



Kommunikation in Echtzeit – die Einflüsse von
kindgerichteter Sprache und kindlicher Aufmerksamkeit auf
den Spracherwerb

Vivien Outters

Kommunikation in Echtzeit – die Einflüsse von kindgerichteter Sprache und kindlicher Aufmerksamkeit auf den Spracherwerb

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsprogramm Behavior and Cognition

der Georg-August University School of Science (GAUSS)

vorgelegt von

Vivien Outters (geb. Radtke)

aus Neubrandenburg

Göttingen, 2021

Betreuungsausschuss

Prof. Dr. Nivedita Mani
Psychologie der Sprache
Georg-August-Universität Göttingen

Dr. Tanya Behne
Kognitive Entwicklungspsychologie
Georg-August-Universität Göttingen

Prof. Dr. Julia Fischer
Kognitive Ethnologie
Deutsches Primatenzentrum

Prüfungskommission

Prof. Dr. Nivedita Mani
Psychologie der Sprache
Georg-August-Universität Göttingen

Dr. Tanya Behne
Kognitive Entwicklungspsychologie
Georg-August-Universität Göttingen

Prof. Dr. Julia Fischer
Kognitive Ethnologie
Deutsches Primatenzentrum

Weitere Mitglieder der Prüfungskommission

Prof. Dr. Hannes Racozy
Entwicklungspsychologie
Georg-August-Universität Göttingen

Prof. Dr. Markus Steinbach
Germanistische Linguistik
Georg-August-Universität Göttingen

Prof. Dr. Annkathrin Schacht
Affektive Neurowissenschaft und Psychophysiologie
Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 28.05.2021

Danksagung

Ich möchte vor allen den Kindern danken, ohne die diese Arbeit unmöglich gewesen wäre, aber auch ihren Eltern, die lange EEG Experimente überstanden haben und immerwieder bereit waren, ihre Kinder bei uns mitmachen zu lassen.

Ein besonderer Dank geht an Nivi, die mich immerwieder weiter voran gebracht hat und nie aufgehört hat an mich zu glauben. Ebenso möchte ich Nivi, Tanya und Julia danken für die Ideen und die Diskussionen, die mich immer weiter gebracht haben und schließlich mit der Abgabe dieser Dissertation enden.

Ein großer Dank geht auch an meine Kollegen, die ein wunderbares Arbeitsklima geschaffen haben. Vor allem möchte ich Sarah und Melli danken, dafür dass sie mit mir Ideen diskutiert haben und mir gezeigt haben, wie passioniert man sein kann.

Die Teilhabe an der RTG „Verstehen von Sozialbeziehungen“ hat mir die Interdisziplinarität und die Vielfalt der Wissenschaft immerwieder vor Augen geführt. Vielen Dank für die Unterstützung, Diskussionen und netten Abende! Ebenso möchte ich mich dafür bedanken, Teil des Science Campus gewesen zu sein.

Ein großer Dank geht auch an meine Familie, die mich in den letzten Jahren immerwieder unterstützt und an mich geglaubt hat. Eure Zuversicht, gute Zureden und Ablenkung haben es möglich gemacht, dass ich diese Arbeit beendet habe. Ich danke Euch dafür!

Abstract

In order to segment, understand and ultimately produce words from the flow of language, a series of learning processes are required. These processes are promoted by various extrinsic factors and intrinsic biases. This dissertation presents three studies that focus on the extrinsic factor of how we, as the environment, address children when we speak (infant-directed speech; IDS) and the children's intrinsic motivation to learn.

The first study examined how the mother's input, the child's preference and the child's involvement in social situations are related to one another. We were able to show that there is a correlation between the mother's input and the preference for a language style in both 6-month-old children and the 13-month-old children. It was also shown that the mother's input in the chosen setting did not vary between the two age groups. Children aged 6 months showed a clear preference for IDS over adult directed speech (ADS) compared to 13-month-old children. When examining the social involvement in connection with IDS and ADS, there was no difference between the age groups: both followed a person's gaze better if the person was previously associated with IDS compared to ADS. Nonetheless, the results highlight a strong correlation between the interaction of child preference and maternal input. This finding could have an impact on our understanding of how children can learn from IDS.

In the second study of this dissertation, the aim was to connect precisely this relationship between children's preference and maternal input with a learning task in the form of a segmentation task. For this purpose, the data collected in the first study was supposed to be combined with EEG data from a segmentation task: We conducted an additional study with the same children to investigate whether there is a connection between maternal input, children's preference and the children's ability to segment words from IDS and ADS. Due to the poor quality and limited amount of data, we rejected to investigate the interactions and only looked at the EEG data. This study revealed that children between the ages of 6 and 13 months showed a response to IDS, regardless of their ability to segment words from a continuous speech stream. This speaks in favor of an attention-directing effect of IDS, which may outshine the learning processes.

In the third study of this dissertation, we wanted to examine children's motivation to learn and determine whether this has an influence on learning performance. Children aged 3 to 4 years participated in the study. The children's eye movements, changes in pupil size and posture were recorded. We were unable to establish a connection between children's motivation to learn and the learning success, nor were we able to provide evidence that the children are aware of their performance or show a change in their posture through better evidence of learning performance. Nonetheless, we were able to show that the attention of the children, characterized by their pupil size, during the test phase correlates with their performance in this phase. This study underlines the important role of children's motivation and attention for good performance in such word learning tasks.

Overall, this dissertation speaks in favour of an attention-grabbing function of IDS, which allows children to communicate with their parents in real time. It was also shown that fast mapping in a word learning task is primarily used for real-time communication and that the fast learning process has no effect on the children's emotions.

Zusammenfassung

Um Wörter aus dem Sprachfluss zu segmentieren, zu verstehen und schließlich zu produzieren bedarf es einer Reihe an Lernprozessen. Diese Prozesse werden durch unterschiedliche extrinsische Faktoren und intrinsische Biases gefördert. Diese Dissertation präsentiert drei Studien, die sich auf den extrinsischen Faktor konzentrieren, wie wir als Umwelt Kinder ansprechen, wenn wir sprechen (kindgerechte Sprache), und auf die intrinsische Lernmotivation der Kinder.

In der ersten Studie wurde untersucht wie der Input der Mutter, die kindliche Präferenz und das kindliche Engagement in sozialen Situationen miteinander in Verbindung stehen. Wir konnten zeigen, dass es eine Korrelation zwischen dem Input der Mutter und der Präferenz für einen Sprachstil sowohl bei den 6-Monate alten Kindern, als auch bei den 13-Monate alten Kindern gibt. Des Weiteren zeigte sich, dass der Input der Mutter nicht zwischen den beiden Altersgruppen variiert. Kinder mit 6 Monaten zeigten eine deutliche Präferenz für kindgerichtete Sprache gegenüber erwachsenengerichteter Sprache im Vergleich zu den 13 Monate alten Kindern. Bei der Untersuchung des Engagements in Verbindung mit KGS und EGS, zeigte sich kein Unterschied zwischen den Altersgruppe: beide folgten der Blickbewegung einer Person besser, wenn diese vorher mit KGS in Verbindung gebracht wurde im Vergleich zu EGS. Die Ergebnisse heben vor allem einen starken Zusammenhang zwischen der Interaktion von kindlicher Präferenz und mütterlichen Input hervor. Dieses Ergebnis könnte einen Einfluss auf unser Verständnis haben, wie Kinder von KGS lernen können.

In der zweiten Studie dieser Dissertation war das Ziel genau diesen Zusammenhang von kindlicher Präferenz und mütterlichem Input mit einer Lernaufgabe, in Form einer Segmentationsaufgabe, in Verbindung zu bringen. Daz

u sollten die in der ersten Studie erhobenen Daten mit EEG Daten in Verbindung gesetzt werden. Wir führten mit den gleichen Kindern eine Studie durch, um zu untersuchen ob es einen Zusammenhang zwischen mütterlichem Input, kindlicher Präferenz und der Fähigkeit der Kinder aus KGS und EGS Wörter zu segmentieren, gibt. Aufgrund der Datenlage verwarfen wir die Untersuchung der Interaktionen und betrachteten lediglich die EEG Daten. In dieser Studie zeigte

sich, dass Kinder im Alter von 6 und 13 Monaten eine Reaktion auf KGS zeigen, losgelöst von ihren Fähigkeiten Wörter aus dem Sprachfluss zu segmentieren. Dies spricht für eine aufmerksamkeitslenkende Wirkung von KGS, die ggf. die Lernprozesse überschattet

In der dritten Studie dieser Dissertation untersuchten wir die kindliche Motivation zu lernen um feststellen, ob diese einen Einfluss auf die Lernleistung hat. An der Studie nahmen Kinder im Alter von 3 bis 4 Jahren teil. Die Augenbewegungen, die Veränderung der Pupillengröße und die Körperhaltung der Kinder wurden aufgenommen. Wir konnten in dieser Studie keinen Zusammenhang von kindlicher Motivation etwas zu lernen und dem Lernerfolg feststellen, ebenso wenig konnten wir Nachweise erbringen, dass die Kinder sich ihrer Performanz bewusst sind und durch einen besseren Nachweis der Lernleistung eine veränderte Körperhaltung zeigen. Nichtsdestotrotz konnten wir zeigen, dass die Aufmerksamkeit der Kinder, gekennzeichnet durch ihre Pupillengröße, während der Testphase mit der Performanz in eben dieser Phase hat korreliert. Diese Studie unterstreicht die wichtige Rolle der kindlichen Motivation und Aufmerksamkeit für die gute Performanz in solchen Wortlernertasks.

Insgesamt spricht diese Dissertation für eine aufmerksamkeitsfördernde Funktion von KGS, die den Kindern eine Kommunikation mit ihren Eltern in Echtzeit erlaubt. Ebenso, zeigte sich, dass *fast mapping* in einer Wortlernaufgabe vor allem für die Echtzeitkommunikation dient und der schnell stattfindende Lernprozess keine Auswirkungen auf die Emotionen der Kinder hat.

Inhaltsverzeichnis

| | |
|---|----|
| Einleitung | 14 |
| Lernen Sprache zu verstehen | 14 |
| Wie Kinder von dem Sprachinput lernen | 15 |
| Lernen Sprache zu produzieren | 16 |
| Überblick Dissertation | 17 |
| Wortlernen: Intrinsische und extrinsische Hinweise zum Wortlernen | 18 |
| Extrinsische Faktoren | 18 |
| Kind-gerichtete Sprache | 18 |
| Gestische Hinweise | 21 |
| Intrinsische Faktoren | 22 |
| Interesse und Neugier | 25 |
| Modelle | 28 |
| Die hier vorliegende Dissertation | 31 |
| Methoden | 32 |
| Aufnahmen der mütterlichen Sprache | 33 |
| Head turn preferenz - ‚Single Screen‘ Central fixation – Methode | 34 |
| Eye tracking | 35 |
| Messung der Pupillengröße | 36 |
| Segmentation – EEG | 37 |
| Aufnahme der Körperhaltung | 39 |
| Untersuchungspunkte der Studien und verwendete Methoden | 42 |
| Maternal input and infants’ response to infant-directed speech | 45 |
| Abstract | 45 |
| Introduction | 46 |
| What is infant directed Speech | 46 |
| IDS preference | 47 |
| The current study | 48 |
| Method | 51 |
| Participants | 51 |
| General Procedure | 51 |
| Recordings | 52 |

| | |
|------------------------------|----|
| Procedure | 52 |
| Data Analysis | 52 |
| Preferential listening task | 53 |
| Data and Design | 53 |
| Procedure | 53 |
| Data analysis | 54 |
| Performance across task | 55 |
| Gaze following | 56 |
| Material and Design | 56 |
| Procedure | 56 |
| Data Analysis | 56 |
| Performance across task | 57 |
| Results | 58 |
| Recording of mothers' speech | 58 |
| Preference | 59 |
| Correlation across task | 60 |
| Gaze following | 63 |
| Correlation across task | 64 |
| Discussion | 65 |
| Performance across task | 67 |
| Conclusion | 69 |
| Infants' word segmentation | 69 |
| Abstract | 71 |
| Introduction | 71 |
| EEG-components | 74 |
| Method | 76 |
| Participants | 76 |
| General Procedure | 77 |
| Material and Design | 77 |
| Stimuli | 77 |
| Procedure | 78 |
| Processing and Data Analysis | 78 |
| Performance across task | 79 |

| | |
|---|-----|
| Results | 80 |
| 300-500 ms | 80 |
| 500-700 ms | 80 |
| Discussion | 82 |
| The influence of children’s successful word segmentation on their internal arousal and body posture | 86 |
| Abstract | 86 |
| Introduction | 87 |
| Curiosity driven learning | 87 |
| Word learning and recognition | 90 |
| Pupil dilation | 90 |
| Body posture | 91 |
| The current study | 93 |
| Method | 94 |
| Participants | 94 |
| Stimuli | 94 |
| Procedure and Design | 95 |
| Training phase | 97 |
| Novel word recognition test phase | 98 |
| Known word recognition test phase | 99 |
| Data Analysis | 99 |
| Word recognition | 99 |
| Pupil dilation | 100 |
| Body posture | 101 |
| Analysis | 102 |
| Word recognition | 102 |
| Word recognition and body posture | 102 |
| Word recognition and pupil dilation | 103 |
| Pupil dilation and body posture | 104 |
| Results | 105 |
| Word recognition | 105 |
| Word recognition and body posture | 106 |
| Word recognition and pupil dilation | 111 |

| | |
|--|-----|
| Training phase | 111 |
| Novel word recognition test phase | 111 |
| Known word recognition test phase | 112 |
| Pupil dilation and body posture | 112 |
| Discussion | 112 |
| Word recognition and pupil dilation | 113 |
| Word recognition and body posture | 116 |
| Conclusion | 117 |
| Diskussion | 119 |
| Welche Rolle spielt die kindliche Aufmerksamkeit | 120 |
| Einordnung der Studien in Modelle | 122 |
| Zusammenfassung | 126 |
| Referenzen | 128 |
| Erklärung | 151 |
| Appendix | 152 |

Einleitung

Kinder erlernen innerhalb von kürzester Zeit Sprache. Vom Neugeborenen, das weder Sprache versteht noch produziert bis zum Kleinkind, welches die Muttersprache ohne Schwierigkeiten versteht und grammatikalische Sätze produzieren kann, vergehen nur 3 Jahre (Kuhl, 2004). In dieser Zeit müssen die Kinder eine große Anzahl von Problemen lösen, um aus dem Sprachinput und den Umgebungsbedingungen lernen zu können.

Lernen Sprache zu verstehen

Ein wichtiger Punkt um Sprache zu lernen, besteht darin, Wörter zu erlernen. Dies bedeutet, dass Kind muss lexikalische Informationen lernen und diese mit Dingen und Konzepten in Verbindung bringen. Dadurch kann auf konkrete oder abstrakte Dinge referiert werden und es können Zusammenhänge gebildet werden. Wörter repräsentieren dabei die kleinste bedeutungstragende Einheit in der Sprache. Um diese Aufgabe zu bewältigen, entwickeln Kinder eine Reihe erstaunlicher Fähigkeiten. Bereits direkt nach der Geburt zeigen Kinder eine Präferenz für Sprache gegenüber nicht sprachlichen Lauten oder rückwärts produzierter Sprache (Dehaene-Lambertz et al. 2002; Pena et al. 2003; Vouloumanos & Werker 2004, 2007). Außerdem bevorzugen Kinder die Stimme der eigenen Mutter (DeCasper & Fifer, 1980) und die eigene Muttersprache (Mehler et al., 1988; Moon et al., 1993) gegenüber fremden Stimmen und Sprachen. Im nächsten Schritt fangen Kinder an, nur die Lauten, die zur eigenen Muttersprache gehören, zu diskriminieren. In den ersten Lebensmonaten können Kinder noch zwischen den Lauten aller Sprachen unterscheiden, innerhalb des ersten Lebensjahres lernen Kinder jedoch die spezifischen Laute ihrer Muttersprache zu erkennen und verlernen die Diskriminierung von Lauten, die nicht in der eigenen Sprache genutzt werden (Rivera-Gaxiola et al., 2005; Werker et al., 1981; Werker & Tees, 1984). Dieser Schritt ist wichtig, um das eigene phonetische Inventar zu bilden und sich auf die Laute zu beschränken, die für die Bildung von Worten der eigenen Muttersprache wichtig sind.

Ebenfalls innerhalb des ersten Lebensjahres lernen Kinder Wörter aus dem kontinuierlichen Sprachfluss zu segmentieren. Bereits mit 5 Monaten können Kinder häufig verwendete Bezeichnungen wie Mama (Bortfeld et al., 2005) oder den eigenen Namen (Mandel et

al., 1995) aus dem Sprachfluss segmentieren. Kurz darauf, mit bereits 7,5 Monaten können Kinder vorher familiarisierte Worte aus einem in Kind-gerichteter Sprache (KGS) präsentierten Sprachfluss segmentieren, jedenfalls in amerikanischem Englisch (Jusczyk & Aslin, 1995). Diese Fähigkeit entwickelt sich innerhalb der nächsten Monate weiter, sodass Kinder bereits im Alter von 16 Monaten unbekannte Wörter aus erwachsenen-gerichteter Sprache (EGS) segmentieren (Mani & Pätzold, 2016).

Mit bereits 6 Monaten haben Kinder ebenfalls erste Bedeutungen für oft benutzte Nomen, wie Essenbezeichnungen oder Körperteile, gelernt (Bergelson & Swingley, 2012). Das immer größere werdende Wissen kann im zweiten Lebensjahr das Lernen beeinflussen. Mani und Plunkett (2010) haben gezeigt, dass phonologisch ähnliche Wörter Kindern mit 18 Monaten helfen können, Target-Wörter zu erkennen, mit 24 Monaten dies jedoch eher hinderlich ist (Mani & Plunkett, 2011). In demselben Alter helfen semantisch ähnliche Wörter den Kindern jedoch wieder Target Wörter zu erkennen (Arias-Trejo & Plunkett, 2010, 2013).

Wie Kinder von dem Sprachinput lernen

Doch wie können Kinder innerhalb von so kurzer Zeit so viel lernen? Eine Möglichkeit Sinn aus der Welt und somit auch aus Sprache zu machen, ist sie zu beobachten und Regularien darin zu finden (z.B. Saffran et al., 1996). Durch das Beobachten immer gleicher Abläufe lernen Kinder schon früh, sinnvolle von nicht sinnvollen Handlungen zu unterscheiden. So sind bereits 5,5 Monate alte Kinder überrascht, wenn eine Haarbürste zum Zähneputzen benutzt wird und schauen dieser Tätigkeit länger zu, als der bekannten Tätigkeit (Bahrick et al., 2002). Auch Sprache unterliegt gewissen Regeln, diese müssen Kinder verstehen und erlernen.

Um Sprache zu verstehen und Wörtern eine Bedeutung beimessen zu können, müssen Kinder lernen, Wörter aus dem Sprachfluss zu segmentieren. Dies ist vor allem deswegen schwer, da Wörter selten allein präsentiert werden und Wortgrenzen oft nicht zu hören sind (Klatt & Stevens, 1973). Die ihnen präsentierte Sprache bietet zwar gewisse akustische Hinweise (vor allem in KGS; für einen Überblick Cristia, 2013), jedoch nicht genügend um die Wortgrenzen im Sprachfluss zu finden.

Eine Möglichkeit für Kinder Wortgrenzen in dem kontinuierlichen Sprachfluss zu finden, besteht darin, das zusammenhängende Auftreten von Silben zu ‚beobachten‘ (z.B. Saffran et al., 1996). Durch die Beobachtung der statistischen Wahrscheinlichkeiten in Äußerungen wie *hübsches Baby* oder *mein Baby* können Kinder somit herausfinden, dass die Silbenkombination *Ba-by* öfter zusammen auftritt als die Silbenkombination *ches-ba* (vgl. Saffran et al., 1996). Zahlreiche Studien konnten nachweisen, dass Kinder bereits ab einem sehr frühen Alter in der Lage sind, solche statistischen Regularien zu erkennen und zu erlernen (z.B. Saffran et al., 1996; L. Smith & Yu, 2008). Doch nicht nur das Erkennen von Wortgrenzen kann auf statistisches Lernen zurückgeführt werden, auch das Problem zu erkennen welches Wort welches Objekt benennt, kann durch die Beobachtung des gemeinsamen Auftretens von Wort und Objekt *berechnet* werden (Yu & Smith, 2007, 2008). Dazu wird von den Kindern kalkuliert, welches Wort, relativ gesehen, häufiger mit einem bestimmten Objekt zusammen auftaucht. Studien im Labor konnten zeigen, dass Kinder schon ab einem Alter von 12 Monaten in der Lage sind, das Auftreten von Wort-Referent-Paaren über einen gewissen Zeitraum wahrzunehmen und das Auftreten der Wort-Objekt-Kombinationen zu analysieren (Smith & Yu, 2008). Smith und Yu (2008) konnten zeigen, dass Kinder nach einer kurzen *cross-situational* designten Lernphase in der Lage waren, eine Wort-Objekt-Assoziation aufzubauen (für gleiche Resultate mit Erwachsenen siehe Yu und Smith, 2007). Zusammenfassend lässt sich sagen, dass das Prinzip der statistisch assoziativen Informationsgewinnung einen großen Beitrag zum Sprachlernen der Kinder liefert.

Lernen Sprache zu produzieren

Die erste Wortproduktion findet meist um den ersten Geburtstag statt und beschränkt sich auf Ein-Wort-Äußerungen. Erst nach dem zweiten Geburtstag kommt es zu Mehr-Wort-Äußerungen. Mit 30 Monaten können Kinder bereits im Durchschnitt 550 Wörter produzieren, wobei die individuellen Unterschiede überwiegen und eine Spannbreite von 250-650 Wörtern normal ist (Frank et al., 2017).

Überblick Dissertation

Doch nicht nur die oben beschriebene Unterschiede in der Anzahl der Wortproduktion ist zu beobachten, Kinder zeigen außerdem eine starke Variation in ihrem Wortschatz, selbst wenn sie im gleichen Haushalt und unter gleichen Bedingungen groß werden. Die Neugier der Kinder auf unterschiedliche Objekte und deren Benennung sowie die intrinsische Motivation der Kinder etwas zu lernen, scheint ebenfalls einen Einfluss auf den Spracherwerb der Kinder zu haben (für eine Rezension siehe Mani & Ackermann, 2018). Eine Vielzahl an Studien hat sich damit beschäftigt, welchen Einfluss die extrinsische Faktoren (z.B.: Sprachinput (Cristia, 2013; Soderstrom, 2007), geteilte Aufmerksamkeit (Carpenter et al., 1989), Zeigegesten (Begus et al., 2014; Begus & Southgate, 2012)) aber auch intrinsische Motivation (Kidd et al., 2012) oder Neugier (Kang et al., 2009) auf den Spracherwerb haben. In der hier vorliegenden Dissertation wird eine Verbindung von selektiver Aufmerksamkeit, gelenkt durch den extrinsischen Faktor Kind-gerichtete Sprache, und kindlichem Interesse hergestellt. Kindliches Interesse wird einerseits durch die Präferenz des Kindes für KGS dargestellt und andererseits wollen wir untersuchen, ob und wie das kindliche Interesse zu lernen, den Lernerfolg beeinflusst. Genauer gesagt, untersucht diese Dissertation das Zusammenspiel von elterlichem Input und der kindlichen Aufmerksamkeit für den Input, sowie das kindliche Interesse von diesem Input zu lernen. Des Weiteren wird untersucht wie Kinder lernen wollen und inwieweit sich die Motivation der Kinder zu lernen erschließen lässt, sowie der Einfluss dieser Motivation auf das Lernen selbst.

Im Folgenden werden zuerst extrinsische und intrinsische Faktoren, die den Spracherwerb beeinflussen könnten, erläutert und Theorien dazu präsentiert. Im Anschluss werden die verwendeten Methoden erläutert, bevor die eigentlichen Studien beschrieben werden. Diese Arbeit endet mit einer Diskussion der Ergebnisse und einer Einordnung der Ergebnisse in die existierenden Modelle und Theorien.

Wortlernen: Intrinsische und extrinsische Hinweise zum Wortlernen

Extrinsische Faktoren.

Kinder wachsen in einer Welt auf, die sich ihnen nicht nur durch Beobachtungen erschließt. Es gibt eine Reihe extrinsischer Faktoren, die das Lernen der Kinder beeinflussen. Unter exotischen Faktoren verstehe ich Faktoren, die den Spracherwerb eines Kindes von außen beeinflussen. Im Folgenden wird auf einen dieser Faktoren genauer eingegangen. Ich betrachte den Sprachinput, der den Kindern präsentiert wird und die Rolle, die seine Variation im Spracherwerb spielt. Des Weiteren werden andere extrinsische Faktoren wie die Blickbewegung und das Zeigen betrachtet.

Kind-gerichtete Sprache. Die menschliche Sprache besteht aus einer Vielzahl von prosodischen und phonologischen Hinweisen, die es erleichtern sollen den semantischen Inhalt zu transportieren (Cristia, 2013). Mit Kindern, die Sprache erlernen, wird eine noch akustisch intensivere Sprache benutzt, KGS.

Dieser Sprachstil unterscheidet sich in prosodischen, phono-lexikalischen und syntaktischen Eigenschaften von dem Sprachstil mit dem Erwachsene untereinander kommunizieren (EGS, Cristia, 2013; Soderstrom, 2007). KGS wird von Eltern, Betreuern aber auch Geschwistern und Fremden verwendet um mit Kindern zu interagieren (Fernald et al., 1989, Rheingold & Adams, 1980; Sachs & Devin, 1976; Shatz & Gelman, 1973; Weppelman et al., 2003). Prosodisch unterscheidet sich KGS hauptsächlich durch eine höhere Tonhöhe von EGS (Cristia, 2013), die über mehrere Jahre des Spracherwerbs der Kinder beibehalten wird (Schreiner, Eiteljörge et al., in prep). Soderstrom (2007) berichtet außerdem von einer erhöhten Variabilität in der Tonhöhe, sowie einer Ausdehnung der Vokale und verlängerten Pausen in KGS, verglichen mit EGS. Des Weiteren benutzt KGS spezifische Konturen, wie zum Beispiel die Hervorhebung von akustischen Eigenschaften für die Satzgrenzen (Cristia, 2013; Soderstrom, 2007). In Bezug auf die phono-lexikalischen und syntaktischen Eigenschaften, wurde KGS als linguistisch einfacher beschrieben, mit mehr Fragen, vermehrten Wiederholungen und einer höheren Anzahl an Wiederholungen einzelner Wörter und Sätze (Grieser & Kuhl, 1988; Soderstrom, 2007; Soderstrom

et al., 2008). Vokale zeigen eine unterschiedlichere Formantenstruktur in KGS als in EGS (Kuhl et al., 1997), was es dem Kind gegebenenfalls erleichtert Kategorien für die Vokale zu erstellen.

KGS wird außerdem mit dem Ausdruck von positiven Emotionen in Verbindung gebracht (Singh et al., 2002), was mit dazu beiträgt die kindliche Aufmerksamkeit zu erregen und die Kommunikation beeinflusst (Soderstrom, 2007). Golinkoff und Kollegen (2015) heben diese Eigenschaft von KGS besonders hervor und verweisen darauf, dass sie einen erheblichen Einfluss auf die kindliche Organisation von Sprache hat.

Es scheint jedoch Variationen zwischen den Eigenschaften von KGS über verschiedene Kulturen hinweg zu geben. Eine Studie mit sechs Sprachen zeigte, dass deutschsprachige Mütter die am wenigsten ausgeprägten Eigenschaften von KGS produzieren, auch wenn man die Eigenschaften von KGS immer noch nachweisen kann (Fernald et al., 1989). Deutschsprachige Mütter zeigten die geringsten Unterschiede zwischen EGS und KGS in Bezug auf die Frequenz der mittleren Tonhöhe, der minimalen und maximalen Tonhöhe und der Variabilität der Tonhöhe, sowie der Länge der Äußerungen und der Länge der Pausen (Fernald et al., 1989).

Des Weiteren scheint es sprachspezifische Unterschiede in den Hinweisen zu phonetischen Kategorien in KGS zu geben (Werker et al., 2007). Dies wirft die Frage auf, ob die Eigenschaften von KGS an die spezifischen linguistischen Bedürfnisse einer Sprache angepasst sind, um dem Kind den Spracherwerb zu erleichtern.

Es scheint außerdem eine Reihe von Variationen zwischen den Eigenschaften von KGS über die Entwicklung der Kinder hinweg zu geben, obwohl es anscheinend nur eine begrenzte Konsistenz zwischen den Sprachen bezüglich bestimmten Eigenschaften gibt, auf die diese Entwicklungsänderungen abzielen. Eine dieser Eigenschaften ist *Voice Onset Time* (VOT). Sundberg und Lacerda (1999) fanden einen Unterschied in der VOT zwischen EGS und KGS. Sie zeigten eine Verkleinerung der VOT in stimmhaften und stimmlosen Konsonanten bei schwedischen Müttern, wenn diese zu ihren 3 Monate alten Kindern sprechen, im Vergleich zu Erwachsenen (Sundberg & Lacerda, 1999). Englund (2005) zeigt hingegen eine längere VOT für alveolare und velare Stopps in KGS, produziert zu 6 Monate alten Kindern, verglichen mit EGS. Es

gibt jedoch auch Studien, die keinen Unterschied in VOT zwischen KGS und EGS fanden (Baran et al., 1977).

Ähnliche Unterschiede finden sich in der Anzahl an Äußerungen in KGS. Henning und Kollegen (2005) berichten von Müttern, die mehr Äußerungen und Wörter mit ihren 3 Monate alten Kindern verwenden, im Vergleich mit 1 Monat alten Kindern. Währenddessen berichtet Kaye (1980) über einen Anstieg der mittleren Länge der Äußerungen (und Wiederholungen) in KGS zu 6 und 26 Wochen alten Kindern. Im Einklang mit diesen Studien, schlagen viele andere Studien vor, dass KGS kürzer als EGS ist (siehe Soderstrom (2007) für einen Überblick). Diese würde im Zusammenhang mit einer weniger komplexen syntaktischen Struktur in KGS stehen. Im Gegensatz dazu argumentieren einige Wissenschaftler dass KGS kürzer, aber nicht weniger Komplex ist als EGS (Newport et al., 1977). Es scheint jedoch der Fall zu sein, dass Mütter in KGS mit jüngeren Kindern mehr Wiederholung benutzen als mit älteren Kindern, was den Sprachinput vereinfachen würde (siehe Soderstrom (2007) für einen Überblick). Zusammenfassend zeigen diese Studien, dass Mütter (und andere Personen) gegenüber dem Kind in ihrer Sprache, abhängig vom Alter und Bedürfnissen der Kinder, variieren.

Des Weiteren müssen Modulationen von KGS über den Entwicklungszeitraum von Kindern nicht nur von Eltern ausgehen. Kinder können den Input der Mutter auch mit ihren Reaktionen zu diesem Input verändern. Einige Studien schlagen vor, dass Mütter deutlicheres KGS benutzen, wenn die Kinder ein positives Feedback geben (Smith & Trainor 2008; Braarud & Stormark, 2008). Kürzlich konnten Studien außerdem zeigen, dass Kinder bereits in einem frühen Alter KGS damit assoziieren, dass mit Kindern kommuniziert wird und EGS damit assoziieren, dass zwischen Erwachsenen kommuniziert wird (Schreiner et al., under review; Soley & Sebastian-Galles, 2020). Solche Ergebnisse geben einen ersten Hinweis darauf, dass Kinder aktive Teilnehmer in einem Kommunikationsprozess sind. Sie könnten gegebenenfalls die Eigenschaften der Kommunikation, die an sie gerichtet wird, lenken und mitbestimmen.

Gestische Hinweise. Neben phonologischen und phonetischen Hinweisen erhält das Kind eine Reihe gestischer Hinweise. Diese Hinweise können die geteilte Aufmerksamkeit, das Folgen der Aufmerksamkeit und das Lenken der Aufmerksamkeit als Ziel haben (Carpenter et al., 1998). Die Aufmerksamkeit zu teilen bedeutet, dass zwei Personen einen Gegenstand gemeinsam betrachten und ggf. mit ihm interagieren. Besonders hervorzuheben hierbei ist, dass das Kind sich der Anwesenheit des Erwachsenen bewusst ist und sich gelegentlich *rückversichert* ob die Aufmerksamkeit des Erwachsenen noch auf dem gemeinsam betrachteten Gegenstand liegt (Carpenter et al., 1998; Trevarthe & Hubley, 1978). Das Folgen der Aufmerksamkeit und das Lenken der Aufmerksamkeit können durch ähnliche gestische Hinweise geschehen. Diese Hinweise können entweder vom Erwachsenen ausgehen und das Kind folgt ihnen, oder später produziert das Kind selbst die Hinweise und lenkt somit die Aufmerksamkeit des Gegenübers (Begus & Southgate, 2012).

Das Folgen der Augenbewegung ist eine dieser gestischen Hinweise. Kinder erhalten, indem sie der Augenbewegung des Gegenübers folgen z.B. Information über den Gegenstand im Fokus des Gegenübers (Lee et al., 1989). Dieses aktive Folgen der Augenbewegung oder der Blickrichtung des Gegenübers ist damit verbunden, dass das Kind beobachtet oder überprüft, worauf sich die Aufmerksamkeit des Erwachsenen richtet (Carpenter et al., 1998). Brooks und Meltzoff (2005) konnten zeigen, dass Kinder erst ab einem Alter von ca. 10 Monaten in der Lage der Augenbewegung direkt zu folgen. Davor unterscheiden sie nicht, ob der Erwachsene die Augen offen oder geschlossen hat (Brooks & Meltzoff, 2005), oder ob sich der ganze Körper/ Kopf dreht (Lempers, 1979). Der direkte Hinweis, der Augenbewegung einer Person zu folgen um Information zu erhalten entwickelt sich also erst später. Nichtsdestotrotz können Kinder bereits ab einem Alter von 3 Monaten der Bewegung einer Person folgen um Informationen zu erhalten (D'Entremont et al., 1997): D'Entremont und Kollegen (1997) zeigten, dass bereits 3 Monate alte Kinder der Blickbewegung einer Person folgen, um ein Objekt in unmittelbare Nähe zu betrachten. Die Kinder wurden zuvor durch Kind-gerichtete Interaktionen und Sprache in die Situation einbezogen. KGS könnte in diesem Zusammenhang das Engagement der Kinder in soziale Interaktionen gesteigert haben. Diese Resultate spiegeln sich auch in einer Studie von Senju und Csibra (2008) wieder. Die

Autoren konnten zeigen, dass Kinder mit 6 Monaten der Blickbewegung einer Person mehr folgen, wenn das Kind zuvor mit KGS adressiert wurde, wohingegen dem Blick der Person weniger gefolgt wird, wenn die Kinder vorher mit EGS adressiert wurden.

Die Fähigkeit der Blickbewegung einer anderen Person zu folgen, ermöglicht es den Kindern bereits früh aktiv an ihrer Umwelt teilzuhaben. Sie erlernen die referentiellen Absichten einer Person zu erkennen und zu deuten, und können dadurch leichter eine Verbindung zwischen den Objekten im Fokus und den verwendeten Wörtern finden. So benennen Eltern meist Objekte oder Tätigkeiten die mit der unmittelbaren Umgebung der Kinder zu tun haben (Tomasello, 2003). Durch das Folgen der Blickbewegung des Erwachsenen können Kinder somit schneller die referentiellen Absichten des Erwachsenen erkennen (e.g. Carpenter et al., 1998).

Intrinsische Faktoren.

Neben äußeren Einflüssen gibt es eine Reihe an Bias, die den Kindern helfen zu lernen. Diese Bias können dadurch beschrieben werden, dass sie die Aufmerksamkeit der Kinder in eine bestimmte Richtung lenken. Unter anderem zeigten eine Reihe von Studien, dass Kinder mit 18 Monaten neue Wort-Objekt Assoziationen lernen, da sie die Strategie benutzen, bereits bekannte Objekte als Referenten ausschließen (Bloom, 2002; Markman & Wachtel, 1988; Markman et al., 2003). Dahinter steht die Annahme, dass jedem Objekt nur ein Name zugeschrieben wird. Des Weiteren verbinden Kinder im Alter von ca. 18 Monaten neue Wörter mit ganzen Objekten und nicht mit deren Eigenschaften oder Teilen (Markman, 1994). Diese Annahmen können helfen neue Wörter für neue Objekte zu lernen. Andere Eigenschaften wie zum Beispiel Farbe (Mani et al., 2013) oder Form (Samuelson, 2002, Smith, 2004) der Objekte, können Kindern ebenfalls helfen das neue Wort mit dem richtigen Objekt in Verbindung zu bringen.

Unterschiedliche neurophysiologische und Verhaltensstudien konnten außerdem zeigen, dass Interesse und Neugier einen Einfluss auf das Verhalten und Lernen der Probanden haben. Castro und Kollegen (2008) gehörten zu den ersten, die die Lernfähigkeit von Probanden durch einen aktiven und passiven Task miteinander verglichen haben. In Kategorie Lernexperimenten zeigten sie, dass Probanden besser in der Lage sind Kategorien zu bilden, wenn sie aktiv die

Informationen aussuchen können, im Vergleich zu einer passiven Präsentation von randomisierten Informationen. Ähnliche Ergebnisse zeigten sich auch mit Kindern im Alter von 7 Jahren in einem Experiment zur Kategorisierung von Tieren (Sim et al., 2015). Kang und Kollegen (2009) konnten durch die Messung der Pupillengröße zeigen, dass Probanden neugieriger sind etwas zu lernen, wenn sie bereits etwas über das Thema wissen, nicht jedoch wenn sie nichts oder schon sehr viel darüber wissen. Anders gesagt, zeigten sie, dass die Neugier etwas zu lernen maximal ist, wenn das Wissen über das Thema mittelmäßig ist. Die Probanden zeigten außerdem erweiterte Pupillen, wenn sie neugierig waren, was die richtige Antwort auf die Frage war, im Vergleich zu Probanden, die nicht neugierig waren. Die Neugier der Probanden an einem Thema hatte ebenfalls Einfluss auf die Fähigkeit die erlernten Informationen abzurufen (Berylne, 1954; Gruber et al., 2014; Fastrich et al., 2018; Kang et al., 2009) Gruber und Kollegen (2014) konnten zeigen, dass Probanden sich Antworten auf Fragen, die sie neugierig machten besser merken konnten als Antworten auf Fragen, die keine Neugier in ihnen hervorriefen. Interessanterweise erinnerten Probanden sich auch an triviale Information, die in einer Phase von großer Neugier präsentiert wurden waren, jedoch nichts mit der eigentlichen Frage und deren Antwort zu tun hatten (Gruber et al. 2014). Vierundzwanzig Stunden nach der eigentlichen Testung wurde die Erinnerungsfähigkeit der Probanden erneut getestet, und es konnte auch hier gezeigt werden, dass Probanden sich besser an Dinge erinnern, über die sie neugierig gewesen waren. Gruber und Kollegen (2014) argumentieren hier, dass die Neugier an einem Thema das Gehirn in ein Stadium versetzt, welches das Lernen erleichtert. Dies erklären sie mit der Aktivierung von Gehirnarealen im Mittelhirn (*medus accumulens* und *substantia nigra/ventral tegmental area complex*, im Folgenden SN/VTA). Diese Areale zeigten eine größere Aktivität, umso mehr die Neugier an einem Thema stieg. Ebenfalls im Mittelhirn liegt das *Striatum*, welches mit der Aktivierung durch Belohnung in Verbindung gebracht wird (Hare et al., 2008; Knutson et al., 2000; McClure et al., 2004; Wittmann et al., 2005). Ripollés und Kollegen (2014) konnten in einer Studie zum Wortlernen eine Verbindung zwischen Wortlernerfolg und der Aktivität des Striatum herstellen. Diese Studien zeigen somit, dass dieselben Areale aktiviert werden, wenn Probanden neugierig über ein bestimmtes Thema sind und wenn Probanden eine (externe) Belohnung erhalten. Bei der Beantwortung einer Frage, über die der Proband neugierig

war, zeigte sich in der Studie von Grube und Kollegen außerdem eine Aktivierung des Hippocampus (Gruber et al., 2014), was in Kombination mit der Aktivierung von SN/VTA in der Literatur als funktionelle Schleife gesehen wird, in der die Regionen des Mittelhirns durch eine Aktivierung des Hippocampus das Lernen modulieren (Düzel et al., 2010; Lisman & Grace, 2005; Lisman et al., 2011). Anders ausgedrückt heißt das, dass die Neugier an einer Frage die Interaktion zwischen SN/VTA und Hippocampus erhöht und ihn somit auf seine bevorstehende Aufgabe (das Behalten der neuen/hervorstechenden Information) *vorbereitet*. Dies versetzt das Gehirn in einen Zustand, indem Dinge (von Interesse aber auch beiläufig präsentierte Dinge) einfacher gelernt werden können (Gruber et al., 2014). Im Einklang mit diesen Ergebnissen fanden Wittmann und Kollegen (2005), dass Probanden signifikant schneller auf Belohnung hervorrufende Bilder reagieren, als auf neutrale Bilder, und diese ebenfalls länger erinnern.

Studien mit Kindern zeigten, dass Kinder bereits in einem frühen Alter (16 Monate) Gesten benutzen um Informationen über Gegenstände zu erhalten (Begus & Southgate, 2012). In einer Folgestudie konnten Begus und Kollegen (2014) zeigen, dass Kinder die Informationen, die sie durch eine Zeigegeste erfragt haben, besser lernen als andere Informationen (Begus et al., 2014). In Übereinstimmung mit der Idee des intermediären Wissenserwerbes, entscheiden sich Kinder dazu, Stimuli zu beobachten, die weder zu komplex noch zu einfach sind (Kidd et al., 2012).

Ruggeri und Kollegen (2016) testeten 6 bis 8 Jahre alte Kinder mit einer einfachen Gedächtnisaufgabe auf einem Tablet. Während für die Hälfte der Zeit die Kinder in der aktiven Lernbedingung die Reihenfolge und Geschwindigkeit des Materials bestimmen konnten, sahen Kinder in der passiven Lernbedingung lediglich die Aufnahmen eines aktiven Teilnehmers. Ruggeri und Kollegen (2016) fanden, dass die Wiedererkennung der Objekte in der aktiven Bedingung besser war als in der passiven. Dieser Effekt konnte auch noch eine Woche nach der Durchführung der Studie nachgewiesen werden. Partridge und Kollegen (2015) zeigten bei Kindern im Alter von 3-4 Jahren ebenfalls, dass selbstbestimmtes Lernen das Kurzzeitgedächtnis für neu erlernte Wort-Objekt Assoziationen positiv beeinflusst, im Gegensatz zu passivem Lernen (siehe aber Ackermann et al., 2020b für andere Ergebnisse in einer Tablet Studie). Dieses Ergebnis

schreiben die Autoren dem größeren Engagement in der selbstbestimmten Lernbedingung zu. Zettersten und Saffran, (2019) ließen Kindern aktiv die Wahl über welches Objekt sie mehr Informationen bekommen wollten. Sie präsentierten Kindern entweder vollständig mehrdeutige, teilweise mehrdeutige oder eindeutige Wort-Objekt-Zuordnungssituationen. Dann konnten die Kinder entscheiden welche Objekte benannt werden sollten, um die Ambiguität aufzuheben. Wurden viele teilweise mehrdeutige Wort-Objekt-Zuordnungssituationen präsentiert, so zeigten Kinder (4-7 Jahre) eine bevorzugte Auswahl der Objekte, die die Ambiguität aufheben. Die Studie zeigt, dass Kinder sich aktiv daran beteiligen können ihre Unwissenheit zu beheben. Ackermann und Kollegen (2020a) zeigten außerdem, dass auch die Neugier von Kindern über gewisse Dinge einen Einfluss auf ihr Lernen hat. Zwei Jahre alten Kinder wurden neue Wort Objekt-Assoziationen aus vier unterschiedlichen Kategorien präsentiert. Um das Interesse der Kinder in einer dieser Kategorien zu messen, wurde die Veränderung der Pupillengröße nach der Präsentation von bekannten Objekten aus jeder Kategorie gemessen. Eine Vergrößerung der Pupille deutet in diesem Fall auf ein größeres Interesse an dieser Kategorie hin. Die Kinder zeigten besseres Lernen von neuen Objekten aus Kategorien, an denen sie Interesse hatte. Das Interesse an einem neuen Objekt selbst zeigte ebenfalls einen gewissen Einfluss auf das Lernen der Kinder.

Interesse und Neugier. Interesse und Neugier sind in der Literatur oft miteinander verbunden. So werden die unterschiedlichen Stadien der Neugier mit einem inneren psychologischen Zustand, dem Interesse, verknüpft (Oudeyer & Lopez, 2016). Um Neugier und Interesse besser voneinander unterscheiden zu können, möchte ich zuerst Neugier und Interesse definieren und Unterschiede und Gemeinsamkeiten herausarbeiten.

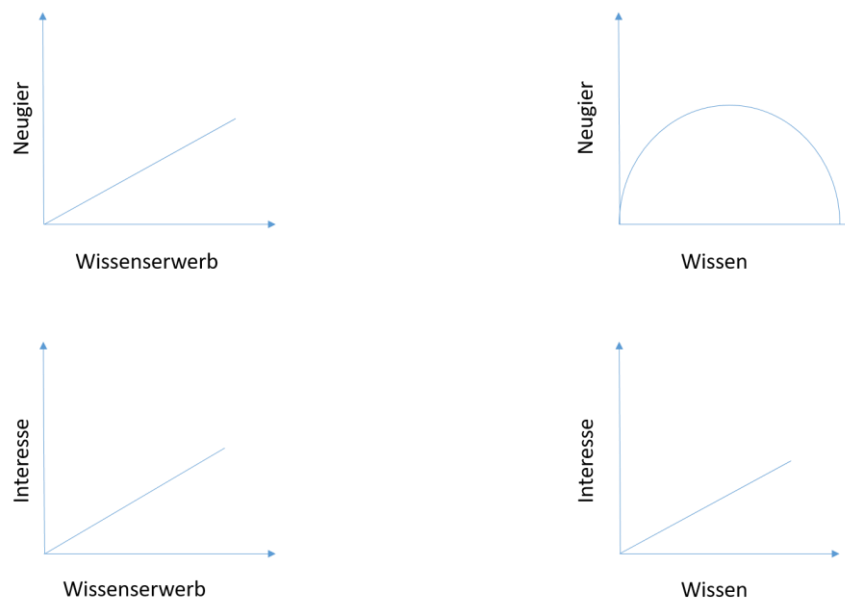
Neugier wird von Löwenstein (S.1; 1994) als entscheidende Motivation für menschliches Verhalten, sowohl positiv als auch negativ, gesehen. Kidd & Hayden (2015) beschreiben Neugier als spezielle Art der Informationsgewinnung und bereits James (1899) spricht davon, dass die kindliche Neugier Kinder dazu antreibt neue Objekte mit unbekanntem Eigenschaften zu erkunden. Berlyne (1954) unterscheidet zwischen zwei Arten von Neugier: perzeptuelle Neugier und epistemische Neugier. Epistemische Neugier wird als Motivation beschrieben Wissen und neue Informationen zu erwerben (Berlyne, 1954; Kidd & Hayden, 2015). Perzeptuelle Neugier hingegen

ist der Drang nach Neuem (Berlyne, 1954; Kidd & Hayden, 2015). Setzt man Neugier dann mit Lernen und Wissen in Verbindung, zeigt sich, dass Lernen durch Neugier vereinfacht und erleichtert wird (Kang et al., 2009). Wissen und Neugier stehen jedoch nicht in einem so linearen Verhältnis. Ist das Wissen über eine Sache sehr gering, ist meist auch die Neugier, die dieser Sache entgegengebracht wird, sehr gering. Auch wenn das Wissen über eine Sache groß ist, ist die Neugier meist klein. Ein mittleres Wissen hingegen steigert die Neugier, bis das Wissen groß genug ist und die Neugier wieder abnimmt (Kang et al., 2009). Dieses Phänomen wird auch als Goldilocks-Effekt bezeichnet (Kidd et al., 2012a).

Interesse hingegen beschreibt einen inneren Zustand (Hidi & Renninger, 2006) und wird durch einen Reiz oder eine Aktivität hervorgerufen (Oudeyer et al., 2016). Dabei kann dieser Reiz etwas Überraschendes, Neues oder Unerwartetes sein. Interesse kann jedoch, wie Neugier auch, durch eine Wissenslücke oder eine nicht zufriedenstellende Kompetenz hervorgerufen werden (Oudeyer et al., 2016). Interesse kann ebenfalls weiter unterteilt werden. Es wird zwischen individuellem und situationsbezogenem Interesse unterschieden (Grossnickle, 2004). Hidi und Renninger (2006) gehen davon aus, dass individuelles Interesse immer aus situationsbedingtem Interesse entsteht. Dabei kann situationsbedingtes Interesse durch einen Reiz oder eine Situation hervorgerufen werden und nur kurz anhalten. Individuelles Interesse hingegen kann sich aus diesem situationsbedingten Interesse entwickeln und zeigt sich dadurch, dass die Person mehr Wissen erlangen will oder eine innere Erregung spürt und sich dadurch intensiver mit dem Thema oder dem Reiz auseinandersetzt. Im Gegensatz zu Neugier nimmt Interesse nicht ab umso mehr Wissen man erwirbt (Grossnickle, 2004; Kang et al., 2009), sondern steigt mit dem Wissen Zuwachs an (siehe Schema 1). Ein weiterer Unterschied besteht im Affekt. Interesse äußert sich immer positiv, wohingegen Neugier auch einen negativen Affekt haben kann, und z.B. in Voyeurismus enden kann.

Interesse und Neugier werden oft in Verbindung mit intrinsischer Motivation gebracht, welche man als Neigung bezeichnen kann, neue Aktivitäten oder Wissen zu erwerben, nur weil man Freude daran hat und nicht wegen einer externen Belohnung

Schema 1.: Zunahme und Abnahme von Neugier und Interesse in Abhängigkeit von Wissen



Um Interesse an einer Situation oder einem Gegenstand zu erlangen, muss die selektive Aufmerksamkeit darauf gerichtet werden. Betrachtet man die Entwicklung von Interesse auf einem Zeitstrahl, so wird die Aufmerksamkeit des Betrachters zuerst durch einen Reiz erregt. Durch diese selektive Aufmerksamkeit kann situationelles Interesse entstehen, woraus sich individuelles Interesse entwickeln kann.

Betrachtet man die aufgeführten Studien und intrinsischen, sowie extrinsischen Faktoren, die das Lernen der Kinder beeinflussen, so fällt auf, dass die Aufmerksamkeit der Kinder stets auf einen bestimmten Aspekt gelenkt ist. Dabei kann die Aufmerksamkeit entweder von außen *gesteuert werden*, in Form von KGS oder der Blickbewegung der Eltern, sie kann aber auch von den Kindern selbst gesteuert werden, entweder durch interne Bias (zum Beispiel *whole Object Bias*) oder das Interesse der Kinder an zum Beispiel einer Kategorie von Dingen. Dies lässt darauf schließen, dass die Aufmerksamkeit, die Kindern Dingen schenken eine wichtige Funktion im Wortlernen einnimmt.

Modelle

Im Folgenden möchte ich drei unterschiedliche Modelle beschreiben, die sich mit kindlichem Lernen und den Faktoren, die dieses beeinflussen, beschäftigen. Dabei möchte ich mit der Theorie der *natural pedagogy* (Csibra & Gergely, 2009, 2011) beginnen, die vor allem die äußeren Einflüsse einbezieht, durch die Kinder einfacher (Wörter) lernen. Im Folgenden möchte ich das *Emergentist Coalition Model* (Hollich et al., 2000) beschreiben, welches unterschiedliche Hinweise in den Prozess des Wortlernens einbezieht und zum Schluss möchte ich das Model von Oudeyer und Smith (2016) genauer betrachten, welches sich vor allem mit selbstbestimmten Lernen und dessen Erfolg beschäftigt.

Eine Theorie, die vor allem auf den Einfluss der äußeren Faktoren eingeht, ist die sogenannte *natural pedagogy*. Csibra und Gergely (2009) gehen davon aus, dass das menschliche Kommunikationssystem speziell dafür entwickelt wurde, um generisches Wissen zwischen Individuen zu vermitteln. Kinder sind dafür mit rezeptiven Fähigkeiten ausgestattet um die anschaulichen Signale der Erwachsenen zu deuten. Erwachsene wiederum vermitteln ihr kontextbezogenes Wissen in anschaulichen Situationen. Als diese anschaulichen Gesten beschreiben Csibra und Gergely (2009) extrinsischen Faktoren, wie das Verwenden von KGS, das Zeigen auf einen Gegenstand oder die Demonstration einer Funktionsweise eines Gegenstandes. Durch diese Gesten kann Information bereits nach einmaliger kommunikativer Demonstration erlernt werden. Csibra und Gergely (2011) gehen weiter davon aus, dass diese Art Wissen zu vermitteln nur bei Menschen auftritt. Was in der Theorie von *natural pedagogy* nicht betrachtet wird, ist die Motivation und das Interesse, die das Kind selbst mitbringt um Sprache zu lernen.

Ein weiteres Model für den Spracherwerb, welches mehrere Hinweise einbezieht ist das *Emergentist Coalition Model* (Hollich et al., 2000). Dieses Model nimmt an, dass unterschiedliche Hinweise, wie Aufmerksamkeit der Kinder, soziale Hinweise und linguistische Hinweise den Kindern zur Verfügung stehen. Unter Hinweisen, die mit der Aufmerksamkeit der Kinder in Verbindung stehen, beschreiben die Autoren zum Beispiel die Domänen-übergreifenden Aufmerksamkeitsfaktoren, zeitliche Kontiguität und das Interesse an Neuem. Unter sozialen

Hinweise verstehen die Autoren zum Beispiel das Folgend der Augenbewegung oder Zeigebewegungen. Hinzu kommen linguistische Hinweise, die den Kindern zum Beispiel helfen Wörter aus dem Sprachfluß zu segmentieren. Solche Hinweise wären z.B. Intonation und Pausen, wie sie verstärkt in KGS vorkommen. Hollich und Kollegen (2000) nehmen an, dass Kinder während des Prozesses des Wortlernens diesen Hinweisen eine unterschiedliche Bedeutung zuschreiben. Die Deutung einiger dieser Hinweise muss auch erst erlernt werden. So schlagen Hollich und Kollegen vor, dass zum Beispiel das Prinzip der perzeptuellen Ähnlichkeit (gelenkt durch die Aufmerksamkeit der Kinder) zu Beginn des Wortlernens eine große Bedeutung beigemessen wird. Diese aufmerksamkeits-gelenkten Hinweise bilden die erste Phase in den Emergentist Coalition Model. In der zweiten Phase realisieren Kinder, dass eine Einbindung der sprechenden Person eine größere Chance bietet das richtige Wort mit dem richtigen Objekt zu kombinieren; Kinder binden die sozialen Hinweise mit ein. Ebenso werden in der zweiten Phase linguistische Hinweise stärker mit eingebunden. Interessant gestaltet sich die Frage wie diese Änderungen über den Lauf der Entwicklung der Kinder passieren. Hollich und Kollegen nehmen an, dass Wortlernen von außen gelenkt wird, durch die Tatsache, dass Kinder einigen Input (zum Beispiel neue Dinge) gegenüber anderen bevorzugen. Dies führt sie zu Beginn des Lernens zu genau den richtigen sprachrelevanten Informationen in ihrer Umgebung. Doch zusätzlich erhalten Kinder Sprachinput, der nicht immer zu den Objekten im Fokus des Kindes passt. Durch Erfahrungen mit seiner Umwelt erlernt das Kind auch soziale Hinweise, wie z.B. das Folgen einer Blickbewegung, effektiv zu nutzen, um Wort-Objekt Assoziationen bilden zu können. Die Auswertung der Verlässlichkeit der Hinweise, die das Kind für wichtig *empfindet* sorgt schließlich dafür, dass das Kind sich von rein perzeptuellen Hinweisen stärker zu sozialen Hinweisen orientiert. Nichtsdestotrotz ist es die Kombination an Hinweisen, die es den Kindern ermöglicht Worte zu lernen. Dieser Prozess, des Auswählens der richtigen Hinweise und deren Wertung, entwickelt sich im Laufe des Wortlernens von einem unreifen zu einem reifen Prinzip, von einem Domän-generellem zu einem Domän-spezifischem Prozess. In dieser Theorie ist das Kind kein rein passiver Betrachter mehr. Es kalkuliert die Verlässlichkeit der Hinweise, die ihm zur Verfügung stehen und passt diese an, basierend darauf an, welche Hinweise es stärker mit einbezieht, um neue

Wort-Objekt Assoziationen zu lernen. Trotzdem wird das Kind stark von den Hinweisen beeinflusst, die in dem Emergentist Coalition Model alle von außen kommen.

Oudeyer und Smith (2016) stellten in ihrem Experiment eine Lernsituation ähnlich eines Spielplatzes mit Robotern nach und entwickelten daraus ein Model, welches die unterschiedlichen Kapazitäten und Mechanismen in Beziehung zu einander stellt. Der Roboter lernt wie er seine Gliedmaßen und seine *Stimme* benutzen kann und welche Auswirkungen diese auf seine Umgebung haben. Die Erkundung seiner Fähigkeiten wird durch die Maximierung des Lernfortschrittes vorangetrieben. Dies geschieht dadurch, dass physische Erfahrungen ausgewählt werden, die die Qualität der Vorhersage in Bezug auf die Konsequenzen seiner Handlungen verbessert. Anders gesagt, der Roboter lernt die Konsequenzen seine eigenen Handlungen hervorzusagen und Fehler in seinen Handlungen zu vermeiden, und berechnet daraus ein internes Belohnungssystem. Dieses Belohnungssystem steht in direktem Zusammenhang mit der Verringerung der Vorhersagefehler. Das Ziel ist es dieses Belohnungssystem durch die Auswahl der Bewegungen zu maximieren. Benutz der Roboter nun von Neugier getriebene Erkundungen, um zu entscheiden welche Bewegung er ausführen will, vermeidet er dadurch Bewegungen, deren Ausgang er schon kennt oder die noch zu schwer sind um ihr Resultat hervorzusagen. Stattdessen *konzentriert* sich der Roboter auf die am schnellsten zu lernende Tätigkeit. Ist diese Tätigkeit ausreichend erlernt, kann der Roboter zur nächsten erfolgsversprechenden Tätigkeit übergehen. So zeigten die Roboter eine systematische Entwicklung durch selbst organisiertes Lernen. Sie begannen mit unorganisierten Körperbewegungen, gefolgt von der Erkundung einzelner motorischer Fähigkeiten. Im nächsten Schritt zeigten die Roboter Bewegungen in Richtung unterschiedlicher Objekte um danach, durch die Erprobung ihrer unterschiedlichen erlernten Fähigkeiten, eine erste Interaktion mit den Gegenständen zu erforschen. Nach einer gewissen Zeit zeigte der Roboter unterschiedliche motorischer Interaktionen mit verschiedenen Objekten. Diese Kompetenzen entstanden nicht durch das Programmieren der Roboter sondern sie entwickelten sie durch selbstständige, neugierige Erkundung der Umgebung. Trotz der Individualität der einzelnen Roboter und ihrer Bewegungen, konnte die oben beschriebene Struktur in der Erkundung der Umgebung erkannt werden. Diese Roboter-Studie zeigte, dass eine dynamische Interaktion

zwischen unterschiedlichen Mechanismen (Neugier, Lernen, Abstraktion), den zur Verfügung gestellten Kapazitäten und der Beschaffenheit des Körpers, sowie die Umgebung und deren Beschaffenheit eine wichtige Rolle in der Entwicklung eines Organismus spielen.

Die Voraussetzungen mit denen die Roboter ausgestattet waren, waren der Drang nach Aktivität (Bewegung) und das generelle Prinzip, dass Bewegungen, die Ungewissheit reduzieren, belohnend sind. Bei diesen Voraussetzungen lassen sich gewisse Parallelen zur Entwicklung von Kindern feststellen. Kinder kommen zur Welt mit der Fähigkeit ihre Muskeln zu bewegen, was entweder zu Bewegung oder zur Produktion von Lauten führt. Durch die Erkundung ihrer Bewegungen und die Beobachtung der Resultate, sowie die Reaktionen der Umwelt und ihrer Mitmenschen, könnten Kinder lernen.

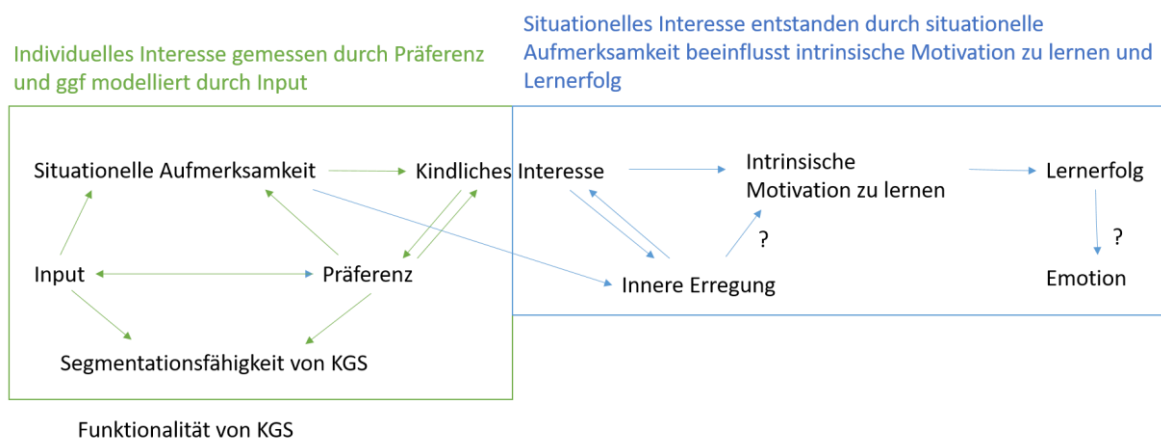
Die hier vorliegende Dissertation

Die hier vorliegende Dissertation betrachtet die selektive Aufmerksamkeit der Kinder, einerseits mit der Interaktion des extrinsischen Faktor KGS und andererseits dem Interesse der Kinder. Genauer gesagt, wird untersucht wie das Zusammenspiel von elterlichem Input und kindlicher Aufmerksamkeit an dem Input ist, und wie das Interesse des Kindes von diesem Input zu lernen das Lernen beeinflusst. Wir untersuchen welchen Einfluss die Präferenz der Kinder und der elterliche Input auf die Verarbeitung von KGS und EGS haben. Dabei untersuchen wir die Zusammenhänge von kindlicher Präferenz und Input. Im Weiteren gehen wir von der selektiven Aufmerksamkeit auf einen Sprachstil auf die selektive Aufmerksamkeit beim Wortlernen ein. Im Detail untersuchen wir inwieweit sich durch die selektive Aufmerksamkeit ein Interesse entwickeln kann und ob aus diesem Interesse die intrinsische Motivation zum Lernen widergespiegelt wird. Steht der Lehrerfolg in Verbindung mit der intrinsischen Motivation zu lernen? Und können Lehrerfolge durch positive Emotionen ausgedrückt werden?

Untersucht wird dieses Zusammenspiel einerseits in einer Studie, die sich den elterlichen Input (KGS und EGS) anschaut und diesen Input mit der Präferenz des Kindes (für KGS und EGS) in Zusammenhang bringt. Außerdem werden die Effekte dieser beiden Faktoren auf das kindliche

Engagement in soziale Interaktionen untersucht. Wir betrachten weiter, ob sich über die Entwicklung im ersten Lebensjahr eine Entwicklung in der Segmentationsfähigkeit von KGS und EGS beobachten lässt. Auf der anderen Seite untersucht eine Studie wie Kinder lernen wollen und ob die Aufmerksamkeit und die Neugier der Kinder einen Einfluss auf deren innere Erregung haben. Wir waren außerdem daran interessiert zu untersuchen wie diese Motivation der Kinder sich auf deren Lernen auswirkt, und ob sich dieser Lernerfolg in einer positiven Emotion widerspiegelt (siehe Schema 2)

Schema 2.: Verbindungen der Möglichen Interaktionen zwischen den Studien in dieser Dissertation



Zur Durchführung der Studien benutzten wir eine Reihe unterschiedlicher Methoden, die im Folgenden näher erläutert werden.

Methoden

Um die angeführten Fragestellungen zu beantworten, haben wir eine Reihe unterschiedlicher Verfahren angewandt: Mütter wurden aufgenommen, wenn sie KGS und EGS produzierten, um die Unterschiede einigen Eigenschaften dieser beiden Sprachstile untersuchen zu können. Die kindliche Präferenz für KGS und EGS wurde gemessen, um herauszufinden welchen Sprachstil die Kinder bevorzugen. Zwei Studien arbeiten mit der Methode des Eye trackings, um entweder die Blickbewegung der Kinder zu verfolgen oder zu untersuchen auf welches Objekt die Kinder, relativ gesehen, mehr schauen und somit das Lernen von neue Wort-Objekt-Assoziationen zeigen können Außerdem wurde der Eyetracker genutzt, um die Veränderungen der Pupillen zu messen. Dies diente dazu, um einen Index der aufgabenbasierten, kindlichen Erregung beim Lernen

von Wörtern zu erhalten. Desweiteren wurde die Methode der Elektroenzephalografie (EEG) genutzt, um die kindliche Fähigkeit Wörter aus dem Sprachfluß zu segmentieren, zu untersuchen. Körperhaltungsdaten wurden mithilfe einer Kinect Kamera aufgenommen, um Veränderungen basierend auf dem Lernerfolg der Kinder messen zu können. Im Folgenden wird auf jede Methode einzeln eingegangen und die Vorteile und Limitationen werden beleuchtet.

Aufnahmen der mütterlichen Sprache. Um die Sprache der Mütter zu analysieren, haben wir Mütter aufgenommen, die eine Geschichte zuerst in EGS erzählt haben und an einem späteren Tag dieselbe Geschichte in Interaktion mit ihrem Kind wiedergeben (KGS). Die Mütter sahen ein kurzes Video, ca. 1 Minute, um sich mit dem Inhalt der Geschichte vertraut zu machen. Die Geschichte handelte von einem Teddybär, der ein Abenteuer erlebte und mit mehreren unterschiedlichen Orten, Personen und Handlungen in Kontakt kam. Um die EGS nicht zu beeinflussen, wurden Mütter immer zuerst in EGS aufgenommen, bevor sie KGS produzierten. Des Weiteren haben wir uns entschlossen die Mütter alleine zu lassen und sich vorzustellen sie würden die Geschichte einem Freund erzählen, um die Sprache nicht durch die Anwesenheit des Experimentators und den damit verbundenen Druck die Geschichte ‚richtig‘ zu erzählen, zu verfälschen. Die Mütter wurden in einen stillen Raum, neben dem Spielraum gebracht, und nahmen auf einem Stuhl vor einem großen Bildschirm Platz. Auf diesem Bildschirm befand sich das lebensgroße Gesicht einer Person. Nachdem die Mütter bequem saßen konnten sie individuell beginnen und die Geschichte soweit erzählen, wie sie sich erinnerten. Das Mikrofon befand sich an den Füßen der Mütter. Um die Geschichte ihrem Kind erzählen zu können, konnten die Mütter sich noch einmal das Video anschauen. Fühlten sich die Mütter bereit, ließen wir sie mit ihrem Kind alleine und schalteten eine Kamera, die ca. 1,5m entfernt stand an, sowie ein Mikrophone, dass sich in unmittelbarer Nähe zur Mutter befand.

Um die Eigenschaften von EGS und KGS zu vergleichen, analysierten wir beide Sprachstile in Bezug auf Tonlage (minimale und maximale Tonhöhe, Bandbreite der Tonlage) und Länge einer einzelnen Äußerung. Dies sind nur einige der Eigenschaften, in denen sich KGS und

EGS unterscheiden, aber sie stellen ein gutes Maß dar, um die Unterschiede von KGS und EGS zu messen (e.g. Fernald et al., 1989).

Head turn preferenz - ‚Single Screen‘ Central fixation - Methode. Eine Studie dieses Dissertation benutzte eine modifizierte Version der klassischen head-turn preference Methode, die single screen central fixation Methode. Die traditionelle Methode benutze drei Orientierungen (zentral, links und rechts) um die kindliche Phonemdiskriminierung (Chambers et al., 2003), Wortsegmentation (Schreiner et al., 2016) und Präferenz für den Sprachinput (Kaplan et al., 1995b) zu untersuchen. Der Trial wird gestartet und ein blinkendes Licht erscheint um das Kind auf den zentralen Punkt zu orientieren. Schaut das Kind gerade aus, startet der Experimentator eins der beiden Lichter (links oder rechts). Orientiert sich das Kind zu dieser Lichtquelle, startet der Experimentator den auditiven Stimulus. Schaut das Kind länger als 2 Sekunden nicht mehr auf das Licht, wird der auditive Stimulus beendet, und die Aufmerksamkeit des Kindes wird wieder nach vorne gelenkt.

Im Vergleich zu der klassischen Methode, wird in dieser Studie die Methode mit zentraler Fixation benutzt, das heißt, das Kind dreht den Kopf nicht nach links und rechts um die Stimuli präsentiert zu bekommen. Stattdessen wird die Aufmerksamkeit des Kindes mit einem zentralen Aufmerksamkeitserreger nach vorne gelenkt und sobald das Kind nach vorne schaut, werden ein bunt blinkendes Schachbrettbild und der auditive Stimulus präsentiert (Cooper et al. 1997, siehe auch TheManyBabyConsortium, 2020). Solange das Kind nach vorne schaut, drückt der Experimentator einen Knopf auf der Tastatur. Schaut das Kind länger als 2 Sekunden nicht mehr auf den zentralen Bildschirm, wird der Bildschirm schwarz und ein neuer Trial beginnt mit der Präsentation des Aufmerksamkeitserregers. Die Annahme dahinter ist, dass das Kind solange nach vorne auf den visuellen Stimuli schaut, wie der auditive Stimulus ebenfalls interessant ist. Nimmt das Interesse ab, beginnt das Kind die Umgebung zu betrachten. In der hier präsentierten Studie diente die Messung der Präferenz dazu, um herauszufinden welchen Sprachstil (KGS, EGS) Kinder im Alter von 6- und 13-Monaten bevorzugen. Wir erwarten, dass Kinder in beiden Altersgruppen eine Präferenz für KGS zeigen.

Eye tracking. Um die Blickbewegungen der Kinder in den Studien dieser Dissertation zu analysieren, wurden Eye tracking Daten erhoben. Eye tracking Daten werden mit einem automatisierten Eyetracker erhoben, der Infrarotlicht auf das Auge richtet. Durch die Spiegelung des Lichtes auf der Hornhaut relativ zur Pupille, lässt sich schätzen wohin der Blick des Kindes gerichtet ist. Unterschiedliche Eyetracker sind auf dem Markt verfügbar. In unserem Labor arbeiten wir mit Tobii X120.

Eine, der hier beschriebenen Studien, beschäftigt sich mit *cross situational word learning*. Das ist eine Aufgabe, in der Kinder lernen welches Wort welches Objekt benennt, indem sie die relative Häufigkeit beobachten mit der das Wort mit einem Objekt zusammen auftritt (Smith & Yu, 2008b). Solche cross situational word learning Studien enthalten generell eine Trainings Phase und eine Test Phase. In jedem Trial der Trainings Phase werden zwei neue Wörter mit zwei neuen Objekten gezeigt. Dem Kind wird jedoch nicht offensichtlich welches Wort auf welches Objekt referiert. Erst über die Dauer des Trainings kann das Kind herausarbeiten, welches Wort welches Objekt präsentiert. In der Test Phase werden dann zwei Objekte mit einem Wort präsentiert und es wird gemessen, ob das Kind die richtige Wort Objekt Assoziation gebildet hat (Smith et al., 2011; Smith & Yu, 2008), indem es auf das Objekt schaut, welches mit dem präsentierten Wort assoziiert. Um den Lernerfolg zu testen, werden die Blickbewegungen des Kindes auf dem Bildschirm gemessen. Die Blickbewegung des Kindes während der Benennung eines Objektes zeigt, ob das Kind zum *Target*-Objekt oder zum *Distraktor*-Objekt schaut (vergleiche Twomey et al., 2018). Der Vergleich zwischen den proportionalen Blicken zum Target oder zum Distraktor gibt dann Aufschluss darüber, ob die Wort-Objekt Assoziationen gelernt wurden. Wenn die Objekt-Wort-Assoziationen erfolgreich gelernt wurden, sollte das Kind insgesamt mehr zum Target als zum Distraktor schauen, so die Annahme dahinter. Wir erwarten, dass Kinder die Wort-Objekt-Assoziationen lernen und dies durch längere Fixationen auf das Target Objekt in der Testphase zeigen.

Eine andere Studie dieser Dissertation beschäftigt sich damit zu untersuchen, ob die Kinder dem Blick einer Person in unterschiedlichen Bedingungen (Blickbewegung kombiniert mit EGS oder KGS) folgen (Senju & Csibra, 2008). Dazu wird der Bildschirm in drei Regionen aufgeteilt.

Die Person, deren Blick gefolgt werden soll, befindet sich in der Mitte und rechts und links von ihr befinden sich zwei Objekte. Zu einem dieser Objekte (Target) wird der Blick der Person in der Mitte sich richten (Senju & Csibra, 2008). Die Blickbewegungen des Kindes werden gemessen, um herauszufinden, ob das Kind proportional mehr zum Target-Objekt schaut, im Vergleich zur Person selbst oder dem Distraktor-Objekt. Der Vergleich zwischen den drei Regionen gibt dann Aufschluss darüber ob das Kind dem Blick der Person folgt. Die Annahme ist hier, dass das Kind dem Blick der Person gefolgt ist, wenn es insgesamt mehr zum Target also zum Distraktor und zur Person selbst geschaut hat. Bei dieser Studie gab es zwei Bedingungen, die untersucht werden sollten, entweder war vor der Blickbewegung der Person KGS oder EGS zu hören. Wir erwarten hier, dass die Kinder generell der Blickbewegung des Models in KGS folgen, wobei die älteren Kinder (13-Monate) gegebenenfalls auch der Blickbewegung in EGS folgen werden.

Messung der Pupillengröße. Der Eyetracker wurde in dieser Dissertation außerdem dazu verwendet die Pupillengröße der Probanden zu untersuchen. Die Pupillengröße wird vom Eyetracker automatisch mit erhoben. Die Einordnung der Veränderung der Pupillengröße erfordert jedoch einen sehr kontrollierten Aufbau, da die Pupille auf unterschiedliche Reize reagiert, wie die Helligkeit des Bildes (Loewenfeld, 1993), aber auch unterschiedliche innere Stadien widerspiegelt, wie Erregung (Bradley et al., 2008; Hess & Polt, 1960; Partala & Surakka, 2003) oder die Anstrengung mit der eine Aufgabe gelöst wird (Kahneman & Beatty, 1966). In der Literatur gibt es zwei unterschiedliche Maße die Veränderungen der Pupille zu messen, phasische Veränderungen und tonische Stadien (Granhölm & Steinhauer, 2004). Diese Studie arbeitet mit den tonischen Stadien der Pupillengröße, um eine Veränderung in der Erregung der Kinder über einen längeren Zeitraum und als Resultat der experimentelle Manipulationen messen zu können (Hepach et al., 2012). Dazu wurden Pupillendaten zu Beginn und zu Ende jeder Phase (Training, Test, Bekannte Objekte) gemessen. Wir verwendeten dafür eine Bubble Video, welches mit Tönen hinterlegt war (Hepach et al., 2012; 2013; 2015). Die Daten wurden mit einer Samplingrate von 40Hz aufgenommen. Die Daten wurden in die weiter Analyse eingeschlossen, wenn für wenigstens 20% des Trials Daten aufgenommen wurden. Daten für jeden Time Bins (Zeitabschnitte) wurden für beide Augen extrahiert und weiter verwendet, wenn es sich um valide Daten einer Fixation

handelte. Im nächsten Schritt wurden die Daten gefiltert: der Unterschied zwischen zwei nebeneinanderliegenden Zeitpunkten wurde kapituliert und die oberen 10% der Daten wurden aus der Analyse ausgeschlossen, da es sich mit großer Wahrscheinlichkeit um Artefakte handelt (siehe auch Ackermann et al., 2020). Die entstandenen Lücken im Datensatz wurden interpoliert (Hepach et al., 2012). Als nächstes wurden die Daten über beide Augen gemittelt, wenn Daten für beide Augen verfügbar waren, ansonsten wurden Daten von einem Auge verwendet. Im letzten Schritt verbinden wir die gesamten Daten die wir über den Zeitraum des Bubble Videos erhoben haben und korrigieren sie, indem wir die Baseline Daten abziehen. Das Maß, welches wir erhalten spiegelt den Unterschied in der Pupillengröße durch die experimentelle Manipulation dar.

Segmentation – EEG. Die Methode der Elektroenzephalografie (EEG) und die der ereigniskorrelierten Potentiale (ERP) sind oft benutzte Methoden um die Gehirnaktivitäten zu messen und in Beziehung zu den kognitiven Entwicklungen der Kinder zu setzen (z.B. Junge et al., 2012; Männel & Friderici, 2011). Der Vorteil von EEG oder ERP Messungen im Vergleich zu verhaltensbezogenen Messmethoden liegt darin, dass sie bereits früh verwendet werden können, da die Kinder noch keine aktive Reaktion zeigen müssen. Die Veränderung der Gehirnaktivität spiegelt die Verarbeitung und Reaktion der Kinder auf einen Reiz wider. Diese Reaktion kann mit Hilfe von EEG und ERP Messungen sichtbar gemacht werden. EEG Messungen bilden dabei die kontinuierliche Aktivität des Gehirns ab, wohingegen ERP Messungen die potentiellen Veränderungen der Gehirnaktivität in Bezug auf einen bestimmten Reiz darstellen. Um eine Reaktion auf einem Stimulus zu untersuchen, werden daher zeitlich gebundene ERPs aufgenommen. Eine Vielzahl von unterschiedlichen Systemen steht auf dem Markt zur Verfügung. Unser Labor benutzt BIOSEMI. Die dehnbare Kappe wird dem Kind, wie eine Mütze, auf den Kopf gesetzt. Die Plastikeinbuchtungen, in denen die Elektroden befestigt werden, werden mit Gel befüllt, bevor die Elektroden eingesteckt werden. Das Gel dient zur Verbesserung der Leitfähigkeit des elektronischen Signals vom Schädel der Kinder zu den Elektroden. Zusätzlich zu den Elektroden auf der Kappe werden zwei Referenz-Elektroden jeweils hinter dem rechten und linken Ohr auf dem Mastoid Knochen platziert. Es wird angenommen, dass diese Referenz-Elektroden keine neuronale Aktivität messen, sondern lediglich Störsignale. Diese Signale werden später

offline von allen Elektroden abgezogen. Während der Studie werden die Stimuli mit Markern im EEG-Signal gekennzeichnet, somit kann die zeitlich gebunden Reaktion auf den Stimulus offline ermittelt werden. Nachdem die Studie durchgeführt wurde, werden die Daten offline verarbeitet. Nachdem alle Elektroden zu den Referenz-Elektroden referenziert wurden, werden die Daten mit einem Bandpassfilter gefiltert (0.1-30Hz) und die Samplingrate wird von 2048 auf 250Hz reduziert. Um die Reaktionen auf die präsentierten Stimuli zu untersuchen, wird das kontinuierliche EEG Signal in kleinen Abschnitte aufgeteilt, beginnend bei 200ms vor dem Beginn des Stimulus und endend 800ms nach Beginn des Stimulus. Die 200ms vor dem Start des Stimulus dienen als Grundlinie. Die Abschnitte werden einer automatisierten Artefaktkorrektur unterzogen. Die Artefaktkorrektur beschränkt sich auf die Elektroden, die für die Studie von Relevanz sind: F7, F3, FC1, FC5, C3, F8, F4, FC2, FC6 und C4 (REF) und schließen Trials aus die den Spannungsgrenzwert von $200\mu\text{V}$ überschreiten. Im Folgenden werden Trials von ein und derselben Bedingungen gemittelt. Um eine repräsentative Anzahl von Trails pro Proband zu haben, werden Probanden mit weniger als 10 Artefakt freien Trials pro Bedingung nicht mit in die finale Analyse einbezogen. Die gemittelten Kurvenverläufe der ERPs jedes Probanden werden danach über die unterschiedlichen Bedingungen gemittelt. Dadurch kann man die einzelnen Bedingungen miteinander vergleichen und Unterschiede basierend auf den Bedingungen feststellen.

Die hier beschriebene Studie beschäftigt sich mit der Fähigkeit der Kinder kurz zuvor familiarisierte Wörter aus KGS oder EGS zu segmentieren. Um zu untersuchen ob Kinder Wörter besser aus KGS oder aus EGS segmentieren können, wird jedes Kind mit ihm unbekanntem Wörtern in KGS oder EGS familiarisiert. Dies geschieht, indem die neuen Wörter in Sätzen präsentiert werden. Direkt nach der Präsentation der Sätze, wird getestet, ob das Kind die neu gehörten Wörter aus dem Sprachfluss der Sätze segmentieren konnte, indem das Target Wort und ein neues unbekanntes Kontrollwort geäußert werden. Wird das Target Wort aus dem Sprachfluss segmentiert, sind Unterschiede in den ERPs zwischen Target und Kontrollwort zu sehen. Zusätzlich sind wir an Unterschieden zwischen den Bedingungen interessiert. Wir erwarten, dass Kinder im Alter von 13 Monaten in der Lage sind Wörter aus KGS zu segmentieren und wir dies durch eine Negativität in der linken frontalen Gehirnhälfte nachweisen können (vgl. Männel &

Friederici, 2013). Für EGS erwarten wir keine direkte Evidenz für die Fähigkeit Wörter zu segmentieren in diesem Alter. Auch für die jüngeren Kinder erwarten wir keine Evidenz dafür zu finden, dass sie Wörter aus EGS segmentieren können. Genauso wenig erwarten wir, basierend auf der Literatur im deutschsprachigen Raum, eine Evidenz für die Segmentationsfähigkeit von KGS in diesem Alter zu finden (Schreiner et al., 2016, 2017).

Aufnahme Körperhaltung. Eine Studie diese Dissertation beschäftigt sich mit der Veränderung der Körperhaltung von Kindern als Indikator für eine emotionale Veränderung. Bereits ab einem Alter von zwei Jahren können Kinder merkbar ihre Körperhaltung verändern. Dies biete die Möglichkeit die Gefühle, z.B. die innere Erregung der Kinder zu untersuchen (Lewis, 1997). In Form von einer aufrechteren Körperhaltung konnten bei Erwachsenen bereits Gefühle gemessen werden, wie zum Beispiel Stolz (Shiota et al., 2003) oder sportlicher Erfolg (Weisfeld & Bresford, 1982). In Studien mit Kindern konnte bereits im Alter von drei Jahren eine aufrechtere Körperhaltung gezeigt werde, wenn die Kinder eine schwere Aufgabe gelöst haben (Lewis et al., 1992). Diese und eine Vielzahl anderer Studien weisen auf einen Zusammenhang zwischen Erfolg und einer aufrechteren Körperhaltung hin (siehe Hepach et al., 2015 für einen Überblick).

Um die Körperhaltung zu messen, werden meist Videos oder Fotos nach einem bestimmten Schema codiert (e.g. Lewis et al., 1992; Shiota et al., 2003; Martens & Tracy, 2013). Neuere Methoden verwenden auch computeranimierten Mannequins oder Lichtpunkte (e.g. Coulson, 2004). In dieser Studie wird eine *Kinect motion capture Camera* (Microsoft Kinect) verwendet. Diese verwendet eine tiefensensorische Bildgebungstechnologie, welche zusätzlich zu Aufnahmen von RGB Bildern noch Informationen über die Entfernung jedes einzelnen Pixels zum Aufnahmegerät gibt. So gibt die Kamera X,Y und Z Koordinaten von bis zu 20 Körperpunkten aus (für eine genauer Beschreibung des Apparates und seiner Funktion siehe Hepach et al., 2015; Shum et al., 2013), wenn sich die Person im Abstand von 1,5 bis 4 m zum Gerät aufhält. Studien haben diese Technologie bereits verwendet um zum Beispiel die Interaktionen von Eltern und Kindern zu untersuchen (Nagai et al., 2012) aber auch um die Veränderungen in der Körperhaltung als

Indikator für einen emotionalen Ausdruck zu messen (Hepach et al., 2015, 2017). Hepach, und Kollegen (2015, 2017) zeigen, dass Kinder eine positivere Emotion zeigen, was durch eine aufrechtere Körperhaltung angezeigt wird, wenn sie ein Ziel erfolgreich erreichen oder wenn sie jemand anderem erfolgreich helfen, ein Ziel zu erreichen. Die Autoren betonen hier besonders den Zusammenhang zwischen Körperhaltung und emotionalem Ausdruck und heben die Möglichkeit hervor, dass der Erfolg einer Aufgabe zu positiveren Emotionen führen kann. Der letztere Vorschlag basiert auf ihren Ergebnissen einer Zunahme der Brusthöhe bei Erwachsenen nach einer positiven Emotion im Vergleich zu einer negativen Emotion (Hepach et al., 2015). In dieser Dissertation wird ebenfalls die Körperhaltung (vor allem die Oberkörperhaltung) als Indikator für einen emotionalen Ausdruck verwendet. Es wurde untersucht inwieweit der Erfolg einer Wortlernaufgabe Veränderungen der Oberkörperhaltung von Kindern als Index für die positiven Emotionen vorhersagt, die durch erfolgreiches Wortlernen ausgelöst werden können. Dazu werden die Kinder zuerst mit der oben beschriebenen Cross situational Wortlernaufgabe präsentiert (in einem separaten Testraum). In einem angrenzenden Raum ist die Kinect Kamera auf einem kleinen Tisch mit einem Computer platziert und wird von einem Studienleiter gestartet sobald das Kind aus dem *Eyetracking-Raum* in den *Kinect-Raum* kommt. Das Kind muss zur Aufnahme der Körperhaltung möglichst gerade auf die Kamera zulaufen, dazu wurde ein Spielzeug direkt neben der Kamera platziert. Bevor das Experiment begann, wurde dem Kind gezeigt, wo sich das Spielzeug befand, und es wurde ihm gesagt, dass es während der Studie mehrmals zu diesem Spielzeug gehen soll. Das Kind ging ungefähr 5 m zwischen der Tür zum Testraum und dem Spielzeug. Die Eltern wurde angewiesen, Abstand zum Kind zu halten, um während der Messung der Körperhaltung des Kindes nicht mit aufgenommen zu werden. Die Messung der Körperhaltung beginnt mit der Aufnahme der Grundhaltung (Baseline), in denen die Kinder zweimal auf die Kinect Kamera zulaufen, ohne eine weitere Aufgabe davor oder danach zu erhalten. Nach diesen Aufnahmen finden die Lern- und Testphase, sowie die Phase mit den bekannten Objekten im Eyetracking Raum statt. Zwischen jeder Phase, also nach der Lernphase, nach der Test-Phase und nach der Phase der bekannten Objekte, läuft das Kind auf das Spielzeug zu und die Körperhaltung wird gemessen. Um die Körperhaltung der Kinder nicht zu beeinflussen, sind die Studienleiter angehalten sämtliche

positiven oder negativen Kommentare in Bezug auf das Eyetracking oder das Ende der Studie zu unterbinden.

Die erhaltenen Daten wurden offline weiter analysiert. In einem ersten Schritt haben wir die Daten gefiltert und bereinigt, um sicherzustellen, dass wir keine Skelette von Erwachsenen und nur Daten in unserem Verfolgungsbereich (3,2 m - 1,2 m Entfernung von der Kinect-Kamera) einbezogen haben. Desweiteren wurden Daten nur eingeschlossen wenn sich die Skelettpunktmitte (also der Skelettpunkt des Rückens) über der Skelettpunkthüftmitte befand, der Skelettpunkt des Kopfes über den Skelettpunkten der Schulter lag, die Skelettpunktmitte des Rückens zwischen den Skelettpunkten links und rechts des Rückens lag und der Skelettpunkt der mittleren Schulter zwischen dem rechten und dem linken Skelettpunkten der Schulter lag. Diese Kriterien stellten sicher, dass nur Datenpunkte in die Analysen einbezogen wurden, die vernünftigerweise als Skelettmessungen des Kindes angesehen werden konnten (siehe Hepach et al., 2016). Die resultierenden Lücken in den Skelettmessungen wurden interpoliert.

Die weiter Analyse konzentrierte sich vor allem auf zwei Körperpunkte, nämlich die y-Koordinate (Höhe) des Brustzentrums, um die Änderung der Oberkörperhaltung abzuschätzen, und die y-Koordinate des Hüftzentrums, um die Änderung der Unterkörperhaltung abzuschätzen (Hepach et al. 2016). Anschließend werden die Messungen der Körperhaltung über die Zeit skaliert, sodass alle Teilnehmer Daten für 20 Datenpunkte in jedem Lauf bereitstellten, unabhängig davon, wie lange sie für jeden Lauf gebraucht haben. Dafür haben wir den Lauf in zwanzig Zeiträume unterteilt und für jeden dieser Zeiträume ein durchschnittliches Haltungsmaß extrahiert. Das Haltungsmaß ist in der Regel weniger genau, je näher der Teilnehmer an der Kinect-Kamera ist. Um sicherzustellen, dass diese weniger genau gemessenen Datenpunkte von der Analyse ausgeschlossen werden, haben wir nur die Daten für die ersten 16 gesammelten Zeitfächer berücksichtigt.

Anschließend Baseline-korrigierten wir die Daten separat für jeden Lauf, der erfasst wurden, das heißt die Daten nach der Lernphase, die Daten nach der Testphase und die Daten nach der Phase mit den bekannten Objekten. Während der Baseline-Korrektur subtrahierten wir die Baseline-

Haltungsdaten von dem Maß, das nach Abschluss jeder Phase für jeden Zeitbereich und jede Phase separat erhalten wurde. Wenn das Kind Daten für beide Baseline-Läufe bereitstellte, wurde der Durchschnitt beider Läufe für die Baselinekorrektur verwendet. War nur ein Baselineauf vorhanden, so wurde dieser verwendet. Diese Baseline korrigierten Daten wurden dann für weitere Analysen verwendet um zu untersuchen, ob das erfolgreiche Lernen und die innere Erregung der Kinder während der Studie einen Einfluss auf ihre Körperhaltung hat.

Zusätzlich werden die von der Kinect gelieferten Bilder manuell codiert. Es wurde zum Beispiel festgehalten wie glücklich die Kinder aussahen. Dies wurde mit den Körperhaltungsdaten korreliert und soll zusätzliche Evidenz für die Messung der Körperhaltung als Index für emotionale Stadien geben.

Untersuchungspunkte der Studien und verwendete Methoden

Die Studien dieser Dissertation verwenden meist eine Kombination aus mehreren Methoden um die Interaktion von kindlicher Aufmerksamkeit, Input und Interesse besser untersuchen zu können.

Die erste hier beschriebene Studie besteht aus drei unterschiedlichen Methoden: Die Mütter produzieren KGS und EGS, die Kinder nehmen an der modifizierten Variante des HPP, der zentralen fixations-Methode; und an einem Eyetracking Experiment, welches das Folgen der Blickbewegung eines Modells untersucht, teil. Durch die Durchführung jedes einzelnen Teils dieser Studie, erhielten wir Informationen über die Eigenschaften von KGS und EGS, das deutschsprachige Mütter mit ihren 6- und 13- Monate alten Kindern verwenden. Die Durchführung des zentralen Fixation Paradigmas erlaubte uns festzustellen welchen Sprachinput 6- und 13- Monate alter Kinder bevorzugen, bzw. welchem Sprachstil sie länger aktiv zuhören. Dies lässt uns Rückschlüsse auf die Präferenz für KGS und EGS in unterschiedlichen Stufen der Entwicklung ziehen. Das Messen der Blickbewegung der Kinder diente dazu, um zu untersuchen, bei welchem Sprachstil die Kinder der Blickbewegung einer anderen Person mehr folgen, bei KGS oder bei EGS. Doch erst das Zusammenspiel all dieser Studien und die Durchführung jeder einzelnen mit den gleichen Kindern in zwei Altersklassen machte diese Studie interessant. Die Untersuchung von

mütterlichen Input in Verbindung mit der Präferenz der Kinder für einen Sprachstil liefert Einblicke in die Interaktion zwischen kindlicher Aufmerksamkeit für KGS und dem verfügbarem Input, in Form von prosodischen Eigenschaften der mütterlichen KGS. Durch die Untersuchung der Blickbewegung der Kinder konnte des Weiteren untersucht werden, ob es eine Verbindung zwischen Input und Präsenz und der allgemeinen Reaktion auf KGS gibt. Diese Studie dient somit vor allem dazu, zu untersuchen ob es eine Verbindung gibt und wie weit diese existiert.

Die zweite Studie dieser Dissertation verbindet zwei der oben beschriebenen Studien (Aufnahmen des mütterlichen Inputs und die Messung der kindlichen Präferenz) mit der Untersuchung der Segmentationsfähigkeit von Kindern durch ereigniskorrelierte Potentiale. Die Entscheidung, die Segmentationsfähigkeit von Wörtern aus fließender Sprache von Kindern mit Hilfe von ERP's zu untersuchen, basiert vor allem auf den Schwierigkeiten, diese bei jungen deutschsprachigen Kindern mit Hilfe von Verhaltensuntersuchungen nachzuweisen (Schreiner & Mani, 2017). Daher wählten wir den sensitiveren Index für die kindliche Segmentation von Wörtern. Dazu untersuchten wir die Unterschiede in den Potentialen von Wörtern, die vorher in KGS oder EGS familiarisiert worden waren. Wie aber bereits in der ersten Studie diese Dissertation ist auch hier vor allem das Zusammenspiel der drei Aufgaben von größerem Interesse. Durch das Untersuchen einer möglichen Interaktion von Input und kindlicher Aufmerksamkeit für KGS mit der kindlichen Fähigkeit Wörter besser in KGS oder EGS zu segmentieren, erschließt sich, inwieweit der Input und die kindliche Aufmerksamkeit einen Einfluss auf das Erlernen von Sprache haben.

Die dritte Studie dieser Dissertation nutzt die Methoden des Eyetrackings, indem es die Blickbewegung der Kinder, aber auch die Pupillengröße misst. Die Blickbewegung ist in dieser Studie interessant, da sie Aufschluss darüber gibt, ob ein Kind eine neue Wort-Objektassoziation erlernt hat oder nicht. Das Messen der Pupillengröße dient in dieser Studie dazu mehr über die intrinsische Motivation der Kinder zu erfahren. Sind Kinder in einen gewissen Zusammenhang mehr erregt und können wir diese innere Erregung mit Hilfe der Pupillengröße in Verbindung mit dem Lehrerfolg bringen? Zusätzlich waren wir daran interessiert mit Hilfe der Messung der oberen

Körperhaltung herauszufinden, ob Kinder eine positive Emotion zeigen, wenn sie einen hohen Lernerfolg zeigen. Um diese Interaktionen zu untersuchen, wurde zuerst die Bildung neuer Wort-Objekt Assoziationen in einer cross-situational Wortlernaufgabe gemessen und mit der Veränderung der Pupillengröße in Verbindung gebracht. Wir erwarteten, dass Kinder, die nach der Testphase der neu erlernten Wort-Objekt Assoziationen eine Vergrößerung der Pupillen zeigen, die neuartigen Wort-Objekt-Assoziationen besser wiedererkennen werden, was auf eine erhöhte Erregung im Zusammenhang mit dem Erfolg der Aufgabe hinweist. Durch das Einbeziehen der Körperhaltung können wir untersuchen ob sich der Lernerfolg der Kinder auch mit ihren Emotionen in Verbindung bringen lässt. Wir stellten die Hypothese auf, dass Kinder, die ein besseres Lernen der neuartigen Wort-Objekt-Assoziationen zeigten, aufgrund ihres Erfolgs bei der Aufgabe positivere Emotionen erfahren würden, die durch die Erhöhung der Oberkörperhaltung, nach Abschluss der Aufgabe, indiziert werden könnten.

Maternal input and infants' response to infant-directed speech

Abstract

Caregivers typically use an exaggerated speech register known as infant-directed speech (IDS) in communication with infants. Infants prefer IDS over adult-directed speech (ADS) and IDS is functionally relevant in infant-directed communication. We examined interactions between maternal IDS quality, infants' preference for IDS over ADS, and the functional relevance of IDS at 6- and 13-months. While 6-month-olds showed a preference for IDS over ADS, 13-month-olds did not. Differences in gaze following behavior triggered by speech register (IDS vs. ADS) were found in both age groups. The degree of infants' preference for IDS (relative to ADS) was linked to the quality of maternal IDS infants were exposed to. No such relationship was found between gaze following behavior and maternal IDS quality and infant IDS preference. The results speaks to a dynamic interaction between infants' preference for different kinds of social signals and the social cues available to them.

Introduction

Infants and young children are typically addressed using a different speech style – referred to as infant-directed speech (hereafter, IDS) – compared to adults. IDS differs from the kind of speech adults use to communicate, i.e., adult-directed speech (hereafter, ADS), with typically higher pitch, greater pitch range and longer pauses between words in IDS compared to ADS (Fernald et al., 1989). Caregivers produce IDS spontaneously, with variation in the quality of IDS across individuals, development, languages and cultures (Cristia, 2013; Soderstrom, 2007). Furthermore, the use of IDS in interactions with infants plays a crucial role in early language development. First, we summarize findings from studies examining the role of IDS in early language development, before outlining the questions under investigation in the current study.

What is infant-directed speech?

IDS and ADS differ along prosodic, phono-lexical, and syntactic dimensions (Cristia, 2013; Soderstrom, 2007). At the prosodic level, the primary difference between IDS and ADS is the higher pitch used in interactions with infants across different age groups (Cristia, 2013), with increased variability in pitch, lengthening of vowels, and pauses in IDS relative to ADS. At the phono-lexical and syntactic levels, IDS is linguistically simpler, including shorter, more redundant utterances, more questions, and increased repetition of single words (Grieser & Kuhl, 1988; Soderstrom, 2007; Soderstrom et al., 2008). Vowels produced in IDS show more disparate formant structures than in ADS (Kuhl et al., 1997). IDS also plays an affective role in early language, such that IDS is typically associated with the expression of positive emotions (Singh et al., 2002).

However, there is considerable variation in the acoustic characteristics of IDS and the use of IDS in communication with infants across different linguistic cultures (Fernald et al., 1989; Werker et al., 2007; Farran et al., 2016; Schieffelin & Ochs, 1983). Research also suggests variation in the characteristics of IDS across development: While mothers' speech increases in amount and complexity in interactions with 3-month-olds compared to 1-month-olds (Henning et al., 2005; see also Kaye, 1980), the mean length of utterance decreases between early infancy and the second year of life (Murray et al., 1990; Sherrod et al., 1977; Stern et al., 1983). At the segmental level, mothers differently exaggerate acoustic contrasts in IDS across development with,

for instance, reduced voice-onset-time overlap between voiced and voiceless stops, in speech to 15-month-olds, while mothers of 6- to 9-month-olds produced these contrasts similarly across IDS and ADS (Malsheen, 1980; see Bernstein-Ratner, 1984; Cristia, 2010 for other differences). Overall, while caregivers spontaneously produce infant-directed speech in communication with infants, there is considerable variation at the individual, developmental, cultural, linguistic level in the characteristics of maternal IDS that infants are exposed to.

IDS preference

Neonates and young infants prefer to listen to IDS compared to ADS (e.g., Cooper & Aslin, 1990; Pegg et al., 1992). This preference remains relatively stable between 4-8 months (Fernald, 1985; Fernald & Kuhl, 1987; Glenn & Cunningham, 1983; Kaplan et al., 1995a, b; Pegg et al., 1992; Schachner & Hannon, 2011; Singh et al., 2002; Werker & McLeod, 1989; but see Hayashi et al., 2001 below) while findings with infants aged 9- to 18-months are mixed (e.g., Glenn & Cunningham, 1983; Hayashi et al., 2001; Newman & Hussain, 2006; Segal & Newman, 2015). There are further differences across studies based on the language being tested, the method employed, the speaker and the type of speech used in the experiments (Cooper et al., 1997; Pegg et al., 1992, Kaplan et al., 1995a, b). While Hayashi and colleagues (2001) propose a U-shaped development in the preference for IDS, recent cross-linguistic data support a preference for IDS across development, with a stronger IDS preference in older infants and in infants from a North American English background, at least when presented with North American English IDS (ManyBabies Consortium, 2020).

The functional relevance of IDS

Infants also learn better from stimuli presented in IDS compared to ADS: Studies suggest that infants segment words better from fluent speech when the speech is more infant-directed (Floccia et al., 2016; Schreiner & Mani, 2017; Singh et al., 2009; Thiessen et al., 2005). Infants are able to segment words from ADS (see Mani & Pätzold, 2016; Schreiner et al., 2016), but they do so more easily (with fewer cues and less exposure) in IDS. Infants also learn words easier from IDS relative to ADS (Graf-Estes & Hurley, 2013; Ma et al., 2011), and the use of IDS in communication with infants correlates with increased vocabulary scores later in life (Fisher &

Tokura, 1996; Vosoughi et al., 2010; Weisleder & Fernald, 2013). Differences in infants' brain responses to IDS and ADS across development (Naoi et al., 2012; Zangl & Mills, 2006) similarly speak to the brain's dynamic and flexible response to IDS across development.

An influence of IDS on language development is typically explained using one of two accounts (Golinkoff et al., 2015). Infants may learn language better from IDS because IDS provides more information, e.g., more exaggerated cues to the boundaries between words, whilst at the same time providing information reduced in complexity (e.g., Fernald, 1989; Fernald et al., 1989). Alternatively, IDS may be more attractive, salient and relevant to the infant and draw attention to the input, thereby bolstering learning (e.g., Zangl & Mills, 2006). It may be difficult to adjudicate between these two approaches in the linguistic domain since the information to be attended to and learned is necessarily presented in IDS (including any potential hyperarticulation that may equally boost learning).

An examination of the influence of IDS on infants' performance in a non-linguistic task may better inform us of the extent to which IDS perceptually attracts infants' attention to the current interaction and, supports learning. Such evidence is provided by work examining the impact of IDS on joint attention (Senju & Csibra, 2008), showing that infants followed the gaze of a model more reliably and quickly when the interaction was cued with IDS relative to ADS provided, highlighting the functional relevance of IDS to interactions with infants (see also Gredebäck et al., 2018).

The current study

These findings reveal that (1) caregivers spontaneously produce IDS albeit with variation across caregivers, cultures, languages and development, (2) infants show a preference for IDS over ADS which also varies across development, (3) IDS has a beneficial impact on language development, and (4) IDS is functionally relevant even in non-linguistic tasks, highlighting the salience of IDS in interactions with young infants.

While the aspects of IDS in parent-infant communication highlighted above have been much studied in isolation, such studies do not allow an examination of the intriguing possibility that they may be related. Recent work highlights a positive relationship between maternal IDS

prosody and later linguistic outcomes and infant attention (see Spinelli et al., 2017 for a review and a recent meta-analysis). These studies suggest that pitch modulation of maternal IDS has a significant influence on infants' later pre-linguistic and linguistic outcomes (e.g., D'Odoricio & Jacob, 2006; Lyasko et al., 2014) including their later joint attention skills (Roberts et al., 2013). Equally, studies suggest that infants respond better in the moment to stimuli that are more prosodically exaggerated (e.g., Niwano & Sugai, 2002a, b; Butler et al., 2014).

Tying these two veins of research together, the current study examines whether there is a relationship between infants' in-the-moment response to IDS and the prosodic properties of maternal IDS they have been exposed to. Such a relationship is presupposed in theories examining the role of IDS in learning. For instance, pedagogical theories of learning suggest that parents employ IDS as an ostensive cue to the infant in order to signal that the interaction is relevant to them and tailor their input to the child in order to provide them with optimal input (Csibra & Gergely, 2009, 2011). Evolutionary developmental theories of language acquisition highlight the importance of IDS in capturing infants' in-the-moment attention in day-to-day interaction, with parents being sensitive to the attentional salience of IDS and modulating their speech based on their infants' responses (McMurray, 2016; Smith & Trainor, 2008). Thus, both these different theoretical accounts imply a close relationship between maternal IDS and infants' response to IDS. Against this background, the current study examines whether and to what extent such a relationship exists, i.e., whether the quality of maternal IDS covaries with infants' attention and response to IDS, in general.

We tested the same infants on their preference for listening to IDS, the functional relevance of IDS, and individual differences in the prosodic characteristics of IDS produced by their caregivers, to examine the interaction between these factors in parent-infant communication. First, we analyzed recordings of mothers producing IDS and ADS to examine whether the quality of German mothers' speech differs across communication with infants and adults. We focused here on the suprasegmental characteristics of IDS given the studies showing an influence of maternal prosody on infants' later language outcomes. Second, a preferential listening task examined whether infants of those mothers prefer listening to IDS over ADS (Cooper et al., 1997). Third, a

gaze following task explored infants' gaze following after someone "speaking" IDS and the same person "speaking" ADS (Senju & Csibra, 2008).

We tested different groups of infants at two specific points, i.e., at 6- and 13-months of age, to examine if these aspects of IDS in early life change across development. We target these age-groups specifically given previous findings of differences in the neural and behavioral responses to IDS and ADS (Naoi et al., 2012; Schreiner & Mani, 2017; Zangl & Mills, 2006), as well as differences in the influence of maternal IDS on learning in younger and older infants (see Spinelli et al., 2017 for meta-analyses). Examining the interaction between these findings at two different ages allows us to investigate how these factors develop across infancy. Given the extensive testing schedule, involving three tasks spread across two days (in each age-group), we collected data from two separate groups of infants at 6- and 13-months.

We expect German mothers to produce systematic prosodic differences in their IDS and ADS at both ages tested (Fernald et al., 1989; Fernald & Simon, 1984). It is less clear whether these differences will vary across the two ages, given the limited research on prosodic differences in IDS across development. Based on previous research, we expect 6-month-olds to show a preference for listening to IDS compared to ADS, while it is less clear whether this preference will remain at 13-months of age (e.g., Newman & Hussain, 2006). Our hypotheses for the gaze following task are similar: Six-month-olds ought to follow the gaze of someone faster in the IDS relative to the ADS condition (Senju & Csibra, 2008) while it remains uncertain how the 13-month-olds might respond. Studies suggest a developmental change in infants' gaze following behavior between 6-months and 13-months (e.g., Carpenter et al., 1998; Woodward, 2003), with the older infants relating the person to the fixated object and showing robust performance in naturalistic settings.

The interaction between the individual tasks allows us to examine (1) the potential factors underlying infants' IDS preference by investigating the extent to which this preference is related to the kind of language input infants are exposed to, (2) the extent to which a preference for listening to IDS is functionally relevant in non-linguistic tasks, such as the gaze following task, and (3) the extent to which this functional relevance of IDS may be related to the kind of input available to the

infant. Taken together, the simultaneous analysis of the three tasks in the same group of infants allows us to examine potential variation in interaction between these aspects of IDS processing across development.

Method

Participants

Eighty-one monolingual German infants were recruited. Forty-two infants at the age of 6 months (5m16d – 7m10d, 19 girls) and 39 infants at the age of 13 months (12m15d – 14m12d, 19 girls) took part in the study. Participants were excluded due to illness on the second day of testing (n=8), fussiness on the part of the baby (n=19), experimental error (n=3), and data loss (n=10) (for a list of participants see Table 1).

Table 1 Number of infants who provided data across tasks and age groups

| Data obtained for each task/ combination of tasks | 6-month-olds (number of infants) | 13-month-olds |
|--|-------------------------------------|---------------|
| Maternal recording | 27 | 27 |
| Preferential listening task | 38 | 36 |
| Gaze following task | 30 | 29 |
| Maternal recording and preferential listening task | 24 | 26 |
| Maternal recording and gaze following task | 23 | 22 |
| Preferential listening task and gaze following task | 27 | 28 |
| All three tasks | 21 | 21 |

General Procedure

The study consisted of three sub-tasks: a preferential listening task, a gaze following task, and recordings of maternal speech. The tasks were split across two days. On the first day, infants participated in the preferential listening task, in one of the conditions of the gaze following task,

and maternal ADS was recorded. After five to 30 days, infants returned to take part in the second condition of the gaze following task and the mother was recorded producing IDS. Additionally, infants also took part in an EEG experiment, whose results are not reported in the current paper. Ethics approval was obtained from the Ethics committee of the University's Psychology Institute.

Recordings

Procedure. Mothers were recorded on two separate days, producing ADS and IDS. At the first appointment, the mother watched a short video (one minute) about a teddy bear, having a small adventure, including different locations and people in order to retell the story shortly after. When the mother felt confident about retelling the story, she walked into a room, sat on a chair facing a woman's face displayed on screen. She was told to imagine retelling the story to a friend. Recordings were obtained via a microphone placed in front of the mothers' feet. While this may have led to differences in the quality of maternal ADS due to parents being left on their own rather than in interaction with another adult, we note that the properties of maternal ADS recordings were similar to those reported in the literature (Fernald et al., 1989). We did not ask the parent to tell the story to another adult in case they felt self-conscious being recorded speaking to an unfamiliar adult, thereby impacting the quality of their speech.

At the end of the second appointment, mothers were recorded telling the same story in IDS. The mother could watch the video again in order to retell the story to her own infant. The experimenter left the room but recorded the interaction between mother and infant via a microphone and a camera placed about 1.5 m away. We did not counterbalance the order of recordings to ensure that the quality of ADS recordings was not influenced by the previous production of similar sentences in IDS. We assumed that the prior production of the stimuli in IDS would be more likely to influence the quality of maternal ADS (based on our experience recording stimuli across numerous studies) than the other way around.

Data Analysis. Average duration of IDS recordings were 152s with an average of 84s analyzable speech (minimum: 28s, maximum: 210s). ADS recordings were on average 118s with

an average duration of 56s analyzable speech (minimum: 25s, maximum: 126s). We calculated the mean, maximum, minimum pitch, pitch range and average duration of utterances separately for the IDS and ADS recordings using PRAAT software (Boersma & Weenink, 2019). Mothers provided, on average, 22 noise-free utterances in IDS ($SD = 11.1$) and 19 noise-free utterances in ADS ($SD = 8.9$). Paired samples t-tests examined whether there was a difference in the prosodic characteristics of maternal IDS and ADS at both age-groups. We also calculated the difference in the prosodic characteristics of maternal IDS and ADS and then compared this difference across the two age groups.

Preferential listening task

Material and Design. A female native speaker of German recorded six short stories for this task (see Appendix 1). Each story was recorded twice, once in IDS and once in ADS. Acoustic analysis using PRAAT software (Boersma & Weenink, 2019) ensured the validity of the stimuli. Paired-samples t-tests run across individual sentences in each story revealed significant differences in the maximum pitch, $t(41) = 15.06, p < 0.001$, minimum pitch, $t(41) = 6.17, p < 0.001$, and mean pitch, $t(41) = 15.14, p < 0.001$, of IDS and ADS stimuli. There were no significant differences in the pitch range, $t(41) = 0.41, p = 0.681$ and duration, $t(41) = 1.310, p = 0.198$ (see Appendix 2.1 for a list of characteristics). The acoustic characteristics of the stimuli were similar to previous reports of naturalistic German IDS and ADS (Fernald et al., 1989).

Procedure. Each infant was tested in a quiet room. Infants sat on the parent's lap or in a car seat about 100 cm away from a monitor. This monitor presented infants with a blinking checkerboard. The auditory stimuli were presented via loudspeakers, located above the television screen. Two cameras directly above the screen recorded the infant's eye movements. Synchronized signals from the two cameras were routed via a digital splitter to provide two separate time-locked images of the infant, which were used for online and offline coding.

Stimuli were presented using the Look software (Meints & Woodford, 2008). During each trial, infants were presented with a blinking checkerboard on the screen. Each trial began with a visual and auditory attention getter to direct the infant's attention to the screen. The auditory stimulus was presented once the infant fixated the screen. An experimenter, seated in an adjacent

room watching a video of the infant, indicated whether the infant was looking at the screen or away by pressing a button on a keyboard. The screen remained black in between trials until the attention getter started the next trial. Infants listened to alternating IDS and ADS trials each with a different story (with a maximum length of 30 seconds). Infants listened to 6 trials, half of which were presented in IDS and half of which were presented in ADS – counterbalanced across infants for the register. Half of the participants were presented with an IDS story in the first and an ADS story in the second trial, and vice versa. Subsequent trials alternated between conditions (IDS-ADS-IDS-ADS-IDS-ADS or ADS-IDS-ADS-IDS-ADS-IDS). The trial continued until the end of the auditory stimulus or until the infant looked away from the screen for more than 2 seconds. A second independent coder evaluated a random sample of 10% of each condition offline. The interrater reliability was high ($r = .937$).

Data analysis. All results refer to the mean duration of the first fixation at the checkerboard averaged for each participant across trials separated by condition. We only included data from infants who provided data for both conditions in the first half of the experiment, i.e., infants needed to provide data for at least one trial in each condition in the first three trials of the experiment. Since looking times drop dramatically across the duration of this task, this criterion avoided differences across conditions, which were driven by the drop in infants' interest towards the latter half of the experiment. This led to exclusion of two infants in each age-group.

Data were analyzed using linear mixed models, ANOVAs and *t*-tests. Due to the consistency across these analyses, we report the results of performance within each task using the ANOVAs and *t*-tests and only report the results of the linear mixed effects models when reporting the correlations across task. The dependent variable was the listening time to IDS and ADS trials. We ran a mixed ANOVA with the within-subjects factor condition (IDS, ADS), and between-subjects factors age (6m, 13m) and order of condition (IDS first, ADS first). The order of presentation of trials was included based on results in the literature (Cooper et al., 1997; Kaplan et al., 1995, but see also Cooper & Aslin, 1990, 1994) suggesting that preferential listening data is strongly influenced by the order of presentation of trials in the study with infants attending more to the early trials in the experiment than the later trials.

Performance across task. We examined the correlations in performance across the three tasks (recordings of maternal input, preferential listening and gaze following) using linear mixed effects models. These models offer the advantage that we can control for performance in one task while examining the relationship in performance between two other tasks.

For the preferential listening task, we ran a linear mixed effects model including condition (IDS, ADS), age (6m, 13m), and order of presentation (IDS first, ADS first) as fixed effects, and subject and item as random effects. We compared this model to a reduced model, excluding condition, to examine the effect of condition on the response. To investigate the performance across tasks, we included variables capturing infants' performance in the gaze following task and maternal IDS quality. Specifically, we included the difference in proportion of target fixations in the IDS and ADS conditions in the gaze following task and the difference in critical prosodic features of maternal IDS and ADS. Thus, the model included the interaction between condition and the prosodic characteristics of maternal IDS, e.g., the difference between maternal mean pitch, pitch range and utterance duration in IDS and ADS and the interaction between condition and the difference in the proportion of target fixations in the IDS and ADS condition in the gaze following task, with subjects, order of presentation and item as random effects (where possible without violating singularity of fit).

A drop1 analysis then examined the key factors that, in interaction with condition (IDS and ADS trials in the preferential listening trials), explained the variance in listening times. This analysis circulated between removing specific variables from the model to estimate the best model fit. We only considered factors that influenced listening times in interaction with condition since main effects would merely indicate that a particular factor impacted listening times overall, and not suggest an effect of this factor on infants' IDS preference.

Finally, following the suggestion of one of the reviewers, we calculated the difference in the acoustic characteristics of maternal IDS and ADS and the acoustic characteristics of IDS and ADS stimuli presented in the preferential listening task $((IDS_{pref} - ADS_{pref}) - (IDS_{mat} - ADS_{mat}))$ and examined whether adding this difference (using a similar drop1 analyses) improved model fit of the preferential listening time data. This analysis allows us to examine how infants' preference

for the IDS presented at test varied as a factor of how similar this IDS was to the IDS they received at home.

Gaze following

Material and Design. The visual stimuli used for the gaze following task were taken from Senju and Csibra (2008). Each video started with a woman, sitting behind a table in the middle of the screen, facing down, with two objects on the table positioned to the left and right side of the model. The objects remained on the table in the same location for the entire duration of the trial. The first part of the video presented infants with the model looking down between the two objects for 2 seconds. Then, an icon appeared on the head of the woman accompanied by a voice saying ‘Hallo’ (“Hello”). The ‘Hallo’ was uttered by a female German speaker in IDS or in ADS. The acoustic characteristics of the auditory stimuli were similar to those presented in Senju and Csibra (2008), see Appendix 2.2. During the critical part of the trial, the icon disappeared and the model turned her head and shifted her gaze to one of the two objects on the table (1 second) and continued to fixate this object for the remaining 5 seconds.

Procedure. Each infant was tested in a quiet room. Infants either sat on the parent’s lap or in a car seat about 65 cm away from a monitor. Each infant was presented with six trials on each day of testing. Each trial contained a unique pair of objects. The direction of the model’s gaze was, as in Senju and Csibra (2008), counterbalanced in ABBABA order. For half of the infants, the model fixated the object to her right in the first trial while she fixated the object to the left for the other half of the infants. After each trial, a jumping teletubby appeared to reorient the infants’ attention to the middle of the screen. In contrast to Senju and Csibra (2008), we used a within-subject design, where each infant was presented with the IDS condition on one day and the ADS condition on another. The assignment of infants to the IDS or ADS condition on the different days was counterbalanced.

Data Analysis. Areas of interest were defined based on the size of the objects on the table (340x340 pixel) and icon on the head of the woman (280x278 pixel). An automated eye-tracker, Tobii X120, provided an estimate of where infants were looking with one data point every 8 ms. Data were included and further analyzed only if the eye-tracker recorded reliable data from one of

the eyes (validity less than 2 on the Tobii scale). Gaze data was further aggregated into 40-ms bins¹. Each of these 40ms bins was coded for whether the infant was looking at the model, the target, or the distractor. These bins were aggregated according to our time window of interest. Our critical time window began at 4000ms with the offset of the icon above the model until the end of the trial (10000ms). We excluded trials where infants did not fixate the model for at least 500ms during the speech presentation, to ensure we only included trials where infants were attending to the model and her subsequent eye-movement. This criterion excluded 8% of all trials.

We analyzed the proportion of target looking (duration of looking to target/ looking to target, distractor, and model) in the critical phase of the trial, i.e., after the model had shifted her gaze and fixated one of the objects, using a repeated-measures ANOVA with the within-subject factor condition (IDS, ADS), and the between-subject factor age (6m, 13m) and order of presentation (IDS first, ADS first).

Performance across task. To investigate the influence of maternal input and preferential listening on the gaze following behavior we run a linear mixed effects analyses included condition (IDS, ADS), and age (6m, 13m) as fixed effects, and subject and item as random effects. We included variables capturing the quality of maternal IDS and infants' performance in the preferential listening task. Performance in the preferential listening task was indexed by the difference in infants' listening times to IDS and ADS. Maternal IDS quality was indexed using the difference between maternal mean pitch, pitch range and utterance duration in IDS and ADS. Drop1 analyses examined the key factors that, in interaction with condition (IDS and ADS trials in the gaze following task), explained the variance in target fixations. As above, we only considered factors that influenced fixations in interaction with condition.

¹ Following the conventions we use in our previous work, we code a child as looking at the target in a particular 40ms bin if the child looked at the target for more than 2/3rds of the number of bins included in any 40ms bin. If the child looked less at the target than this threshold the child was coded as not looking at the target or looking wherever the child did look for at least 2/3rds of the bin (e.g. Eiteljörge, Adam, Elsner, Mani, 2019).

Results

Recordings of mothers' speech

We analyzed the acoustic characteristics of mothers' speech in both IDS and ADS, with regard to her maximum pitch, minimum pitch, mean pitch (each in Hz), pitch range (in semitones), and the mean duration of utterances (in seconds per utterance, see Table 2). In all cases and at both age-groups, the prosodic characteristics of maternal IDS differed significantly from the characteristics of maternal ADS (see Table 1). We also compared the difference in these acoustic characteristics in IDS and ADS, e.g., maximum pitch in IDS - maximum pitch in ADS, across the two ages and found no differences in these acoustic characteristics across IDS and ADS between the two ages tested (see Table 2). Thus, mothers did not produce more exaggerated IDS (relative to their ADS) between the two age-groups tested.

Table 2 Acoustic characteristics of mothers' IDS and ADS and difference in prosodic characteristics of maternal IDS and ADS. Mean, maximum, and minimum f_0 are reported in Hz, duration in seconds and f_0 range in semitones per second.

| | Acoustic characteristics - mothers of 6-months-olds | | | | | | Comparing IDS and ADS in mothers of 6-month-olds | |
|------------|---|-----------|----------|-----------|-----------|-----------|--|----------|
| | IDS | | ADS | | IDS - ADS | | Statistic | |
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>t</i> | <i>p</i> |
| max f_0 | 345 | 46 | 306 | 27 | 39 | 58 | 3.79 | <.001 |
| min f_0 | 183 | 46 | 106 | 16 | 77 | 49 | 8.16 | <.001 |
| mean f_0 | 256 | 35 | 203 | 17 | 53 | 38 | 6.96 | <.001 |
| duration | 2.27 | .40 | 2.81 | .47 | -0.53 | 0.51 | -4.45 | <.001 |

| | | | | | | | | |
|-------------|-------|------|-------|------|-------|------|-------|-------|
| f_0 range | 11.62 | 5.14 | 18.77 | 2.71 | -7.15 | 5.96 | -6.40 | <0.01 |
|-------------|-------|------|-------|------|-------|------|-------|-------|

| | Acoustic characteristics - mothers of 13-month-olds | | | | | | Comparing IDS and ADS in mothers of 13-month-olds | | Comparing the relative IDS-ADS difference between mothers of 6- and 13-month-olds | |
|-------------|---|-----------|----------|-----------|---------------|-----------|---|----------|---|----------|
| | IDS | | ADS | | IDS minus ADS | | Statistic | | | |
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> |
| max f_0 | 353 | 45 | 320 | 44 | 33 | 45 | 2.74 | <.001 | 0.41 | 0.68 |
| min f_0 | 176 | 34 | 117 | 22 | 58 | 35 | 7.41 | <.001 | 1.62 | 0.11 |
| mean f_0 | 259 | 30 | 209 | 22 | 49 | 28 | 6.79 | <.001 | 0.34 | 0.73 |
| duration | 1.99 | .32 | 2.61 | .50 | -0.62 | 0.56 | -5.37 | <.001 | 0.61 | 0.55 |
| f_0 range | 12.80 | 3.93 | 17.82 | 3.29 | -5.01 | 4.58 | -5.08 | <.001 | -1.48 | 0.15 |

Preference

A mixed ANOVA with the dependent variable listening time, the within-subjects factor condition (IDS, ADS), and between-subjects factors age (6m, 13m), and order of condition (IDS first, ADS first) found a significant main effect of condition, $F(1, 69) = 4.16, p = .045, \eta_G^2 = .01$, and a significant interaction between condition, age, and order, $F(1, 69) = 5.27, p = .025, \eta_G^2 = .01$, but no other main effects. Pivoting on age, we found that 6-month-olds listened longer to IDS

compared to ADS trials, $F(1,36) = 8.86, p = .005, \eta_G^2 = .01$, with an interaction between condition and order of presentation of condition, $F(1,36) = 17.98, p < .001, \eta_G^2 = .04$. With 13-month-olds, we found no main effect of condition and only an interaction between condition and order of presentation of condition, $F(1,33) = 25.89, p < .05, \eta_G^2 = .16$. Splitting these interactions at both age-groups, we found that 6-month-olds who received the IDS trial first in the experiment showed an IDS preference, $t(18) = 4.48, p < .001, 95\% CI: 2.06, 5.7$, while those who received an ADS trial first did not, $t(18) = -1.06, p = .301, 95\% CI: -2.02, .67$. Thirteen-month-olds showed a preference for ADS when the experiment began with an ADS trial, $t(17) = -3.82, p = .001, 95\% CI: -7.39, -2.13$, and a preference for IDS when the experiment started with an IDS trial $t(16) = 3.39, p = .004, 95\% CI: 1.72, 7.49$. Figure 1 plots the mean listening times to each condition separated by age and condition presented first.

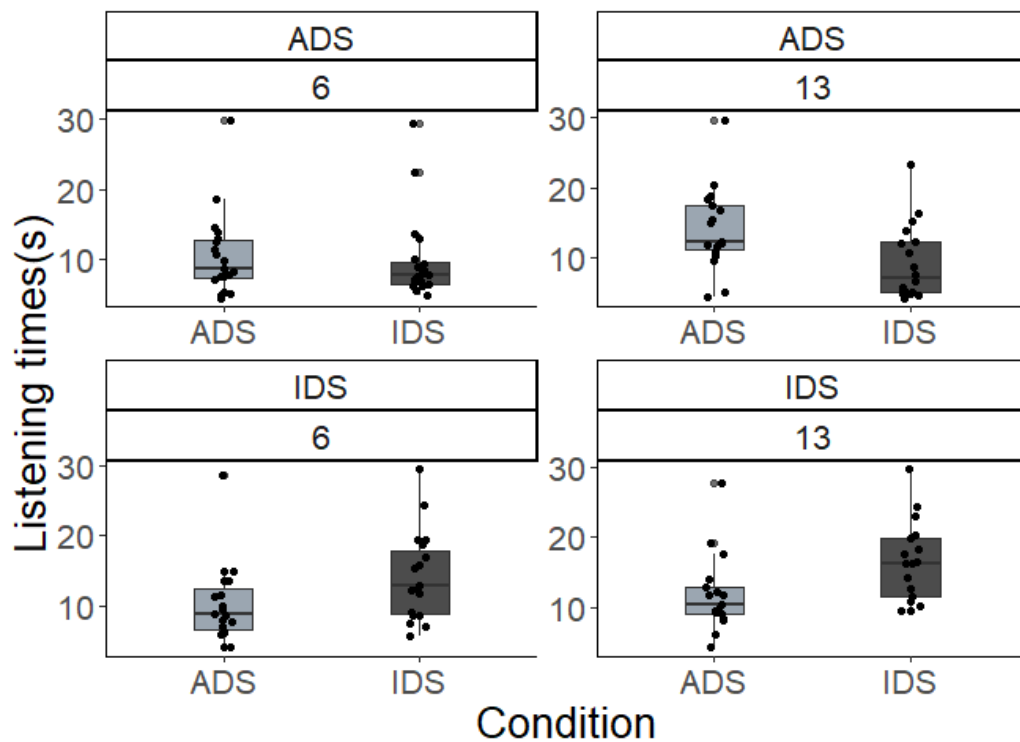


Figure 1. Listening times to IDS and ADS trials (Condition) at 6- and 13-months of age split according to order of trial presentation (ADS first (upper panel) and IDS first (lower panel)).

Correlation across tasks. We ran a linear mixed effects model on the data from the preferential listening task, i.e., listening times, with the interaction between condition (IDS, ADS) and the acoustic characteristics of the difference in maternal IDS and ADS (i.e., , mean pitch difference, difference in pitch range and utterance duration difference) and the difference in target looking across the IDS and ADS condition in the gaze following task (see below for further details on this measure) as fixed effects and random effects on subject, order of presentation and item (where possible given problems of singular fit when controlling for variance at the item level). Drop1 analysis found a significant interaction between condition and mean pitch difference at 6-months of age, $\chi^2(1) = 5.34, p = .021$, and an interaction between condition and duration of utterances at 13-months of age, $\chi^2(1) = 6.99, p = .008$ (see Table 2 for estimates). This pattern was corroborated by correlation tests between listening times in the IDS condition and the mean pitch difference ($r = .48, p = .026$) between maternal IDS and ADS. Note, this pattern changed when considering the total duration of listening time to IDS and ADS trials in the preferential listening task: Drop1 analysis found no significant interaction between condition and mean pitch difference at 6-months of age (although a larger model including maximum and minimum pitch did find an interaction between pitch difference and condition, $\chi^2(1)=6.32, p=.011^2$), but an interaction between condition and duration of utterances at 13-months of age, $\chi^2(1) = 8.29, p = .003$. While the data entered into the model were continuous, for ease of depiction and interpretation, Figure 4 presents the averaged data for the interaction between condition mean pitch (4A) at 6-months of age and the interaction between condition and duration at 13-month of age. The estimates given in Table 3 provide a picture of the variance explained by these factors at each age-group.

Table 3 Model examining listening times in the preferential listening task with interaction between condition (IDS, ADS) and difference in mean pitch, duration, pitch range across maternal IDS and ADS; as well as difference in fixations to the target in IDS and ADS trials in the gaze following task as fixed effects and random intercept on subject split by age (6m,

² We initially ran larger models including maximum, minimum, mean pitch difference, pitch range and utterance duration. We do not report the results of these larger models given potential issues of collinearity across these measures.

13m): lmer(Listening times ~ condition*(Meanpitch + Duration + pitchrange + PTL) +
(1|first)+(1 | subject)

| | 6-month-olds | | | 13-month-olds | | |
|-------------------------------|--------------|----------------|--------------|---------------|----------------|------------------|
| Predictors | Estimates | CI | <i>p</i> | Estimates | CI | <i>p</i> |
| Intercept | 5.00 | -0.44 – 10.43 | 0.070 | 17.14 | 9.49 – 24.79 | <0.001 |
| condition:Meanpitch | 0.05 | 0.01 – 0.10 | 0.014 | -0.01 | -0.11 – 0.10 | 0.878 |
| condition:Duration | -0.38 | -4.44 – 3.68 | 0.855 | -7.41 | -12.43 - -2.39 | 0.004 |
| condition:pitchrange | 0.26 | -0.05 – 0.56 | 0.099 | --0.05 | -0.64 – 0.55 | 0.880 |
| condition:target fixations | -2.27 | -18.22 – 13.68 | 0.781 | 8.81 | -21.56 – 39.19 | 0.570 |

We also calculated the difference between the acoustic characteristics of IDS and ADS of the maternal recordings and the acoustic characteristics of IDS and ADS stimuli presented in the preferential listening task $((IDS_{pref} - ADS_{pref}) - (IDS_{mat} - ADS_{mat}))$. The greater this difference was, the more exaggerated was the difference between IDS and ADS stimuli presented in the preferential listening task relative to the difference between the mothers' own IDS and ADS. A drop1 analysis found that adding this difference improved model fit at 6 month of the preferential listening time data with regards to mean pitch, $\chi^2(1) = 5.77, p = 0.016$. At 13 months we found improved model fit with regards to duration $\chi^2(1) = 5.12, p = .023$. In particular, infants showed a familiarity preference at 6 months, with infants, whose mothers similarly exaggerated IDS relative to ADS as was the case in the stimuli presented in the preferential listening task, showing a greater

IDS preference. In simplified terms, infants preferred listening to IDS that was more similar to their mothers' IDS.

Gaze following

Figure 2 plots the proportion of target looking after the model had initiated her gaze to the target object separated by condition (IDS and ADS). We analyzed the proportion of looking to the target after the model had shifted her gaze and fixated one of the objects presented on the table in front of her. A repeated-measures ANOVA with the within-subject factor condition (IDS, ADS), and the between-subject factor age (6m, 13m) and order of condition (IDS first, ADS first) revealed a significant main effect of condition, $F(1,57) = 7.31, p = 0.009, \eta_G^2 = 0.03$, indicating increased target looking in IDS compared to ADS (see Figure 3, with a split by age-group). No other main effects or interactions were found. These results suggest no differences across age-groups with regard to the influence of IDS on infants' gaze-following behavior.

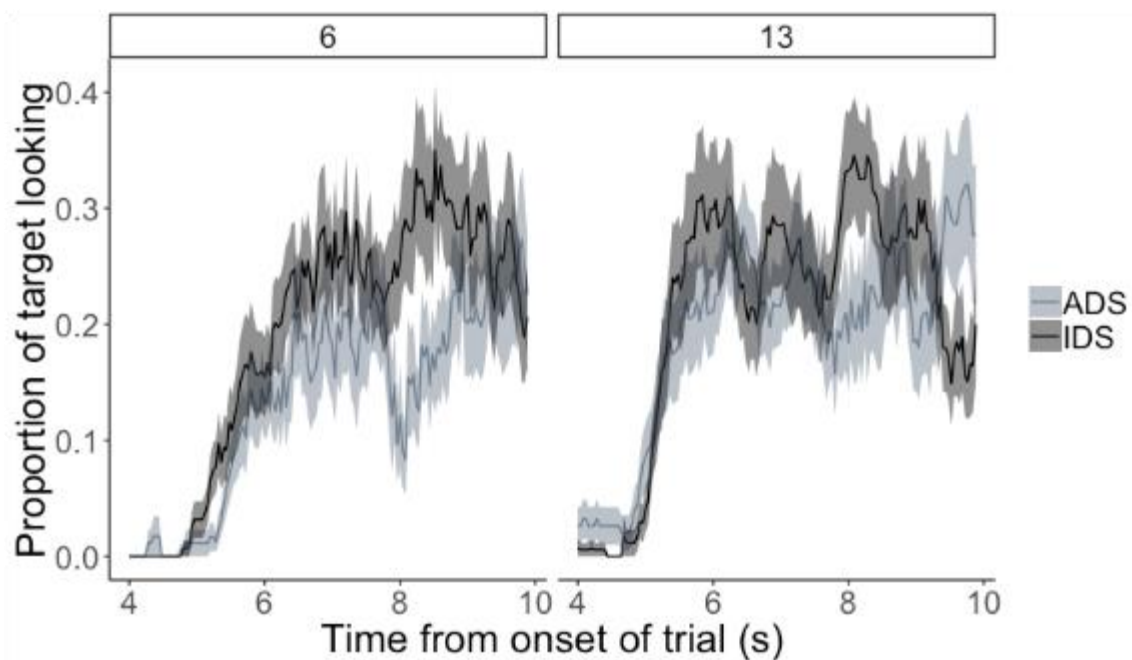


Figure 2. Proportion of target fixations after the model initiated her gaze (at 4s) to the target object in 6-month-olds and 13-month-olds across conditions (IDS and ADS).

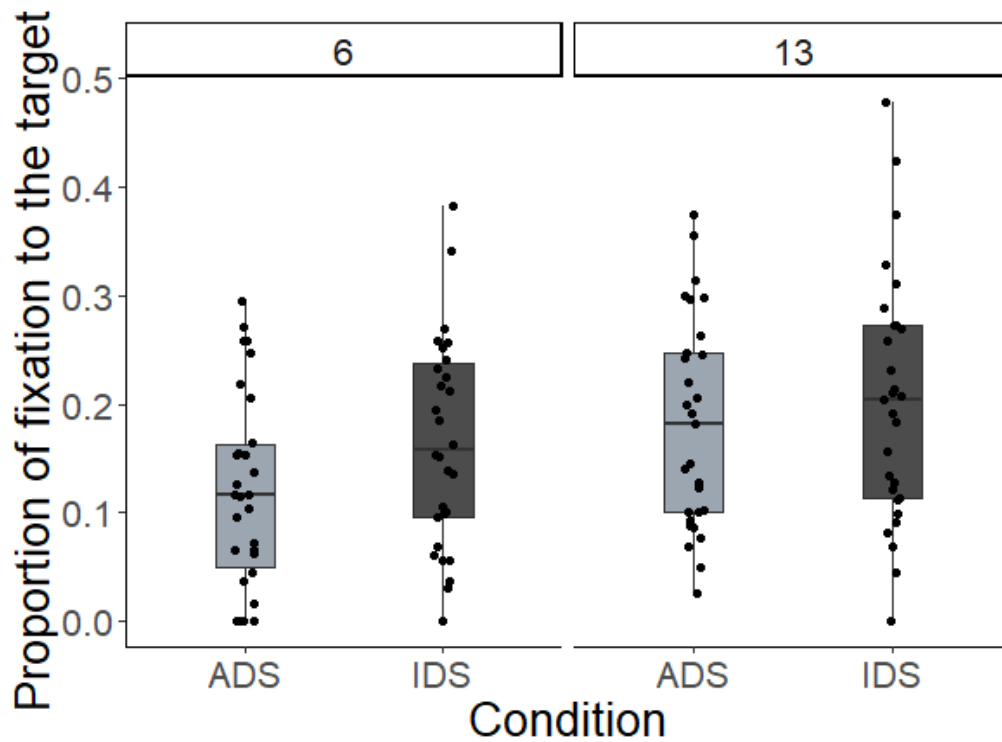


Figure 3. Proportion of target fixations after the model initiated her gaze (at 4s) to the target object in 6-month-olds and 13-month-olds across conditions (IDS and ADS) averaged across the critical window.

Correlation across tasks. We ran a linear mixed effects model on the fixations to the target, with the interaction between condition (IDS, ADS) and the acoustic characteristics of the difference in maternal IDS and ADS (i.e., mean pitch difference, utterance duration and difference in pitch range) and the normalized difference in listening times to IDS and ADS trials in the preferential listening task as fixed effects and random effects on subject and item. There were no significant interactions between condition and any of the above factors in this analysis, $p_s > .112$.

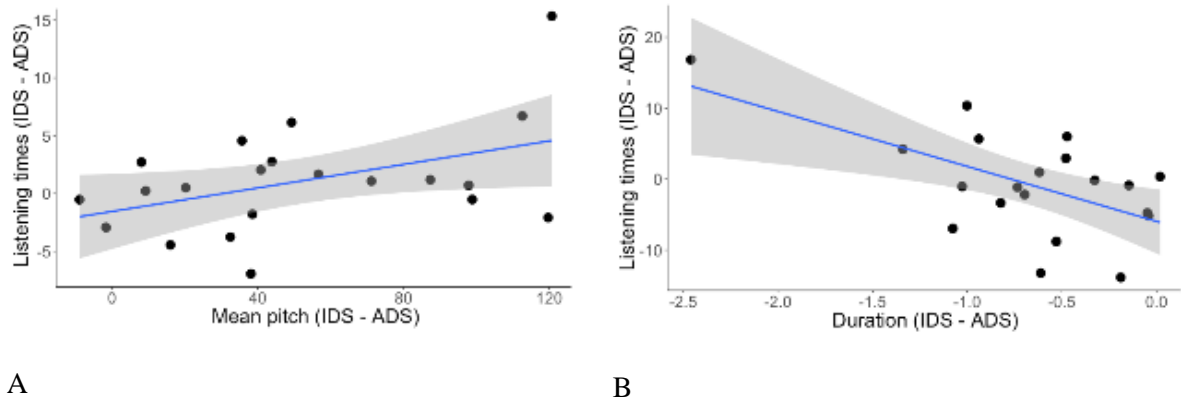


Figure 4. (A) Correlation between data from preferential listening task at 6-months and difference in mean pitch of maternal IDS and ADS (B) Correlation between data from preferential listening task at 13-months and difference in duration of utterances in maternal IDS and ADS. Note that IDS typically has utterances of shorter duration than ADS, hence the negative fit.

Discussion

The aim of the current study was to investigate the interaction between the prosodic characteristics of maternal IDS, infants' preference for IDS over ADS, and the functional relevance of IDS in infant-directed communication, i.e., the extent to which IDS impacts infants' responding. We hypothesized that these three facets of IDS may influence one another, such that variation in one of these factors co-occurs with variation in the other. In particular, we were interested in (1) whether there are prosodic differences in the IDS and ADS produced by German mothers, (2) whether German infants show a preference for IDS compared to ADS, (3) whether German infants follow someone's gaze faster when this person was previously presented speaking IDS compared to ADS, (4) how these characteristics vary across development, and (5) if and how these factors interact with one other. We were interested in the potential roots of infants' preference for IDS and the extent to which this may be related to the quality of maternal IDS, and the extent to which infants' preference for IDS may itself be related to the functional relevance of IDS in communication.

Similar to other studies (e.g., Fernald et al., 1989), there were significant differences in the prosodic characteristics of German mothers' IDS compared to their ADS at both age-groups. Mothers of infants at 6- and 13-months produced utterances with higher mean, maximum, and minimum pitch in IDS compared to ADS. They produced utterances of shorter duration in IDS relative to ADS. While utterances in IDS tend to be longer due to the lengthening of vowels and longer pauses, the total duration of utterances in naturalistic IDS tends to be shorter due to the fewer number of words per utterance (see Soderstrom, 2007, for a review). They also produced utterances with a narrower pitch range in IDS compared to ADS. This may be because mothers produced utterances with higher minimum pitch in IDS, thereby potentially reducing the range possible in this register.

We found no developmental differences in IDS produced by mothers of 6- and 13-month-olds. In other words, 13-month-olds hear IDS which is prosodically similar to that heard by 6-month-olds. Admittedly, these findings are in the context of mothers' retelling a story to their infant, which may differ from the prosodic characteristics of their other interactions with their infant. However, ongoing data from our laboratory examining natural interactions between mothers and infants at the same ages similarly find no differences in the prosodic characteristics of maternal IDS across early infancy (c.f., Schreiner et al., in prep).

A preference for IDS over ADS was only found at 6-months of age. This finding is in line with previous studies reporting an IDS preference between 4- and 8-months (Cooper et al., 1997; Fernald, 1985; Kaplan et al., 1995a, b; Newman & Hussain, 2006; Singh et al., 2002; Thiessen et al., 2005; Werker & McLeod, 1989). At 13-months we did not find a preference for IDS over ADS. This is likely attributable to changes in the speed of infants' language processing from 6- to 13-months of age, with older infants being able to process a broader range of stimuli (c.f., Schreiner & Mani, 2017), including less infant-directed stimuli, leading to 13-month-olds not showing as much as a preference for IDS. Similar findings are reported in Newman and Hussain (2006).

These findings contrast with findings of a decrease in preference between 7- and 10-months, and a later increase in preference after this age (Hayashi et al., 2001). These also contrast

with the findings from the ManyBabies project on IDS preference, with a more robust IDS preference past 12 months of age (ManyBabies Consortium, 2020). This preference was more robust in North American English (NAE) infants relative to infants from non-NAE backgrounds and was tested using NAE IDS, which differs starkly from German IDS. We presented infants with German IDS and ADS, with the possibility that the reduced prosodic differences between German IDS and ADS (Fernald et al., 1989; Schreiner & Mani, 2017) lead to a reduced IDS preference at this later age.

Inline with previous findings (Cooper et al., 1997; Kaplan et al., 1995a; ManyBabies Consortium, 2020; see also Cooper & Aslin, 1990, 1994), we found that the difference in listening times varied with order of presentation of trials. This was particularly so with the 13-month-olds who listened longer to whichever register they were presented with first. The 6-month-olds did not show increased listening times to ADS trials even when the first trial presented was in ADS. With all the constraints associated with interpreting null effects, this suggests conflicting influences from the order of presentation of trials and a potentially “real” IDS preference at 6-months. This is in line with our results, showing an overall IDS preference only at 6-months.

Despite the effect appearing smaller in the older age group relative to the younger age-group, we do not find statistical differences in the performance of the gaze following task across the two age groups. Here, we found a significant effect of condition that did not differ across age groups, with 6- and 13-month-olds fixating the target more when cued by IDS than by ADS. This finding replicates Senju and Csibra’s (2008) findings with 6-month-olds (see also Hernik & Broesch, 2019) and extends this IDS effect to 13-month-olds, speaking overall to the functional relevance of IDS in communication with young infants. IDS appears to attract infants’ attention more and drives them to engage with others in social interactions; IDS is attentionally salient. Since the task performed by infants was non-linguistic – and could not therefore be influenced by hyperarticulation in IDS - the findings highlight the attentional salience of IDS in driving infant responding and learning.

Performance across tasks

Before discussing these results, we highlight two features of this analysis which are key to interpretation. First, we only took the difference in the acoustic characteristics of individual mothers' IDS and ADS into the analyses, allowing us to tap into the infant-directedness of maternal input, and not the prosodic characteristics of the mothers' speech, per se. Second, we controlled for variance in the other tasks when examining performance in any one task, i.e., in analyzing the contribution of variation in maternal input on performance in the preferential listening task, we controlled for performance in the gaze following task, and vice versa.

Overall, we found that the acoustic characteristics of maternal IDS (relative to ADS) influenced performance most. Differences in the mean pitch and duration of utterances in maternal IDS (relative to ADS) influenced performance in the preferential listening task. At 6 months, infants whose mothers produced more exaggerated prosody (with regard mean pitch) in IDS (relative to ADS) showed a greater preference for IDS (relative to ADS) in the preferential listening task. This finding highlights a keen link between the prosodic characteristics of maternal IDS and infants' IDS preference. Examining the difference between maternal IDS and the IDS stimuli presented in the preferential listening task, we found, that infants whose mothers' IDS was similar to the IDS in the preferential listening task showed a more robust IDS preference, suggesting a familiarity preference underlying the IDS effect reported here. This raises questions as to the extent to which a similar familiarity preference underlies the IDS preference in the studies reported to-date, especially against the background of studies suggesting that infants show improved segmentation of words from exaggerated IDS even when the IDS typical to their native language is not similarly exaggerated (Floccia et al., 2016; Schreiner & Mani, 2017). Given the exploratory nature of these analyses, we do not comment on this issue further but highlight that different factors may underlie the reported IDS preference, including but not limited to the quality of the IDS that children are exposed to, and the quality of the stimuli presented to infants in a particular task.

At 13 months, we found that infants whose mothers produced shorter utterances in IDS (relative to ADS) showed a greater preference for IDS (relative to ADS) in the preferential

listening task. Shorter utterances have been highlighted as one of the key features of maternal IDS and appear to interact with infants' IDS preference at 13-months, as was the case with the pitch of maternal IDS at 6-months. We raise, however, the concern that the results with 13-month-olds need to be treated with considerable caution, given the presence of outliers that may have driven this effect and the fact that there was no independent IDS preference found at this age – we only found an interaction between condition and order of presentation at this age.

We found no influence of performance in the gaze following task on performance in the preferential listening task and vice versa. One reason why we found no influence of maternal IDS quality on gaze following performance may be due to the increased social cues provided in this task. The gaze cue, in conjunction with the increased attentional salience of IDS, may leave little room for modulation of infants' performance by maternal IDS quality, due to the relevance of these cues. In other words, the increased social cues in the gaze following task may outweigh individual variation caused by infants' preference for IDS (over ADS) or the acoustic characteristics of maternal IDS (relative to ADS). This explanation highlights the dynamic interaction between multiple cues in infant responding to IDS, with differential weighting of social cues, maternal input and IDS salience in the same infants in different tasks across early language development.

Conclusion

Our findings highlight developmental differences in infants' preference for IDS over ADS, with 6-month-olds showing a preference while 13-month-olds do not. We do not find developmental differences in the functional relevance of IDS across development, nor any in the quality of maternal IDS. Across tasks, the findings highlight the strong link between infants' response to IDS and the quality of maternal IDS, which may have consequences for our understanding of infants' learning from IDS. We note that the direction of this effect remains unclear due to the cross-sectional nature of the study. In other words, the link between maternal IDS and infant responding can either be explained by mothers being conditioned to produce more exaggerated IDS when their infants respond more actively to this register (c.f., Smith & Trainor, 2008), or infants responding more actively to this register because they have more exposure to exaggerated IDS in their daily input. At present, we cannot adjudicate between these two

possibilities. We suggest that the repeated findings of a link between maternal input and infants' response to IDS in the current study highlight the shared underpinnings of maternal IDS and infants' response to IDS. The current study, therefore, speaks to a dynamic interaction between infants' preference for different kind of social signals and the social cues available to them.

Infants' word segmentation

Abstract

Caregivers typically use an exaggerated speech register known as infant-directed speech (IDS) in communication with infants. Infants are better segmenting words from IDS compared to ADS. We wanted to examine interactions between maternal IDS quality, infants' preference for IDS over ADS, and infant's ability to segment word from a fluent speech stream at 6- and 13-months. Due to the small amount of data consisting of all three tasks, we focused on analyzing the EEG data. We measured frontal brain activity in two time windows, namely between 350-500ms and 500-700ms after stimulus onset at 6- and 13-months of age. We found no evidence for infants' ability to segment words from fluent speech. Interestingly, we found evidence for differences in infants' brain activity across registers, regardless of condition. In particular, we found at 6 month an increased frontal positivity in IDS blocks relative to 13-month-old children. Furthermore, the 13-month-old response in ADS blocks was more localized to the left hemisphere, with 13-month-olds showing a more negative response to IDS relative to ADS altogether. These results speak for a sensitivity to IDS both early and later in development.

Introduction

In learning a language, infants face the difficult task of extracting individual words from a fluent speech stream. Nevertheless infants as early as 4.5 - 6 months of age infants are able to detect highly frequent words, like their own name or the parents' name, in continuous speech (Bortfeld et al., 2005; Mandel et al., 1995). One of the first studies giving evidence for word segmentation in natural speech with infrequent words showed that 7.5-month-old infants listen longer to passages containing words they have previously been familiarized with before in isolation (Jusczyk & Aslin, 1995). This pattern was also observed if infants were first presented with sentences containing the words to be segmented and then tested on isolated tokens of the previously familiarized words (relative to unfamiliar words). This study provided early evidence of infants being able to detect words and word boundaries in fluent speech (Jusczyk & Aslin, 1995). However, this ability seems to depend on the environmental settings, e.g. the gender of the speaker

(Houston & Jusczyk, 2000; but see van Heugten & Johnson, 2012), the language that the child is exposed to (Schreiner & Mani, 2017) or the emotional cues provided by the speaker (Sing, Morgan, & White, 2004).

Another factor that seems to have an impact on infants' segmentation skills is the kind of speech style used in training and test. Infant directed speech (hereafter IDS) - the speech register used to communicate with infants- has been shown to facilitate language acquisition (Ma et al., 2011; Thiessen et al., 2005). IDS is characterized by higher pitch, longer pauses between utterances and more repetition (Cristia, 2013). IDS facilitates language acquisition on different levels: syntactic parsing, phoneme perception and boundary detection (Cristia, 2013; Soderstrom, 2007). Furthermore, infants' detection of syntactic boundaries has also been shown to be facilitated by IDS (Kemler Nelson et al., 1989), IDS is linked with an expanded vowel space (Kuhl et al., 1997), which is further associated with improved phoneme discrimination (Werker, et al., 2007), and IDS has been shown to facilitate segmentation abilities in artificial language (Thiessen et al., 2005) and natural language (Schreiner & Mani, 2017). Furthermore IDS has been shown to facilitate word learning in infants (Ma et al., 2011).

Regarding the ability to segment words, Thiessen, Hill and Saffran (2005) showed that American English 6.5 to 7.5-month-old infants showed longer listening times to familiarized words in IDS (relative to part-words spanning word boundaries). Importantly, a similar difference between words and part-words was not found when infants were familiarized in the speech register adults use in talking to one another, namely, adult-directed speech (henceforth, ADS). Similarly, Zangl and Mills (2007) demonstrated in a EEG-study a boost in neural activity to words familiarized in IDS compared to ADS at 6 months. 13-month-olds, meanwhile showed a boost in both familiarized and unfamiliarized words in IDS only. These findings underscore, in general, the sensitivity of the infants' brain to IDS relative to ADS and, in particular, infants' improved segmentation of words from IDS relative to ADS.

However, studies investigating other languages document difficulties replicating segmentation in that early age using behavioral measures (Schreiner et al., 2016; Nazzi et al., 2014;

Floccia et al., 2016). Recently, two studies investigating German speaking infants at 9 months of age (Schreiner & Mani, 2017) and British English speaking infants at 10, 5 month of age (Floccia et al., 2016) found evidence for infants' segmentation of words from fluent speech only when the stimuli were presented in exaggerated IDS. Furthermore, German 7.5-month-olds did not show evidence of segmenting words from fluent speech even when presented with exaggerated IDS (Schreiner & Mani, 2017).

While there is considerable evidence for the use of IDS in speech to infants across a range of cultures (Fernald et al., 1989; Nwokah, 1987; Shatz & Gelman, 1973; Weppelman et al., 2003), studies also report differences in the characteristics and exaggeratedness of IDS across caregivers speaking different languages (Fernald et al., 1989). German speaking parents, for instance, produce less exaggerated IDS in comparison to American English speaking parents. Additionally, parents might vary in the exaggeratedness of their IDS on an individual basis, with differences in maternal IDS predicting infants' preference for listening to IDS relative to ADS (Outters et al., 2020). Thus, Outters et al. (2020) report that infants of mothers who used less exaggerated IDS showed reduced attention to IDS (relative to ADS) relative to infants whose mothers used more exaggerated IDS at 6-months of age. Extending this finding further, the current study examines the extent to which the infant-directedness of maternal input directed to infants, predicts their ability to segment words from fluent speech at 6- and 13-months of age.

Indeed, infants' word learning and, relatedly, their ability to segment words from fluent speech may themselves be influenced by differences in infants' attention to speech in the infant-directed register. While, infants generally show a preference for IDS relative to ADS at an early age (Cooper & Aslin, 1990; Fernald, 1985; Fernald & Kuhl, 1987; Glenn & Cunningham, 1983; Kaplan et al., 1995, ManyBabies Consortium, 2020; Owren, & Cooper, 1995; Pegg et al., 1992; Schachner & Hannon, 2011; Singh et al., 2002; Werker & McLeod, 1989), the findings are more mixed towards the end of the second year of life (e.g., Glenn & Cunningham, 1983; Hayashi et al., 2001; Newman & Hussain, 2006; Segal & Newman, 2015). If infants attend more to IDS relative to ADS, then this may also impact their segmentation of words from IDS relative to ADS, such that

infants who show less of a preference for IDS relative to ADS may also show reduced segmentation of words from IDS. Against this background, the current study examined the shared underpinnings of infants' preference for IDS and their segmentation of words from IDS to shed further light on the question of how individual differences in infants' response to IDS may impact their learning from IDS.

Given difficulties tapping into German infants' segmentation of words from fluent speech using behavioural tasks at a young age (Schreiner & Mani, 2017), the current study examined a potentially more sensitive index of infants' segmentation of words from fluent speech, namely, their event-related potential response to words previously familiarized in IDS relative to ADS. We also presented infants with a preferential listening task and recorded their mothers producing IDS and ADS, in order to examine the shared underpinnings of infants' preference for IDS, their segmentation of words from IDS and the quality of maternal IDS (see Outters et al., 2020 for further details of preferential listening task and maternal recordings). To examine differences in the interplay between these three tasks across development, we tested all three tasks at two different age groups, 6- and 13-months respectively. We chose these two age groups given previous research documenting differences in infants' sensitivity to IDS and ADS at the two ages (Männel & Friederici, 2013; Outters et al., 2020; Zangle & Mills, 2007), and research suggesting that the ability to segment speech develops gradually between 6- and 12-months of age (Jusczyk, 1997; Saffran et al., 2006).

EEG components

Studies using EEG measures highlight differences in the time-course and topography of segmentation effects across the ages tested, which vary based on the kind of stimuli children are familiarized with (words or sentences). In what follows, we detail the results of EEG studies investigating segmentation abilities in Dutch speaking, French speaking and German speaking infants.

Kooijman, Hagoort and Cutler (2005) examined infants' recognition of previously familiarized words in continuous speech. Thus infants were familiarized with isolated tokens of

words and then tested on their segmentation of these words in continuous speech at test. 10-month-olds' ability to recognize the words in fluent speech was indexed by a more negative response over the left hemisphere to familiarized words, compared to unfamiliarized words between 350 and 500ms. Results with younger infants at 7-months of age are mixed, with a significant effect of familiarity in a posthoc analyses with a subset of electrodes (Kooijman et al., 2013) between 350-450 ms and between 430-530 ms. Furthermore, there were individual differences across infants, with a subset of children showing an overall positivity and a smaller number of children showing a negativity similar to older children. This group of children displaying a negativity over the right frontal and left posterior electrodes showed better language skills at the age of three. Furthermore, when exposed to a more difficult task, to recognize words (presented in isolation) which children had been familiarized with in fluent speech earlier, only 10-month-olds who possessed larger vocabularies later on (at 12-months) show a segmentation response (Junge et al., 2012) indicated by a more negative ERP component between 200-650 ms for familiarized versus unfamiliarized words.

French speaking 12-month-olds show recognition of words (familiarized in isolation) when the words are presented in continuous speech at test (Goyet et al., 2010), indexed by a negative ERP component between 350 - 500 ms indicating segmentation of familiarized words in comparison to unfamiliarized words. Investigating German speaking infants' segmentation abilities, Männel and Friederici (2013) presented 6-, 9- and 12-month-olds with the more difficult version of the task, where infants were familiarized with words embedded in sentences and tested on their recognition of the words in isolation at test. The target word was either accentuated or spoken normally within the sentences presented at familiarization. Infants at 6 months showed segmentation only for accentuated familiarized words, indicated by a broadly distributed positive ERP response, starting around 500 ms. At 9 months, infants showed segmentation independent of accentuation, indexed by a negative ERP component starting at 400 ms. Additionally a late negative ERP component was found only for words that had been accentuated at familiarization. A negative ERP component starting at 350ms indexed word segmentation independent of accentuation at 12 month.

Regarding the current study, we expect to find evidence of infants segmenting words from fluent IDS at 13-months. Based on previous research, we expect this effect to be indexed by a negativity in the left frontal electrodes (F7, F3, FC1, FC5, C3) from 350ms (Männel & Friederici, 2013). We do not expect to find evidence for segmentation of words from ADS at this age. At the younger ages, given the results reported above, with German 6-month-olds showing segmentation of accentuated familiarized words alone and results with 7-month-olds being mixed, we do not expect to find strong evidence for segmentation. If anything, we expect to find a broadly distributed positivity to words familiarized in IDS starting from around 400ms, with this effect indexing immature segmentation in earlier research (Kooijmann et al., 2013; Männel & Friederici, 2013). In order to be able to compare segmentation abilities across age-groups, we will examine the same frontocentral electrodes which have previously indexed early segmentation (F7, F3, FC1, FC5, C3, F8, F4, FC2, FC6, and C4) and time windows across age-groups (350-500ms; 500-700ms), while splitting the data across age-groups for further exploratory analyses.

Method

Participants

Eighty-one monolingual German infants were recruited for this study. Forty-two infants at the age of 6-months (5m16d – 7m10d, 19 girls) and 39 infants at the age of 13-months (12m15d – 14m12d, 19 girls) took part in the study. Due to infants' unwillingness to participate in the EEG study, we only obtained 33 EEG-Datasets for 6-month-olds and 25 EEG Datasets for 13-month-olds. Data from additional infant was lost during preprocessing due to their not providing a sufficient number of artefact-free trials. Thus a subset of 28 participants at 6-months and 19 participants at 13-months were entered into further analyses. See table 4 for an overview of number of participants in each part of the study.

Table 4 Number of infants who provided data across tasks and age groups

| Data obtained for each task/ combination of tasks | 6-month-olds (number of infants) | 13-month-olds |
|--|-------------------------------------|---------------|
| EEG recording | 28 | 19 |

| | | |
|---|----|----|
| Maternal recording | 27 | 27 |
| Preferential listening task | 38 | 36 |
| EEG recording and Maternal recording | 23 | 14 |
| EEG recording and preferential listening task | 26 | 19 |

General Procedure

This study was part of a bigger dataset including data from infants' participation in a number of other tasks. Infants were, therefore, also presented with a preferential listening task, where we examined their preference for listening to IDS as opposed to ADS using the central fixation procedure, a gaze following task, where we examined how quickly and reliably they followed the gaze of a model when previously exposed to IDS and ADS. In addition, we also obtained recordings of maternal IDS and ADS. The tasks were split across two days. The preferential listening task was done on the first day. Maternal ADS recordings were done on the first day, and maternal IDS recordings were recorded the second day. The EEG experiment took part on the second day. Ethics approval was obtained from the Ethics committee of the University's Institute for Psychology.

Material and Design

The stimuli and further details of the design of the preferential listening task and the maternal recordings can be found in Outters et al., 2020. Here, we focus on the stimuli and design of the EEG study.

Stimuli. Critical stimuli consisted of 40 monosyllabic German words that did not belong to the productive or receptive vocabulary of children under 36-months-of-age (Grimm and Doil, 2006). Words were recorded both in isolation and embedded in sentences. We created eight grammatically and syntactically correct German sentences for each word. Word position within the sentences was varied so that the critical word appeared twice in the first, second, third, and final

positions of the sentence. All 320 sentences contained between four and eight words. The full list of words can be found in Appendix 3. The words and sentences were recorded by a native German female speaker for the entire experiment. Every sentence was recorded once in ADS and once in IDS. For ADS recordings, the speaker was instructed to produce the sentences as she would when speaking to another adult. For IDS sentences, the speaker was instructed to speak as she would when speaking to an infant. Words in isolation were produced clearly and precisely with a neutral intonation.

Procedure. Each infant was tested in a quiet room. The infant sat on the lap of the mother about 60 cm away from a laptop monitor on which a silent child-friendly movie was played. One experimenter stayed in the room with the infant and played silently with the infant during the entire time of the experiment. The auditory stimuli were presented via a loudspeaker placed above the laptop in front of the infant.

The 40 critical words were randomly paired-up with one another to produce 20 word pairs. For each word pair, one word served as the familiarized word and was presented in the eight sentences during the familiarization phase and twice in isolation during the test phase. The other word in the word pair, the control word, was presented twice in isolation in the test phase. For each word pair, the assignment of words as familiarized and control words was counterbalanced across participants. In addition, presentation of each word pair in ADS and IDS was counterbalanced across participants.

The study consisted of 40 blocks. Each block consists of a familiarization phase and a test phase. The familiarization phase included 8 sentences with the critical familiarized words embedded into the sentences. The test phase presented infants with four words, two familiarized and two unfamiliarized control words. Each infant heard 10 IDS and 10 ADS blocks, where the familiarization stimuli were either in IDS or ADS across blocks. The 20 blocks were repeated once for a total of 40 blocks per experiment. The whole experiment lasted about 24 minutes.

Preprocessing and Data analysis. Data were recorded using a Biosemi Active Two Amplifier System. Brain caps according to the 10-20 international convention recorded 32Ag/Ag-

CL electrodes with a sampling rate of 2048 Hz. During postprocessing, the sampling rate was reduced to 250 Hz. Impedances were kept under 30 k Ω . We re-referenced electrodes offline to the average of the right and left mastoid and filtered offline using 0.1 Hz high-pass and 30 Hz low-pass filter. The data was baseline corrected to pre-stimulus activity from 200 ms before stimulus onset to stimulus onset. Automatic artefact rejection was conducted for each trial excluding trials where the voltage threshold exceeded 200 μ V for left and right frontal electrodes (F7, F3, FC1, FC5, C3, F8, F4, FC2, FC6, and C4) from 200 ms before the stimulus onset to 800 ms after the stimulus onset. Data was only included if the infant contributed at least 10 trials per condition. This resulted in 32 datasets for 6-month-old infants and 22 datasets for 13-month-old infants. In addition, based on visual inspection of the data, we excluded an additional 7 ERP sets (three 13-month-old and four 6-month-olds) because the data from a channel to be included in the analyses was unusable. ERPs were time-locked to the onset of the words in the test phase and averaged across familiarized and control words separately for the IDS and ADS condition. Based on the literature reviewed above, we focused our analysis on two time windows, between 350-500 ms and 500-700 ms after word onset. We aggregated our electrodes according to right (F7, F3, FC1, FC3, C3) and left (F8, F4, FC2, FC4, C4) hemisphere electrodes. The data for each time window was analyzed using linear mixed effects models with the interaction between condition (familiar, control), register (IDS, ADS), hemisphere (left, right) and age (6-months, 13-months) as fixed effects and random effects of subject, age and register with the slope of interaction between condition and register across subject and age and the slope of condition across register. Follow-up models then investigated significant interactions between the fixed effects. Fixed effects were sum-coded to allow for interpretation of the interactions.

Performance across tasks. Unfortunately, there were too few datapoints at each age to examine performance across tasks (between 13 to 22 participants provided data for pairs of tasks at the different ages). We therefore did not compute correlations across tasks.

Results

300-500 ms

There were no significant effects of any of the fixed effects or interactions between the fixed effects in this time window. The results of the model are presented in Appendix 4.

500-700 ms

Models including the interaction between condition, register, hemisphere and age revealed significant interactions between age and register and between age, register and hemisphere. No significant interactions with condition were revealed, suggesting no privileged response to familiarized words (relative to control words) at any age-group in any register. In other words, we found no evidence for infants' segmentation of words from fluent speech. The results of the main model and follow-up models are presented in Appendix 5. Follow-up models examined the interaction between age and register and found a significant difference across age-groups in infants' response to IDS and an interaction between age and hemisphere in infants' response to ADS. Thus, 6-month-olds showed more of a frontal positivity in the IDS condition relative to 13-month-olds (see Figure 7). In the ADS condition, 13-month-olds showed more of a negativity across the left hemisphere (see Figure 6). Furthermore, summary estimates suggest a significant effect of register at 13-months of age, with 13-month-olds showing a more negative response in the IDS condition relative to the ADS condition (see Figure 5). None of these effects interacted with condition, highlighting only infants' response to both control and familiar words overall.

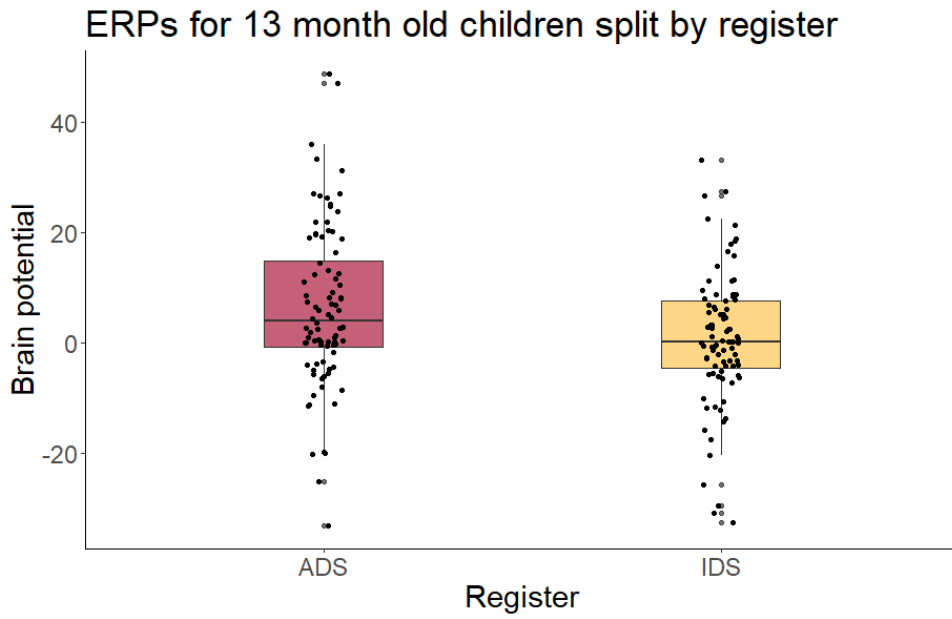


Figure 5: Average prain potential at 500-700ms split by register at 13 month.

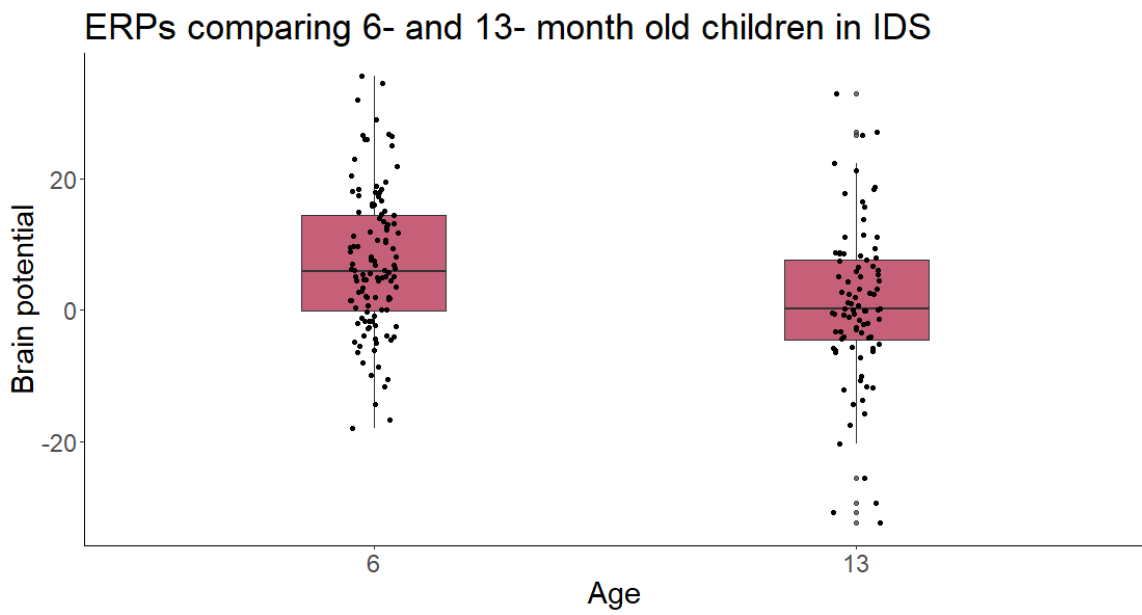


Figure 6: Average prain potential at 500-700ms split by age for IDS

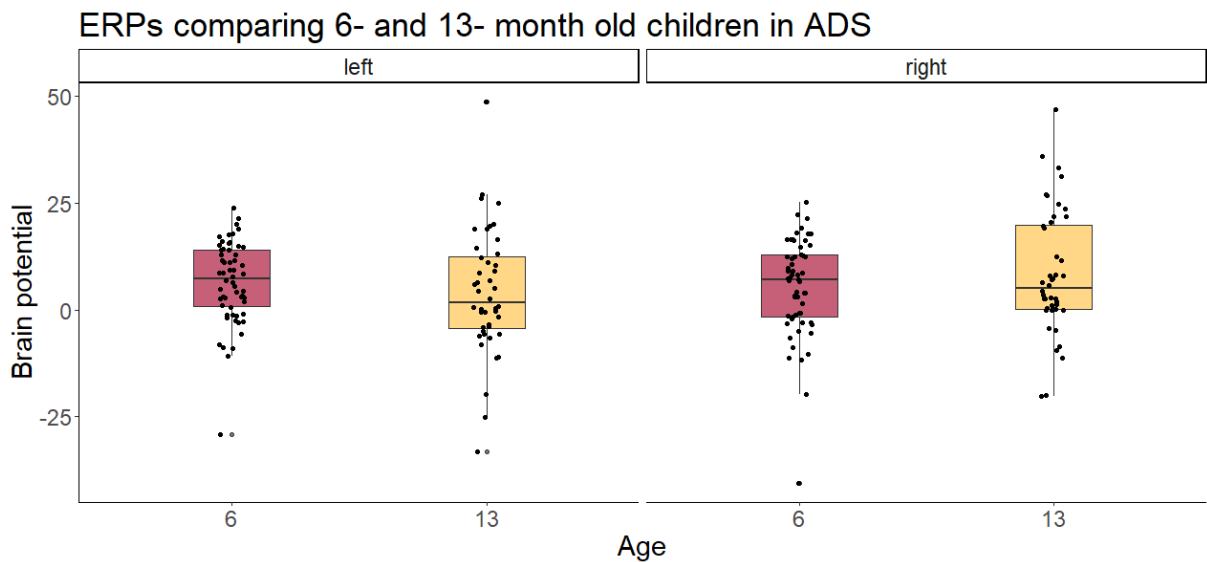


Figure 7: Average brain potential at 500-700ms split by hemisphere and age for ADS.

Discussion

The aim of the current study was to examine German learning infants' segmentation of words from fluent speech at the ages of 6- and 13-months. We were also interested in potential interactions between the quality of maternal speech and infants' segmentation abilities as well between infants' preference for IDS and their segmentation from IDS and ADS. However, given the limited data we could obtain from infants across the tasks presented, we chose to not examine the above interactions further and focus more on infants' segmentation. Indeed, the findings in the EEG study further support this decision since we found no evidence of infants' being able to segment words from fluent speech, regardless of the register of speech (IDS or ADS) or infants' age at time of testing.

We measured frontal brain activity between 350-500 ms and 500-700 ms after stimulus onset at 6- and 13-months of age. Across the two time windows and the age-groups tested, we found no evidence for infants' segmentation of words from fluent speech, inasmuch as we found no evidence for differences in brain activity at test to words infants had been previously familiarized with and novel control words.

How can we explain this complete failure to find evidence for infants' segmentation of words from fluent speech? First, we note that, to a certain extent these results overlap with previous

results examining German infants' segmentation of words from fluent speech. Thus, earlier work finds that 7.5-month-old and 9-month-old German infants show no evidence of segmentation of words from fluent IDS, and that 9-month-olds only show such evidence when IDS is exaggerated (Schreiner & Mani, 2017; see similar evidence from British-speaking infants at 10-months of age, Floccia et al., 2016). Furthermore, this is also in keeping with our findings that it is only at 16-months of age that we find evidence of German speaking infants segmenting words from IDS and ADS (Mani & Pätzold, 2012).

We also note that we presented infants with the more difficult version of the task, where infants were familiarized with words embedded in sentences and tested on the previously familiarized words in isolation. Previous studies examining this more difficult version of the task also find no evidence for segmentation at 10-months (only in infants with larger vocabularies at 12-months, and only across a posthoc analyses with a subset of electrodes, Junge et al., 2012). Similarly, French infants demonstrate segmentation of words from fluent speech when familiarized with the words in isolation and tested on their response to the words in fluent speech (Goyet et al., 2010; but see Nazzi et al., 2014). Indeed, the only ERP study reporting evidence of segmentation in this version of the task also found evidence for segmentation only when the words were accentuated early in development, and regardless of the accentuation of words later in development (Männel & Friderici, 2016). This finding is also in keeping with previous work reporting segmentation at this young age only in exaggerated IDS. Taken together, the current study along with previous findings of a similar nature highlight the fragile beginnings of early infant segmentation of words from fluent speech, with infants requiring multiple cues to segment words from fluent speech.

We further note that our analysis of the data as a whole may also not be tapping into the greater variability that has been hinted at in infants' segmentation of words from fluent speech. Thus studies suggest not only that infants' segmentation skills early in life correlate with later language output but also that the polarity of the infant segmentation response varies, particularly at younger age-groups (Junge et al., 2012; Junge & Cutler, 2014; Kidd et al., 2018; Kooijman et al.,

2013). Given that studies also showed evidence for the influence of language output on the segmentation abilities at 10 months (Junge et al., 2012; Kidd et al., 2018), it might be possible that our group data does not adequately capture potential individual differences in early segmentation abilities.

Interestingly, we do find evidence for differences in infants' brain activity across registers, regardless of condition. In particular, we found an increased frontal positivity in 6-month-olds in IDS blocks relative to 13-month-olds. Furthermore, the 13-month-old response in ADS blocks was more localized to the left hemisphere, with 13-month-olds showing a more negative response to IDS relative to ADS altogether (across both hemispheres). While none of these effects interacted with condition, they are highly reminiscent of previous work examining infants' response to IDS and ADS at 6- and 13-months of age (Zangl & Mills, 2007). In that study, 13-month-olds also showed a more negative response (between 600 to 800 ms) to IDS relative to ADS across both hemispheres regardless of whether the words were familiar or unfamiliar to them. The authors interpret this finding, as we do, with regards to the attention-grabbing spotlight function of IDS, leading to increased neural activity to IDS relative to ADS. The authors discuss the underpinnings of this response relative to increased attentional resources or an arousal or affective response to IDS which increases infants' attention. Importantly, in both our study and Zangl & Mills (2007), familiarity with the stimuli presented did not influence this attention-grabbing response, potentially highlighting the overpowering effect of IDS on brain activity. While this may provide an additional reason why we find no evidence for segmentation in the current study, we note that these findings contrast with reports of segmentation from exaggerated IDS. Furthermore, the increased positivity to IDS in 6-month-olds relative to 13-month-olds is also reminiscent of previous work highlighting differences in the polarity of infants' response to IDS; with a positivity being related to less mature language skills later in life (Kooijman et al., 2013). We note that there was no difference across ages with regards to infants' brain activity in ADS blocks.

Overall, we suggest that the current task underscores previous findings of the development of infants' response to IDS, with a privileged sensitivity to IDS both early and later in

development. IDS grabs infants' attention so that infants show increased neural activity in blocks where they had previously been exposed to IDS relative to blocks where they had been previously exposed to ADS. This increase neural activity is likely to have positive effects on functioning, although we were unable to find such evidence with regards to word segmentation in the current task. Sadly, due to data loss, we were also unable to relate this to the features of infants' environment with regards to quality of maternal IDS or to infants' preference for IDS as indexed by their preference for listening to IDS relative to ADS. Given the attention-grabbing function of IDS attested to in the current study and earlier work, we suggest that this is a critical avenue for future research.

The influence of children's successful word recognition on their internal arousal and body posture

Abstract

Active learning approaches highlight the role of children in their own language learning. Against this background we designed a study examine the extent to which children's internal state of arousal in a word learning task predicts their word learning success. Additionally, we also examined the extent to which word learning success itself may predict the positive emotions experienced by children following successful learning. We presented children (N=50) with a cross-situational word learning task and measured their pupil dilation during and following the completion of the task. Learning success was measured by their eye movements during a word recognition test phase. The changes in body posture provide insights into the emotional state of the children. Results showed successful recognition of both the trained word-object associations that children were introduced to in the cross-situational word learning paradigm as well as already known word-object associations. Further we found positive evidence of an association between increase in pupillary dilation of the novel word recognition task and word recognition success in this task. We discuss these results by taking different accounts like attentional processes, effort or interest into account. Regarding changes in body posture we couldn't find evidence that word learning success shows an influence of the child's emotion. Over all our results speak for the important role of children's motivation and engagement in performance in a word learning task.

Introduction

Language is acquired in social settings, and exposure to linguistic input from caregivers and peers in social interactions is critical to language learning success. Sociocultural and pedagogical approaches to language learning highlight the importance of appropriate language input and rich social interactions with caregivers on language learning, where, children, learn from models who are already in possession of the information they are after (Vygotsky, 1978; Csibra & Gergely, 2009, 2011) While there is undoubtedly an influence of input quality and quantity on language learning (see Mani & Ackermann, 2018, for a review), children also actively shape their learning experience by selectively sampling from the information presented to them. Curiosity-driven or active learning approaches (Kidd & Hayden, 2015; Oudeyer & Smith, 2016, Piaget, 1952) highlight the role of children who, like little scientists, choose when and what to learn from whom (Mani & Ackermann, 2018 for a review).

In keeping with such active learning approaches, we examine, first, how the internal state of the child, with regard to their (emotional) state of arousal during and after a word learning task, predicts their learning of novel word-object associations. Second, we examine whether successful learning predicts the subsequent internal state of the child with regard to positive emotions that may be triggered by their success at the task. In what follows, we briefly review the literature on curiosity-driven learning and intrinsic motivation with regards to word learning, followed by a review of the literature on how intrinsic motivation and emotional state are assessed via changes in children's pupil dilation and body posture.

Curiosity driven learning

Curiosity may be defined as a state of interest or arousal triggered by input that is potentially novel, deviant from one's expectations or, in some cases, just beyond one's repertoire of knowledge (Berlyne, 1954, Kidd & Hayden, 2015; Oudeyer et al., 2016). It is often intertwined with discussions of intrinsic motivation (Berlyne, 1960; Oudeyer et al., 2016; Ryan & Deci, 2000), which itself may be operationalized as a drive to participate in an activity or acquire knowledge due to the pleasure that is derived from doing so, rather than due to any external motivation triggered

by extrinsic (e.g., material) rewards. Historically, theorists initially tied curiosity to the novelty of a stimulus, with Berlyne (1954) suggesting that curiosity was increased when people experience an intermediate state of familiarity with a newly presented stimulus. Thus, if the stimulus is too far removed from what one knows, there is reduced conflict between the presented information and one's representation of the world, leading to reduced curiosity. The same is the case if the newly presented information overlaps considerably with concepts already familiar to e.g., the child. Echoes of this suggestion remain in so-called "knowledge gap" theories, which suggest that curiosity is triggered by a need to reduce uncertainty in one's knowledge about a particular topic (Kinney & Kagan, 1976; Lowenstein, 1994).

Empirical support for such suggestions is provided by studies reporting an inverted-U-shape in participants' curiosity with regard to their knowledge about a certain topic. For instance, in one study, participants reported being maximally curious about answers to questions they knew something about, while being less curious about questions they knew nothing or too much about (Kang et al., 2009). Participants' curiosity was predicted by the pupillary response prior to the answer being revealed, with more dilated pupils in participants who reported being more curious about the answer to a question (Kang et al., 2009). Increased curiosity also predicts retrieval of information, with participants better recalling the answer to a question when they were more curious about (Berlyne, 1954; Gruber et al., 2014; see Fastrich et al., 2018 for results with a large scale database and review). Indeed, here, even stimuli that are 'accidentally' presented in a high curiosity phase were retrieved with greater success relative to stimuli presented in a low curiosity phase (Gruber et al., 2014). Similar findings are reported in studies examining neurocomputational models of learning, with maximal learning reported when the model is allowed to select the stimuli in a curious manner, i.e., by maximizing the novelty of the stimuli to be presented with regard to the model's prior experience (Twomey & Westermann, 2018).

Curiosity-driven learning in early child development is not well researched, but the evidence there is suggests that learning is similarly motivated by children's interest in particular stimuli in their environment and that children actively choose whom to learn from (Bergus et al.,

2014; Begus & Southgate, 2012; Birch et al., 2010; Koenig & Harris, 2005; Ma et al., 2011) , as well as when and what to learn (Ackermann et al., 2020a; Gerken et al., 2011; Lucca & Wilbourn, 2019; Partridge et al., 2015; see Mani & Ackermann, 2018 for a review). Thus, from as early as 16-months of age, infants use pointing gestures to elicit information and retain this information better when it is provided in response to their elicitation (Begus et al., 2014; Begus & Southgate, 2012). In keeping with the knowledge gap hypotheses outlined above, infants are also selective with regards to the complexity of the stimuli they attend to, focusing more on stimuli that are neither too complex or too simple (Kidd et al., 2012). With regards to word learning in particular, children retain novel word-object associations better when they were interested in the category to which the objects belonged, such that children who showed greater interest in animals were better able to retain a novel animal name relative to a novel vehicle name (Ackermann et al., 2020a).

The studies reviewed above have examined the influence of children's and adults' interest in particular stimuli on subsequent learning and highlight the important role of specific interest and curiosity on learning success. However, children's internal state of arousal during a task or motivation to complete a task, as well as the positive feelings associated with a task may similarly modulate learning success. Evidence here is provided in the literature on older children's learning, typically, in classroom situations. This work suggests that children's motivation to learn is a key factor in learning success, with engagement in a task (Salmela-Aro et al., 2016; Skinner et al., 2008) and the positive emotions that are triggered through engagement in a task predicting subsequent learning outcomes (e.g., Pekrun, 2006; Alexander et al., 2014).

Furthermore, in keeping with suggestions that learning may be rewarding, in and of itself, and associated with positive emotions, neuroimaging research has recently examined the link between the reward circuitry and learning circuits in the brain. Based on the increasing number of studies suggesting the involvement of the ventral striatum in reward processing (Hare, et al., 2000; McClure et al., 2004), recent findings that the level of activity in the striatum is linked to participants' success at word learning (Ripolles et al., 2014) and their interest in learning (Kang et al., 2009; Gruber et al., 2014; see Kidd & Hayden, 2015 for a review) suggests that there may be an

intrinsic component to participants' motivation to learn. The possibility that this applies to language acquisition in early development is speculated by Mangardich & Sabbagh (2018; see also Mani & Ackermann, 2018), who suggest that hearing the label of a presented stimulus and inferring the intention of the speaker may be a high value inference for children in referential communication. Mangardich and Sabbagh (2018) further suggest that the pace of early word learning may in part be explained by the notion that word learning in children shares features with reward based learning.

Taken together, previous studies suggest that one of (many) factors driving early word-object association learning is that word-learning may be rewarding in and of itself for children, generating positive emotions associated with reward, potentially because children acquire information relevant for future referential communications with each new learned word-object association. Thus far, however, there are no studies examining the emotional and motivational state of young children following successful word learning. Against this background, we examine the extent to which children's state of arousal in a task during and following completion of a task predict word learning success in children, and furthermore, the extent to which word learning success may predict the positive emotions experienced by children following successful learning. In what follows, we will review the literature on our chosen indices for learning, arousal and reward, namely, fixation behavior in word learning tasks, pupillary arousal and body posture during and after task completion, respectively, to examine the extent to which they reliably index the factors under consideration in the current study.

Word learning and recognition.

Word learning in young children can be assessed by presenting children with images of as yet unfamiliar objects and training them on the labels for these objects using the intermodal preferential looking paradigm (Golinkoff et al., 1987). Following such training, children are then presented with two (or sometimes four) images of the previously trained objects and hear the label for one of these objects. Increased fixations to the labelled objects following presentation of the label for this object is typically interpreted as evidence for children's learning and recognition of

the trained word-object associations (see Westermann & Mani, 2019 for a review of early word learning research in young children; Bergelson & Swingley, 2009; Tincoff & Jusczyk, 2012; Mani & Plunkett, 2008, Taxitari et al., 2019; Ackermann et al., 2020a). In the cross-situational variant of the word learning task employed here (Smith & Yu, 2008; Yu & Smith, 2011), children are trained on temporarily ambiguous word-object mappings, where the ambiguity can be resolved when children consider the information presented across the entire training session. In particular here, children are presented with two unfamiliar images in training and hear two different novel labels, without being unambiguously told which object is referred to by which of the two labels. In subsequent trials, however, the labels and objects presented are controlled so that every time children see a particular object, they hear the label for this object, so that they use the frequency with which a particular label co-occurs with a particular object to infer the correct word-object mappings. At test, children are then presented with two of the trained objects and hear the label for one of these objects. As noted above, increased fixations to the labelled object is interpreted as evidence of children having learned the word-object mappings. Children's eye-movements can, therefore, provide us with further information of their learning of novel word-object mappings in such tasks.

Pupil dilation. The human pupil dilates in response not just to the physical properties of a stimulus but also indexes mental and emotional processing load during exposure to a stimulus (see Laeng et al., 2012 for a review of the literature in the 60 decisive years of pupillometry research). Following initial suggestions of pupillary changes reflecting changes in arousal state (Hess & Polt, 1960), subsequent studies documented increased pupil dilation during states of cognitive load (Kahneman, 1973; Kahneman & Beatty, 1966). Based on their review of the research on pupillometry, Laeng and colleagues further suggest that the pupil may offer “a window to the preconscious”, by indexing changes in participants' allocation of attention (Dehaene et al., 2006), engagement (Bradley et al., 2008), motivation and sustained processing (Hepach and colleagues, 2012, 2016, 2017), and how rewarding or important a particular stimulus might be (Bijleveld et al., 2009; Laeng & Falkenberg, 2007). Some of these findings have been replicated with developmental populations. For instance, Hepach and colleagues recorded children's tonic pupil

dilation during presentation of neutral videos as an index of their emotional arousal in response to previously presented emotionally valent stimuli (Hepach et al., 2016), and suggest that increases in pupil dilation reflect children's motivation to help others attain their instrumental goal (Hepach et al., 2012, 2016). With regards to the current study, the findings briefly reviewed here suggest that the pupillary response during a task may reliably index participants' arousal during and after a task, potentially even when participants do not display awareness or sensitivity to the stimuli being presented (Weiskrantz, 1990, 1998; Tamietto et al., 2009).

Body posture. Studies examining changes in participants' body posture have suggested a link between a person's emotional state and their body posture (see Hepach et al., 2015 for a review). These studies report an association between a person's feeling of shame or pride with their upper body posture, with a more upright body posture and erected head indexing pride and a lower body posture with lowered head indexing shame (Lewis et al., 1992). Similarly, success at a difficult task or failure at an easy task have been shown to induce changes in children's body posture at three years of age (Lewis, 1992), highlighting consistency in the findings across the life span. Hepach and colleagues (2015, 2016) show that children display increased positive emotions, indicated by more upright body posture, when they successfully achieve a goal or when they successfully help someone else achieve a goal. The authors particularly stress here the link between body posture and emotional expression and highlight the possibility that task success could trigger increased positive emotions. The latter suggestion is based on their findings of an increase in adult chest height following a positive emotion, relative to a negative emotion (Hepach et al., 2015). Corroborating evidence in these tasks is delivered by behavioral coding. Manual coding of the valence of the participants, e.g., how happy the person looked, and how much they smiled correlated with participants' body posture, providing additional evidence for the use of body posture as an index of participants' emotional state. Against this background, the current study will examine the extent to which success at a word learning task predicts changes to children's upper body posture as an index of the positive emotions that may be triggered with successful word learning. Such positive emotions may then underlie some of children's intrinsic motivation to learn words, with word learning being rewarded and associated with positive emotions, in and of itself.

The current study

The current study examined the extent to which children's internal state of arousal (indexed by pupil dilation) during and following completion of a word learning task predicts their word learning success. Furthermore, we also examined the extent to which word learning success itself (indexed by their eye-movements during the cross-situational word learning task) may predict the positive emotions (indexed by changes in body posture) experienced by children following successful learning. We presented children with a cross-situational word learning task (Smith & Yu, 2008), where children could learn six novel word-object associations based on the frequency of co-occurrence of a particular label with an object across trials. Children have been shown to learn novel word-object associations from a very early age (Bergelson & Swingley, 2009; Taxitari et al., 2019; Tincoff & Jusczyk, 2006). We specifically employed a potentially more difficult cross-situational word learning task given our interest in the degree of variability in word learning success.

We measured children's looking behavior, pupillary arousal and body posture at different points in the study. Looking behavior was recorded over the entire eye tracking part of the study. We expected that children would show increased fixations to the labelled object as an index of their learning and recognition of the trained word-object associations. Children's body posture were measured at several points during the session, namely, at the beginning of the study, right before, during and after training of the novel word-object associations, as well as before and after testing their recognition of the newly learned word-object associations, and finally, before and after testing their recognition of a series of known word-object associations (see schema 1). Changes in body posture from before to after test was our critical index of children's emotional state following task completion. We hypothesized that children who showed improved learning of the novel word-object associations would experience more positive emotions due to their success at the task, which would be indexed by elevation of upper body posture following task completion.

In addition, we measured children's pupillary arousal before, during and after the training phase, as well as before and after the testing their recognition of the newly learned word-object

associations and finally, before and after testing their recognition of the known word-object associations. Pupillary changes from before to after testing children's recognition of the newly learned word-object associations was our index of children's emotional state of arousal following task completion. We hypothesized that children who showed greater pupil dilation following the novel word recognition test phase would show better recognition of the novel word-object associations, highlighting increased arousal related to task success.

Method

Participants

Fifty monolingual German children between 36 to 48month of age (average 3y; 2m; 28d; range 3y, 0m, 4d – 3y, 11m, 30d, 29 girls) were recruited for this study. Two children were excluded from the study due to experimental error (1) and due to their knowing one of the novel objects prior to the study (1). There were 48 children in the final sample, but not all children provided data for both measures. For the looking measure, 47 children provided eye movement and pupillometry data - data collection failed at an early stage in the task for for one child. Regarding the body posture measure, 40 children provided posture data. No data could be recorded for the other eight children due to technical problems. Overall, 39 children provided data for both measures. Parents gave informed written consent for their child's participation in the study and children received a book in appreciation for their participation. Ethics approval was given by the Psychology Institute's Ethics committee (176_2).

Stimuli

Children were trained on six novel word-object associations. The chosen novel words were – akan, basa, sibu, modi, isot, upos. The novel objects were six objects of roughly the same size which were assumed to be unfamiliar to the children (see Fig. 8). Parents were asked after the experiment whether any of the objects were familiar to the child. A female native German speaker was recorded saying the words in isolation in infant directed speech. During the training phase, a single token of two different novel words were spliced together such that one of the novel words was presented 500ms into the trial, and a different novel word was presented 500ms after the end of

the first word. This compiled auditory stimulus was then merged with still images of two of the novel objects, presented to the left and right side of the screen with images measuring approximately 680 and 890 pixels separated by a distance of 40 pixels. During the test phase, individual tokens of the novel words were spliced together to create separate auditory stimuli for each novel word, where, participants heard four repetitions of the same word at 0 ms, 2000 ms, 4000 ms and 6000 ms into the trial, e.g., akan, akan, akan, akan. The compiled auditory track was merged with still images of two of the novel objects (used during training), one of which was the named target object, presented to the left and right side of the screen.

A separate set of stimuli was used for the known word recognition test phase. Here we presented children with pairs of images of familiar objects (dog, duck, ball, frog and car) as they heard the labels for one of these objects repeated four times (as in the novel word recognition test phase). For details of counterbalancing of labels and objects across all phases, refer to the description of the procedure below.



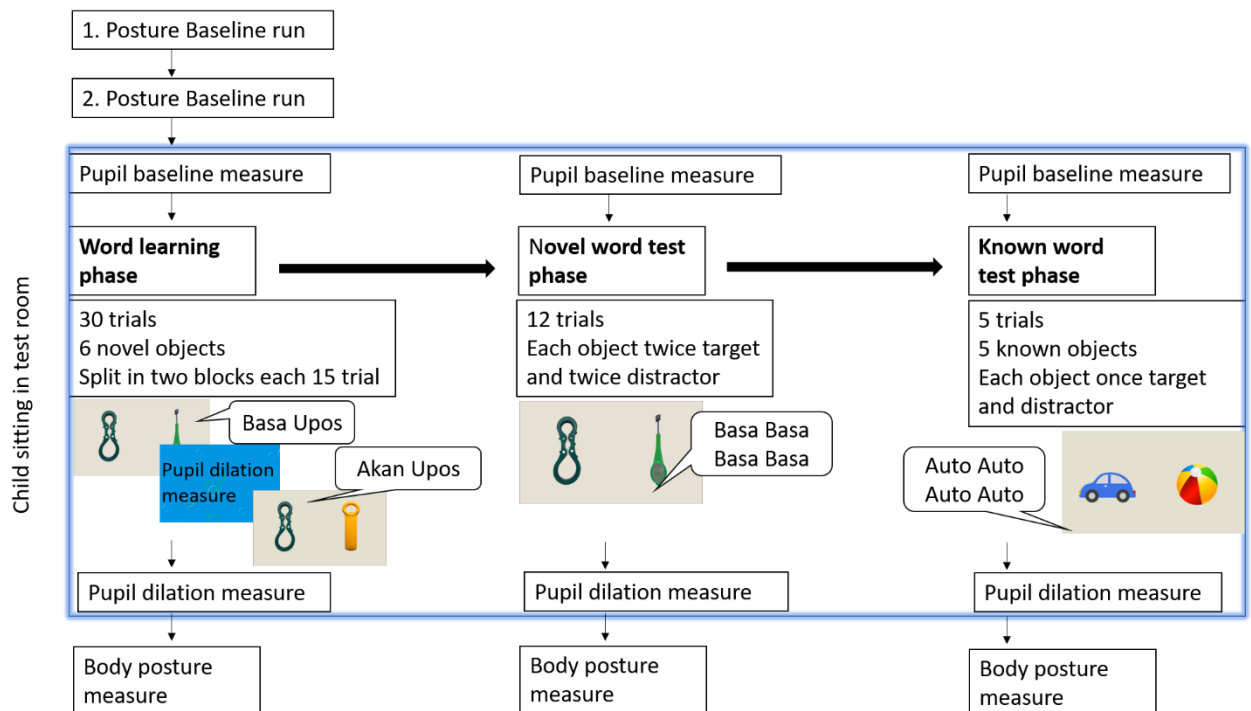
Figure 8: 6 Unknown objects presented during the study.

Procedure and Design

We recorded both eye tracking data (eye movements and pupillary measures) as well as body posture data (change in body posture) at different points during the study. For the eye tracking task, eye movements and pupil diameter were captured using a Tobii eye tracker (X120) with a gaze sampling rate of 120 Hz and a pupil diameter sampling rate of 120 Hz. The eye tracking data was collected in a quiet room with only artificial lighting. The eye tracker was placed on a platform underneath the TV screen (1920x1080 pixel) on which the stimuli were presented.

The child sat in a car seat about 65 cm away from the monitor. The body posture data was recorded using a Kinect motion capture camera (Microsoft Kinect), connected to a separate laptop. The Kinect camera was placed on a small table next to a toy, in a room adjacent to the room where testing took place. Before the experiment started, the child was shown where the toy was and was told that they would be asked to walk to this toy multiple times during the study. The child walked a distance of approximately 5 m between the door to the testing room and the toy. The caregiver was told to keep their distance from the child so as not to be tracked during measurement of the child's body posture.

The eye tracking study is split into three phases, namely a training phase, a novel word recognition test phase, and a known word recognition test phase (see Schema 3). The child provided eye-movement data during each of the phases. Pupillary data was acquired before and after the training, novel word test phase and known word test phase, as well as in the middle of the training phase. Body posture data was acquired before the learning phase as well as before and after the training, novel word test and known word test phase. We will first describe acquisition of the body posture data followed by description of the acquisition of pupillary and eye-tracking data.



Schema 3: Order of study phases with example stimuli and measure for each phase

Once the child felt comfortable with the experimenters, one experimenter and the child walked to the testing room where the child received a building block. The child was told to walk to the adjacent room and place the building block in a toy located directly next to the Kinect camera, which the experimenter had previously pointed to. This procedure was performed twice during the baseline run capture. Children were told to walk to the toy again at the end of the training phase, after completion of the novel word recognition test phase, and at the end of the experiment, i.e., after completion of the known word recognition test phase.

Following the baseline walks, the child was escorted to the testing room and seated in a child-seat in front of a TV screen, with the eye-tracker placed directly underneath the TV screen. Calibration was conducted using a 5-point calibration procedure. Once calibration was completed, we recorded the first pupillary measure of children's baseline tonic pupillary response as they watched a 10 second video of bubbles floating around the screen accompanied by chimes (see Hepach et al., 2012, 2013, 2015). This neutral video was repeated in the middle (after 15 training trials) and at the end of the training phase, at the beginning and end of the novel word recognition test phase and at the beginning and end of the known word recognition test phase, allowing us to measure children's pupil dilation (as an index of arousal) at critical points throughout the experiment.

Training phase. The training phase presented children with 30 trials, where each trial consisted of two of the novel objects presented to the left and right side of the screen respectively as the child heard the labels for each of these objects repeated once. The objects were presented in silence for 500 milliseconds, followed by the onset of the label for one of the objects. The onset of the label for the other object was 500ms after the offset of first label. Each trial lasted four seconds. Thus, while children could not infer the respective word-object associations from a single trial (since they saw two objects and heard two labels in each trial), across trials, they would hear the label for a particular object every time they saw this object onscreen. Children could use the co-occurrence of label and object across trials to infer the intended word-object pairings. Across the experiment, each object was labelled 10 times. The pairing of objects across trials was

counterbalanced such that each object appeared 10 times, five times to the left and five times to the right of the screen, paired with each of the other five objects. The order of the words in each trial was counterbalanced such that each word appeared 10 times (accompanied by the object it was representing), being mentioned first and second in the trial an equal number of times. Importantly we ensured that there was no correlation between the side of presentation of the object and the position of the word in the trial. Thus, if the word was presented first in the trial, the object could be either on the right or on the left side of the screen, depending the trial. A fixation cross appeared between each trial to ensure that the child was looking at the screen before the next trial started. Counterbalancing of the assignment of labels to particular objects resulted in six different lists and participants were randomly assigned to one of the lists. Thus across the lists, each of the objects was associated with each of the labels. The entire training phase lasted five minutes. Fifteen trials into the training phase, children were presented with the neutral bubble video to allow measurement of their pupillary dilation halfway through the experiment. We similarly presented children with the bubble video at the end of the training phase to measure their pupillary dilation at the end of the training phase. Then, children were given a building block and told to walk by themselves to the toy in the next room, to allow measurement of their body posture.

Novel word recognition test phase. After they had played with the toy for a short while, children were asked to return to the testing room. They were seated in front of the eye tracker and presented with the bubble video again to allow measurement of their pupillary dilation before the onset of the novel word recognition test phase. Then, children were presented with 12 trials. In each trial, children saw two of the previously presented novel objects to the left and right side of the screen for eight seconds and heard the label for one of these objects repeated four times at 0, 2000, 4000 and 6000ms into the trial. Across trials, each object appeared four times, half of the time as the target and the other half of the time as a distractor, and half of the time to the left and the other half of the time to the right side of the screen. The bubble video was presented at the end of the novel word recognition test phase, allowing us to measure children's pupil dilation at the end of the novel word recognition test phase. Next, children were given a building block and asked to

walk to the toy in the adjacent room, allowing us to measure their body posture at the end of the novel word recognition test phase.

Known word recognition test phase. After a brief period playing with the toy and the building block, children were asked to return to the testing room and seated on the chair. The bubble video was presented once more allowing us to measure children's pupil dilation at the beginning of the known word recognition test phase. Then, children were presented with five known word recognition trials where they saw two known objects and heard the label for one of the objects onscreen. The label was presented at 0, 2000, 4000, 6000 ms into the trial, with the entire trial lasting 8000 ms. Each of the five known objects (dog, duck ball, frog, and car) appeared twice across the entire phase, once as the target object and once as the distractor object, i.e., children heard the label for this object in one trial and heard the label for the other object in the other trial. The bubble video was presented one last time to allow for the measurement of children's pupil dilation at the end of the known word recognition phase. The children then received a building block for the last time and were asked to walk to the toy in the adjacent room, allowing us to measure their body posture at the end of the known word recognition test phase.

The children were then allowed to play for a while with the toy before receiving a book in appreciation for their participation in the experiment while the caregiver was informed about the purpose of the study.

Data Analysis

For each child, we examined their word learning and recognition, pupillary dilation and body posture across different phases of the experiment. In what follows, we describe the data preprocessing for each of these datasets.

Word recognition. The eye tracking output from Tobii during the novel and known word recognition test phases was extracted and further analysed in combination with the stimulus information for each trial. The eye tracker provides an estimate of where children was looking every 8 ms. Only data which were coded as reliable, i.e., with a validity less than 2, were included in the analyses. These data were further aggregated across 40-ms bins, with each of these 40-ms

bins coded for whether the child was looking at the target or the distractor³. Based on the location and size of the objects on the left and right side of the screen, we defined areas of interest and coded these for whether they were the target or a distractor in the trial. Data from a particular trial was only included and further analyzed if we obtained data for at least 20% of the trial (see Ackermann et al. 2020a), leading to an exclusion of 46 trials (7.73 %) during the novel word recognition test phase and 1 trial (0.12%) in the known word recognition test phase. We required that participants provided data for at least two trials in each test phase to be included in the analyses. This led to no exclusions. We calculated the proportion of target looking (PTL) for each trial based on the total amount of time children spent looking at the target over the amount of time children spent looking at the target and the distractor across the entire trial. This measure was then aggregated across trials for each child, resulting in one mean PTL value over all trials for each child. We entered this PTL value into further analyses described below. Forty seven children provided data for the novel word recognition test phase, while 44 children provided data for the known word recognition test phase (two children did not return to the eyetracker booth for the known word recognition test phase and one child did not provide enough data during this phase).

Pupil dilation. Pupillary data collected during presentation of the bubble video at the beginning and end of the training phase, the novel word recognition test phase and the known word recognition test phase were extracted from the Tobii output. Pupil diameter was sampled at 120 Hz. Data from both eyes was extracted for each time bin. Time bins were only included in the analysis if they were part of a fixation and valid eye tracking data (average over both eyes less than 2 on a Tobii scale) Data from a trial was only included if we had data for at least 20% of that trial. Data was then filtered by removing data from timepoints where the difference in pupil size between adjacent time points was in the upper 10% of the difference in pupil size between all adjacent time points were removed (Ackermann et al., 2020a). This ensured that large differences in pupil size between two adjacent samples, which are likely to be artifacts, are excluded from further analysis.

³ Following the conventions employed in our previous work, we code a child as looking at the target in a particular 40ms bin if the child looked at the target for more than 2/3rds of the number of bins included in any 40ms bin. If the child looked less at the target than this threshold, the child was coded as not looking at the target or looking wherever the child did look for at least 2/3rds of the bin (e.g. Eiteljörge, Adam, Elsner, & Mani, 2019).

Missing data was interpolated across every four samples, equivalent to a 70ms sliding window (Hepach et al., 2012). Filtered, interpolated data from each eye was then averaged together, if data from both eyes was provided, or using only the data from one eye when this was not the case. For the novel word recognition test phase, 40 participants provided data both before and after the onset of the test phase (one child did not provide any data for the novel word test phase, six children did not provide data on both measurements, before and after the novel word test phase). For the known word recognition, 43 children provided data both before and after the known word recognition test phase. We aggregated the pupillary data across the entire duration of the bubble video and subtracted the pupillary measure before the onset of the training phase and each test phase from the measure obtained after completion of each phase respectively. This difference score was then entered into subsequent analyses (described below).

Body posture. Data was preprocessed using a separate MATLAB script. In a first step, we filtered and cleaned the data to ensure that we did not include skeletons of adults and we only included data 1) that was in our tracking range (3,2 m- 1,2 m distance from the Kinect camera), 2) where the skeleton point center back was above the skeleton point hip center, 3) the skeleton point of the head was above the skeleton points of the shoulder, 4) the skeleton point center hip was between the skeleton points left and right hip, 5) the skeleton point center shoulder is between the skeleton points right and left shoulder. These criteria ensured that only datapoints that could reasonably be considered skeletal measurements of the child were included in the analyses (see Hepach et al., 2016). Resulting gaps in skeletal measurements were then interpolated. Following these criteria, we excluded an additional 6 participants. Two did not provide valid data for baseline walks (for one participant only the adult was tracked, one participant turned back in the middle of the walks), three participants whose data violated the checks outlined above. We further excluded the data from one participant who incorrectly received the test phase before the learning phase. An additional eleven single walks were removed (1: child walking backwards, 2: child walking wrong way, 3: child turned back, 3: child fell or grabbed something on the ground, 1: child and adult tracked as one skeleton and 1: child was walking with big toy). We focused on two body points in particular, namely the y-coordinate (height) of the chest's center to estimate the upper body posture

change, and the y-coordinate of the hip center to estimate lower body posture change (Hepach et al. 2016). We scaled the posture measurements over time, such that all participants provided data for 20 data points in each run, regardless of how long they spent on each run. Thus, we divided the run into twenty time bins and extracted an average posture measure for each of these time bins. The posture measure is typically less accurate the closer the participant is to the Kinect camera. To ensure that these less accurately measured data points are excluded from the analysis, we only considered the data for the first 16 time bins collected. Then, we baseline corrected the data acquired after completion of the novel word recognition test or known word recognition test phase. During baseline correction, we subtracted the baseline posture data from the measure obtained after completion of each phase for each time bin and each phase separately. If the child provided data for both baseline walks, baseline correction was performed using the average of both walks, if the child did not provide data for the first baseline walk, the second baseline walk was taken for the baseline correction. This baseline corrected difference score was then entered into subsequent analyses (described below).

Analysis

Word recognition.

First, we examined whether children had learned the novel word-object associations and recognized the known word-object associations presented. Here, we ran two one-sample t-tests to examine whether the PTL was above chance (50%) in the novel word recognition and known word recognition phases, separately.

Word recognition and body posture. To examine the relationship between word recognition and body posture, we examined the data from 34 children who provided both body posture data and eye movement data from both the novel word recognition test phase and the known word recognition test phase. We ran a generalized linear mixed model on the posture data including Time and its interaction with Word recognition and Setting, as well as fixed effects of Age (in days)⁴

⁴ We z transformed age to improve model fit.

and Gender and random intercept for subject and random slope for time. Time refers to the 16 time bins in each run, where the data in each bin for the critical run was baseline corrected to corresponding data in each bin in the baseline run (or runs depending on whether the child provided data for a single or both baseline runs). Word recognition was the PTL score from the novel word recognition test phase or the known word recognition test phase, depending on which phase was being analyzed, indicated by setting. Breaking this model down into the different settings, we run a model on the novel word test phase and the known word test phase separately. The dependent measure was either the change in the children's chest height or hip height across separate models.

We collected additional data on the valence of the child's emotion, provided by two coders. These two adult coders were blind to the study's hypotheses and settings. They rated the emotions of the child on an individual trial level, based on the frame by frame recordings of each trial. Each coder was asked to rate how positive the emotion was, the child was experiencing (most positive = 3, least positive = 1). We calculated the mean rating for each phase of the study (baseline, training, novel word recognition and known word recognition) and calculated the difference between baseline and all other phases, to receive a value expressing the change in the child's emotion (positive value = more positive emotion compared to baseline, negative value = less positive emotion compared to baseline). We merged the two coders' values and ran a linear mixed model on the valence value including Word recognition and its interaction with Setting, as well as fixed effects for Age (in days)⁵, Gender, Subject and Coder.

Word recognition and pupil dilation. To examine the relationship between word recognition and pupillary arousal, we examined the data from 40 children who provided both pupillary data (for the training and the novel word recognition test phase) and eye movement data from the novel word recognition test phase, and 38 children who provided both pupillary data and eye movement data from the known word recognition test phase.

We ran a generalized linear mixed model on the pupillary arousal data including Condition and its interaction with Word recognition⁵, as well as fixed effects of Age (in days)², Gender and

⁵ We z transformed age to improve model fit.

subject. The dependent measure was the pupil measure, which was the difference in pupil dilation from before the novel test phase or the known test phase phase to after this phase (known test/novel test – Baseline), using either the data collected around the novel or known word recognition test phase, indicated by condition. The PTL measure used in the analysis was either the PTL score for each trial from the novel word recognition test phase or the known word recognition test phase, again indicated by condition.

Additionally, we ran separate spearman's correlations on the pupil dilation and PTL data. The pupil measure used in the analysis was the difference in pupil dilation from before the test phase to after this phase, using either the data collected around the novel or known word recognition test phase, depending the phase being analyzed. The PTL measure was the PTL score for each trial from the novel word recognition test phase or the known word recognition test phase, across the different correlations. Additionally we run Spearman's correlations on the pupil dilation data during the learning phase and PTL data during the novel word recognition phase. The pupil dilation measure was the difference in pupil dilation from before to mid and after learning.

Pupil dilation and body posture

To examine the relationship between body posture and pupillary arousal, we included data from 28 children in the novel word recognition phase and 28 children in the known word recognition phase who provided both pupillary and body posture data for these phases.

We ran a generalized linear mixed model on the posture data including Time (16 time bins) and its interaction with the baseline corrected Pupillary arousal and Condition, as well as fixed effects for Age (in days)⁶ and Gender and a random intercept for subject and a random slope for time. The pupil measure used in the analyses was the difference in pupil dilation from before the test phase to after the test phase (test – baseline), using either the data collected around the novel or known word recognition test phase, indicated by condition. The dependent measure was the change in the children's chest height.

⁶ We z transformed age to improve model fit.

Results

Word recognition.

First, we examined whether children had learned the novel word-object associations and recognized the known word-object associations presented, i.e., showed increased fixations to the labelled object relative to the other distracter object in the novel and known word recognition test trials. One sample t-tests, comparing the PTL to chance (50%), found that the proportion of fixations to target was above chance in both the novel word recognition, $t(46) = 3.44$, $p = .001$ 95% *CI*: 0.51, 0.55 (see Fig.9a), and known word recognition phases, $t(43)=8.41$, $p < 0.001$, 95% *CI*: 0.57, 0.61. (see Fig.9b).

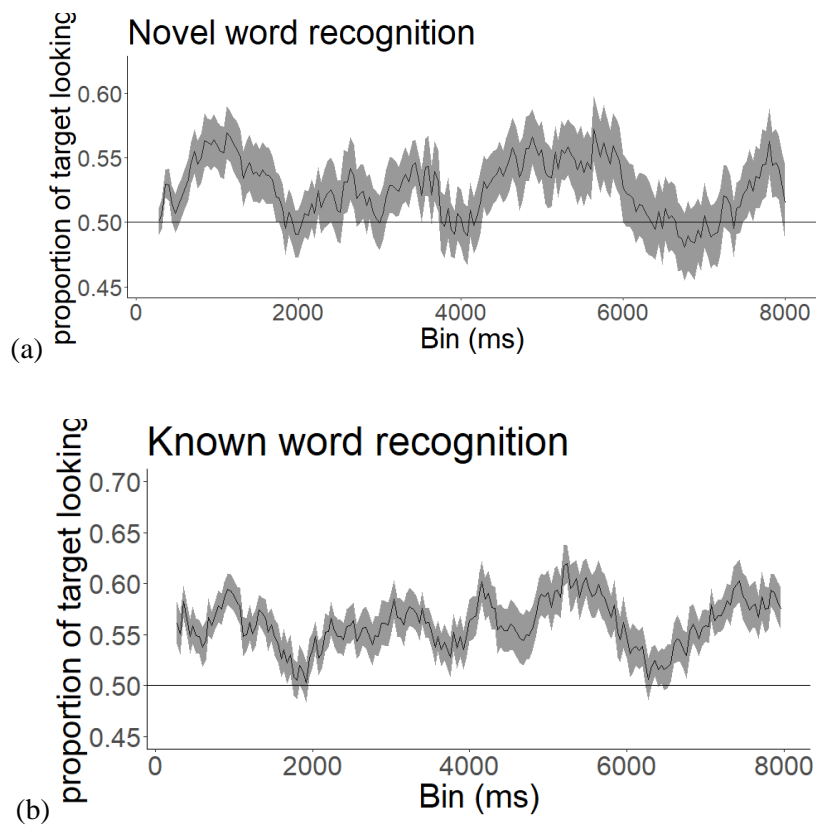


Figure 9. Proportion of target looking during the novel word recognition test phase (a) and the known word recognition test phase (b).

Word recognition and body posture

Investigating the model including Time and its interaction with Word recognition and Setting, as well as fixed effects of Age (in days)⁷ and Gender and random intercept for subject and random slope for time, we found a change in posture across the trial which differs across the two settings, with higher upper body posture following familiar word recognition relative to novel word recognition. Upper body posture drops across time following novel word recognition. (see Fig. 10 & Table 5). We also found an interaction between time, Proportional target looks and setting (novel, familiar words; Figure 11). Breaking this down, we didn't find an interaction with learning in the test phase or the known phase, individually. There is only a difference across time in upper body posture following novel word recognition but not following familiar word recognition. (see Table 6)

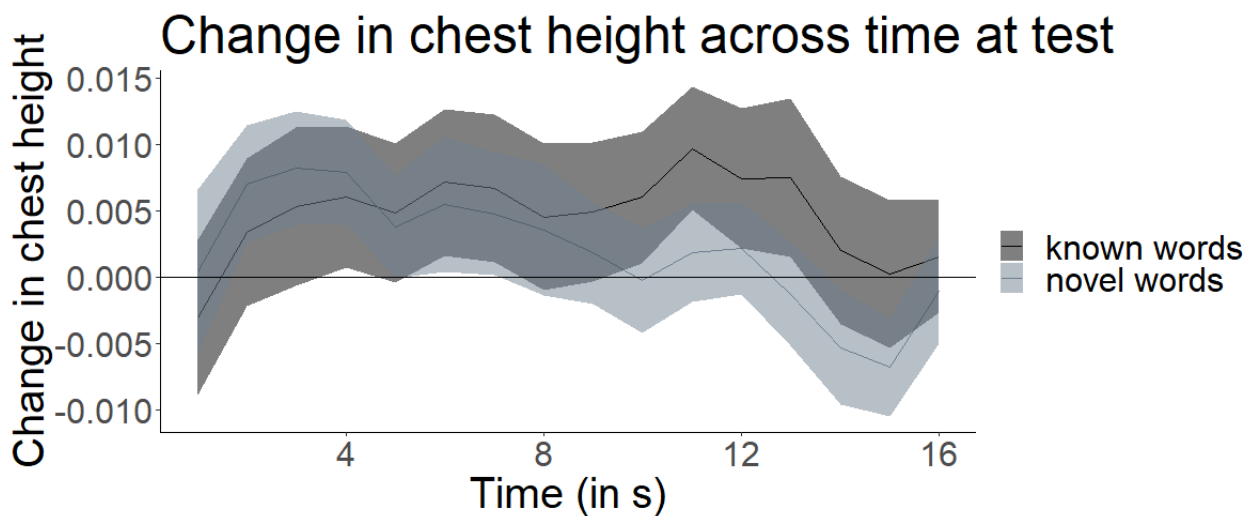


Figure 10. Changes in upper body posture (chest height) over time separated in novel word recognition test phase and known word recognition test phase.

⁷ We z transformed age to improve model fit.

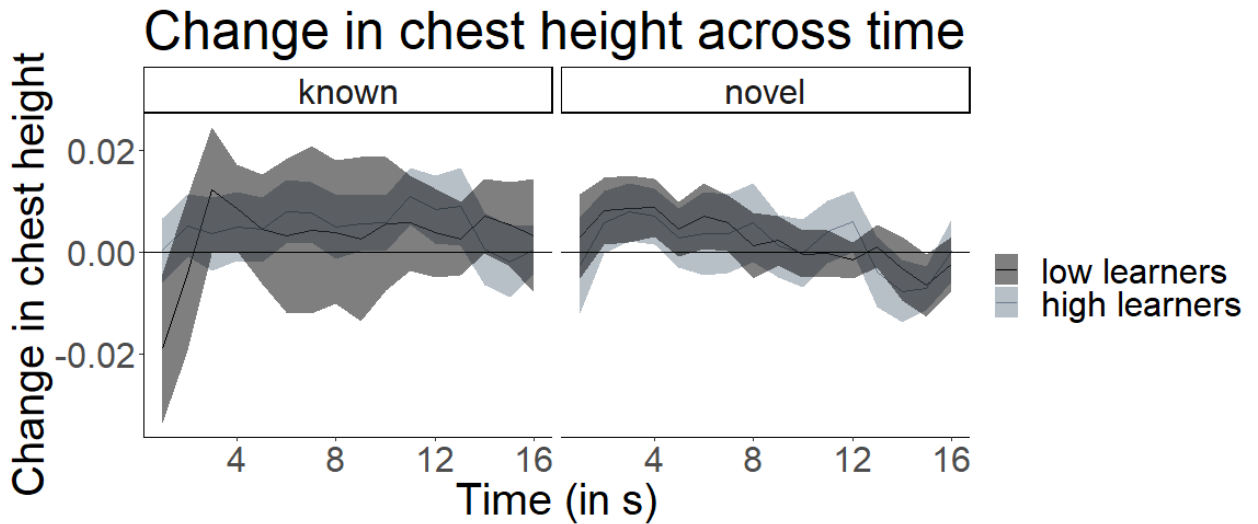


Figure 11. Changes in upper body posture (chest height) over time separated in novel word recognition test phase and known word recognition test phase, split into high and low learners, indicated by the median split of proportional target looks.

Table 5. Model examining the upper posture data including Time and its interaction with Word recognition (proportional target looks) and Setting (novel words, known words), as well as fixed effects of Age (in days) and Gender and random intercept for subject and random slope for time:
 $\text{lmer}(\text{upper body posture} \sim \text{time} * \text{Proportional target looks} * \text{setting} + (1 | \text{gender}) + (1 | \text{age.ds}) + (1 + \text{time} | \text{id}))$

| Predictors | Estimates | Std Error | t-value | p-values |
|--|-----------|-----------|---------|-------------|
| Intercept | 0.002 | 0.004 | 0.610 | .541 |
| time | 0.0003 | 0.0003 | 1.037 | .299 |
| Proportional target looks | 0.042 | 0.037 | 1.149 | .250 |
| Setting novel word recognition | 0.007 | 0.003 | 2.215 | .026 |
| Time : Proportional target looks | -0.011 | 0.003 | -2.959 | .003 |
| Time : setting novel word recognition | -0.001 | 0.0003 | -3.262 | .001 |
| Proportional target looks : setting novel word recognition | 0.030 | 0.048 | 0.615 | .538 |

| | | | | |
|------------------------------------|-------|-------|-------|-------------|
| Time : Proportional target looks : | 0.011 | 0.004 | 2.298 | .021 |
| setting novel word recognition | | | | |

Table 6. Model examining the posture data including Time and its interaction with Word recognition (proportional target looks, as well as fixed effects of Age (in days) and Gender and random intercept for subject and random slope for time split by setting (novel word recognition, known word recognition): lmer(upper body posture ~ time* Proportional target looks + (1 |gender) + (1 |age.ds) + (1+time|id))

| Predictors | Novel word recognition phase | | | familiar word recognition phase | | |
|----------------------------------|------------------------------|-----------|------------------|---------------------------------|-----------|------------------|
| | Estimates | Std Error | <i>p</i> -values | Estimates | Std Error | <i>p</i> -values |
| Intercept | 0.008 | 0.005 | .120 | >0.001 | 0.005 | .670 |
| time | -0.0007 | 0.0003 | .048 | >0.001 | 0.0003 | .885 |
| Proportional target looks | 0.035 | 0.070 | .612 | >-0.001 | 0.079 | .956 |
| Time : Proportional target looks | 0.0006 | 0.005 | .900 | ->-0.001 | 0.005 | .243 |

We found exactly the same pattern for hip high. There is a change in posture across trial which differs across the two conditions with higher lower body posture following familiar word recognition relative to novel word recognition. Lower body posture drops across time following novel word recognition (see Figure 12, Table 7)

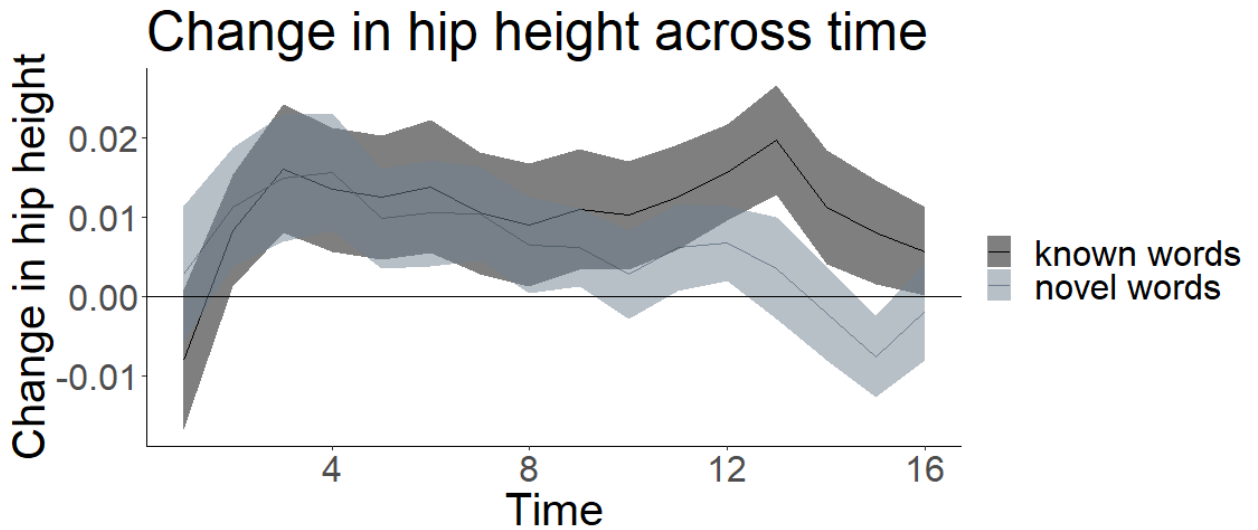


Figure 12. Changes in lower body posture (chest height) over time separated in novel word recognition test phase and known word recognition test phase.

| Predictors | Estimates | Std Error | <i>t</i> -value | <i>p</i> -values |
|--|-----------|-----------|-----------------|------------------|
| Intercept | 0.005 | 0.007 | 0.834 | .404 |
| time | 0.0007 | 0.0005 | 1.1485 | .137 |
| Proportional target looks | 0.058 | 0.048 | 1.189 | .234 |
| Setting novel word recognition | 0.008 | 0.004 | 1.935 | .053 |
| Time : Proportional target looks | -0.014 | 0.005 | -2.88 | .004 |
| Time : setting novel word recognition | -0.001 | 0.0004 | -3.793 | >.001 |
| Proportional target looks : setting novel word recognition | 0.105 | 0.063 | -1.671 | .094 |
| Time : Proportional target looks : setting | 0.021 | 0.006 | 3.298 | >.001 |

Table 7. Model examining the lower posture data including Time and its interaction with Word recognition (proportional target looks) and Setting (novel words, known words), as well as fixed effects of Age (in days) and Gender and random intercept for subject and random slope for time:

$\text{lmer}(\text{lower body posture} \sim \text{time} * \text{Proportional target looks} * \text{setting} + (1 | \text{gender}) + (1 | \text{age.ds}) + (1 + \text{time} | \text{id}))$

Investigating the children's emotion indicated by the valence coding, we linear mixed model including PTL scores and its interaction with setting, we find an influence of setting, indicating children being more happy after the known word recognition test phase, compared to the novel word recognition test phase (see Table 8 and Figure 13)

| Predictors | Estimates | Std Error | t-value | p-values |
|--|-----------|-----------|---------|-----------------|
| Intercept | 0.581 | 0.142 | 4.076 | <.001 |
| Proportional Target looks | -0.019 | 1.200 | -0.016 | .987 |
| Setting novel word recognition | -0.407 | 0.104 | -3.913 | <.001 |
| Proportional target looks: setting novel word recognition | -2.344 | 1.514 | -1.547 | .121 |

Table 8. Model examining the valence data including Word recognition and its interaction with Setting (novel words, known words), as well as fixed effects of Age (in days), Gender, Subject and Coder: $\text{lmer}(\text{valence measure} \sim \text{PTL} * \text{setting} + (1 | \text{gender}) + (1 | \text{age.ds}) + (1 | \text{id}) + (1 | \text{coder}))$

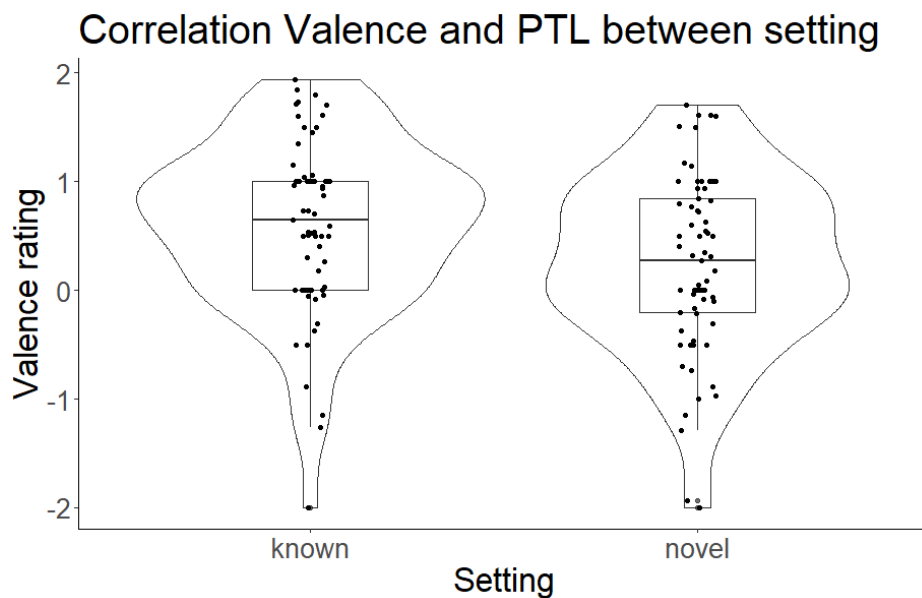


Figure 13. Difference in valence rating between novel word recognition test phase and known word recognition test phase.

Word recognition and pupil dilation.

Training phase. Spearman's correlation investigating the relation between pupil dilation during learning (difference between before to mid training and before to after training) and performance during the novel word recognition test phase showed no correlation between pupillary arousal during training and performance at test ($ps > 0.05$).

Novel word recognition test phase. Spearman's correlation investigating the relationship between pupil dilation from before to after the novel word recognition test phase with the PTL at novel word recognition test phase. Results showed a strong correlation between pupillary arousal following test and performance at test, $r(38) = .41, p = .009$ (see Figure 14)

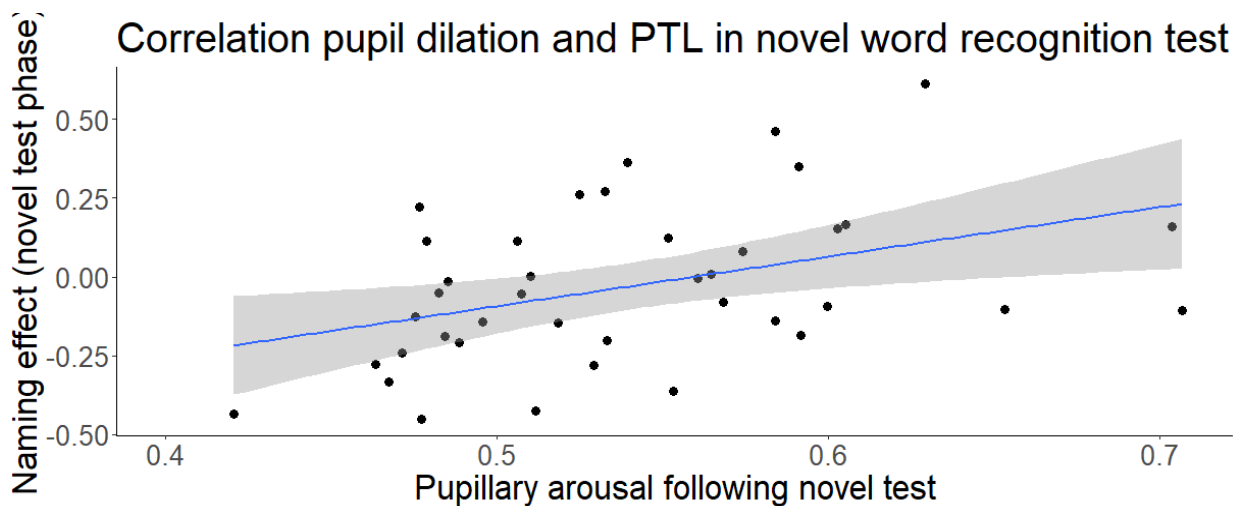


Figure 14. Correlation between pupil dilation (indicated by the difference between pupil dilation from before to after novel word recognition) and the infants word recognition during the novel word recognition test phase (indicated by the PTL score during test).

Known word recognition test phase. Investigating the relationship between the pupil dilation and the PTL during the known word recognition test phase, we correlated pupil dilation data from before to after the known word recognition test phase with the PTL scores at the known word recognition test phase. Results showed no correlation between pupillary arousal and performance, $r(36) = .19, p = .24$.

Pupil dilation and body posture

Investigation the relationship between pupillary arousal and body posture we didn't find any effect of pupillary arousal (or interaction with time or interaction with novel or known word recognition phase) on upper body posture ($ps > .05$) or indeed individually separated by novel or known word recognition test.

Discussion

Against the background of suggestions that learning, in adults, may be rewarding in and of itself, i.e., that we may be intrinsically motivated to learn (Ripolles et al., 2014; Kang et al., 2009; Gruber et al., 2014), the current study set out to examine the relationship between successful word learning in children and the positive emotions triggered by such word learning success, indexed by pupillary arousal and upper body posture. In particular, we hypothesized a positive relationship between word learning success and pupillary arousal during task completion, indexing increased arousal related to task success, as well as a positive relationship between word learning success and upper body posture following task completion, indexing positive emotions related to task success.

First, we examine children's success at word learning in the cross-situational word learning task presented to them. We found that children showed successful recognition of both the trained word-object associations that they were introduced to in the cross-situational word learning paradigm as well as the known word-object associations. In both cases, children looked more at the image of the labelled object relative to the unlabeled distractor object. Since objects were counterbalanced so that they were targets as often as they were distractors, increased looking at an object when presented with the label for this object can be reasonably interpreted as evidence for novel word-object association learning and recognition, as well as known word recognition. Next,

we examined the extent to which successful word learning and word recognition was associated with first, increased pupillary arousal during task completion and second, positive emotions, as indexed by changes to upper body posture following task completion as well as valence codings.

Word recognition and pupil dilation

We also examined changes in pupillary arousal from before to after task completion with regards to the training phase, the novel word recognition test phase and the known word recognition test phase. We found no evidence for an association between pupillary arousal following completion of the training phase and performance in the novel word recognition task. Neither did we find evidence for an association between pupillary arousal following completion of the known word recognition test phase and performance in the known word recognition task. However, we did find positive evidence of an association between increase in pupillary dilation (from before to after) following completion of the novel word recognition task and word recognition success in this task.

As noted earlier, the human pupil dilates in response to a range of cognitive processes including attention and resource allocation (Laeng et al., 2012; Dehaene et al., 2006; Hepach & Westermann, 2016), memory load (Ahern & Beatty, 1981; Bijleveld et al., 2009), cognitive load (Just et al., 2003) as well as emotional factors such as motivation and reward (Bijleveld et al., 2009), engagement (Bradley et al., 2008; Hepach & Westermann, 2016) and perseverance (Siegle et al., 2008; Hepach and colleagues, 2012, 2016, 2017).

Indeed, the positive association between pupillary arousal following the novel word recognition task and performance in this task could be explained by increased resource or effort allocation during the task. It is important to distinguish here between resource or effort allocation during encoding and during recognition, since we did not find any evidence for an association between pupillary arousal during the training phase and novel word recognition performance. Such a finding could have been interpreted as evidence for improved attentional allocation during training leading to improved encoding and subsequent recall of the word-object associations at test. We will not interpret the null effect in this analysis and will therefore leave open the question

regarding the association between resource/effort allocation and encoding. However, we did find an association between performance at test and pupillary arousal at test, which we take to suggest that children who invested more cognitive resources or effort during the task showed improved recognition of the novel word-object associations. In other words, those children who were more on task showed improved task performance (see Rondeel et al., 2015; Gross & Dobbins, 2020 arguing for pupil dilation as an index of effort exertion).

This interpretation would also be in keeping with studies showing an association between pupil size and interest or engagement in a task (Kang & Wheatley, 2017; Laeng et al., 2016; Ackermann et al., 2020a). For instance, pupillary arousal to musical excerpts has been found to vary with participants' engagement in music in their daily lives or their interests in particular kinds of music (Gingras et al., 2015; Laeng et al., 2016), suggesting associations between long-term individual interests and pupillary responses to particular types of stimuli. Similarly, pupillary arousal to categories of familiar objects has been shown to predict children's learning of novel word-object associations for objects from categories individual participants showed increased arousal to (Ackermann et al., 2020a). Pupillary arousal has also been shown to predict participants' prosocial behaviour, with participants who showed increased pupillary arousal also being more willing to help those in need (Hepach et al., 2016), potentially therefore investing more in the situation. Against this background, and in concord with the interpretation of the association between pupillary arousal and task performance as indicative of effort or resource allocation, these findings could also suggest that those children who were more engaged or aroused by the task showed improved recognition performance. Admittedly, the directionality of this association is, however, difficult to ascertain. For instance, participants could be more aroused or engaged in a task due to their being able to complete the task or they may perform better in the task because they are more engaged in the task. We will, therefore, steer clear of attributing directionality either way and highlight here merely our interpretation of the association between pupillary arousal and task performance as indexing the association between increased engagement or effort allocation and improved task performance.

Of particular interest is the fact that we only found an association between pupillary arousal and performance in the novel word recognition test but not the known word recognition test. One explanation for our failure to find evidence for such an association in the known word recognition task might tap into the ease of this task – thus there may have been fewer demands on resources or less of a need to engage further in the task in order to show successful performance in this task. In other words, children may have been able to complete this task without engaging too much in the task or investing greater cognitive resources in the task, thereby allowing for less variation in pupil size during task completion. In contrast, the novel word recognition task may have called for greater allocation of resources leading to the reported association between pupil size and task performance in this task.

Finally, we highlight one additional interpretation of the reported increase in pupil size related to task performance that is more in keeping with the aims of the current study. In particular, there has been some suggestion that increase in pupil size may be related to emotional factors such as motivation and reward. In particular, Biljeveld et al. (2009) suggest that pupillary dilation may be increased for more (monetarily) rewarding tasks but only in cases where the task requires additional effort. While this interpretation is closely tied to participants' motivation or engagement to complete a task (as discussed above), it adds a further perspective in terms of the association between how rewarding and how demanding a task is. Thus, our finding of the association between increased pupillary arousal and performance in the novel word recognition task but not the known word recognition task, could potentially be interpreted as tapping into this nexus between reward and task difficulty, with the former being difficult enough to engage participants in the task and being potentially more rewarding. This brings us, however, to the analyses examining word recognition and body posture that we discuss next.

Word recognition and body posture.

We had hypothesized that children who showed improved word recognition performance would also show increased upper body posture following task completion as an index of the positive emotions generated by successful task completion. This was motivated by studies showing children's and adult's success at a task is similarly associated with their emotional state and more upright body posture (see Hepach et al., 2015 for a review; Lewis, 1992; Hepach et al., 2016). Thus, for instance, children display increased positive emotions, indicated by more upright body posture, when they successfully achieve a goal or when they successfully help someone else achieve a goal. However, we found no evidence that those children who showed improved task performance showed more upright body posture. In other words, we found no evidence that children who did better at the task showed increased positive emotions.

On the one hand, this result appears to contradict with the findings on body posture reviewed above, as well as suggestions of the connectivity between motivational and learning circuits in the brain (Syal & Finley, 2011; Ripolles et al., 2016, 2018). On the other hand, the lack of a significant association can owe to a number of factors which we briefly discuss here. For instance, studies highlighting the link between learning and motivation and/or reward circuitry in adults have typically tapped into more implicit measures of reward, e.g., increased fMRI activation in the ventral striatum in cases of successful learning (Ripolles et al., 2016, 2018). Thus, it is possible that our measure of body posture and its relation to participants' positive emotional state may not adequately tap into how rewarding word learning may be for young children. Alternatively, it is possible that children may not be as aware of the fact that they have learned the words in the current task as they may be of other aspects of their behavior. For instance, this body posture measure has been shown to index positive emotions upon successful goal completion in other tasks, e.g., in helping others (Hepach et al., 2015, 2016). The contrast between these findings would then either suggest that word learning is not as intrinsically rewarding to children as prosocial behavior or that children may not be as aware of their successful word learning relative to their having helped someone, leading to increased positive emotions following such prosocial behavior. Speaking in favor for this assumptions are our findings in our second measure to

investigate children's emotion, indicated by valences scores, more positive emotions after testing already known words. In line with this finding, we found a drop in body posture across time following new learned words but not already known words. This result was independent from the performance of the children during the word recognition test phases. These findings might suggest that children are happier after performing an easier task, with already known words compared to the novel word recognition test phase and might be an indicator that word learning in this task was not as intrinsically rewarding and easy to measure as other tasks. Thus, tasks making children more aware of their word learning success may be better able to tap into the reward-related aspects of word learning. Finally, we note that children may have been aware of having successfully learned words, but any positive emotions that may have been generated as a result of this may be more explicit at other times during the experiment. Thus, children may have experienced positive emotions at some other point during the task, e.g., when they successfully recognized the words relative to the point at which we examined changes in body posture as our chosen index of such positive emotions. Thus, considerable more work is needed to further investigate the relationship between positive emotions, expressed in changes in body posture, and word learning success.

Conclusion

The current study set out to examine the role of children's internal emotional state on their word learning success. Against the background of active learning approaches that highlight the contribution of the child to learning success, we found the children who were more engaged in or aroused by the task showed improved word recognition performance. While we found no evidence of greater engagement during encoding being associated with word recognition performance, the correlation between engagement at test and performance at test, highlights the important role of children's motivation and engagement in performance in such tasks. These findings are in keeping with the extensive literature on children's classroom learning highlighting the role of motivation and engagement in classroom performance (Salmela-Aro et al., 2016; Skinner et al., 2008). Notably, we did not find any evidence that children experienced positive emotions upon successfully learning words, at least as indexed by changes to their upper body posture. While not entirely in keeping with the literature on the affective state of word learning in adults, we suggest a

number of reasons why children may be affectively involved in word learning and avenues for future research.

Diskussion

Kinder werden während der Zeit ihres Spracherwerbs mit einer Vielzahl an äußeren Einflüssen konfrontiert. Aus diesen müssen sie lernen zwischen wichtigen und unwichtigen Informationen zu unterscheiden und welche Informationen sie nutzen wollen, um optimal zu lernen. In Bezug auf Sprache müssen Kinder zuerst das phonetische Lautinventar ihrer eigenen Muttersprache lernen, um anzufangen Wörter aus dem Sprachfluss zu segmentieren. Erst danach können Kinder Wort-Objekt Assoziationen bilden, um danach komplexere Zusammenhänge verstehen zu können. Um diesen Prozess zu meistern, nutzen die Kinder eine Reihe von unterschiedlichen inneren Bias, ihre eigene Motivation etwas zu lernen aber auch unterschiedliche extrinsische Faktoren, wie KGS.

In der ersten Studie dieser Dissertation (Kapitel 2) war das Ziel die Wechselwirkungen zwischen den prosodischen Eigenschaften von mütterlichem KGS, der Präferenz von Kindern für KGS und der funktionellen Relevanz von KGS in kindgerechter Kommunikation zu untersuchen. Wir hatten angenommen, dass diese drei Facetten von KGS sich gegenseitig beeinflussen können und über die Entwicklung der Kinder variieren. Um dies zu untersuchen, nahmen wir den mütterlichen Input in Form von KGS und EGS auf, maßen die Präferenz von Kindern für KGS und EGS und untersuchten ob Kinder einer Blickbewegung mehr folgten, wenn diese mit KGS oder EGS präsentiert wurde. Wir konnten einen Unterschied in den Eigenschaften von EGS und KGS sowohl zu 6 als auch zu 13 Monate alten Kinder feststelle. Eine Präferenz für KGS ließ sich nur im Alter von 6 Monaten nachweisen. Das Folgen der Blickbewegung variierte zwischen beiden Altersgruppen nicht, und zeigte, dass die Kinder dem Blick des Modells besser in KGS als in EGS gefolgt sind. In Bezug auf die Interaktionen zwischen den Studien stellten wir fest, dass die akustischen Eigenschaften des mütterlichen KGS (im Vergleich zum EGS) mit der kindlichen Präferenz für EGS und KGS korrelieren. Wobei wir zum jetzigen Zeitpunkt nicht wissen ob die Kinder den elterlichen Input durch ihre Präferenz modifizieren, oder ob die Eltern durch ihren Input die Präferenz der Kinder beeinflussen.

In der zweiten Studie dieser Dissertation (Kapitel 3) waren wir daran interessiert die Wechselwirkungen zwischen den prosodischen Eigenschaften von mütterlichem KGS, der Präferenz der Kinder und der Fähigkeit der Kinder Wörter aus einem Sprachfluss zu segmentieren, zu untersuchen. Wie in Studie 1 wurden 6- und 13-Monate alte Kinder getestet. Die Segmentationsfähigkeit der Kinder wurde mit Hilfe von EEG Messungen untersucht. Aufgrund der Datenmenge entschieden wir uns jedoch nur für die Auswertung der EEGs. Die Auswertung der EEG Messung ergab in beiden Altersklassen keine Evidenz dafür, dass Kinder aus EGS oder KGS Wörter segmentieren können. Es zeigte sich jedoch in beiden Altersklassen eine Reaktion auf KGS. Diese Reaktion war von der Kondition (familiarisierte oder neue Wörter) unabhängig. Diese Ergebnisse sprechen für eine Modulierung der kindlichen Aufmerksamkeit durch KGS.

In der dritten Studie dieser Dissertation (Kapitel 4) waren wir daran interessiert den Einfluss von kindlicher Erregung an einer Aufgabe in Zusammenhang mit dem Lernerfolg der Kinder zu untersuchen und herauszufinden, ob ein positives Lernergebnis die Emotionen der Kinder beeinflusst und man diese messen kann. Dazu untersuchten wir den Lernerfolg der Kinder in einer *cross-situational*-Wortlern Situation. Allgemein konnten wir hier nachweisen, dass Kinder Wort Objekt Assoziationen lernen und wiedererkennen können. Das Einbeziehen der Veränderung der Pupillengröße zeigte Evidenz dafür, dass es einen Zusammenhang zwischen Zunahme der Pupillengröße während der Worterkennungsaufgabe und dem Erfolg in dieser Aufgabe gab. Dieses Ergebnis gab Grund zur Annahme, dass Kinder, die mehr in die Aufgabe involviert oder mehr von ihr begeistert waren, auch eine bessere Performanz zeigten.

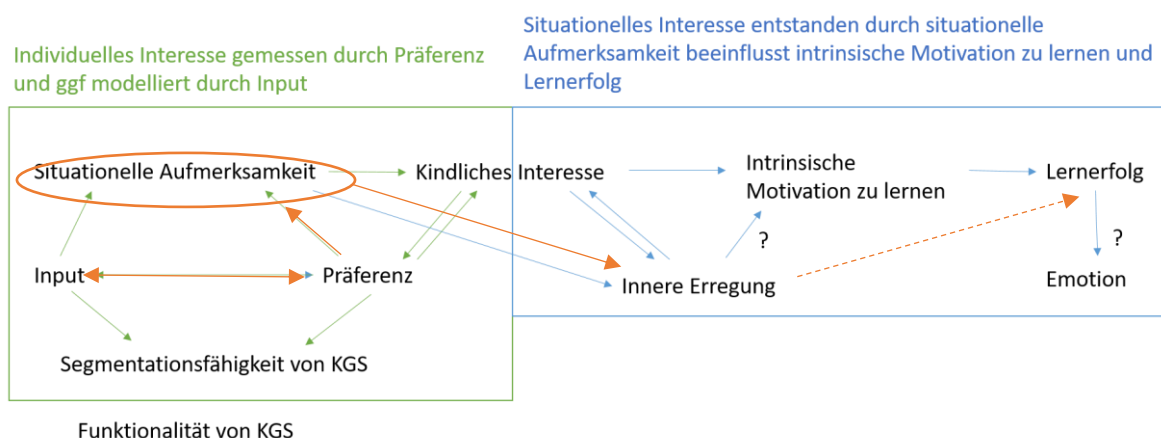
Zusammenfassend stellen diese Studien einen Zusammenhang zwischen kindlicher Aufmerksamkeit, extrinsischen Faktoren und der inneren Erregung der Kinder her.

Welche Rolle spielt die kindliche Aufmerksamkeit?

In allen Studien konnte gezeigt werden, dass die kindliche Aufmerksamkeit eine entscheidende Rolle beim Spracherwerb spielt. Wir fanden, dass Kinder der Blickbewegung einer Person besser folgen, wenn diese mit KGS kombiniert ist, als mit EGS. Des Weiteren zeigten wir, dass die Präferenz der Kinder mit dem elterlichen Input interagiert. Auch wenn aus der hier präsentierten

Studie nicht hervorgeht in welche Richtung diese Beziehung geht, also ob der Input die Präferenz moduliert oder vice versa, so zeigt sie doch, dass die kindliche Aufmerksamkeit, die einem Sprachstil entgegengebracht wird über die Entwicklung der Kinder hinweg variiert und mit dem Input interagiert. In der EEG Studie fanden wir eine erhöhte frontale Positivität bei 6-Monate alten Kindern in KGS-Blöcken im Vergleich zu 13-Monate alten Kindern. Darüber hinaus war bei 13-Monate alten Kindern eine negativere Reaktion auf KGS im Vergleich zu EGS insgesamt, also über beide Hemisphären hinweg, zu erkennen. All dies spricht für eine aufmerksamkeitsmodulierende Funktion von KGS. Diese Funktion von KGS konnte bereits in anderen Studien nachgewiesen werden (Kooijman et al., 2013; Senju & Csibra, 2008; Zangl & Mills, 2007). Wir konnten jedoch nicht nachweisen, dass die Aufmerksamkeit, die die Kinder der KGS entgegenbringen, ihnen dabei hilft, Wörter aus dem Sprachfluss zu segmentieren. Damit konform, konnten wir in der dritten Studien (Kapitel 4) eine Korrelation zwischen Erregung während der Aufgabe und der Performanz der Aufgabe nachweisen. Keine Evidenz fanden wir jedoch für die Vergrößerung der Pupillen während der Lernphase und deren Interaktion mit der Performanz in der Wiedererkennung von neuen Wort-Objekt Assoziationen. Diese hätte gezeigt, dass die Aufmerksamkeit der Kinder ihren Lernerfolg moduliert (Schema 4)

Schema 4: Verbindungen die durch die Resultate der Studien dieser Dissertation aufgezeigt wurden



Einordnung der Studien in Modelle

Im Folgenden möchte ich die Studien dieser Dissertation in die unterschiedlichen, oben beschriebenen Modelle einordnen.

Betrachtet man die Theorie von Csibra und Gergely (2009, 2011) in Bezug auf die hier aufgeführten Studien, so stellt sich zuerst die Frage nach der Rolle des Inputs. Die Natural Pedagogy Theorie fokussiert sich auf äußere Hinweise, die den Kindern dabei helfen sollen einfacher zu lernen. Als ein Beispiel wurde KGS angeführt. Wir konnten die Ergebnisse von (Senju & Csibra, 2008) replizieren und zeigen, dass Kinder besser der Blickbewegung nach KGS folgen, im Vergleich zu EGS. Betrachtet man die Funktion, die Csibra und Gergeley (2009) KGS zuschreiben, so wird deutlich, dass KGS in dieser Theorie dazu dient dem Kind zu verdeutlichen, dass die Kommunikation mit ihm stattfindet und sie dadurch ggf. etwas lernen können. Unsere Daten zeigen inline mit dieser Theorie, dass Kinder auf KGS anders reagieren als auf EGS. Die Aufmerksamkeit, die beiden Sprachstilen entgegengebracht wird, ist deutlich größer in KGS als in EGS. Allerdings konnten wir nicht nachweisen, ob diese Aufmerksamkeit durch die kindliche Präferenz oder den elterlichen Input entsteht.

Ordnen wir die Theorie in die neue Literatur zur aufmerksamkeitsmodulierenden Wirkung von KGS ein, so erkennt man, dass KGS den Kindern anzeigt, die bevorstehende Kommunikation ist für sie vorgesehen (Schreiner et al., under review; Soley & Sebastian-Galles, 2020). Nichtsdestotrotz, erweist es sich noch als schwierig, den von Csibra und Gergely hervorgesagte Verbesserung in der Lernfähigkeit der Kinder nachzuweisen, jedenfalls in einem frühen Alter. Es scheint in der hier vorgestellten EEG Studie eher der Fall zu sein, dass die aufmerksamkeitsfördernde Funktion von KGS das Lernen überschattet.

Die Ergebnisse der dritten Studie lassen sich hingegen nur schlecht mit der naturale pedagogy Theorie erklären. Die innere Erregung, die den Kindern beim Wiedergeben der Wort-Objekt Assoziationen hilft, ist egal in welchem möglichen Szenario nicht durch äußere Hinweise entstanden. Dies zeigt einen Punkt auf, der in dieser Theorie nicht weiter betrachtet wird: das Kind als aktiver Lerner, der selbst Entscheidungen trifft.

Dies bringt mich zu dem Coalition model. Das Coalition model beschreibt das Kind bereits als aktiven Teilnehmer in seinem eigenen Lernprozess, da es unterschiedliche Hinweise (äußere aber auch innere) gegeneinander abwägt und darauf basierend kalkuliert welchen es mehr Bedeutung schenkt um den maximalen Informationsgewinn zu erzeugen. Dieses Model erklärt die Ergebnisse der Studien insoweit, als dass Kinder eine Entwicklung vollziehen und dem Hinweis KGS eine unterschiedliche Bedeutung zuweisen. Dies zeigt sich durch die Präferenz mit 6 Monaten aber der Abwesenheit einer Präferenz für KGS über EGS mit 13 Monaten. Es zeigt sich jedoch auch, was ebenfalls durch die Theorie des Coalition Models unterstützt wird, dass Kinder den Hinweis der KGS weiterhin nutzen. Sie schenken ihm mehr Aufmerksamkeit, selbst im späteren Alter, wenn er mit einer Blickbewegung einhergeht. Die Korrelation zwischen Input und Präferenz der Kinder könnte für das Coalition Model komplizierter zu erklären sein. Falls der elterliche Input die Präferenz der Kinder moduliert, würde das Model die Daten erklären können. Die Kinder würden in diesem Fall ihren Lernprozess anpassen (angezeigt durch die Präferenz) und somit den Hinweis der Eltern optimal für sich nutzen, indem sie ihm mehr Aufmerksamkeit schenken. Die andere Variante, dass die Eltern ihren Input dem Kind gegenüber anpassen, fällt schwerer mit dem Coalition Model zu erklären. Die Anpassung der Hinweise wird in diesem Model nicht behandelt, zwar können die Kinder andere Hinweise mit einbeziehen (z.B. im höheren Alter ADS), auf eine Anpassung der Hinweise durch das Verhalten der Kinder geht diese Theorie jedoch nicht ein.

Die Einordnung der hier vorliegenden Ergebnisse in das Model von Oudeyer ist schwer. Unsere Studien zeigen keine Evidenz für eine intrinsische Motivation der Kinder, da wir keinerlei Interaktion zwischen der Lernphase und der Erregung der Kinder in dieser Phase, mit der Performanz in der Testphase finden. Ebenso konnten wir keine Evidenz für eine Veränderung der kindlichen Körperhaltung basierend auf dem Lernerfolg der Kinder finden. Die hier aufgeführte Dissertation hat somit keine Übertragungseffekte des Lernens gefunden. Eine Einordnung in eine motivationsbedingte Theorie fällt daher schwer.

Eine Theorie, die die Ergebnisse dieser Dissertation gut erklärt, ist die Theorie über Echtzeit Kommunikation von McMurray (2016). McMurray (2016) sieht zwei Mechanismen, die

den kindlichen Spracherwerb beeinflussen. Einerseits müssen Kinder die Verbindungen von Wörtern und Objekten, aber auch die Kategorisierung von Vokalen langsam und genau lernen, andererseits brauchen Kinder Techniken um in Echtzeit Probleme verarbeiten zu können und schnell auf neue Situationen zu reagieren, in denen z.B. ein neues Wort präsentiert wird. Um dies genauer zu erläutern, betrachtet McMurray (2016) zum Beispiel die Verwendung von KGS. Die Verwendung von KGS wird in der Forschung als Sprachvereinfachend angesehen (Kuhl et al., 1997). Es zeigte sich jedoch in zahlreichen Studien, dass der Input es nicht unbedingt erleichtert die Verteilung von Vokalen zu erlernen (z.B. McMurray et al., 2013). Vielmehr scheint KGS durch seine Eigenschaften, wie Sprachgeschwindigkeit, Prosodie und Emotion die Wahrnehmung zu beeinflussen. Genau diese Eigenschaften stehen in der Literatur auch dafür, die kindliche Aufmerksamkeit für diesen Sprachstil positiv zu beeinflussen (Golinkoff et al., 2015). McMurray (2016) sieht darin die Reaktion der Eltern auf Signale der Kinder. Somit bedienen sich Eltern KGS um die direkte oder Echtzeit Kommunikation zwischen sich und dem eigenen Kind zu vereinfachen, indem sie den verwendeten Sprachstil anpassen. Diese Anpassung geschieht mit dem Ziel die Aufmerksamkeit, die Erregung und die Emotionen der eigenen Kinder zu beeinflussen und dadurch die Kommunikation mit dem Kind zu vereinfachen. Somit entsteht der evolutionäre Druck KGS zu verwenden nicht durch die Notwendigkeit Laute in der eigenen Sprache zu diskriminieren, sondern durch die Notwendigkeit, die die Eltern darin sehen mit ihrem Kind in diesem Moment zu kommunizieren. KGS vereinfacht somit die Echtzeit-Kommunikation zwischen Eltern und Kindern. Genau diesen Aspekt der Modulierung der kindlichen Aufmerksamkeit konnten die hier präsentierten Studien zeigen. Wir fanden, bei der Untersuchung der KGS, dass diese die kindliche Aufmerksamkeit auf sich zieht, sowohl in einer funktionalen, nicht linguistischen Aufgabe, aber auch in einer linguistischen Aufgabe. Die Interaktion zwischen Input und kindlicher Präferenz spricht dabei für die von McMurray (2016) beschriebene Echtzeit Kommunikation. Die Kinder reagieren durch bevorzugtes Zuhören auf die Modulation des Elterlichen Inputs, oder die Eltern reagieren auf unterschiedliche Tonlagen auf die Kinder. Beide Interpretationen werden von der Theorie von McMurray getragen. Des Weiteren spricht für die Theorie der Echtzeit Kommunikation, dass in der EEG Studie eine allgemeine Reaktion auf KGS gezeigt werden

konnte. Als Reaktion wurde eine aufmerksamkeitsmodulierende Komponente sowohl bei den 6 als auch bei den 13-Monaten alten Kindern gezeigt. Interessant ist dabei, dass keine bessere Lernperformanz durch KGS gezeigt werden konnte. All diese Resultate sprechen für die aufmerksamkeitslenkende Funktion von KGS, die der Kommunikation der Mutter mit ihrem Kind dient.

Weiter beschreibt McMurray (2016), dass Kinder, um Wörter mit Objekten in Verbindung bringen können, diese langsam und allmählich lernen müssen. Einerseits, da unterschiedliche Dinge gleiche Namen haben können, andererseits können die Kinder nur durch die immer wiederkehrende Wiederholung von Objekt und Namen eine sichere Verbindung zwischen beiden aufbauen. Nichtsdestotrotz, können Kinder Wort-Objekt-Assoziationen auch durch *fast mapping* bilden. Fast mapping sieht McMurray (2016) nicht als Lernprozess an, sondern als Mittel um in Echtzeit kommunizieren zu können. Dabei nimmt er an, dass fast mapping erst durch die langsamen Lernprozesse erlernt werden kann. In Bezug auf die hier vorgestellten Studien, geht seine Annahme, dass die Echtzeit Verarbeitung von Wort-Objekt Situationen, dem Kind hilft um in diesem Moment zu handeln, aber keine langfristigen Lernerfolge mit sich bringt, wenn die Wort-Objekt-Assoziationen nicht über einen längeren Zeitraum wiederholt werden. Wir konnten in unserer Studien zeigen, dass die Performanz von Kindern im Test der neu erlernten Wort-Objekt-Assoziationen größer ist, wenn die Kinder mehr involviert/ engagiert in der Aufgabe sind. Allerdings konnten wir keinerlei Übertragungseffekte des Lernens finden, die sich entweder in einer Interaktion zwischen der Erregung in der Lernphase und der Performanz in der Testphase hätten widerspiegeln müssen, oder aber in der Veränderung der Körperhaltung der Kinder. Eine Veränderung der Körperhaltung hätte dafür gesprochen, dass die Kinder sich ihres Lernens und ihrer guten Performanz *bewusst* waren. Wir haben in dieser Studie allerdings eine weitere Unterstützung für die Theorie von McMurray (2016), da wir genau wie in den anderen zwei Studien dieser Dissertation vor allem Anhaltspunkte für Mechanismen für eine Echtzeit-Kommunikation zwischen Müttern und Kindern aufzeigen konnten. Diese Mechanismen für die kindliche und mütterliche Kommunikation sind, wie von McMurray (2016) beschrieben, die Anpassung der Mutter und die Reaktion des Kindes auf den verwendeten Sprachstil, sowie die

kindliche Fähigkeit neue Wort-Objekt-Assoziationen schnell zu bilden (fast mapping), die Kindern vor allem dann gelingt, wenn sie der Aufgabe mehr Aufmerksamkeit entgegen bringen. Wir konnten, genau wie McMurray (2016) es beschreibt, in dieser Dissertation nicht nachweisen, dass die Verwendung von KGS dazu dient das Lernen von Sprache zu vereinfachen, sondern zeigten vor allem die Reaktion der Kinder auf KGS, welche dafür spricht, dass KGS zur Echtzeit Kommunikation genutzt wird. Auch die Ergebnisse in Bezug auf die Bildung von Wort-Objekt Assoziationen sprechen für ein schnelles Erlernen eben dieser, um die Aufgabe zu lösen (und kommunizieren zu können) und nicht für ein langfristiges Lernen. Dies erklärt auch, warum wir keine Hinweise darauf finden konnten, dass Kinder sich ihrer guten Performanz in der Aufgabe bewusst waren, wie es in anderen Bereichen der Fall war (Hepach et al. 2012, 2017). Im Einklang mit McMurrays (2016) Annahme, KGS dient der Echtzeit Kommunikation konnten Studien zeigen, dass Kinder bereits ab einem frühen Alter (6 Monate) KGS mit einer Kind-gerichteten Kommunikation verbinden. Wohingegen sie EGS mit einer Kommunikation zwischen Erwachsenen assoziieren (Schreiner et al., under review). Dies spricht dafür, dass Kinder KGS früh als Hinweis wahrnehmen, der ihnen anzeigt, dass die folgende Kommunikation für sie bestimmt ist, wodurch sie womöglich mehr Aufmerksamkeit auf diesen Sprachstil lenken, um somit die Kommunikation zwischen den Eltern und sich selbst zu vereinfachen.

Zusammenfassung

Zusammenfassen lässt sich sagen, dass die hier aufgeführten Studien gezeigt haben, dass es einen Zusammenhang zwischen kindlicher Aufmerksamkeit und der Kommunikation zwischen Mutter und Kind im Allgemeinen gibt. Spezieller konnten wir zeigen, dass es eine Interaktion zwischen kindlicher Präferenz und elterlichem Input gibt, eine Reaktion auf KGS im EEG messbar ist, diese aber keine Hinweise darauf gibt, dass aus KGS besser segmentiert werden kann. Somit konnten wir vor allem die aufmerksamkeitsmodulierende Wirkung von KGS aufzeigen, jedoch keinen Zusammenhang zwischen kindlicher Aufmerksamkeit und dem kindlichen Lernerfolg in der frühen Entwicklung der Kinder zeigen. Auch in Bezug auf Wortlernen und das kindliche Bewusstsein über den eigenen Lernerfolg und die kindliche Motivation zu lernen, konnten wir

keine Hinweise auf eine Interaktion zwischen Lerneffekten und kindlicher Motivation finden, die für einen bewussten Lernprozess spricht. Insgesamt sprechen die hier präsentierten Daten für McMurrays (2016) Annahme, dass KGS und fast mapping vor allem dazu dient, die Echtzeit Kommunikation von Kinder und Eltern zu vereinfachen.

Referenzen

- Ackermann, L., Hepach, R., & Mani, N. (2020). Children learn words easier when they are interested in the category to which the word belongs. *Developmental Science*, 23(3).
<https://doi.org/10.1111/desc.12915>
- Ackermann, L., Lo, C. H., Mani, N., & Mayor, J. (2020). Word learning from a tablet app: Toddlers perform better in a passive context. *PLOS ONE*, 15(12), e0240519.
<https://doi.org/10.1371/journal.pone.0240519>
- Ahern, S., & Beatty, J. (1981). Physiological evidence that demand for processing capacity varies with intelligence. In *Intelligence and learning* (pp. 121-128). Springer, Boston, MA.
- Alexander, P. A., Pekrun, R., & Linnenbrink-Garcia, L. (2014). International handbook of emotions in education.
- Arias-Trejo, N., & Plunkett, K. (2010). The effects of perceptual similarity and category membership on early word-referent identification. *Journal of Experimental Child Psychology*, 105(1-2), 63–80. doi: 10.1016/j.jecp.2009.10.002
- Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, 128(2), 214–227. doi: 10.1016/j.cognition.2013.03.008
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., and Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception* 33, 717–746. doi: 10.1068/p5096
- Bahrack, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and Memory for Faces and Actions in Infancy: The Salience of Actions over Faces in Dynamic Events. *Child Development*, 75(6), 1629–1643.
- Baran, J. A., Laufer, M. Z., & Daniloff, R. (1977). Phonological contrastivity in conversation: A comparative study of voice onset time. *Journal of Phonetics*, 5(4), 339–350.
[https://doi.org/10.1016/S0095-4470\(19\)31204-5](https://doi.org/10.1016/S0095-4470(19)31204-5)

- Begus, K., Gliga, T., & Southgate, V. (2014). Infants Learn What They Want to Learn: Responding to Infant Pointing Leads to Superior Learning. *PLoS ONE*, *9*(10), e108817.
<https://doi.org/10.1371/journal.pone.0108817>
- Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function: Infant pointing serves an interrogative function. *Developmental Science*, *15*(5), 611–617.
<https://doi.org/10.1111/j.1467-7687.2012.01160.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>
- Berlyne, D. E. (1954). *A theory of human curiosity*. Nom de l'éditeur non disponible.
- Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language*, *11*(3), 557-578.
- Bijleveld, E., Custers, R., & Aarts, H. (2009). The unconscious eye opener: Pupil dilation reveals strategic recruitment of resources upon presentation of subliminal reward cues. *Psychological Science*, *20*(11), 1313-1315.
- Birch, S. A. J., Akmal, N., & Frampton, K. L. (2010). Two-year-olds are vigilant of others' non-verbal cues to credibility. *Developmental Science*, *13*(2), 363–369.
<https://doi.org/10.1111/j.1467-7687.2009.00906.x>
- Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
- Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer. Computer program, version 6.0. 46.
- Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and Me: Familiar Names Help Launch Babies Into Speech-Stream Segmentation. *Psychological Science*, *16*(4), 298–304. <https://doi.org/10.1111/j.0956-7976.2005.01531.x>
- Braarud, H. C., & Stormark, K. M. (2008). Prosodic modification and vocal adjustments in mothers' speech during face-to-face interaction with their two-to four-month-old infants: A double video study. *Social Development*, *17*(4), 1074-1084.

- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602-607.
- Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>
- Butler, S. C., O'Sullivan, L. P., Shah, B. L., & Berthier, N. E. (2014). Preference for infant-directed speech in preterm infants. *Infant Behavior and Development*, *37*(4), 505-511. <https://doi.org/10.1016/j.infbeh.2014.06.007>
- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, *63*(4), i. <https://doi.org/10.2307/1166214>
- Castro, R. M., Kalish, C., Nowak, R., Qian, R., Rogers, T., & Zhu, X. (2008). Human active learning. In *Advances in neural information processing systems*, 241-248.
- Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), 69-77. [https://doi.org/10.1016/s0010-0277\(02\)00233-0](https://doi.org/10.1016/s0010-0277(02)00233-0).
- Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants & apes; preference for motherese. *Infant Behavior and Development*, *20*(4), 477-488.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for Infant-Directed Speech in the First Month after Birth. *Child Development*, *61*(5), 1584. <https://doi.org/10.2307/1130766>
- Cooper, R. P. & Aslin, R. N. (1994). Developmental Differences in Infant Attention to the Spectral Properties of Infant-Directed Speech. *Child Development*, *65*(6), 1663-1677
- Coulson, M. (2004). Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *J. Nonverbal Behav.* *28*, 117–139. doi: 10.1023/B:JONB.0000023655.25550.be
- Cristia, A. (2010). Phonetic enhancement of sibilants in infant-directed speech. *The Journal of the Acoustical Society of America*, *128*, 424. <http://doi.org/10.1121/1.3436529>

- Cristia, A. (2013). Input to Language: The Phonetics and Perception of Infant-Directed Speech: The Phonetics and Perception of Infant-Directed Speech. *Language and Linguistics Compass*, 7(3), 157–170. <https://doi.org/10.1111/lnc3.12015>
- 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>
- Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1149–1157. <https://doi.org/10.1098/rstb.2010.0319>
- D'Entremont, B., Hains, S. M., & Muir, D. W. (1997). A demonstration of gaze following in 3-to 6-month-olds. *Infant Behavior and Development*, 20(4), 569-572.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208(4448), 1174-1176.
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in cognitive sciences*, 10(5), 204-211.
- Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences*, 21(8), 355–361. doi: 10.1016/s0166-2236(98)01263-6
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *science*, 298(5600), 2013-2015.
- D'Odorico, L., & Jacob, V. (2006). Prosodic and lexical aspects of maternal linguistic input to late-talking toddlers. *International Journal of Language & Communication Disorders*, 41(3), 293-311.
- Düzel, E., Penny, W. D., & Burgess, N. (2010). Brain oscillations and memory. *Current opinion in neurobiology*, 20(2), 143-149.

- Eiteljoerge SFV, Adam M, Elsner B, Mani N. 2019 Consistency of co-occurring actions influences young children's word learning. *R. Soc. open sci.* 6: 190097.
<http://dx.doi.org/10.1098/rsos.190097>
- Englund, K. T. (2005). Voice onset time in infant directed speech over the first six months. *First Language*, 25(2), 219–234. <https://doi.org/10.1177/0142723705050286>
- Farran, L. K., Lee, C. C., Yoo, H., & Oller, D. K. (2016). Cross-cultural register differences in infant-directed speech: an initial study. *PloS one*, 11(3), e0151518.
<https://doi.org/10.1371/journal.pone.0151518>
- Fastrich, G. M., Kerr, T., Castel, A. D., & Murayama, K. (2018). The role of interest in memory for trivia questions: An investigation with a large-scale database. *Motivation Science*, 4(3), 227–250. <https://doi.org/10.1037/mot0000087>
- Federmeier, K. D., & Kutas, M. (2002). Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, 40, 730-747.
- Federmeier, K. D., Mai, H., & Kutas, M. (2005). Both sides get the point: Hemispheric sensitivities to sentential constraint. *Memory & Cognition*, 33(5), 871-886.
- Fernald, A. (1989). Intonation and Communication Intent in Mothers' Speech to Infants: Is the Melody the Message? *Child Development*, 60(6), 1497-1510.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10(3), 279–293.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477–501.
<https://doi.org/10.1017/S0305000900010679>
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, 67(6), 3192-3218.

- Floccia, C., Keren-Portnoy, T., DePaolis, R., Duffy, H., Delle Luche, C., Durrant, S., White, L., Goslin, J., & Vihman, M. (2016). British English infants segment words only with exaggerated infant-directed speech stimuli. *Cognition, 148*, 1–9.
<https://doi.org/10.1016/j.cognition.2015.12.004>
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language, 44*(3), 677–694.
doi: 10.1017/s0305000916000209
- Gerken, L., Balcomb, F. K., & Minton, J. L. (2011). Infants avoid ‘labouring in vain’ by attending more to learnable than unlearnable linguistic patterns: Infants attend more to learnable patterns. *Developmental Science, 14*(5), 972–979. <https://doi.org/10.1111/j.1467-7687.2011.01046.x>
- Gingras, B., Marin, M. M., Puig-Waldmüller, E., & Fitch, W. T. (2015). The eye is listening: Music-induced arousal and individual differences predict pupillary responses. *Frontiers in human neuroscience, 9*, 619.
- Glenn, S. M., & Cunningham, C. C. (1983). What do babies listen to most? A developmental study of auditory preferences in nonhandicapped infants and infants with Down’s syndrome. *Developmental Psychology, 19*(3), 332.
- Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (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/10.1177/0963721415595345>
- 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.
- Goyet, L., de Schonen, S., & Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: An ERP study. *Brain Research, 1332*, 75–89.
<https://doi.org/10.1016/j.brainres.2010.03.047>

- Graf-Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy* 18(5), 797-824.
- Granholm, E. E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology*.
- Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: a critical test of natural pedagogy. *Child Development*, 89(6), 2091-2098.
- Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental psychology*, 24(1), 14.
- Gross, M. P., & Dobbins, I. G. (2020). Pupil dilation during memory encoding reflects time pressure rather than depth of processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Grossnickle, E. M. (2016). Disentangling curiosity: Dimensionality, definitions, and distinctions from interest in educational contexts. *Educational Psychology Review*, 28(1), 23-60.
- Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, 84(2), 486-496.
- Hare, T. A., O'doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of neuroscience*, 28(22), 5623-5630.
- Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research*, 44(6), 1189-1200.
- Henning, A., Striano, T., & Lieven, E. V. M. (2005). Maternal speech to infants at 1 and 3 months of age. *Infant Behavior and Development*, 28(4), 519-536.
<https://doi.org/10.1016/j.infbeh.2005.06.001>
- Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological science*, 23(9), 967-972.

- Hepach, R., Vaish, A., & Tomasello, M. (2013). A new look at children's prosocial motivation. *Infancy, 18*(1), 67-90.
- Hepach, R., Vaish, A., & Tomasello, M. (2015). Novel paradigms to measure variability of behavior in early childhood: Posture, gaze, and pupil dilation. *Frontiers in Psychology, 6*.
<https://doi.org/10.3389/fpsyg.2015.00858>
- Hepach, R., Vaish, A., & Tomasello, M. (2017). The fulfillment of others' needs elevates children's body posture. *Developmental Psychology, 53*(1), 100–113.
<https://doi.org/10.1037/dev0000173>
- Hernik, M., & Broesch, T. (2019). Infant gaze following depends on communicative signals: An eye-tracking study of 5-to 7-month-olds in Vanuatu. *Developmental Science, 22*(4), e12779.
- Hess, E. H. (1965). Attitude and pupil size. *Scientific american, 212*(4), 46-55
- Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143*(3611), 1190-1192.
- Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., Rocroi, C., & 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, 65*(3), i-vi+1-135.
- Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26*(5), 1570–1582. <https://doi.org/10.1037/0096-1523.26.5.1570>
- Hubley, P., & Trevarthen, C. (1979). Sharing a task in infancy. *New Directions for Child and Adolescent Development, 1979*(4), 57-80.
- James, W. (1899). On a certain blindness in human beings. *Pragmatism and other writings, 269*.
- Junge, C., & Cutler, A. (2014). Early Word Recognition and Later Language Skills. *Brain Sciences, 4*(4), 532–559. <https://doi.org/10.3390/brainsci4040532>

- Junge, C., Cutler, A., & Hagoort, P. (2012). Electrophysiological evidence of early word learning. *Neuropsychologia*, *50*(14), <https://doi.org/10.1016/j.neuropsychologia.2012.10.012>.
- Junge, C., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development: Rapid recognition. *Developmental Science*, *15*(4), 463–473. <https://doi.org/10.1111/j.1467-7687.2012.1144.x>
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive psychology*, *29*(1), 1–23.
- Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 56-88.
- Kahneman, D. (1973). *Attention and effort* (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583-1585
- Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The Wick in the Candle of Learning: Epistemic Curiosity Activates Reward Circuitry and Enhances Memory. *Psychological Science*, *20*(8), 963–973. <https://doi.org/10.1111/j.1467-9280.2009.02402.x>
- Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology: General*, *146*(4), 569.
- Kaplan, P. S., Bachorowski, J. A., Smoski, M. J., & Hudenko, W. J. (2002). Infants of depressed mothers, although competent learners, fail to acquire associations in response to their own mothers' IDS. *Psychological Science*, *13*, 268-271.
- Kaplan, P. S., Goldstein, M. H., Huckleby, E. R., & Cooper, R. P. (1995a). Habituation, sensitization, and infants' responses to motherse speech. *Developmental Psychobiology*, *28*(1), 45–57.

- Kaplan, P. S., Goldstein, M. H., Huckleby, E. R., Owren, M. J., & Cooper, R. P. (1995b). Dishabituation of visual attention by infant-versus adult-directed speech: Effects of frequency modulation and spectral composition. *Infant Behavior and Development*, 18 (2), 209-223.
- Kaye, K. (1980). Why we don't talk baby talk to babies. *Journal of Child Language*, 7, 489-507.
- Kaye, K. (1980). Why we don't talk 'baby talk' to babies. *Journal of Child Language*, 7(3), 489-507. <https://doi.org/10.1017/S0305000900002804>
- Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, 16(1), 55-68. <https://doi.org/10.1017/S030500090001343X>
- Kidd, C., & Hayden, B. Y. (2015). The Psychology and Neuroscience of Curiosity. *Neuron*, 88(3), 449-460. <https://doi.org/10.1016/j.neuron.2015.09.010>
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLoS ONE*, 7(5), e36399. <https://doi.org/10.1371/journal.pone.0036399>
- Kidd, E., Junge, C., Spokes, T., Morrison, L., & Cutler, A. (2018). Individual Differences in Infant Speech Segmentation: Achieving the Lexical Shift. *Infancy*, 23(6), 770-794. <https://doi.org/10.1111/infa.12256>
- Kinney, D. K., & Kagan, J. (1976). Infant attention to auditory discrepancy. *Child development*, 155-164.
- Klatt, D. & Stevens, K. (1979). On the automatic recognition of continuous speech: Implications from a spectrogram-reading experiment. *IEEE Transactions on Audio and Electroacoustics*, 21(3), 210-217, doi: 10.1109/TAU.1973.1162453
- Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage*, 12(1), 20-27.
- Koenig, M. A., & Harris, P. L. (2005). Preschoolers Mistrust Ignorant and Inaccurate Speakers. *Child Development*, 76(6), 1261-1277. <https://doi.org/10.1111/j.1467-8624.2005.00849.x>

- Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, 24(1), 109–116. <https://doi.org/10.1016/j.cogbrainres.2004.12.009>
- Kooijman, V., Junge, C., Johnson, E. K., Hagoort, P., & Cutler, A. (2013). Predictive Brain Signals of Linguistic Development. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00025>
- Krinsky, R., & Nelson, T. O. (1981). Task difficulty and pupillary dilation during incidental learning. *Journal of Experimental Psychology: Human Learning and Memory*, 7(4), 293–298. <https://doi.org/10.1037/0278-7393.7.4.293>
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843. <https://doi.org/10.1038/nrn1533>
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684–686.
- Kuhl, Patricia. K., Andruski, J. E., Chistovich, I. A., Christovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science*, 277(5326), 684–686. <https://doi.org/10.1126/science.277.5326.684>
- Laeng, B., & Falkenberg, L. (2007). Women's pupillary responses to sexually significant others during the hormonal cycle. *Hormones and behavior*, 52(4), 520-530.
- Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious?. *Perspectives on psychological science*, 7(1), 18-27.
- Lee, K., Eskritt, M., Symons, L. A., & Muir, D. (1989). Children's Use of Triadic Eye Gaze Information for „Mind Reading“. *Developmental Psychology*, 34(3), 525–539.
- Lempers, Jacques D. (1979). Young children's production and comprehension of nonverbal deictic behaviors. *The journal of Genetic Psychology*, 135(1), 93–102.
- Lewis, M. (1997). The self in self-conscious emotions. *Ann. N. Y. Acad. Sci.* 818, 119–142. doi: 10.1111/j.1749-6632.1997.tb48251.x

- Lewis, M., Alessandri, S. M., and Sullivan, M. W. (1992). Differences in shame and pride as a function of children's gender and task difficulty. *Child Dev.* 63, 630–638. doi: 10.2307/1131351
- Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, 46(5), 703-713.
- Lisman, J., Grace, A. A., & Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in neurosciences*, 34(10), 536-547.
- Loewenfeld, I. E. (1993). The pupil. *Anatomy, physiology, and clinical applications*.
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1), 75–98
- Lucca, K., & Wilbourn, M. P. (2018). Communicating to Learn: Infants' Pointing Gestures Result in Optimal Learning. *Child Development*, 89(3), 941–960.
<https://doi.org/10.1111/cdev.12707>
- Lucca, K., & Wilbourn, M. P. (2019). The what and the how: Information-seeking pointing gestures facilitate learning labels and functions. *Journal of Experimental Child Psychology*, 178, 417–436. <https://doi.org/10.1016/j.jecp.2018.08.003>
- Lyakso, E. E., Frolova, O. V., & Grigorev, A. S. (2014). Infant Vocalizations at the First Year of Life Predict Speech Development at 2-7 Years: Longitudinal Study. *Psychology*, 5(12), 1433.
- Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word Learning in Infant- and Adult-Directed Speech. *Language Learning and Development*, 7(3), 185–201.
<https://doi.org/10.1080/15475441.2011.579839>
- Malsheen, B. J. (1980). Two hypotheses for phonetic clarification in the speech of mothers to children. *Child Phonology*, 2, 173-184.
- Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants's recognition of the sound pattern of their own namees. *Psychological Science*, 6(5), 314–317.

- Mangardich, H., & Sabbagh, M. A. (2018). Children remember words from ignorant speakers but do not attach meaning: evidence from event-related potentials. *Developmental science*, *21*(2), e12544.
- Mani, N., & Ackermann, L. (2018). Why Do Children Learn the Words They Do? *Child Development Perspectives*, *12*(4), 253–257. <https://doi.org/10.1111/cdep.12295>
- Mani, N., Johnson, E., McQueen, J. M., & Huettig, F. (2013). How yellow is your banana? toddlers' language-mediated visual search in referent-present tasks. *Developmental Psychology*, *49*(6), 1036–1044. doi: 10.1037/a0029382
- Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental science*, *11*(1), 53-59.
- Mani, N., & Plunkett, K. (2010). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological Science*, *21*(7), 908–913. doi: 10.1177/0956797610373371
- Mani, N., & Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, *121*(2), 196–206. doi: 10.1016/j.cognition.2011.06.013
- Mani, N., & Pätzold, W. (2016) Sixteen-Month-Old Infants' Segment Words from Infant- and Adult-Directed Speech, *Language Learning and Development*, *12*(4), 499-508, DOI: 10.1080/15475441.2016.1171717
- Männel, C., & Friederici, A. D. (2011). Intonational phrase structure processing and syntactic knowledge in childhood: ERP studies in 2-, 3-, and 6-year-old children. *Developmental Science*, *14*(4).
- Männel, C., & Friederici, A. D. (2013). Accentuate or repeat? Brain signatures of developmental periods in infant word recognition. *Cortex*, *49*(10), 2788–2798. <https://doi.org/10.1016/j.cortex.2013.09.003>
- Männel, C., & Friederici, A. D. (2016). Neural correlates of prosodic boundary perception in German preschoolers: If pause is present, pitch can go. *Brain research*, *1632*, 27-33.

- ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science*, 3(1), 24-52.
- 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)
- ManyBabies Consortium (2020). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science*. 3(1), 24-52.
- 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.
- 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.
- Martens, J. P., and Tracy, J. L. (2013). The emotional origins of a social learning bias does the pride expression cue copying? *Soc. Psychol. Pers. Sci.* 4, 492–499.
doi: 10.1177/1948550612457958
- McClure, S. M., York, M. K., & Montague, P. R. (2004). The neural substrates of reward processing in humans: the modern role of FMRI. *The Neuroscientist*, 10(3), 260-268.
- McMurray, B. (2016). Language at Three Timescales: The Role of Real-Time Processes in Language Development and Evolution. *Topics in Cognitive Science*, 8(2), 393–407.
<https://doi.org/10.1111/tops.12201>
- McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence?. *Cognition*, 129(2), 362-378.
- 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.
- Meints, K., & Woodford, A. (2008). Lincoln infant lab package 1.0: A new programme package for ipl, preferential listening, habituation and eyetracking [Computer software manual].
- Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant behavior and development*, 16(4), 495-500.

- Murray, A. D., Johnson, J., & Peters, J. (1990). Fine-tuning of utterance length to preverbal infants: Effects on later language development. *Journal of Child Language*, *17*, 511–525.
- Nagai, Y., Nakatani, A., Qin, S., Fukuyama, H., Myowa-Yamakoshi, M., & Asada, M. (2012, November). Co-development of information transfer within and between infant and caregiver. In *2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)* (pp. 1-6). IEEE.
- Newman, R., & Hussain, S. (2006). Changes in preference for infant-directed speech in low and moderate noise by 4.5-to 13-month-olds. *Infancy*, *10*(1), 61-76.
- Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and noneffects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children*, 109–149. Cambridge, UK: Cambridge University Press
- Niwano, K., & Sugai, K. (2002a). Intonation contour of Japanese maternal infant-directed speech and infant vocal response. *The Japanese Journal of Special Education*, *39*(6), 59-68.
- Niwano, K., & Sugai, K. (2002b). Acoustic determinants eliciting Japanese infants' vocal response to maternal speech. *Psychological Reports*, *90*(1), 83-90.
<http://dx.doi.org/10.2466/PRO.90.1.83-90>
- Naoi, N., Minagawa-Kawai, Y., Kobayashi, A., Takeuchi, K., Nakamura, K., Yamamoto, J. I., & Shozo, K. (2012). Cerebral responses to infant-directed speech and the effect of talker familiarity. *Neuroimage*, *59*(2), 1735-1744.
- Oudeyer, Pierre-Yves, & Smith, L. B. (2016). How Evolution May Work Through Curiosity-Driven Developmental Process. *Topics in Cognitive Science*, *8*(2), 492–502.
<https://doi.org/10.1111/tops.12196>
- Oudeyer, P.-Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning. In *Progress in Brain Research* (Bd. 229, S. 257–284). Elsevier.
<https://doi.org/10.1016/bs.pbr.2016.05.005>
- Outters, V., Schreiner, M. S., Behne, T., & Mani, N. (2020). Maternal input and infants' response to infant-directed speech. *Infancy*, *25*(4), 478-499.

- Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International journal of human-computer studies*, 59(1-2), 185-198.
- Partridge, E., McGovern, M. G., Yung, A., & Kidd, C. (2015). *Young Children's Self-Directed Information Gathering on Touchscreens*. 6.
- Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development*, 15(3), 325-345.
- Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Pena, M., Maki, A., Kovačić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, 100(20), 11702-11705.
- Piaget, J., & Cook, M. T. (1952). The origins of intelligence in children.
- Rheingold, H. L., & Adams, J. L. (1980). The significance of speech to newborns. *Developmental Psychology*, 16(5), 397–403.
- Ripollés, P., Ferreri, L., Mas-Herrero, E., Alicart, H., Gómez-Andrés, A., Marco-Pallares, J., ... & Rodríguez-Fornells, A. (2018). Intrinsically regulated learning is modulated by synaptic dopamine signaling. *Elife*, 7, e38113.
- Ripollés, P., Marco-Pallares, J., Alicart, H., Tempelmann, C., Rodríguez-Fornells, A., & Noesselt, T. (2016). Intrinsic monitoring of learning success facilitates memory encoding via the activation of the SN/VTA-Hippocampal loop. *Elife*, 5, e17441.
- Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H.-J., Rodríguez-Fornells, A., & Noesselt, T. (2014). The Role of Reward in Word Learning and Its Implications for Language Acquisition. *Current Biology*, 24(21), 2606–2611. <https://doi.org/10.1016/j.cub.2014.09.044>

- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental science*, 8(2), 162-172.
- Roberts, S., Fyfield, R., Baibazarova, E., van Goozen, S., Culling, J. F., & Hay, D. F. (2013). Parental speech at 6 months predicts joint attention at 12 months. *Infancy*, 18(S1), 1-15.
- Rondeel, E., Van Steenbergen, H., Holland, R., & van Knippenberg, A. (2015). A closer look at cognitive control: Differences in resource allocation during updating, inhibition and switching as revealed by pupillometry. *Frontiers in human neuroscience*, 9, 494.
- Ruggeri, A., Markant, D., Gureckis, T., & Xu, F. (2016, January). Active control of study leads to improved recognition memory in children. In *Proceedings of the 38th annual conference of the Cognitive Science Society*. Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, 67.
- Sachs, J., & Devin, J. (1976). Young children's use of age-appropriate speech styles in social interaction and roleplaying. *Journal of Child Language*, 3, 81–98
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Salmela-Aro, K., Moeller, J., Schneider, B., Spicer, J., & Lavonen, J. (2016). Integrating the light and dark sides of student engagement using person-oriented and situation-specific approaches. *Learning and Instruction*, 43, 61–70. <https://doi.org/10.1016/j.learninstruc.2016.01.001>
- Samuelson, L. K. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15-20-month-olds. *Developmental psychology*, 38(6), 1016.
- Schachner, A., & Hannon, E. E. (2011). Infant-directed speech drives social preferences in 5-month-old infants. *Developmental Psychology*, 47(1), 19-25.
- Schieffelin, B.B., & Ochs, E. (1983). A cultural perspective on the transition from prelinguistic to linguistic communication. In R.M. Golinkoff (Ed.), *The transition from prelinguistic to linguistic communication*, 115–131. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Schreiner, M. S., Altvater-Mackensen, N., & Mani, N. (2016). Early Word Segmentation in Naturalistic Environments: Limited Effects of Speech Register. *Infancy*, *21*(5), 625–647. <https://doi.org/10.1111/infa.12133>
- Schreiner, M. S., Eiteljoerge, S. F. V., Behne, T., Mani, N. (in prep). The quality and quantity of infant-directed speech across children's development.
- Schreiner, M. S., & Mani, N. (2017). Listen up! Developmental differences in the impact of IDS on speech segmentation. *Cognition*, *160*, 98–102. <https://doi.org/10.1016/j.cognition.2016.12.003>
- Schreiner, M., S., Outters., V., Palm., R., & Mani, N. (under review). Infant associations of infant- and adult-directed speech
- Segal J., & Newman, R. S. (2015). Infant preferences for structural and prosodic properties of infant-directed speech in the second year of life. *Infancy*, *20*(3), 339–351. <https://doi.org/10.1111/infa.12077>
- 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>
- Sherrod, K., Friedman, S., Crawley, S., Drake, D., & Devieux, J. (1977). Maternal language to prelinguistic infants: Syntactic aspects. *Child Development*, *48*(4), 1662-1665. doi:10.2307/1128531
- Shatz, M., & Gelman, R. (1973). The development of communication skills: Modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development*, *38*(5), 1–38
- Shiota, M. N., Campos, B., and Keltner, D. (2003). The faces of positive emotion. *Acad. Sci. 1000*, 296–299. doi: 10.1196/annals.1280.029
- Shum, H. P., Ho, E. S., Jiang, Y., and Takagi, S. (2013). Real-time posture reconstruction for Microsoft Kinect. *IEEE Trans. Cybern.* *43*, 1357–1369. doi: 10.1109/TCYB.2013.2275945

- Sim, Z. L., Tanner, M., Alpert, N. Y., & Xu, F. (2015). Children Learn Better When They Select Their Own Data. *Cog.Sci.*
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, 3(3), 365–394.
- Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, 51(2), 173-189.
- Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, 14(6), 654–666.
- Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, 100(4), 765–781. <https://doi.org/10.1037/a0012840>
- Smith, L. B. (2005). Shape: A Developmental Product. In *Language and Space Workshop: Defining Functional and Spatial Features, Jun, 2001, University of Notre Dame, South Bend, IN, US; This volume is based on edited papers presented at the aforementioned conference..* Oxford University Press
- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13(4), 410-420.
- Smith, K., Smith, A. D. M., & Blythe, R. A. (2011). Cross-Situational Learning: An Experimental Study of Word-Learning Mechanisms. *Cognitive Science*, 35(3), 480–498. <https://doi.org/10.1111/j.1551-6709.2010.01158.x>
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. <https://doi.org/10.1016/j.cognition.2007.06.010>
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27(4), 501–532. <https://doi.org/10.1016/j.dr.2007.06.002>

- Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, *35*(04), 869–902.
- Soley, G., & Sebastian-Galles, N. (2020). Infants' expectations about the recipients of infant-directed and adult-directed speech. *Cognition*, *198*, 104214.
- Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, *44*, 1-18.
- Stern, D., Spieker, S., Barnett, R., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, *10*(1), 1-15.
doi:10.1017/S0305000900005092
- Sundberg, U., & Lacerda, F. (1999). Voice Onset Time in Speech to Infants and Adults. *Phonetica*, *56*(3–4), 186–199. <https://doi.org/10.1159/000028450>
- Syal, S., & Finlay, B. L. (2011). Thinking outside the cortex: Social motivation in the evolution and development of language. *Developmental science*, *14*(2), 417-430.
- Taxitari, L. (2019, December). Early grammatical development in Cyprus: A CDI study. In *International Symposium on Monolingual and Bilingual Speech 2019* (p. 136).
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5
- Tomasello, M. (2003). The key is social cognition. *Language in mind: Advances in the study of language and thought*, 47-57.
- Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., Weiskrantz, L., & de Gelder, B. (2009). Unseen facial and bodily expressions trigger fast emotional reactions. *Proceedings of the National Academy of Sciences*, *106*(42), 17661-17666.
- Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*(4), 432-444.

- Twomey, K. E., Ma, L., & Westermann, G. (2018). All the Right Noises: Background Variability Helps Early Word Learning. *Cognitive Science*, 42, 413–438.
<https://doi.org/10.1111/cogs.12539>
- Twomey, K. E., & Westermann, G. (2018). Curiosity-based learning in infants: A neurocomputational approach. *Developmental Science*, 21(4), e12629.
<https://doi.org/10.1111/desc.12629>
- van Heugten, M., & Johnson, E. K. (2012). Infants Exposed to Fluent Natural Speech Succeed at Cross-Gender Word Recognition. *Journal of Speech, Language, and Hearing Research*, 55(2), 554–560. [https://doi.org/10.1044/1092-4388\(2011/10-0347\)](https://doi.org/10.1044/1092-4388(2011/10-0347))
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental science*, 7(3), 270-276.
- 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.
- Vosoughi, S., Roy, B., Frank, M., Roy, D. (2010). Contributions of prosodic and distributional features of caregivers' speech in early word learning. *Proceedings of the 5th International Conference on Speech Prosody*. Chicago, IL.
- Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, 23(3), 34-41.
- Weisfeld, G. E., and Beresford, J. M. (1982). Erectness of posture as an indicator of dominance or success in humans. *Motiv. Emot.*, 6, 113–131.
[doi:10.1007/BF00992459](https://doi.org/10.1007/BF00992459)
- Weiskrantz, L. (1990). The Ferrier Lecture, 1989-Outlooks for blindsight: explicit methodologies for implicit processes. *Proceedings of the Royal Society of London. B. Biological Sciences*, 239(1296), 247-278.
- Weiskrantz, L. (1998). Consciousness and commentaries. *International Journal of Psychology*, 33(3), 227-233.

- Weppelman, T. L., Bostow, A., Schiffer, R., Elbert-Perez, E., & Newman, R. S. (2003). Children's use of the prosodic characteristics of infant-directed speech. *Language & Communication*, 23(1), 63–80. [https://doi.org/10.1016/S0271-5309\(01\)00023-4](https://doi.org/10.1016/S0271-5309(01)00023-4)
- Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 43(2), 230-246.
- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, 103(1), 147–162. <https://doi.org/10.1016/j.cognition.2006.03.006>
- Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child development*, 349-355.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, 7(1), 49-63.
- Westermann, G., & Mani, N. (Eds.). (2017). *Early word learning*. Routledge.
- Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H. J., & Düzel, E. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, 45(3), 459-467.
- Wlotko, E. W. & Federmeier, K. D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, 45, 3001-3014
- Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, 6(3), 297-311.
- Yu, C., & Smith, L. B. (2007). Rapid Word Learning Under Uncertainty via Cross-Situational Statistics. *Psychological Science*, 18(5), 414–420. <https://doi.org/10.1111/j.1467-9280.2007.01915.x>
- Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental science*, 14(2), 165-180.

Zangl, R., & Mills, D. L. (2007). Increased Brain Activity to Infant-Directed Speech in 6- and 13-Month-Old Infants. *Infancy*, *11*(1), 31–62. https://doi.org/10.1207/s15327078in1101_2

Zettersten, M., & Saffran, J. R. (2019). Sampling to learn words: Adults and children sample words that reduce referential ambiguity. *Developmental Science*, e13064.

Erklärung

Hiermit erkläre ich, dass alle Teile der Dissertation unabhängig voneinander verfasst wurden. Die Unterstützung Dritter wurde nur akzeptiert, wenn dies im Hinblick auf die Prüfungsordnung wissenschaftlich vertretbar und akzeptabel war und alle Quellen angegeben wurden passend. Teile dieser Arbeit und einige Abbildungen wurden im Folgenden verwendet Artikel.

Outters, V., Schreiner, M. S., Behne, T., & Mani, N. (2020). Maternal input and infants' response to infant-directed speech. *Infancy*, 25(4), 478-499.

Vivien Outters,

Göttingen, 2021

Appendix

Studie 1.

Appendix 1 Stimuli recorded for preferential listening task

Short stories taken from: <http://www.kidsnet.at/Deutsch/erlebnisbaum.htm>

1. Mein Bruder und ich waren alleine zu Hause. Nach dem Abendessen las ich noch ein wenig und legte mich ins Bett. Draußen regnete es stark und der Wind pfiff um das Haus. Ein wenig ängstlich war mir schon zu Mute. Alles war so still irgendwie furchterregend. Thomas mein Bruder arbeitete ganz oben in seinem Zimmer am Computer. Bevor ich das Licht ausmachte, schaltete (End of IDS stimuli) ich die Alarmanlage ein. Bald schlief ich tief und fest. Kurz danach machte auch Thomas das Licht aus und schlief (End of ADS stimuli) auch ein.
2. Gestern war ich alleine zu Hause. Alle waren in ein Konzert gegangen nur ich hatte keine Karte mehr bekommen. Nach dem Essen guckte ich fern und las in einem Buch. Eine Stunde später legte ich mich ins Bett und versuchte zu schlafen, aber es gelang mir nicht. Inzwischen war ein Gewitter aufgezogen und es war schrecklich dunkel. In der Diele knarrte es, Türen quietschten und ein (End of IDS stimuli) Fenster schlug auf und zu. Plötzlich wurde die Haustüre geöffnet und ich vernahm leise Schritte. Was wäre, wenn da ein Einbrecher ist (End of ADS stimuli).
3. Letzten Sommer beschlossen meine Freunde aus der Wohnstraße und ich ein Baumhaus zu bauen. In der Nähe unserer Straße entdeckten wir einen passenden alten Baum. Wir brauchten nicht einmal eine Woche und schon war das Baumhaus fertig. Jeden Tag nach den Hausaufgabe trafen wir uns dort zum Spielen. Eines Tages passierte etwas Schreckliches. Ich war alleine beim Baumhaus (End of IDS stimuli). Da stolperte ich über einen naheliegenden Stacheldraht, verlor das Gleichgewicht und versuchte mich

mit der rechten Hand am Boden abzustützen. „Au“, schrie ich (End of ADS stimuli).

4. Im August machten meine Eltern und ich eine Urlaubsrundfahrt. Am ersten Tag fuhren wir ins Burgenland, wo wir eine große Burg besichtigten. An der Kasse kauften wir eine Familienkarte. Mit ihr bekamen wir ein Prospekt überreicht und konnten auch an einer Führung teilnehmen. Zuerst marschierten wir also über den Innenhof in den Rittersaal, der noch gut erhalten war (End of IDS stimuli). Der Führer erzählte uns schmunzelnd, dass die Stühle deshalb eine hohe Lehne hatten, damit die Ritter nach einer fröhlichen Feier nicht nach hinten kippen konnten. Weiter ging es dann (End of ADS stimuli) in die ehemaligen Schafräume, die Kemenaten.
5. An einem herrlichen, wunderschönen Sommertag waren meine Familie und ich bei Freunden zu Besuch. Wir hatten viel Spaß mit unserer Freundin Mareike und die Zeit verging wie im Flug. Nach dem Abendessen überkam uns aber ein wenig die Langeweile. Da hatte Michaela eine tolle Idee. Sie erzählte meinem Bruder und mir von einem alten (End of IDS stimuli) verfallenen Haus ganz in der Nähe. Obwohl es schon dunkel war, machten wir uns mit Taschenlampen auf den Weg. Als wir das Gebäude erreichten, überkam uns aber ein ungutes Gefühl. Das Haus sah richtig (End of ADS stimuli) gruselig aus.
6. Vor gar nicht allzu langer Zeit, kurz bevor ich einschlief, hörte ich ein lautes Krachen. Erschreckt setzte ich mich im Bett auf, meine Schwester schlief noch ganz fest. Da hörte ich es das Klirren nochmals. Ängstlich suchte ich die Anderen. Gemeinsam schauten wir was geschehen war. Im Bad war alles in Ordnung. Langsam (End of IDS stimuli) gingen wir die Treppe hinunter. Auch im Erdgeschoß war alles ruhig. Trotzdem waren

meine Hände leicht feucht und ich war froh, dass ich nicht alleine war
(End of ADS stimuli).

Studie 1: Appendix 2. Acoustic characteristic of IDS and ADS stimuli in preferential listening task and gaze following

A2.1 Acoustic characteristics of IDS and ADS stimuli in preferential listening task.

| Preferential listening task | | |
|-----------------------------|--------|--------|
| | IDS | ADS |
| max pitch (Hz) | 404.81 | 268.60 |
| min pitch (Hz) | 139.77 | 102.38 |
| mean pitch (Hz) | 241.66 | 195.31 |
| pitch range (st) | 18.73 | 18.32 |
| duration (s) | 2.92 | 3.30 |

A2.2 Acoustic characteristics of IDS and ADS stimuli in gaze following task.

| Gaze following task (adapted German version and original English version) | | | | |
|---|------------|------------|-------------|-------------|
| | German IDS | German ADS | English IDS | English ADS |
| max pitch (Hz) | 494.12 | 291.56 | 538.56 | 272.14 |
| min pitch (Hz) | 256.21 | 123.54 | 259.82 | 181.72 |
| mean pitch (Hz) | 401.72 | 248.63 | 434.76 | 220.27 |
| pitch range (st) | 11.37 | 14.86 | 12.61 | 6.98 |

Studie 2: Appendix 3.

| Predictors | value | | |
|-------------------------------------|-----------|--------------|----------------|
| | Estimates | CI | p |
| (Intercept) | 5.27 | 3.66 – 6.87 | < 0.001 |
| condition | -0.41 | -1.80 – 0.99 | 0.568 |
| age | 1.30 | -0.31 – 2.90 | 0.114 |
| hemisphere | 0.06 | -0.61 – 0.73 | 0.864 |
| register | 0.82 | -0.95 – 2.59 | 0.363 |
| condition age | 0.23 | -1.16 – 1.62 | 0.746 |
| condition hemisphere | -0.02 | -0.69 – 0.65 | 0.955 |
| age : hemisphere | 0.54 | -0.13 – 1.21 | 0.114 |
| condition : register | -0.47 | -2.03 – 1.09 | 0.553 |
| age : register | -0.39 | -2.16 – 1.38 | 0.667 |
| Hemisphere : register | -0.65 | -1.31 – 0.02 | 0.058 |
| condistion : age : hemisphere | 0.19 | -0.48 – 0.85 | 0.584 |
| Cond : age : register | -0.31 | -1.87 – 1.25 | 0.697 |
| condition : hemisphere: register | -0.17 | -0.84 – 0.50 | 0.621 |

| | | | |
|---|-------|--------------|-------|
| age : hemisphere : register | 0.47 | -0.20 – 1.14 | 0.169 |
| condition : age: hemisphere : register | -0.23 | -0.90 – 0.44 | 0.502 |

Study 2: Appendix 4 Main model

| Predictors | value | | |
|---------------------------------|-----------|---------------|------------------|
| | Estimates | CI | p |
| (Intercept) | 4.89 | 3.20 – 6.58 | <0.001 |
| condition | -0.62 | -1.85 – 0.61 | 0.325 |
| age | 1.36 | -0.32 – 3.05 | 0.113 |
| hemisphere | -0.10 | -0.82 – 0.62 | 0.792 |
| register | 1.05 | -0.29 – 2.39 | 0.124 |
| condition : age | 0.46 | -0.77 – 1.69 | 0.465 |
| condition : hemisphere | 0.03 | -0.69 – 0.75 | 0.928 |
| age : hemisphere | 0.61 | -0.11 – 1.33 | 0.095 |
| condition : register | -0.26 | -1.80 – 1.29 | 0.744 |
| age : register | -1.72 | -3.06 – -0.37 | 0.012 |
| hemisphere : register | -0.72 | -1.43 – 0.00 | 0.051 |
| condition : age : hemisphere | 0.10 | -0.62 – 0.82 | 0.785 |

| | | | |
|--------------------------|-------|--------------|--------------|
| condition : age : | -1.01 | -2.55 – 0.53 | 0.200 |
| register | | | |
| Condition : hemisphere : | -0.40 | -1.12 – 0.32 | 0.280 |
| register | | | |
| age : hemisphere : | 0.74 | 0.02 – 1.46 | 0.045 |
| register | | | |
| condition : age : | -0.03 | -0.75 – 0.69 | 0.945 |
| hemisphere : register | | | |

Study 2 Appendix 5 focusing on IDS

| Predictors | value | | |
|------------------------|-----------|--------------|------------------|
| | Estimates | CI | p |
| (Intercept) | 3.84 | 1.84 – 5.84 | <0.001 |
| condition | -0.36 | -2.21 – 1.50 | 0.706 |
| age | 3.08 | 1.09 – 5.08 | 0.002 |
| hemisphere | 0.62 | -0.41 – 1.64 | 0.237 |
| condition : age | 1.47 | -0.39 – 3.33 | 0.121 |
| condition : hemisphere | 0.43 | -0.60 – 1.45 | 0.412 |
| age : hemisphere | -0.12 | -1.15 – 0.90 | 0.813 |
| condition : age : | 0.13 | -0.90 – 1.15 | 0.810 |
| hemisphere | | | |

