

Simon Pröll / Jan-Claas Freienstein / Oliver Ernst

Exemplarbasierte Annäherungen an das Silbengelenk

Abstract

Using elicited experimental data from 35 kindergartners, this article investigates the ability of children aged 3-6 to deal with German ambisyllabic consonants. In this age group – yet without structured tuition in reading and writing, the development of the phonologic system still in progress – both intra- and inter-subject variation appears to be greater than previous studies conducted on school children revealed. It is argued that this variability supports the view of an exemplar based theory on language acquisition; stable analytic knowledge on this specific syllable structure emerges as a function of age and frequency. Simultaneously, the data suggest a second kind of underlying variation, stemming from the contact of two conflicting typological syllable systems.

1. Einleitung
2. Daten und Verfahren
 - 2.1 Vorgängerstudien
 - 2.2 Technik und Versuchsaufbau
 - 2.3 Lexemauswahl
3. Auswertung
 - 3.1 Variablenanalyse
 - 3.2 Kindzentriert
 - 3.3 Silbentypologische Interpretation
 - 3.4 Exemplarbasierte Interpretation
4. Zusammenschau
5. Literatur

1. Einleitung

Ziel dieses Beitrags ist es, empirisch zu ermitteln, wo – und wie – Kinder explizit Silbengrenzen im sogenannten „Silbengelenk“ ziehen, wenn die entsprechende Artikulation exponiert wird. Als „Silbengelenk“ (oder „ambisilbische Konsonanten“¹, vgl. Hall 2000, 263) werden einzelne Konsonanten bezeichnet, die offenbar zu zwei Silben gleichzeitig gehören: In *Mutter* [mʊtɐ] fungiert der Dentalplosiv diesem Modell nach sowohl als Koda der ersten als auch als Onset der zweiten Silbe. Schulkinder lernen im Zuge des Schriftspracherwerbs, diese doppelt geschriebenen Konsonanten durch „Klatschen“ auf zwei Silben „aufzutrennen“. Die Frage ist hier nicht nur, ob die Silbierfähigkeit von Kindern genuin gegeben ist (Kinder also überhaupt bewusst auf Silben zugreifen können), sondern auch, wie Kinder das Problem, nur einen Konsonanten in zwei Silben zu realisieren, lösen. Ausgehend von der Prämisse, dass der

¹ Ambisilbische Konsonanten sind nicht identisch mit Geminaten. Sie haben weder die entsprechende Länge noch zwangsläufig die Etymologie, auch wenn sich laut Ramers (1992, 273) „ambisilbische Konsonanten in phonologischen Prozessen“ aber „wie Geminaten“ verhalten, während Geminaten vermutlich gleichzeitig „immer ambisilbisch“ seien (vgl. auch Hulst 1985 mit der Überlegung, dass Varietäten genau deshalb nie Geminaten und ambisilbische Konsonanten zugleich haben).

Erwerbsprozess zum Zeitpunkt des Schuleintritts bereits bedeutende Umstrukturierungen erfahren hat – und dass die Kenntnis des Schriftsystems zu veränderten Antwortmustern führt –, wurden für die vorliegende Studie experimentelle Daten mit Kindergartenkindern erhoben (Abschnitt 2). Die Auswertung (Abschnitt 3) zeigt zunächst, dass sich bei der Realisierung von Konsonanten und Affrikaten stärkere interindividuelle Unterschiede als in späteren Altersgruppen nachweisen lassen (Abschnitte 3.1 und 3.2). Dies liefert Evidenzen für zwei unterschiedliche Arten von Variation: Zum einen findet der Erwerb offenbar in einer Konkurrenzsituation zweier unterschiedlicher typologischer Systeme statt (Abschnitt 3.3); zum anderen zeigt die spezifisch altersabhängige Variation unter den Kindern Muster auf, die für ein exemplarbasiertes Modell des Phonologieerwerbs der Erstsprache sprechen (Abschnitt 3.4).

2. Daten und Verfahren

2.1 Vorgängerstudien

Es existieren einige wenige empirische Vorarbeiten zur Realisierung von Silbengelenken bei Kindern, die zum einen zu unterschiedlichen Ergebnissen (und in der Folge Interpretationen) kommen und zum anderen unserer Ansicht nach teilweise methodische Schwächen aufweisen bzw. für den nötigen Interpretationsumfang die Rohdaten nicht klar genug offenlegen. Wir skizzieren die relevanten Vorgängerstudien kurz.

Risel (1999) erfragt Daten von 42 Drittklässlern und 28 Viertklässlern und kommt zu folgendem generellen Eindruck, der sich für „[m]ehrere“ (Risel 1999, 77) Kinder ergibt: „Von intuitivem Verfügen über die Einheit Silbe konnte in den vorliegenden Fällen nicht gesprochen werden.“ Interessant ist die (dort eher beiläufig erwähnte) Beobachtung durch Risel (1999, 77), dass vier der 42 Drittklässler auch bei der Trennung von *Birne* (einem Kontrolllexem) den Silbenübergang überbrücken („*Birn-ne*“).² Durch Risel (2002, 77) ist festgehalten, es sei im Vergleich von Zweit-, Dritt- und Viertklässlern „eine Tendenz zur orthographie-nahen Artikulation festzustellen, dies lässt sich unschwer als Ergebnis von Unterricht interpretieren. Ein Beispiel ist *Tasse*. In Klasse 2 trennt kein Kind [tas.sə], in 3 sind es 36 %, in 4 schon 57 %.“

Huneke (2000) erhebt Daten von 38 Kindern im Frühjahr vor dem Schuleintritt. Die Wahrscheinlichkeit, dass die Kinder zu diesem Zeitpunkt bereits mit Sprachspielen im Kindergarten in Berührung gekommen sein könnten, wird nicht berücksichtigt. Als Methode wird angeführt, dass die Kinder gebeten wurden, die zu erhebenden Lexeme „silbisch (‘in Robotersprache’ oder ‘wie beim Auszählen’) zu sprechen“ (Huneke 2000, 7). Ob die Roboteranalogie hier nicht zu unerwünschten Nebenerscheinungen führt, kann von außen schwer abgeschätzt werden.³ Die Analyse wird nicht für einzelne Konsonantenphoneme aufgeschlüsselt; Huneke (2000) liefert die Rohdaten nicht mit, eine leicht erweiterte Fassung des Textes führt sie jedoch im Anhang auf (Huneke 2002, 100–102). Aus den nicht unbedingt einheitlichen Resultaten folgert Huneke (2000, 14), dass Grundschul Kinder einen „intuitiven Zugang zum Silbengelenk“ hätten, sie könnten „beim Syllabieren die Silbengelenke in zwei Komponenten aufspalten“. Aus einer kritischen Re-Analyse derselben Daten schlussfolgert Hanke (2002, 66) dagegen, dass es sich im Gegenteil eher zeigen würde, dass die dort untersuchten Kinder „gesprochene Sprache auf unterschiedliche Weise (d. h. silbisch, phonetisch, phonologisch,

² Stefan Kleiner verdanken wir den nützlichen Hinweis, dass dies eventuell auf /t/-Vokalisierung zurückzuführen sein könnte.

³ So ähnlich kritisch auch Risel (2002, 78–79) in der Nachlese einer entsprechenden studentischen Arbeit, in der „Robotersprache mehr Gelenkartikulation brachte“: „Die Gefahr von *Methodenartefakten* ist in diesem Bereich offenbar recht groß [Hervorhebung im Original]“.

morphemorientiert, ...)“ gliederten. Sie sieht weiterhin einen Zusammenhang zwischen Schrifterfahrung und der Anwendung analytischer Verfahren. In einer ebenfalls durch Hanke (2002) umgesetzten Vergleichsstudie mit 33 frisch eingeschulter Kindern überwiegen dagegen klar offene Silben.

Die Studien postulieren also unterschiedlich klare Ergebnisse, die sich dazu noch widersprechen. Das sollte stutzig machen. Ziel ist es im Folgenden nicht nur, die Methodik der Studien zu verbessern bzw. die Reproduzierbarkeit ihrer Ergebnisse zu testen, sondern auch, mit jüngeren Kindern zu arbeiten. Das geschieht aus zwei Gründen: Zum einen gilt es, den möglichen Einfluss des Schriftspracherwerbs auszuschließen; zum anderen spricht einiges dafür, dass sich um den Eintritt ins schulfähige Alter systematische Änderungen im Spracherwerb abspielen, was die Repräsentation und den Abruf phonetisch-phonologischer Information angeht.

Dass durchweg alle vorangegangenen Studien auf Kinder kurz vor oder bereits nach der Einschulung als Versuchspersonen zurückgreifen, ist aus dem Erklärungsinteresse der Studien heraus bindend: Im Fokus steht jeweils (sei es aus deskriptiver oder aus didaktischer Perspektive) der Schriftspracherwerb des Kindes. Hier stehen aber der Phonologieerwerb sowie die Silbentypologie des gesprochenen Deutschen im Zentrum. Wir wählen daher Kindergartenkinder im Alter von 3 bis 6 Jahren als Probanden. Ziel eines (verbesserten) experimentellen Vorgehens ist es,

- a) die Kinder zu einer möglichst langen (Schall-)Silbenübergangsphase zu bringen, dabei aber
- b) die Aufmerksamkeit so weit von der Sprachproduktion wegzubewegen, dass die Proben so natürlich wie möglich bleiben.

Klar muss aber sein, dass es keine realweltliche Kommunikationssituation gibt, in der die Silbengrenze derart exponiert artikuliert wird. Auch wenn durch Aufmerksamkeitsverlagerung versucht wird, die „Natürlichkeit“ der Artikulation hoch zu halten, hat man es bei den resultierenden Daten letztlich mehr mit einer Art „Belastungstest“ der Silbenstruktur als mit natürlicher Sprache zu tun.

2.2 Technik und Versuchsaufbau

Die Datenerhebung fand in einem Augsburger Kindergarten statt. Die Kinder wurden einzeln in einem ihnen bekannten Raum des Kindergartens an einen Tisch geführt, an dem sie einer der beiden Versuchsleiterinnen⁴ gegenüber saßen. Dort sollten sie Handschuhe anziehen, um das Klatschgeräusch auf der Aufnahme zu dämpfen und die Aufmerksamkeit des Kindes von der Sprachproduktion wegzulenken (s.o.). Anschließend wurde ihnen erklärt, dass sie an einem Spiel teilnehmen dürfen, bei dem es darum geht, den Gegenstand zu erraten, dessen Bild die Versuchsleiterin hochhält. Die Kinder sollten die Antwort dann laut rhythmisierend aussprechen und dazu mit den Handschuhen klatschen. Die Aufnahmen wurden mit einem portablen digitalen Audiorecorder (in 44,1 kHz/24bit) angefertigt.

2.3 Lexemauswahl

Das Beispiellexem, anhand dessen den Kindern das Prinzip erklärt und vorgemacht wurde, war *Banane*. Es wurde bewusst ein dreisilbiges Beispielwort mit offenen Tonsilben ausgewählt, um eine reine Imitationsleistung der Kinder identifizieren zu können. Die nicht mehr durch die Versuchsleiterin vorgeführten, sondern durch die Kinder aktiv anhand von Bildkarten erkannten Lexeme decken die fünf Kategorien Plosive ([p], [t], [k], [b], [d], [g]), Nasale

⁴ Die Autoren danken herzlich Katharina Reiser und Sara Huber, die sich zur Durchführung der Erhebung bereit erklärt haben.

([n], [m], [ŋ]), Liquide ([l])⁵, Frikative ([f], [s], [ʃ], [χ]) und Affrikaten ([ts], [pf], [ks])⁶ ab. Es wurden Lexeme gewählt, die ikonisch einfach dargestellt werden konnten und bei denen gewährleistet war, dass sie Kindern dieser Altersklasse bekannt sein können (vgl. Abbildung 1).

Kategorie	Laut	Lexem
Plosive	[p]	Treppe
	[t]	Schlitten
	[k]	Schnecke
	[b]	Robbe
	[d]	Teddy
	[g]	Bagger
Nasale	[n]	Sonne
	[m]	Hammer
	[ŋ]	Finger
Liquide	[l]	Teller
Frikative	[f]	Affe
	[s]	Messer
	[ʃ]	Flasche
	[χ]	Knochen
Affrikaten	[ts]	Katze
	[pf]	Apfel
	[ks]	Hexe

Abbildung 1: Laute und Lexeme.

Die Lexeme sind so parallel wie möglich gehalten: Alle sind zweisilbig und initialbetont, der potentiell ambisilbische Laut folgt als einzelner Konsonant auf akzentuierten Kurzvokal. In einigen Fällen benannten die Kinder das Bild des Teddys mit *Teddybär* (15 mal), das der Robbe mit *Seerobbe* (8 mal) und das der Katze mit *Miezekatze* (3 mal).

Die fünf Kategorien (Plosive, Nasale, Liquide, Frikative, Affrikaten) wurden randomisiert zu einer Reihe gefügt und diese Reihe dann gespiegelt (wobei die Spiegelung nicht vollständig sein konnte, da mehrere Kategorien eine ungerade Anzahl von Elementen haben).⁷ Um Positionseffekte (vgl. Huber 2005, 171–181; im vorliegenden Kontext sind vor allem Übungseffekte relevant) dokumentieren zu können, wurden die 37 Kinder in vier Experimentalgruppen aufgeteilt, für die jeweils per unvollständigem Ausbalancieren andere Reihenfolgen der Kategorien konstruiert wurden. Die konkrete Realisierung durch einzelne Lexeme wurde weiter randomisiert (vgl. Abbildung 2 für eine Aufstellung der Reihenfolge der Lexeme in den einzelnen Reihen).

⁵ [r] wurde ausgeklammert, da Kinder dieser Altersklasse /r/ individuell sehr verschieden realisieren, was die Zuordnung der erwerbssprachlichen Approximationen zu einer Lautkategorie schwierig gemacht hätte; zusätzlich ist die Realisierung von /r/ im Großraum Augsburg momentan stark im Wandel. Insgesamt war daher nicht mit sauber auswertbaren Ergebnissen zu rechnen.

⁶ Ob [ks] als Affrikate des Deutschen anzusehen ist, ist umstritten (vgl. zur Diskussion etwa Wiese 1988, 62; Prinz / Wiese 1991; Altmann / Ziegenhain 2007, 150; Ramers 2001, 26–27). Solange diese Frage offen ist, erscheint es uns jedoch naheliegend, in Datenerhebungen wie dieser vorerst eher auf weitere als auf restriktivere Definitionen zurückzugreifen. Andere Lautungen (wie z.B. [ps]), die im Fall von weiten Definitionskriterien als Affrikaten gesehen werden könnten, kommen nicht an Silbengrenzen nominaler zweisilbiger Konkreta des Erbwortschatzes vor und sind daher im hier gewählten Kontext nicht erfragbar.

⁷ Da sowohl die gesamte Zahl als auch die der Nasale ungerade ist, bildet die Kategorie Nasal jeweils den Spiegelpunkt.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
R ₁	Treppe	Affe	Schlitten	Sonne	Katze	Hexe	Schnecke	Messer	Hammer	Flasche	Robbe	Teller	Apfel	Finger	Teddy	Knochen	Bagger
R ₂	Hammer	Bagger	Flasche	Teddy	Teller	Robbe	Apfel	Knochen	Finger	Messer	Katze	Schnecke	Hexe	Schlitten	Affe	Treppe	Sonne
R ₃	Messer	Finger	Hexe	Robbe	Teddy	Knochen	Katze	Treppe	Sonne	Schnecke	Apfel	Flasche	Schlitten	Bagger	Teller	Hammer	Affe
R ₄	Katze	Teller	Schnecke	Flasche	Bagger	Finger	Treppe	Affe	Hammer	Knochen	Schlitten	Sonne	Teddy	Messer	Robbe	Hexe	Apfel

Abbildung 2: Untersuchungsreihen.

In manchen Fällen liegen mehrere Realisierungen für ein Lexem vor. Dies wurde durch Wunsch der Versuchsleiterin nach Wiederholung initiiert, meist um einer besseren Verständlichkeit auf der Aufnahme willen, wenn es zu Störgeräuschen kam oder sich die Gesprächspassagen von Kind und Versuchsleiterin überlappen. Allerdings gibt es auch Fälle, in denen die Kinder in ihrer Realisierung erkennbar unsicher waren – auch hier wurde meist ein weiterer Versuch unternommen. Falls die einzelnen Antworten auf denselben Stimulus wechseln, wird grundsätzlich die erste Realisierung gewertet. Markante Wechselbewegungen werden im Laufe der Analysen mit angeführt.

Ein Kind war den gesamten Versuchsablauf hindurch stumm; ein anderes war nur zum regulären, linearen Sprechen bereit. Insgesamt sind 35 Datensätze in die detailliertere Auswertung eingegangen.

3. Auswertung

Alle Aufnahmen wurden mittels Praat (Boersma / Weenink 2014) akustisch-phonetisch analysiert. Da es Ziel der Untersuchung war, die Art des Übergangs eines vokalischen Segments in ein nachfolgendes konsonantisches im Bereich einer Silbengrenze zu erfassen, wurde durch die Erhebungssituation gezielt angestrebt, gerade diesen temporalen Aspekt der Artikulation künstlich zu verlängern und zu exponieren. Was wir direkt beobachten können, ist also nicht das Silbengelenk selbst, sondern die Strategie zur Realisierung eines überlangen Silbenübergangs.

3.1 Variablenanalyse

Zunächst sollen die Daten rein deskriptiv aus einem lexem- bzw. lautbasierten Blickwinkel heraus ausgewertet werden. Abbildung 3 zeigt die Auszählung aller Daten zu kontinuierlich bildbaren Konsonanten im „Silbengelenk“.

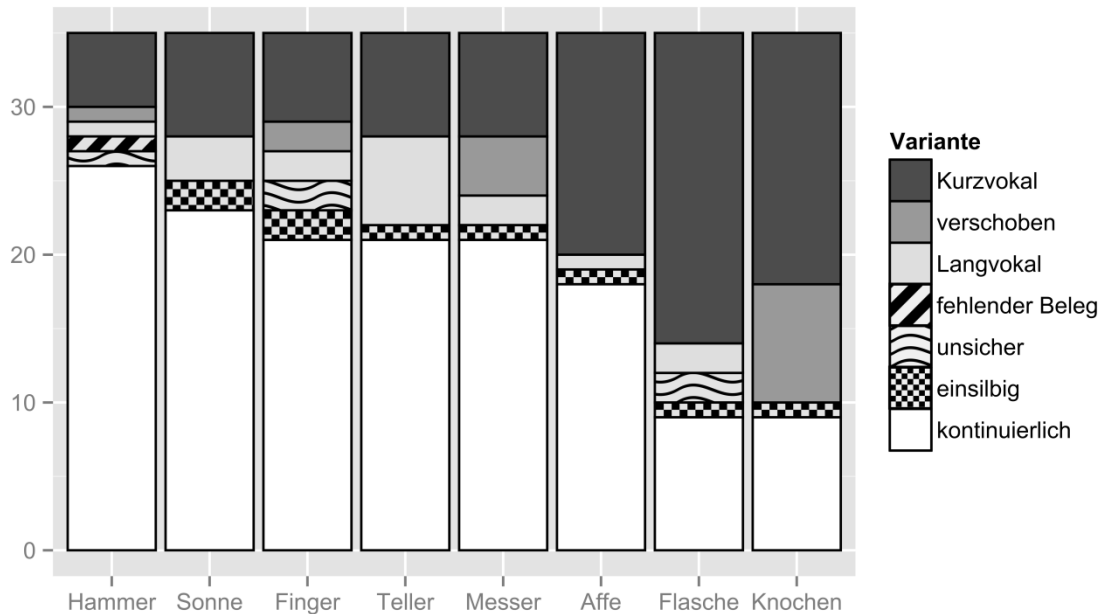


Abbildung 3: Realisierungen kontinuierlich bildbarer Laute im Silbengelenk.

Die weißen Balken zeigen die jeweilige Anzahl der Kinder an, die den entsprechenden Laut über die Silbengrenze hinweg artikulieren (Typ „kontinuierlich“), also sowohl in der Koda der ersten Silbe als auch im Onset der zweiten. Für den größten Teil dieser Belege gilt, dass die Artikulation dabei niemals abbricht. Das steht im Gegensatz zum Postulat Hunekes (2000, 9), nach dem „die Kinder das Silbengelenk beim silbischen Sprechen in zwei Komponenten zerlegen. Dabei kommt es im Silbengelenk zu zwei Lautstärkemaxima, es wird eine Sprechpause eingeschoben oder es tritt eine Verzögerung relativ zur Sprechgeschwindigkeit auf.“ Eine Unterbrechung des Schallstroms tritt nur in 12 Instanzen (das entspricht 8,1 % der Gelenkbelege) auf. Die anderen kontinuierlichen Belege entsprechen dem Muster, das in Abbildung 4 beispielhaft (*Sonne*, Kind 33) erfasst ist: Der Nasal ertönt durchgängig.

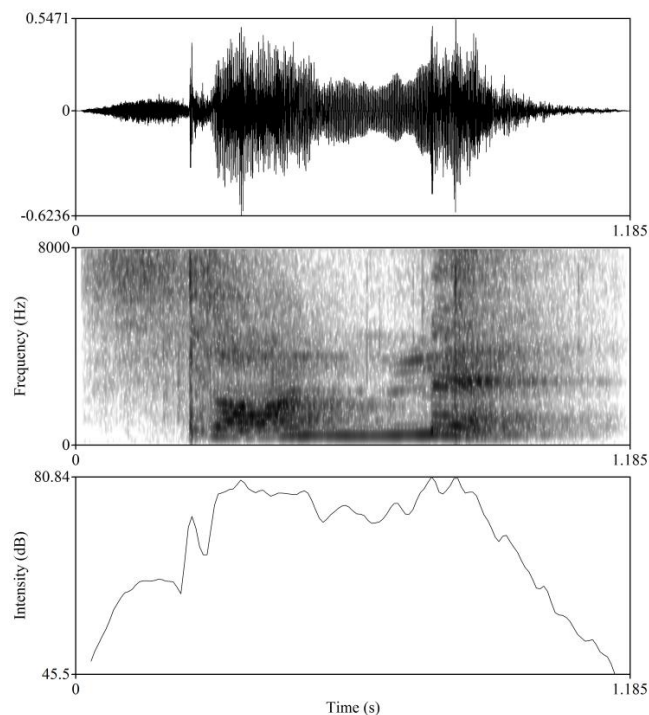


Abbildung 4: Oszillogramm, Spektrogramm und Intensitätsverlauf von *Sonne*, Kind 33.

Definierendes Kriterium der zweithäufigsten Variante („Kurzvokal“, dunkelgraue Balken) ist, dass die Artikulation nach einem kurzen Vokal in der ersten Silbe abbricht; der Konsonant eröffnet dann die zweite Silbe (wie in [tɛ.lɐ]). Realisierungen, in denen der fragliche Laut komplett der ersten Silbe zugeschlagen wird, sind als „verschoben“ gekennzeichnet (Hanke 2002, 66 wertet diesen Belegtypus als „morphemorientiert“). Sie treten mit acht Instanzen am häufigsten bei *Knochen* auf (die zweite Silbe beschränkt sich dann auf den Nasal), fast so häufig wie die 9 Fälle, in denen /χ/ in beiden Silben realisiert wird.⁸ Davon unterschieden sind die deutlich selteneren Fälle, in denen die erste Silbe offen auf Langvokal endet ([fla:ʃə], Typ „Langvokal“) oder in denen das Wort klar einsilbig artikuliert wurde. Einige wenige Realisierungen sind (von Seiten der Kinder) sehr zaghaft und lassen sich nicht sinnvoll einem Typ zuordnen („unsicher“).

Weitere Auffälligkeiten im Einzelnen: Bei *Sonne* wechseln zwei Kinder auf die Aufforderung zur Wiederholung von kontinuierlichen Realisierungen zu offenen Silben (kurz), zwei nach einsilbiger Realisierung zur kontinuierlichen Realisierung (in die Grafik gingen immer die Erstmeldungen ein). Ein Kind beginnt die zweite Silbe von *Finger* mit einem hörbaren Lösen des nasalen Verschlusses, eines segmentiert [fɪŋ.ɡʊ], also mit einem abgesetzten Velarplosiv als Onset der zweiten Silbe. Bei zwei Kindern kommt die Artikulation an der Silbengrenze zum Stillstand; Silbe eins lautet mit [ɪ] aus, Silbe zwei setzt mit einem [ɪ] im Anlaut ein. Die einsilbige Realisierung von *Affe* wird bei Bitte zur Wiederholung kontinuierlich artikuliert.

In Abbildung 5 sind die Resultate für Laute im „Silbengelenk“, die nicht kontinuierlich artikulierbar sind (wofür im Deutschen nur pulmonale Plosive in Frage kommen), dargestellt. Hier überwiegen klar die Belege, in denen die erste Silbe auf Kurzvokal endet. Mit „Stimmton“ sind diejenigen Fälle gekennzeichnet, in denen die Silbengrenze mit dem Stimmtone des Plosivs überbrückt wird.

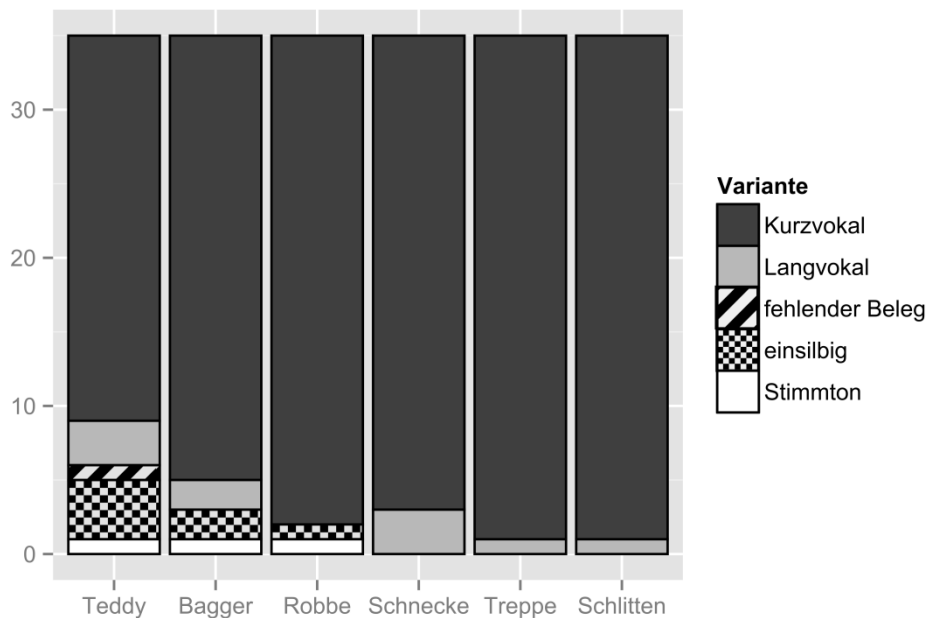


Abbildung 5: Realisierungen von Plosiven im Silbengelenk.

Abbildung 6 visualisiert die Ergebnisse zu den Lexemen mit Affrikate in Gelenkposition. Auch hier überwiegt die Kategorie „Kurzvokal“ ganz deutlich; es zeigt sich dazu eine weitere

⁸ In zwei davon variiert jedoch die phonetische Gestalt der Realisierung: Während der Auslaut von Silbe eins mit [χ] besetzt ist, beginnt der Anlaut von Silbe zwei mit einem [h] (Beispiel, Kind 19, 6 Jahre: [knoχ.hɛn]).

Option: Das plosivische Element bzw. die ganze Affrikate schließt die erste Silbe, das frikativische Element tritt aber auch an den Onset der zweiten Silbe, also [k^hat'.sə], wobei [t'] auditorisch und messphonetisch der ersten Silbe zuzuweisen ist, weil im Onset der zweiten Silbe die Plosion nicht nachweisbar ist, oder [həkʰsə], mit der Silbengrenze „im“ [s]. Auch diese Fälle werten wir als „kontinuierlich“.

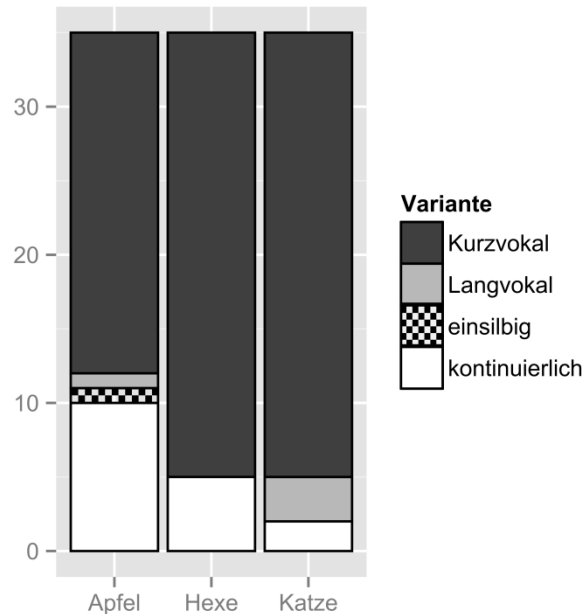


Abbildung 6: Realisierungen von Affrikaten im Silbengelenk.

Aus einem lexemzentrierten Blickwinkel zeigt sich also klar, dass zum Teil deutliche Unterschiede in der Realisierungspraxis bestehen. Die Lexeme, die einen stimmhaften Laut mit (relativ) großer Schallfülle in der potentiellen Gelenkposition aufweisen, bieten sich für die Kinder scheinbar am deutlichsten für eine kontinuierliche Realisierung an.

Dazu unterstreichen diese Resultate, was eigentlich längst selbstverständlich sein sollte: Die immer wieder anzutreffende Annahme, dass der im Gelenk stehende Laut in der Überlautung quasi „verdoppelt“ werden könnte (etwa [ʃlit'.t^hən]) – ein Umstand, den man schließlich „hören könne“ –, ist naiv und offensichtlich eine Idee, die mit Ágel (2003) eindeutig als *skriptizistisch* angesehen werden muss: Über die (implizite) Teilidentifikation sprachsystematischer Phänomene mit ihrer Schreibung (und Ignoranz bzw. Fehlinterpretation ihres Ursprungs) hat die angebliche „Doppelung“ ihren Irrweg in manche Teile der Didaktik/Pädagogik angetreten. In keiner der hier ausgewerteten 304 Realisierungen von Plosiven oder Affrikaten lässt sich eine Verdoppelung eines Plosivs feststellen. Bei Kindern, die nicht entsprechend instruiert worden sind, ist diese Interpretation intuitiv nicht angelegt.

3.2 Kindzentriert

Knapp ein Drittel der Kinder haben (bezogen auf die Laute, die dies artikulatorisch ermöglichen) ein weitgehend durchgängig artikuliertes Silbengelenk. Ein Kind (Kind35, m, 4 Jahre) zeigt nicht nur ein klares, eingipfliges Silbengelenk in allen kontinuierlich bildbaren Fällen (*Affe, Sonne, Messer, Hammer, Flasche, Teller, Finger, Knochen*), sondern überbrückt auch bei den Lenisplosiven (*Robbe, Teddy, Bagger*) die Silbengrenze mit dem Stimmtton (negative voice onset time). Es zeigt sich aber auch, dass bei einer großen Anzahl an Belegen das Silbengelenk nicht durch einen kontinuierlichen Laut überbrückt wird – die Kinder realisieren also offensichtlich einen Silbentyp, den es aus typologischer Perspektive im Standard-

deutschen nicht geben dürfte (vgl. Wiese 2000, 45–47; 277–280): offene Silben, die auf ungespannten Kurzvokal enden.

Andere Belege sind deutliche Reflexe der Unsicherheit der Kinder: Kind04 (m, 5 Jahre) generalisiert das dreisilbige Muster des Beispielworts *Banane* auf 14 der 17 zweisilbigen Untersuchungslexeme, indem es – trotz späterer Aufforderungen durch die Versuchsleiterin, nur zweimal zu klatschen – eine dritte Silbe des Typs [hə] oder [hä] anhängt. Kind05 (w, 4 Jahre) wechselt mit dem elften Lexem *Robbe* auf einsilbige Interpretationen und kehrt nur kurz bei *Teddybär* zur Mehrsilbigkeit zurück. Für Kind13 (w, 3 Jahre) ist nach drei Lexemen, die zweisilbig realisiert werden, ebenfalls *Teddybär* der Umschlagpunkt: Ab dort findet es nicht mehr in das zweisilbige Schema zurück und bildet fortan sehr zögerlich unter Verdopplung und Verdreifachung von CV-Silben Ketten wie [ǰnə.hə.hə.xən]. Keines der Lexeme zeigt einen Hinweis auf Ambisilbizität.

Worauf ist die Varianz der Daten insgesamt, die Unsicherheit einzelner Kinder sowie der Unterschied zu den Ergebnissen älterer Studien (vgl. Abschnitt 2.1) zurückzuführen? Die Vermutung liegt nahe, dass hier zwei Variationsszenarien gleichzeitig zu berücksichtigen sind: Zum einen gilt es, die typologischen Implikationen der Befunde theoretisch einzuordnen und zu hinterfragen. Zum anderen decken die hier erhobenen Daten aber auch einen Altersabschnitt der Kinder ab, in dem noch verstärkt Erwerbsprozesse ablaufen: Erwartungsgemäß sind z.B. segmentalphonetisch erwerbssprachliche Realisierungen und Approximationen gängig; so ist etwa /s/ in *Messer* bei 7 von 35 Kindern näher an [θ] realisiert.⁹

Die nachfolgenden beiden Abschnitte beschäftigen sich jeweils einzeln mit den beiden angenommenen Variationsachsen in den Daten. Abschnitt 3.3 widmet sich zunächst einer typologischen Interpretation der Ergebnisse, bevor in Abschnitt 3.4 die erwerbsorientierte Perspektive näher beleuchtet wird.

3.3 Silbentypologische Interpretation

Theoretischer Dreh- und Angelpunkt der oben durchgeführten Analysen ist, was unter „Silbe“ verstanden wird – es gibt jedoch für die Fragestellung dieses Beitrags (mindestens) zwei relevante Möglichkeiten, wie eine Silbe definiert werden kann: Zum einen in Gestalt der „Schallsilbe“, zum anderen als „Drucksilbe“.¹⁰ Abbildung 7 (basierend auf Maas 2002, 16) illustriert den Unterschied mittels des Minimalpaares *Rate* / *Ratte*.

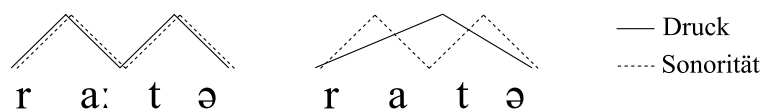


Abbildung 7: Silbenschemata *Rate* und *Ratte*.

⁹ Auch der Grad, wie erfolgreich die Kinder es bewerkstelligen, begleitend zum Sprechen zu klatschen, variiert stark, wurde hier aber nicht differenzierter ausgewertet. Rainbow (1977), Rainbow / Owen (1979) und Jarjisian (1989) widmen sich der Entwicklung rhythmischer Fertigkeiten eingehender.

¹⁰ Vereinfacht lässt sich die Drucksilbe mit Intensität gleichsetzen; was dagegen Schallfülle akustisch-phonetisch ist, mag gegenwärtig als offene Frage gesehen werden. Neben Akzent und Frequenz spielt sicherlich die Sonorität eine wichtige Rolle. Kohler (1995, 74) weist jedoch auf ein wesentliches Problem der sonoritätsbasierten Silbendefinition hin: Die Sonoritätshierarchie der Laute des Deutschen ist kein empirischer Wert, sondern anhand der typischen deutschen Silbenstruktur konstruiert. Entgegen dieser Hierarchie hat aber z. B. ein [i] (als Silbenkern) in signalphonetischen Messungen (und auch perzeptorisch) weniger Schallintensität als ein [ʃ], das in einer zunehmenden oder abnehmenden Position stehen muss. Definiert man nun eine Silbe des Deutschen als Element, das dieser angenommenen Kurve von Einzelintensitäten entspricht, ist das a) ein Zirkelschluss und b) falsch. Dazu treten Fälle wie der Einsilbler *Obst*, der die Sonoritätskurve eines zweisilbigen Worts aufweist (hier offenbar als Reflex einer historischen Zweisilbigkeit, vgl. mhd. *obez*).

In der standarddeutschen Realisierung von *Ratte* [ratə] weicht die Gliederung in Schallsilbe und Drucksilbe voneinander ab (vgl. auch Röber-Siekmeyer 2002, 127–128). Maas (2002, 16–17) sieht in diesen Fällen einen Hinweis auf die wahrnehmungstechnische Dominanz der „Drucksilbe“:

Durch die Orthographie noch nicht beeinflusste Kindergartenkinder oder Schulanfänger, die man zur Silbengliederung anhält, behandeln Formen wie *Ratte* häufig als einsilbig – zum Leidwesen einer Didaktik, die, wenn sie überhaupt mit lautstrukturellen Konzepten wie dem der Silbe umgeht, dann allein am Konzept der Sonoritätssilbe orientiert ist.

Auch in den obigen Analysen finden sich entsprechende Fälle (insbesondere beim Lexem *Teddy*), die diese Annahme stützen können. Ein elegantes Modell zur Erklärung (nicht nur) dieser Befunde liefert die Hypothese des „Silbenschnitts“, die meist mit Sievers (1876) sowie Trubetzkoy ([1939] 1977) in Verbindung gebracht wird. Über den Zeitraum seiner Anwendung hat diese immer wieder Modifikationen erfahren (vgl. detailliert Restle 2003, 3–37), für die folgenden Betrachtungen sind die Details der Theorie bzw. ihre konkurrierenden Ansätze aber nicht weiter relevant; wir beschränken uns daher auf einen groben Überblick des Grundprinzips. Der vokalische Nukleus einer betonten Silbe stellt das Energiemaximum dar, nach dem ein „Decrescendo“ stattfinden muss.¹¹ Entweder leistet dies ein konsonantischer Endrand (Typ geschlossene Silbe, Beispiel *Tan-te* [tan.tə]), wobei der Konsonant den vokalischen Nukleus „scharf“ in die Form eines Kurzvokals¹² schneidet – so in der standarddeutschen Realisierung von *Ratte* in Abbildung 7 –, oder das Energieniveau sinkt in Form eines „sanften Silbenschnitts“ im Vokal ab (offene Silbe, Beispiel *Na-se* [na:.zə]) – der Vokal erscheint lang. Ein Silbengelenk tritt ein, wenn das Decrescendo des Nukleus mittels des Onsets der Folgesilbe stattfinden muss: Dieser Konsonant hat dann eine Doppelfunktion, zum einen schneidet er scharf den Nukleus der ihm vorausgehenden Silbe, zum anderen eröffnet er die darauffolgende Silbe.¹³ Es ist nicht ganz klar, was das für die Lokalisierbarkeit der Silbengrenze bedeutet. Berg (1992, 46–47) hält es für denkbar, dass „die Silbengrenze (und folglich die Silbe) unscharf, d.h. nicht notwendigerweise zwischen zwei phonemischen Einheiten zu setzen“ sei.

Steht in der Silbengelenksposition ein Laut, der kontinuierlich artikulierbar ist, wird der scharfe Silbenschnitt so indirekt nachweisbar, weil die *Anschlusskorrelation* (vgl. Restle 2003) eine Artikulation direkt in der Koda der ersten Silbe „erzwingt“. Bricht nach dem Vokal die Artikulation ab, liegt kein Anschluss vor; eine offene Silbe auf Kurzvokal entsteht. Dieser Silbentypus ist durch den Silbenschnitt nicht motivierbar und dem standarddeutschen System fremd – nicht aber den alemannischen und zum Teil bairischen Basismundarten, die als Substrat für den Regiolekt im Großraum Augsburg dienen. Der (mit 30 % der kontinuierlich bildbaren Laute und 79 % der Affrikatenrealisierungen durchaus häufige) empirische Beleg dieses Typs einer offenen, kurzen Silbe zeigt demnach, dass es nicht problemfrei möglich ist, alle vorliegenden Daten widerspruchlos mittels eines silbenschnittbasierten Modells zu erfassen.

¹¹ Becker (2002) prägt für diese Erscheinung den Terminus „Implosion“.

¹² Natürlich argumentiert die Silbenschnitttheorie gegen die phonologische Relevanz von Quantität. Dessen ungeachtet nutzen wir hier „kurz“ oder „lang“, zunächst als rein phonetische Merkmale, ohne damit eine Aussage darüber treffen zu wollen, ob es sich auf phonologischer Ebene um die Quelle oder lediglich eine Begleiterscheinung einer Distinktion handelt.

¹³ Spiekermann (2002) stellt Kriterien auf, wie der Silbenschnitt über die Intensität akustisch messbar sein soll: Ein sanft geschnittener Vokal – im klassischen Modell ein Langvokal – hat ein oder mehrere Energiemaxima in der Vokalmittte, ein scharf geschnittener nur eines am Vokalbeginn. Ausschlaggebend soll aber „das Halten eines hohen Energieniveaus“ (Spiekermann 2002, 195) sein. Differenzierend dazu merkt Maas (2002, 25) an: „Bei den Anschlussverhältnissen handelt es sich um eine komplexe perzeptive Kategorie, bei der nicht zu erwarten ist, dass sie gewissermaßen monoton auf ein für alle Arten von Segmentübergängen einheitliches Syndrom im Signal reduzierbar ist.“ Zur artikulatorischen Sicht siehe Hoole / Mooshammer / Tillmann (1994) sowie Kroos / Hoole / Kühnert / Tillmann (1997).

Es ist evident, dass sich in Bezug auf die Möglichkeiten der Silbierung in den Daten der Kinder ein regional-typologischer Gegensatz manifestiert.

Für Plosive ist die Situation weniger eindeutig: Sollte hier ein direkter Anschluss vorliegen, wäre lediglich denkbar, dass der Verschluss auf die Koda der ersten Silbe fällt, das „Gelenk“ über eine verlängerte Verschlussphase „überbrückt“ wird und das Explosionsgeräusch den Onset der zweiten Silbe bildet. Maas (2002, 20) betont jedoch explizit, dass die Ambisyllabizität nicht einfach nur eine feinere zeitliche Auflösung auf der Beobachtungsebene sei: Die Vorstellung, dass die Silbengrenze quasi „durch“ das Element laufe (also etwa die Implosion zur vorigen und die Explosion zur späteren Silbe gehöre), sei abzulehnen. Der Unterschied zwischen einer (im Standarddeutschen sanktionierten) betonten offenen Silbe, die auf Kurzvokal endet oder einer tatsächlichen Verschlussbildung ist allerdings auf Aufnahmen mit klatschenden Kindern durch Messung¹⁴ nicht zu treffen.

Aussagekräftig sind hier daher insbesondere die Realisierungen, in denen Silben auf offenen Langvokal enden. Auch hier ist über die Silbenschnitthypothese keine befriedigende Interpretation in Richtung eines Gelenks möglich: Offensichtlich ist der Vokal nicht scharf geschnitten, sonst erschiene er kurz. Ein sanfter Schnitt wiederum spräche gegen ein Silbengelenk an dieser Stelle (siehe oben). Die Vermutung liegt nahe, dass die Systeme der betroffenen Kinder hier keine Möglichkeit für einen Silbenschnitt aufweisen, aber auch keine offenen Silben auf Kurzvokal zulassen – es kommt daher in der Belastungssituation zur Dehnung des Vokals.

Dass sich lediglich ein Teil der in 3.1 und 3.2 ermittelten Resultate schlüssig per Silbenschnittmodell interpretieren lässt, zeigt, dass unter den 35 hier analysierten Kindern offenbar typologische Variation vorliegt: Der Erstspracherwerb der Kinder findet in einer Kontakt- bzw. Konfliktsituation zwischen dem standarddeutschen System¹⁵ und einem typologisch abweichenden Regionalsystem statt – Ersteres lässt sich über die Silbenschnitthypothese befriedigend erfassen, Letzteres nicht.

3.4 Exemplarbasierte Interpretation

Daneben ist bei Kindern dieser Altersklasse aber auch denkbar, dass systematische Variation in den Daten auf erwerbsbasierte Prozesse zurückzuführen sein könnte. Die frühen Phasen der Sprachwahrnehmung und des Erwerbs des phonologischen Systems der Erstsprache sind experimentell breit dokumentiert; am besten ist hier davon zu sprechen, dass sich generelle zu sprachspezifischen Mustern verschieben bzw. einengen (vgl. Walley 2005, 449). Am Ende des ersten Lebensjahres (9 bis 12 Monate) ist bereits eine Form der kontrastiven Phonologie

¹⁴ Das akustische Korrelat des Silbenschnitts, das Spiekermann (2002) anführt, lässt sich hier nicht reproduzieren. Ein alternatives Vorgehen wäre, die schwache bzw. intendierte Artikulation eines Verschlusses in der ersten Silbe über einen Formantenverlauf des Vokals (der eine koartikulatorisch bedingte Bewegung hin zum Artikulationsort des nachfolgenden Plosivs andeuten würde; zur Methode allgemein siehe auch Foulkes / Docherty / Jones 2011) nachzuweisen. Diese Probleme fügen sich in das Bild einer grundsätzlich als methodisch unterspezifiziert anzusehenden Phonetik mit Kindern als Forschungsobjekt; für die eingeschränkte Nutzbarkeit und Vergleichbarkeit etablierter Verfahren sind insbesondere die hohe Grundfrequenz kindlicher Stimmen, die zu einer niedrigeren Auflösung der Formanten führt, sowie die im Vergleich zu Erwachsenen verzerrte Relation zwischen der Größe der Resonanzräume und der Zunge verantwortlich (vgl. Khattab / Roberts 2011, 169).

¹⁵ Schmidt / Vennemann (1985, 167) konstatieren, der Silbenschnitt im Standarddeutschen sei eine „Entlehnung einer phonotaktischen Eigenschaft“ aus dem Niederdeutschen, Maas (2002, 32) betont, dem Silbenschnitt des Standarddeutschen würden „konkurrierende Silbenmodelle gegenüberstehen, insbesondere der aggregierende Typ, der wohl den alemannischen Raum bestimmt.“ Dies fügt sich in die Hypothese, dass das moderne standarddeutsche phonologische System a) zumindest teilweise als Kontaktphänomen mit dem Niederdeutschen herzuleiten sowie b) als typologisch von einem beträchtlichen Teil seiner ihm verwandten Varietäten (die es scheinbar „überdacht“) substantiell abweichend zu sehen ist (vgl. Pröll / Kleiner i.E.).

(im Sinne einer Diskriminierungsfähigkeit von Einzellauten) ausgeprägt. Danach, wenn der Lexikerwerb in seine quantitativ stärkste Phase tritt, klafft jedoch im Forschungsstand zum Phonologieerwerb eine Lücke, was den Eindruck erwecken kann, dass keine phonologisch relevanten Entwicklungen mehr stattfinden würden (vgl. Walley / Flege 1999, 308; Walley 2005, 450, 457). Im *Lexical Restructuring Model* (nach Metsala / Walley 1998) wird dagegen davon ausgegangen, dass der massive Lexikerwerb die Repräsentation und Verarbeitung auch phonologischer Muster – allein über die große quantitative Änderung – klar beeinflusst (vgl. Walley 2005, 450).

Offenbar sind sprachliche Repräsentationen von Kleinkindern nicht segmental und detailliert präsent, sondern holistisch und unterspezifiziert (vgl. zum Forschungsüberblick Walley 2005), ein expliziter Zugriff auf Segmente entwickelt sich erst später; er ist in diesem Alter noch nicht vorhanden und vermutlich auch nicht nötig, wie Walley (2005, 455–456) ausführt:

For the more demanding task of linking words with objects, infants/toddlers may use just enough available information to avoid confusing a new word with the few words in their lexicons. Importantly, individual segments may not be needed to distinguish the small number of words that are known to the older infant/toddler.

Während gröbere, holistische Repräsentationen in einem eher losen Netzwerk eines kleinen Lexikons also ausreichen, sind für engere Netzwerke mehr Details erforderlich, um Einheiten zuverlässig voneinander unterscheiden zu können (vgl. Nittrouer / Miller / Crowther / Manhart 2000; Walley 2005). Mit dem Wachstum des Lexikons werden auch die entsprechenden Netzwerke dichter; dies erfordert also einen zusätzlichen Grad an Information. Es scheint so, als ob Kinder ab etwa 5 Jahren durch ihr wachsendes Lexikon in den Bereich einer Netzwerkdichte kommen, die feinere phonetische Information und analytische Fertigkeiten zur Itemdiskriminierung erfordert (vgl. Walley 2005, 457). Die Ergebnisse von Charles-Luce / Luce (1990, 1995) legen nahe, dass Kinder noch bis ins Alter von 7 Jahren ein Lexikon aufweisen können, in dem eine große Zahl der Einheiten sich rein durch holistische Information unterscheiden lassen. In etwa diesem Alter beginnt im Regelfall die alphabetische Phase des Schriftspracherwerbs (vgl. zur Übersicht über die einzelnen Stufen Helbig / Kirschhock / Martschinke / Kummer 2005, 32–35); Walley (2005, 450) rechnet mit einem weiteren entscheidenden Einfluss des Schriftsystems auf die phonologische Verarbeitung.

Die Annahme einer *exemplarbasierten phonologischen Repräsentation* (etwa durch Johnson 2007) ist im Grunde eine Weiterentwicklung der Prototypentheorie und eine Annäherung an das in einem breiteren psychologischen Kontext vorherrschende Modell der Beschaffenheit von Erinnerungen generell: Jede einzelne Wahrnehmung hinterlässt eine Spur („trace“), ein Exemplar, das nicht lediglich auf abstrakte/kontrastive Merkmale reduzierbar ist. Kategorien entstehen aus der Anhäufung von (mehr oder weniger ähnlichen) Einzelerfahrungen; Sprachstrukturen sind somit emergente Resultate des Gebrauchs, Lexikon und phonologisches System sind auf die Akkumulation von phonetisch detailliert repräsentierten Tokens zurückzuführen (vgl. Cangemi 2014, 5–11; Johnson 2007, 28). Während das Lexikon wächst – sowohl in Bezug auf den Wortschatz als auch auf die Anzahl der jeweils zugrundeliegenden Exemplare – bilden sich allmählich phonematische Strukturen heraus (vgl. Walley / Flege 1999, 310).

Ein exemplarbasierter Zugang, in dem ein reiches Bündel an phonetischem Detail zur phonologischen Kategorisierung beiträgt, ist im Maximalfall das Gegenteil eines „abstraktionistischen“ (vgl. terminologisch so etwa Cangemi 2014) Zugangs, in dem phonologische Information einzelphänomenbasiert isoliert wird – mindestens aber eine Erweiterung.

Abbildung 8 zeigt den Zusammenhang zwischen dem Alter der Kinder und ihren Realisierungen der kontinuierlich bildbaren Konsonanten (absolute Zahlen der Altersgruppen: 7 Dreijährige, 9 Vierjährige, 10 Fünfjährige, 8 Sechsjährige).¹⁶

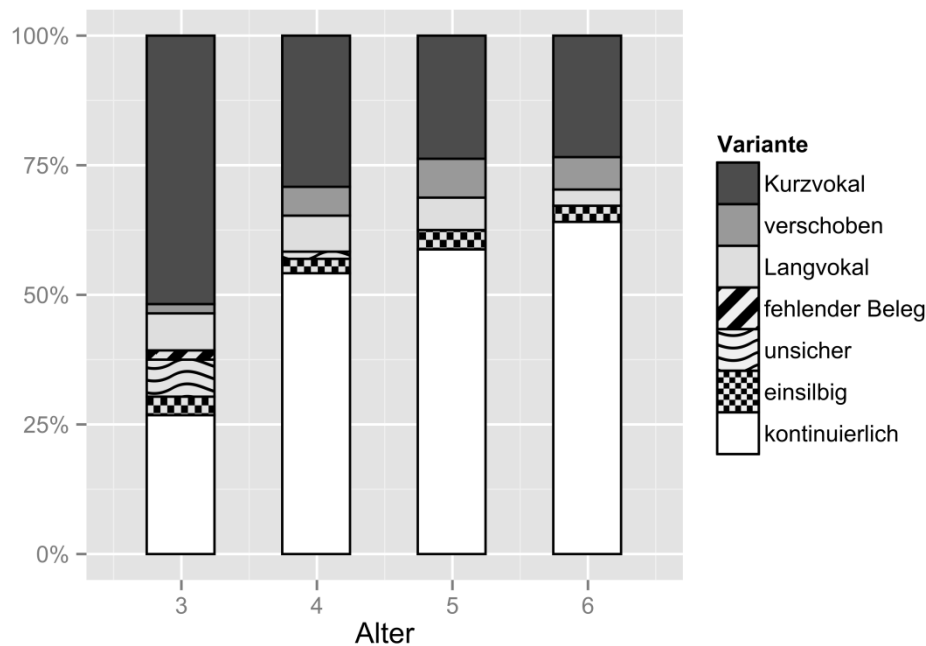


Abbildung 8: Altersverteilung der Realisierungen kontinuierlich bildbarer Laute im Silbengelenk.

Die Entwicklung ist eindeutig: Die offenen Silben gehen von über 50 % auf knapp unter 25 % zurück, während die kontinuierliche, überbrückende Realisierung sich zur klar dominierenden Variante entwickelt.¹⁷ Das ist als Reflex eines phonologischen Erwerbsprozesses zu werten, in dem eine holistische Repräsentation nach und nach durch einen analytischen Zugang ergänzt wird.

Empirische Evidenzen für die Phonologieentwicklung in dieser Altersstufe sind bislang allerdings nicht nur sehr sporadisch gesucht worden, sondern außerdem praktisch durchgängig rein segmentbezogen. Als seltene Ausnahme ist durch Nittrouer / Miller / Crowther / Manhart (2000) belegt, dass Siebenjährige besser als Fünfjährige in der Lage sind, auch detaillierte segmentale Information zur Diskriminierung zu nutzen; die Fünfjährigen dagegen besitzen noch nicht in diesem Grad die Fähigkeit, die Stimulusfokussierung auf diese Art zu gewichten. Sie sind daher noch stärker darauf angewiesen, sich auf die (gröbere) Suprasegmentalstruktur zu verlassen: „5-year-olds have not yet had enough experience [...] to show this kind of flexibility“ (Nittrouer / Miller / Crowther / Manhart 2000, 282).

Dass es Variation in der Realisierung gibt, ist also einem Entwicklungsunterschied der Kinder geschuldet.¹⁸ Die jüngeren Kinder greifen noch stärker auf gröbere, holistische Informationen zurück, auf „information distributed throughout the speech waveform“ (Walley

¹⁶ Für ein Kind lag das Alter nicht vor, seine Realisierungen wurden deshalb nicht mit einbezogen.

¹⁷ Mittels eines Chi-Quadrat-Tests lässt sich zeigen, dass die verschiedenen Altersstufen hochsignifikant voneinander abweichen ($\chi^2 \approx 21,25$; $p < 0,01$; die Varianten wurden dabei zur besseren Handhabung in die drei relevanten Gruppen „kontinuierlich“, „Kurzvokal“ und „sonstige“ zusammengefasst). Da der Chi-Quadrat-Test lediglich zeigt, dass sich Gruppen unterscheiden, nicht aber wie stark, wurde zusätzlich Cramér's V (die Effektstärke) berechnet: Bei einem Wert von $V \approx 0,198$ kann man von einem belastbaren Zusammenhang ausgehen.

¹⁸ Auch Szagun (2000, 240–251) geht – außerhalb eines exemplarbasierten Erwerbsmodells – davon aus, dass im Erstspracherwerb „[i]ndividuelle Unterschiede als Spracherwerbsstile“ existieren, wobei sie zwei unterschiedliche Stile, „analytisch“ und „holistisch“ unterscheidet.

/ Flege 1999, 310). Die Folgerung von Nittrouer / Miller / Crowther / Manhart (2000, 282), dass es den jüngeren Kindern ihrer Studie an Erfahrung mangle, impliziert jedoch bereits, dass Detail auch ein Resultat von Frequenz ist – aus zunehmender Frequenz folgt bessere Analysierbarkeit (so auch Bybee 2001, 121; Ota / Green 2013). Das gilt sowohl für Types als auch für Tokens, vgl. Pierrehumbert (2003, 118): „The phonological system appears to be initiated bottom-up from surface statistics over the speech stream, but refined using type statistics over the lexicon.“ Irreguläre und niederfrequente Phänomene werden später vollständig erworben als reguläre hochfrequente.¹⁹ Dass das Netzwerk des Lexikons erst im Lauf des Erwerbs dichter wird – und, wie oben illustriert, viele Einträge bis ins schulfähige Alter hinein die einzigen in ihrer zukünftigen, phonetisch detaillierten Kategorie sein können, kann für Fälle wie das Silbengelenk nach sich ziehen, dass z.B. *Teddy* die einzige Instanz einer Struktur mit einem Lenisplosiv im Silbengelenk ist. Die oben genutzten Lexeme sind daher zunächst nicht als einfache Träger der jeweiligen Phoneme im Silbengelenk zu sehen, sondern als voneinander relativ unabhängige Grobstrukturen. Für jüngere Kinder ist deshalb wohl stärker als für die älteren für jedes einzelne Lexem je nach Erwerbskontext, Expositions- und Gebrauchsfrequenz sowie dem Grad der Vernetzung im wachsenden Lexikon eine individuelle Repräsentationsdichte zu erwarten. Die explizite Überbrückung der Silbengrenzen im Silbenklatschen, eine im alltäglichen Kommunikationsverhalten nicht existierende „Belastungssituation“ für das segmental-suprasegmentale Zusammenwirken, ist für die Kinder ohne Zugriff auf detaillierte phonetische Information offenbar schwerer. Erst ab einem gewissen Alter bildet sich auch der analytische Zugriff, der erst die Erschließung von Einheiten unterhalb der Silbe ermöglicht (vgl. Maas 2002, 12).²⁰ Im Schulalter ist dann bei vielen – wenn auch nicht allen – Kindern die sog. *Phonologische Bewusstheit* ausgeprägt (vgl. Galliker 2013, 222); nicht als Resultat des Schriftspracherwerbs, sondern als nötige Vorläuferfähigkeit.

4. Zusammenschau

Die experimentelle Erhebung und Analyse von Realisierungen silbengelenkverdächtiger Positionen, die von Kindergartenkindern artikuliert werden, adressiert mehrere Forschungslücken: Zum einen liegen erstmals Daten zur Silbierfähigkeit am Silbengelenk im Alter von 3 bis 6 Jahren vor. Es konnte gezeigt werden, dass die Silbe als Struktur in dieser Altersklasse meist abrufbar ist, aber starke individuelle Unterschiede in der Fähigkeit bestehen, eine einheitliche Strategie in der Realisierung des Silbengelenks zu bilden. Insgesamt zeigen die Daten zwei unterschiedliche Variationsachsen: Zum einen tritt das suprasegmentalphonologische System des Standarddeutschen im Erwerb in Konkurrenz mit dem typologisch abweichenden System der süddeutschen regionalen Umgangssprachen. Zum anderen legen die Daten auch nahe, dass sich im beobachteten Zeitfenster ein Erwerbsprozess abspielt, der sich durch die Theorie einer exemplarbasierten Phonologie erfassen lässt: Vereinzelte, holistische Repräsentationen werden durch Frequenzeffekte und wachsendes Lexikon mit zunehmendem Alter durch phonetisches Detail, das analytische Zugänge ermöglicht, ergänzt. Weiterhin erhärtet sich der Verdacht, dass die bisherige Betrachtungsweise des Silbengelenks (und der Hypothese vom Silbenschnitt) von ihrer Natur her im Grunde skriptizistisch – also durch eine implizit schriftsystembasierte Blickverengung beeinflusst – sein könnte.

¹⁹ Vgl. auch Pierrehumbert (2003, 117): „[T]he knowledge of unproductive, irregular, or opaque alternations found only in erudite vocabulary continues to develop into adulthood.“

²⁰ Auch in Bezug auf die Struktur der Silbe gibt es allerdings Salienz-Präferenzen im Erwerb: Jusczyk / Goodman / Baumann (1999) zeigen, dass Säuglinge (Alter: neun Monate) im englischsprachigen Raum Ähnlichkeiten im Silbenanlaut registrieren, nicht aber im Silbenauslaut; auch der Zusammenhang dieser Fokussierung mit dem Aufbau des Lexikons und dem Informationsabruf scheint deutlich (so auch Walley 2005, 455).

5. Literatur

- AADG = Kleiner, Stefan (2011ff.): Atlas zur Aussprache des deutschen Gebrauchsstandards. Unter Mitarbeit von Ralf Knöbl. Online: <http://prowiki.ids-mannheim.de/bin/view/AADG/> (31.03.2016)
- Ágel, Vilmos (2003): Prinzipien der Grammatik. In: Lobenstein-Reichmann, Anja / Reichmann, Oskar (Hrsg.): Neue historische Grammatiken. Zum Stand der Grammatikschreibung historischer Sprachstufen des Deutschen und anderer Sprachen. Tübingen: Niemeyer. 1–46.
- Altmann, Hans / Ziegenhain, Ute (2007): Phonetik, Phonologie und Graphemik fürs Examen. Göttingen: Vandenhoeck & Ruprecht.
- Becker, Thomas (2002): Silbenschnitt und Silbenstruktur in der deutschen Standardsprache der Gegenwart. In: Auer, Peter / Gilles, Peter / Spiekermann, Helmut (Hrsg.): Silbenschnitt und Tonakzente. Tübingen: Niemeyer. 87–101.
- Berg, Thomas (1992): Umrisse einer psycholinguistischen Theorie der Silbe. In: Eisenberg, Peter / Ramers, Karl Heinz / Vater, Heinz (Hrsg.): Silbenphonologie des Deutschen. Tübingen: Narr. 45–99.
- Boersma, Paul / Weenink, David (2014): Praat: doing phonetics by computer. Version 5.3.72, 17. April 2014, online: <http://www.praat.org/>
- Bybee, Joan (2001): Phonology and language use. Cambridge: Cambridge University Press.
- Cangemi, Francesco (2014): Prosodic detail in Neapolitan Italian. Berlin: Language Science Press.
- Charles-Luce, Jan / Luce, Paul A. (1990): Similarity neighbourhoods of words in young children's lexicons. In: *Journal of Child Language* 17 (1), 205–215.
- Charles-Luce, Jan / Luce, Paul A. (1995): An examination of similarity neighbourhoods in young children's receptive vocabularies. In: *Journal of Child Language* 22 (3), 727–735.
- Foulkes, Paul / Docherty, Gerard / Jones, Mark J. (2011): Analyzing stops. In: Di Paolo, Marianna / Yaeger-Dror, Malcah (Hrsg.): *Sociophonetics*. London, New York: Routledge. 58–71.
- Galliker, Mark (2013): *Sprachpsychologie*. Tübingen: Francke.
- Hall, T. Alan (2000): *Phonologie. Eine Einführung*. Berlin, New York: Walter de Gruyter.
- Hanke, Petra (2002): Interdisziplinäre Betrachtungen zur Bedeutung sprachlicher Strukturen beim Schriftspracherwerb. In: Tophinke, Doris / Röber-Siekmeyer, Christa (Hrsg.): *Schärfungsschreibung im Fokus*. Baltmannsweiler: Schneider-Verlag Hohengehren. 56–70.
- Helbig, Paul / Kirschhock, Eva-Maria / Martschinke, Sabine / Kummer, Ursula (2005): *Schriftspracherwerb im entwicklungsorientierten Unterricht*. Bad Heilbrunn: Klinkhardt.
- Hoole, Philip / Mooshammer, Christine / Tillmann, Hans G. (1994): Kinematic analysis of vowel production in German. In: *Proceedings of the ICSLP '94, Yokohama*. Vol. 1. 53–56.
- Huber, Oswald (2005): *Das psychologische Experiment*. 4. Auflage. Bern: Huber.
- Hulst, Harry van der (1985): Ambisyllabicity in Dutch. In: Bennis, Hans / Beukema, Frits (Hrsg.): *Linguistics in the Netherlands 1985*. Dordrecht: Floris. 57–66.
- Huneke, Hans-Werner (2000): Intuitiver Zugang von Vorschulkindern zum Silbengelenk. In: *Didaktik Deutsch* 8, 4–17.
- Huneke, Hans-Werner (2002): Intuitiver Zugang von Vorschulkindern zum Silbengelenk – eine Grundlage für den Erwerb der Schärfungsschreibung? In: Tophinke, Doris / Röber-Siekmeyer, Christa (Hrsg.): *Schärfungsschreibung im Fokus*. Baltmannsweiler: Schneider-Verlag Hohengehren. 85–104.
- Jarjisian, Catherine (1989): Beat keeping and in-tune singing in preschool children. In: *Update* 7 (2), 24–27.

- Johnson, Keith (2007): Decisions and Mechanisms in Exemplar-based Phonology. In: Solé, Maria-Josep / Beddor, Patrice Speeter / Ohala, Manjari (Hrsg.): *Experimental Approaches to Phonology*. Oxford: Oxford University Press. 25–40.
- Jusczyk, Peter W. / Goodman, Mara B. / Baumann, Angela (1999): Nine-Month-Olds' Attention to Sound Similarities in Syllables. In: *Journal of Memory and Language* 40 (1), 62–82.
- Khattab, Ghada / Roberts, Julie (2011): Working with children. In: Di Paolo, Marianna / Yaeger-Dror, Malcah (Hrsg.): *Sociophonetics*. London, New York: Routledge. 163–177.
- Kohler, Klaus J. (1995): *Einführung in die Phonetik des Deutschen*. 2., neubearbeitete Auflage. Berlin: Erich Schmidt Verlag.
- Kroos, Christian / Hoole, Philip / Kühnert, Barbara / Tillmann, Hans G. (1997): Phonetic evidence for the phonological status of the tense-lax distinction in German. In: *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM)* 35, 17–25.
- Maas, Utz (2002): Die Anschlusskorrelation des Deutschen im Horizont einer Typologie der Silbenstruktur. In: Auer, Peter / Gilles, Peter / Spiekermann, Helmut (Hrsg.): *Silbenschnitt und Tonakzente*. Tübingen: Niemeyer. 11–34.
- McKean, Cristina / Letts, Carolyn / Howard, David (2013): Functional reorganization in the developing lexicon: separable and changing influences of lexical and phonological variables on children's fast-mapping. In: *Journal of Child Language* 40 (2), 307–335.
- Metsala, Jamie L. / Walley, Amanda C. (1998): Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In: Metsala, Jamie L. / Ehri, Linnea C. (Hrsg.): *Word Recognition in Beginning Literacy*. Hillsdale, NJ: Lawrence Erlbaum. 89–120.
- Nittrouer, Susan / Miller, Marnie E. / Crowther, Court S. / Manhart, Mary Jane (2000): The effect of segmental order on fricative labeling by children and adults. In: *Perception & Psychophysics* 62 (2), 266–284.
- Ota, Mitsuhiro / Green, Sam J. (2013): Input frequency and lexical variability in phonological development: a survival analysis of word-initial cluster production. In: *Journal of Child Language* 40 (3), 539–566.
- Pierrehumbert, Janet B. (2003): Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. In: *Language and Speech* 46 (2/3), 115–154.
- Prinz, Michael / Wiese, Richard (1991): Die Affrikaten des Deutschen und ihre Verschriftung. In: *Linguistische Berichte* 133, 165–189.
- Pröll, Simon / Kleiner, Stefan (i.E.): Silbengrenzen im Gebrauchsstandard. *Empirie, Theorie und Typologie*. Erscheint in: *Deutsche Sprache*.
- Rainbow, Edward L. (1977): A longitudinal investigation of the rhythmic abilities of pre-school aged children. In: *Bulletin of the Council for Research in Music Education* 50, 55–61.
- Rainbow, Edward L. / Owen, Diane (1979): A progress report on a three-year investigation of the rhythmic ability of preschool age children. In: *Bulletin of the Council for Research in Music Education* 59, 84–86.
- Ramers, Karl Heinz (1992): Ambisilbische Konsonanten im Deutschen. In: Eisenberg, Peter / Ramers, Karl Heinz / Vater, Heinz (Hrsg.): *Silbenphonologie des Deutschen*. Tübingen: Narr. 246–283.
- Ramers, Karl-Heinz (2001): *Einführung in die Phonologie*. München: Fink.
- Restle, David (2003): *Silbenschnitt – Quantität – Kopplung. Zur Geschichte, Charakterisierung und Repräsentation der Anschlußprosodie*. München: Fink.
- Risel, Heinz (1999): Können Kinder Wörter problemlos in Silben gliedern? In: *Grundschule* 31 (7/8), 76–77.

- Risel, Heinz (2002): Zur Silbierkompetenz von Grundschulkindern. In: Tophinke, Doris / Röber-Siekmeyer, Christa (Hrsg.): Schärfungsschreibung im Fokus. Baltmannsweiler: Schneider-Verlag Hohengehren. 71–84.
- Röber-Siekmeyer, Christa (1993/1995): Die Schriftsprache entdecken. Rechtschreiben im offenen Unterricht. Weinheim, Basel: Beltz.
- Röber-Siekmeyer, Christa (2002): Prosodisch orientierte Untersuchungen zur Wahrnehmung von Schärfungswörtern von Kindern am Schriftanfang. In: Tophinke, Doris / Röber-Siekmeyer, Christa (Hrsg.): Schärfungsschreibung im Fokus. Baltmannsweiler: Schneider-Verlag Hohengehren. 106–143.
- Schmidt, Oskar / Vennemann, Theo (1985): Die niederdeutschen Grundlagen des neuhochdeutschen Lautsystems. In: Beiträge zur Geschichte der deutschen Sprache und Literatur 107, 1–20 & 157–173.
- Sievers, Eduard (1876): Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen. Leipzig: Breitkopf & Härtel.
- Spiekermann, Helmut (2000): Silbenschnitt in deutschen Dialekten. Tübingen: Niemeyer.
- Spiekermann, Helmut (2002): Ein akustisches Korrelat des Silbenschnitts: Formen des Intensitätsverlaufs in Silbenschnitt- und Tonakzentsprachen. In: Auer, Peter / Gilles, Peter / Spiekermann, Helmut (Hrsg.): Silbenschnitt und Tonakzente. Tübingen: Niemeyer. 181–199.
- Szagun, Gisela (2000): Sprachentwicklung beim Kind. Weinheim, Basel: Beltz Taschenbuch.
- Trubetzkoy, Nikolaj S. ([1939] 1977): Grundzüge der Phonologie. Göttingen: Vandenhoeck & Ruprecht.
- Walley, Amanda C. (2005): Speech Perception in Childhood. In: Pisoni, David / Remez, Robert (Hrsg.): The Handbook of Speech Perception. Malden, MA [u.a.]: Blackwell, 449–468.
- Walley, Amanda C. / Flege, James E. (1999): Effects of lexical status on the perception of native and nonnative vowels: A developmental study. In: Journal of Phonetics 27 (3), 307–332.
- Wiese, Richard (1988): Silbische und lexikalische Phonologie. Studien zum Chinesischen und Deutschen. Tübingen: Niemeyer.
- Wiese, Richard (2000): The phonology of German. Oxford: Oxford University Press.

Adressen der Verfasser:

Dr. Simon Pröll, Ludwig-Maximilians-Universität München, Germanistische Linguistik, Schellingstr. 3 RG, D-80799 München.

E-Mail: simon.proell@lmu.de

Dr. Jan-Claas Freienstein / Dr. Oliver Ernst, Universität Augsburg, Lehrstuhl für Deutsche Sprachwissenschaft, Universitätsstr. 10, D-86159 Augsburg.

E-Mail: freienstein@phil.uni-augsburg.de; oliver.ernst@phil.uni-augsburg.de