilder / Eye Tracking Videos auf unserer Webseite / Facebook-Seite / unserem Newsletter veröffentlichen dürfen, werden wir vor einer Veröffentlichung Rücksprache mit Ihnen halten und das Bild- oder Videomaterial vorher zeigen.

Aufwandsentschädigung

Wir würden uns sehr freuen, wenn Sie und Ihr Kind an unserer Studie teilnehmen. Als Zeichen unserer Dankbarkeit möchten wir Ihrem Kind ein Spielzeug oder ein Buch schenken.

Recht auf Löschung

Die Teilnahme an dieser Studie ist freiwillig. Ihre Entscheidung über die Teilnahme wird Ihre zukünftigen Kontakte zur Universität Paderborn in keiner Weise beeinflussen. Wenn Sie sich zu einer Teilnahme entschließen, steht es Ihnen jederzeit frei, Ihr Einverständnis zurückzuziehen und Ihre Teilnahme vorzeitig zu beenden.

Sie haben jederzeit das Recht auf eine Auskunft über die zu Ihrer Person gespeicherten Daten zu erhalten sowie das Recht auf Löschung dieser Daten. Eine Löschung der Daten kann ab dem Zeitpunkt der Anonymisierung nicht mehr verlangt werden.

Auskunfts- und Beschwerderecht

Falls Sie jetzt Fragen zum Forschungsprojekt haben, äußern Sie diese bitte direkt. Wenn Sie später inhaltliche Fragen zur Forschung oder zu damit verbundenen Themen (z. B. dem Ablauf der Studie oder den Ergebnissen) haben, kontaktieren Sie bitte Frau Prof. Dr. Katharina J. Rohlfing. Fragen rund um den Datenschutz können Sie an die Datenschutzbeauftragten der Universität Paderborn richten (<https://www.uni-paderborn.de/universitaet/datenschutz/kontakt/>). Sie haben auch das Recht sich an die LDI NRW zu wenden (<https://www.lidi.nrw.de>), um mehr über datenschutzrechtliche Bestimmungen zu erfahren oder Beschwerden zu melden.

Einverständniserklärung (bitte behalten Sie diese Seite mit Informationen und Vereinbarung!)

Mir ist bekannt, dass es das Ziel dieser Studie ist, zu untersuchen, ob Kleinkinder eher auf interessante Objekte oder auf soziale Hinweise achten und welche dieser Lernstrategien sie für den Spracherwerb nutzen. Die Teilnahme meines Kindes an dieser Studie ist freiwillig.

Die Daten werden für dieses Forschungsvorhaben verwendet. Des Weiteren bin ich darüber in Kenntnis gesetzt worden, dass ich die Teilnahme meines Kindes an der Studie zum Wortlernen ohne Angabe von Gründen jederzeit beenden kann. Ich kann innerhalb der nächsten 10 Jahre meine Einwilligung der Teilnahme, ohne Angabe von Gründen, mit Wirkung für die Zukunft für mein Kind widerrufen, ohne dass mir oder meinem Kind dadurch Nachteile entstehen. Meine personenbezogenen Daten bzw. die personenbezogenen Daten meines Kindes werden in diesem Falle umgehend gelöscht.

Hiermit erkläre ich, dass ich die Informationen in dieser Einverständniserklärung gelesen habe und dass ich Gelegenheit hatte, Fragen zur Studie zu stellen. Ich erkläre mich damit einverstanden, dass

_____ an der Studie teilnimmt und dass:

(Vorname/ Nachname des Kindes)

- ☐ die erhobenen Daten (Videoaufnahmen und Eye Tracking Daten) meines Kindes wie beschrieben verarbeitet werden,
-
- ☐ einzelne Video- bzw. Tonsequenzen zu Lehr-, Trainings- und Vortragszwecken Dritten vorgeführt und veröffentlicht werden
- ☐ nach Rücksprache einzelne *Videosequenzen/ Bilder / Eye-Tracking Videos* auf unserer *Webseite / Facebook-Seite / auf Flyern / Newslettern* veröffentlicht werden dürfen
- ☐ diese Daten lediglich im Rahmen dieses Kooperationsvertrages an Dritte weitergegeben werden können (z. B. zum Zweck für die Verfassung von Bachelor oder Masterarbeiten).

Name des Elternteils:

Unterschrift:

Unterschrift der Experimentatorin:

Datum:

Einverständniserklärung

Mir ist bekannt, dass es das Ziel dieser Studie ist, zu untersuchen, ob Kleinkinder eher auf interessante Objekte oder auf soziale Hinweise achten und welche dieser Lernstrategien sie für den Spracherwerb nutzen. Die Teilnahme meines Kindes an dieser Studie ist freiwillig.

Die Daten werden für dieses Forschungsvorhaben verwendet. Des Weiteren bin ich darüber in Kenntnis gesetzt worden, dass ich die Teilnahme meines Kindes an der Studie zum Wortlernen ohne Angabe von Gründen jederzeit beenden kann. Ich kann innerhalb der nächsten 10 Jahre meine Einwilligung der Teilnahme, ohne Angabe von Gründen, mit Wirkung für die Zukunft für mein Kind widerrufen, ohne dass mir oder meinem Kind dadurch Nachteile entstehen. Meine personenbezogenen Daten bzw. die personenbezogenen Daten meines Kindes werden in diesem Falle umgehend gelöscht.

Hiermit erkläre ich, dass ich die Informationen in dieser Einverständniserklärung gelesen habe und dass ich Gelegenheit hatte, Fragen zur Studie zu stellen. Ich erkläre mich damit einverstanden, dass

_____ an der Studie teilnimmt und dass:
(Vorname/ Nachname des Kindes)

- ☐ die erhobenen Daten (Videoaufnahmen und Eye Tracking Daten) meines Kindes wie beschrieben verarbeitet werden,

- ☐ einzelne Video- bzw. Tonsequenzen zu Lehr-, Trainings- und Vortragszwecken Dritten vorgeführt und veröffentlicht werden
- ☐ nach Rücksprache einzelne *Videosequenzen/ Bilder / Eye-Tracking Videos* auf unserer *Webseite / Facebook-Seite / auf Flyern / Newslettern* veröffentlicht werden dürfen
- ☐ diese Daten lediglich im Rahmen dieses Kooperationsvertrages an Dritte weitergegeben werden können (z. B. zum Zweck für die Verfassung von Bachelor oder Masterarbeiten).

Name des Elternteils:

Unterschrift:

Unterschrift der Experimentatorin:

Datum:

Anhang D: Demographische Daten**Demographische Daten**

Diese Daten werden aus statistischen Gründen erhoben, ihre Angabe ist freiwillig. Die Speicherung Ihrer in der Studie erhobenen Daten erfolgt jedoch selbstverständlich anonym. Vielen Dank für Ihre Angaben!

Geburtsdatum: _____

Geschlecht: männlich ☐ weiblich ☐

1. Schulbildung der Mutter (Abschluss)

Hauptschulabschluss	<input type="radio"/>
Realschulabschluss	<input type="radio"/>
Abitur/Fachabitur	<input type="radio"/>
Fachhochschulabschluss	<input type="radio"/>
Hochschulabschluss	<input type="radio"/>
Promotion	<input type="radio"/>

Wie viel Zeit verbringen Sie täglich ca. mit Ihrem Kind (bitte ankreuzen)?

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bis zu 1,5 Std.	3 Std.	4-7 Std.	den ganzen Tag

2. Schulbildung des Vaters (Abschluss)

Hauptschulabschluss	<input type="radio"/>
Realschulabschluss	<input type="radio"/>
Abitur/Fachabitur	<input type="radio"/>
Fachhochschulabschluss	<input type="radio"/>
Hochschulabschluss	<input type="radio"/>
Promotion	<input type="radio"/>

Wie viel Zeit verbringen Sie täglich ca. mit Ihrem Kind (bitte ankreuzen)?

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bis zu 1,5 Std.	3 Std.	4-7 Std.	den ganzen Tag

Bitte beantworten Sie nun noch einige Fragen zu Ihrem Kind:

1. Ist Ihr Kind früher geboren? Ja o Nein o

Falls ja, würden Sie sein Geburtsgewicht (ca.) angeben und den eigentlichen
Geburtsstermin?

Geburtsgewicht: _____

Geburtsstermin: _____

2. Sind Seh- und Hörvermögen normal ausgeprägt?

Ja o Nein o Falls nein, welches? _____

Vielen Dank für das Ausfüllen des Fragebogens!

Anhang E: Studienprotokoll



Protokoll

SPRACHSPIELLABOR

TAG 1:

ID _____ Rec _____

Studie _____

Randomisierungsgruppe _____

Datum _____

Uhrzeit _____

ET-Experimentatorin _____

IA-Experimentatorin _____

1) Informationen über die Versuchsperson

Geburtsdatum _____

Alter zum Zeitpunkt des Experiments _____

Geschlecht ☐ männlich ☐ weiblichGruppe ☐ Interaktion ☐ keine Interaktion

2) Eye Tracking - Kalibrierung

Positionierung	Kalibrierung (Anzahl der Versuche)	Anzahl der kali- brierten Punkte	Datenqualität	Pausen
<input type="checkbox"/> auf dem Schoß	1	1 2 3 4 5	<input type="checkbox"/> sehr gut	0 1 2 3
<input type="checkbox"/> im Hochstuhl	2	1 2 3 4 5	<input type="checkbox"/> gut	
	3	1 2 3 4 5	<input type="checkbox"/> befriedigend*	
	4	1 2 3 4 5	<input type="checkbox"/> ausreichend*	
			<input type="checkbox"/> keine*	

*Grund: ☐ Fehler seitens des Experimentators, ☐ Verhalten des Kindes, ☐ Verhalten des Elternteils, ☐ technischer Fehler, ☐ Sonstiges _____

Im Labor befanden sich

☐ Mutter ☐ Vater ☐ Experimentatorin 1 ☐ Experimentatorin 2 (Student/in)☐ Sonstige _____

ID _____ Rec _____

Wurde das Experiment frühzeitig beendet?

☐ Nein ☐ Ja Grund _____

3) Eye Tracking

Veränderung der Sitzposition*	Bewegung	Weinen	Lachen	Schnuller	Hält etw. in der Hand	Zeigt auf den Monitor	Vokalisation
<input type="checkbox"/> JA	<input type="checkbox"/> WENIG	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA
<input type="checkbox"/> NEIN	<input type="checkbox"/> VIEL	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN

*Wie wurde die Position verändert? _____

Phase	weighted gaze samples (%)	gaze samples (%)
S 1.		
T 1.		
S 2.		
T 2.		

4) Sonstiges

Geschenk: ☐ Zertifikat
☐ Buch
☐ Spielzeug _____

Weitere Kommentare:

TAG 2:

ID _____

Rec _____

Datum _____

Uhrzeit _____

ET-Experimentatorin _____

IA-Experimentatorin _____

Im Labor befanden sich

☐ Mutter ☐ Vater ☐ Experimentatorin 1 ☐ Experimentatorin 2 (Student/in)☐ Sonstige _____

Wurde das Experiment frühzeitig beendet?

☐ Nein ☐ Ja Grund _____**Eye Tracking**

Veränderung der Sitzposition*	Bewegung	Weinen	Lachen	Schnuller	Hält etw. in der Hand	Zeigt auf den Monitor	Vokalisation
<input type="checkbox"/> JA	<input type="checkbox"/> WENIG	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA	<input type="checkbox"/> JA
<input type="checkbox"/> NEIN	<input type="checkbox"/> VIEL	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN	<input type="checkbox"/> NEIN

*Wie wurde die Position verändert? _____

Phase	weighted gaze samples (%)	gaze samples (%)
T1.		
T2.		

3) Schlaf

Wie oft hat das Kind in der Zeit zwischen der ersten und zweiten Erhebung geschlafen?

Wie lange hat das Kind insgesamt in der Zeit zwischen der ersten und zweiten Erhebung geschlafen?

4) Sonstiges

Geschenk: ☐ Zertifikat
☐ Buch
☐ Spielzeug _____