

---

# Lernbarkeitstheorie

---

## Grundlagen einer Lernbarkeitstheorie

### Avant Propos

“We will not be concerned for the moment with justifying the particular definitions of the concepts we introduce. In fact, a theme throughout much of this book will be that certain assumptions are wrong, that they won’t work for the solution of the learnability problem.”

(Wexler & Culicover 1980, 40)

## Eine elementare Lernbarkeitstheorie

- Eine Theorie des Erstspracherwerbs kann aufgefasst werden als ein Tripel  $\langle \mathcal{G}, \mathcal{I}, LP \rangle$ , wobei
  1.  $\mathcal{G}$  eine Klasse von Grammatiken ist (die **möglichen** Grammatiken),
  2.  $\mathcal{I}$  eine Klasse von möglichen Mengen von Inputs  $I(G)$  (wobei  $G \in \mathcal{G}$ ) ist,
  3.  $LP$  eine Prozedur ist, die Mengen von Inputs auf Grammatiken abbildet.
- Eine solche Theorie ist korrekt, wenn
  1.  $\mathcal{G}$  eine beschreibungsadäquate Grammatik  $G$  für jede natürliche Sprache enthält,
  2.  $\mathcal{I}$  alle Inputmengen aus einer Sprache enthält, die dem Lerner zugänglich sind; d.h. wenn  $G$  aus  $I(G)$  gelernt werden kann, dann gilt  $I(G) \in \mathcal{I}$ ,
  3. jedes  $G \in \mathcal{G}$  durch  $LP$  aus jeder Menge  $I(G) \in \mathcal{I}$  gelernt werden kann.

## Eine elementare Lernbarkeitstheorie 2

- Eine Lernbarkeitstheorie muss (die Elemente von)  $\mathcal{G}, \mathcal{I}$  spezifizieren,  $LP$  angeben und ein Kriterium  $C$  enthalten, das angibt, wann  $LP$  erfolgreich sein muss.
  1. Der Lerner muss die Zielgrammatik aufgrund einer **endlichen** aber ausreichend großen Menge von Sätzen (auch **primäre Daten** genannt) erschließen.
  2.  $LP$  ist eine **Funktion**  $f$ , die solche primären Daten  $D$  auf Grammatiken abbildet und dabei das Kriterium  $C$  erfüllt.
  3.  $C$  besagt, dass jede Inputmenge  $D'$ , die  $D$  enthält, es erlaubt bei der Zielgrammatik anzukommen, wenn  $D$  dies erlaubt.
- Eine Funktion von einer Menge  $A$  nach einer Menge  $B$  ist eine binäre Relation  $R$  zwischen  $A$  und  $B$  mit folgender Eigenschaft: Für jedes  $a \in A$  gibt es genau ein  $b \in B$ , so dass  $(a, b) \in R$ .

## Mengenlernbarkeit

- Diese Bedingungen werden zusammen ausgedrückt im Begriff der **Mengenlernbarkeit**.
- Abkürzungen:
  1. Sei  $\Sigma^*$  die Menge aller möglichen Ketten über einem endlichen **Terminalvokabular**  $\Sigma$ .
  2. Die Sprache, die von  $G$  charakterisiert wird, bezeichnet man mit  $L(G)$ .
- Sei  $\mathcal{G}$  eine Menge von Grammatiken über einem festen  $\Sigma$ .  $\mathcal{G}$  heißt **mengenlernbar**, wenn eine Funktion  $f$  von endlichen Teilmengen von  $\Sigma^*$  nach  $\mathcal{G}$  existiert, so dass für jedes  $G \in \mathcal{G}$  eine endliche Menge  $D \subseteq L(G)$  existiert, so dass
  1.  $f(D) = G'$ , wobei  $L(G') = L(G)$ , und
  2. für alle endlichen  $D'$  mit  $D \subseteq D' \subseteq L(G)$  gilt:  $f(D) = f(D')$ .

## Schwache und starke Äquivalenz von Grammatiken

- Zwei Grammatiken  $G_1$  und  $G_2$  heißen **schwach äquivalent**, wenn  $G_1$  und  $G_2$  genau dieselbe Menge an Ketten erzeugen.
- $G_1$  und  $G_2$  heißen **stark äquivalent**, wenn sie schwach äquivalent sind und wenn sie den generierten Ketten die selben Ableitungsbäume zuordnen.
- Beachte:
  1. Mengenlernbarkeit verlangt lediglich, dass  $LP$  ein  $D$  auf eine Grammatik abbildet, die schwach äquivalent zur Zielgrammatik ist (siehe 1.).
  2. Wenn allerdings die Zielgrammatik (oder eine schwach äquivalente) einmal gefunden wurde, dann ändert sie sich nicht mehr, auch wenn weitere Daten präsentiert werden (siehe 2.).

## Eine mengenlernbare Klasse $\mathcal{G}$

- Sei  $\mathcal{G} = \{G_1, G_2, \dots\}$  eine Klasse von Grammatiken, die folgende Sprachen charakterisieren ( $a^i$  ist eine Kette aus  $a$ s der Länge  $i$ ):
  1.  $L(G_1) = \{a, aa, aaa, \dots\}$
  2.  $L(G_2) = \{aa, aaa, \dots\}$
  3. ...
  4.  $L(G_i) = \{a^i, a^{i+1}, \dots\}$
  5. ...
- $\mathcal{G}$  ist mengenlernbar:
  1. Sei  $D$  eine endliche Menge endlicher  $a$ -Ketten, die von einem  $G$  aus  $\mathcal{G}$  generiert werden.
  2. Sei  $j$  die Länge der kürzesten Kette aus  $D$ .
  3. Dann kann  $f$  definiert werden als:  $f(D) = G_j$ .
  4. Für jedes  $G_i \in \mathcal{G}$  existiert so ein  $D : \{a^i\}$ .
  5. Wenn  $D$  eine endliche Menge, so dass  $\{a^i\} \subseteq D \subseteq L(G_i)$ , dann gilt auch  $f(D) = G_i$ , denn jede Kette aus  $L(G_i)$  hat mindestens die Länge  $i$ .

## Eine nicht-mengenlernbare Klasse $\mathcal{H}$

- Sei  $\mathcal{H} = \{H_0, H_1, H_2, \dots\}$  eine Klasse von Grammatiken, die folgende Sprachen charakterisieren:
  1.  $L(H_0) = \{a, aa, aaa, aaaa, \dots\}$
  2.  $L(H_1) = \{a\}$
  3.  $L(H_2) = \{a, aa\}$
  4. ...
  5.  $L(H_i) = \{a, aa, aaa, \dots, a^i\}$
  6. ...
- $\mathcal{H}$  ist nicht mengenlernbar:
  1. Für jedes endliche  $D$  kann  $f$  **immer** entweder eins aus unendlich vielen  $H_{i>0}$  wählen, oder  $H_0$ .
  2. Wählt  $f$  immer ein  $H_{i>0}$ , dann wird es nie  $H_0$  wählen, auch nicht bei Input von  $L(H_0)$ .
  3. Wählt  $f$  immer  $H_0$ , dann wird es nie ein  $H_{i>0}$  wählen, auch wenn Daten aus  $L(H_{i>0})$  präsentiert werden.
  4. Gemischte Strategien funktionieren auch nicht.

## Exkurs: Beweis durch Widerspruch

- Ein wichtiges Beweisschema ist der **Beweis durch Widerspruch**.
- Beweis von Hypothese  $H$  durch Widerspruch:
  1. Nimm an, es gelte die Negation von  $H$ :  $H'$
  2. Leite dann unter dieser Annahme durch logische Schlußfolgerungen einen Widerspruch ab.
  3. Das einzige, was in dieser Ableitung zum Widerspruch geführt haben kann, ist die Annahme, dass  $H'$  gilt.
  4. Daher muss  $H'$  falsch sein ( $H'$  kann nicht richtig sein, und eine dritte Möglichkeit gibt es nicht).
  5. Wenn  $H'$  falsch ist, dann ist  $H$  wahr.      q.e.d.
- Das Schema beruht auf der Annahme, dass  $H$  (oder  $H'$ ) nicht gleichzeitig wahr und falsch sein können.

## Eine nicht-mengenlernbare Klasse $\mathcal{H}$ 2

- Beweis durch Widerspruch:
  1. Angenommen es gäbe eine Lernbarkeitsfunktion  $f$  für  $\mathcal{H}$ . Nach Mengenlernbarkeit muss es für jedes  $H \in \mathcal{H}$  ein endl.  $D$  geben mit  $f(D) = H$ .
  2. Sei  $D_0$  eine solche Menge für  $H_0$ .
  3. Dann gilt nach Mengenlernbarkeit für jedes  $D$  mit  $D_0 \subseteq D \subseteq L(H_0)$ :  $f(D) = H_0$ .
  4. Sei  $H_{i>0}$  eine Grammatik mit  $L(H_i) \supseteq D_0$  (solche  $H_i$ s existieren:  $D_0$  ist endlich und es gibt unendlich viele  $H_i$ s, deren Sprachen wachsen).
  5.  $L(H_i)$  ist eine endliche Datenmenge. Nach Mengenlernbarkeit muss  $L(H_i)$  eine Teilmenge enthalten, die als Input für  $H_i$  dient. Es folgt, dass  $f(L(H_i)) = H_i$ .
  6. Aber da  $D_0 \subseteq L(H_i) \subseteq L(H_0)$  gilt, folgt nach Mengenlernbarkeit, dass  $f(L(H_i)) = H_0$ .
  7. Das ist ein Widerspruch zu 5., da  $f$  Funktion sein soll. Es folgt, dass  $f$  nicht existiert; damit ist  $\mathcal{H}$  nicht mengenlernbar.

## Intuitionen über Lernbarkeit

- Lernbarkeit folgt nicht unbedingt dem, was man vielleicht intuitiv als “leichter” oder “schwerer” lernbar empfinden würde.
- Die Klasse  $\mathcal{G}$  ist lernbar, obwohl sie eine unendliche Zahl von unendlichen Sprachen generiert.
- Die Klasse  $\mathcal{H}$  ist nicht lernbar, obwohl sie nur eine unendliche Sprache (und unendliche viele endliche Sprachen) generiert.

## Mengenlernbarkeit und Zeit

- Mengenlernbarkeit ignoriert komplett den **Faktor Zeit**: Es wird so getan als ob dem Lerner alle Daten auf einmal zur Verfügung stünden.
- Tatsächlich hört der Lerner aber einen Satz nach dem anderen. Das soll in die Lernbarkeitstheorie integriert werden.
- Konsequenzen:
  1. Jedesmal, wenn der Lerner einen neuen Satz hört, ändert er möglicherweise seine Hypothese.
  2.  $LP$  soll nun eine **berechenbare** Funktion sein, da es in der Zeit operiert.
- Eine Funktion  $f$  ist berechenbar, wenn es ein endliches **Verfahren** (einen **Algorithmus**) gibt, das  $f$  bei definierter Eingabe in einer endlichen Anzahl von Schritten berechnet (und sonst immer weiterrechnet).

## Vom richtigen Zeitpunkt

- Es muss einen geeigneten Zeitpunkt geben, an dem der Lerner genug Input gesehen hat, um die Zielgrammatik zu finden.
- Wie kann man dies in der Theorie ausdrücken? Es genügt nicht zu sagen, dass dem Lerner genügend Zeit gegeben wird.
- Problematisches Szenario:
  1. Angenommen, die Zielsprache enthält alle Ketten über dem Wort  $a$ .
  2. Weiter angenommen, es würde zu jedem Zeitpunkt nur der Satz  $aaa$  präsentiert.
  3. Unter solchen Umständen kann man nicht verlangen, dass der Lerner die Zielgrammatik findet.
- Lösung: Für jeden Satz  $S$  der Sprache muss es einen Zeitpunkt  $t$  geben, an dem  $S$  in der Inputsequenz auftaucht.

## Vom richtigen Zeitpunkt 2

- Um das Problem zu illustrieren, betrachte man nochmals die Klasse  $\mathcal{G}$ :
  1.  $L(G_1) = \{a, aa, aaa, \dots\}$
  2.  $L(G_2) = \{aa, aaa, \dots\}$
  3.  $\dots$
  4.  $L(G_i) = \{a^i, a^{i+1}, \dots\}$
  5.  $\dots$
- Angenommen, es würden dem Lerner nur  $a$ -Ketten der Länge größer 10 dargeboten.
- Dann wäre die Lernprozedur niemals in der Lage, die Grammatiken  $G_1, G_2, \dots, G_{10}$  voneinander zu unterscheiden.

## Textlernbarkeit

- Eine **Informationssequenz** (auch: **Text**)  $I(L)$  einer Sprache  $L$  ist eine Sequenz von Sätzen aus  $L$ , so dass jeder Satz darin vorkommt.
- Sei  $I(L) = a_1, a_2, \dots$  eine Informationssequenz. Dann ist  $S_t(I) = \{a_1, \dots, a_t\}$  eine **Auswahl** von  $I$  zur Zeit  $t$  (also eine ungeordnete Menge).
- $LP$  ist eine berechenbare Funktion von der Menge der Auswahlen nach der Menge der Grammatiken.
- Eine Klasse von Grammatiken  $\mathcal{G}$  ist **textlernbar**, wenn  $LP$  existiert, so dass für jedes  $G_i \in \mathcal{G}$  und jede Informationssequenz  $I(L(G_i))$  ein  $\tau$  existiert, so dass gilt: Wenn  $t > \tau$ , dann
  1.  $LP(S_t(I)) = LP(S_\tau(I))$  und
  2.  $L(LP(S_\tau(I))) = L(G_i)$ .

## Textlernbarkeit 2

- Bedingung 1. der Textlernbarkeit verlangt, dass zu einem bestimmten Zeitpunkt  $\tau$  sich die Hypothese des Lerners nicht mehr ändert.
- Bedingung 2. verlangt, dass die Sprache, die von der zum Zeitpunkt  $\tau$  gefundenen Grammatik generiert wird, (schwach) äquivalent ist zur Sprache, die von der Zielgrammatik generiert wird.

## Theoreme

- **Charakterisierungstheorem:** Sei  $\mathcal{K}$  eine abzählbare Klasse von Grammatiken, die die Sprachklasse  $K$  generiert. Dann ist  $\mathcal{K}$  mengenlernbar genau dann, wenn  $K$  keine unendliche Teilmenge von Sprachen  $K' = \{L_0, L_1, L_2, \dots\}$  enthält, so dass
  1.  $L_0 = \bigcup_{i=1}^{\infty} L_i$  und
  2. für jede endliche Teilmenge  $F$  von  $L_0$  gibt es eine unendliche Zahl an  $L_i \in K'$ , so dass  $F \subseteq L_i$ .
- Theorem der **superfiniten Klassen** (Gold 1967): Keine Klasse von Grammatiken, die alle endlichen Sprachen über einem Vokabular  $\Sigma$  generiert plus eine unendliche Sprache, ist textlernbar.

## Theoreme 2

- **Obermengentheorem** (für Mengen und Texte): Wenn eine Klasse von Grammatiken  $\mathcal{G}$  nicht lernbar ist, und wenn eine Klasse von Grammatiken  $\mathcal{H}$  eine Menge von Sprachen  $L(\mathcal{H})$  generiert, so dass  $L(\mathcal{H})$  eine Obermenge von  $L(\mathcal{G})$  ist, dann ist  $\mathcal{H}$  auch nicht lernbar.
- Lemma (namenlos): Wenn eine Klasse von Grammatiken nicht mengenlernbar ist, dann ist sie auch nicht textlernbar.

## Lernbarkeit kontextfreier Grammatiken

- Behauptung: Die Klasse der kontextfreien (ktf) Grammatiken  $\mathcal{C}$  ist nicht textlernbar.
- Beweis:
  1.  $\mathcal{H}$  war nicht mengenlernbar. Man kann die Elemente von  $\mathcal{H}$  als ktf Grammatiken schreiben:
    - (a)  $H_1: S \rightarrow a$   
 $H_2: S \rightarrow a, S \rightarrow aA_1, A_1 \rightarrow a$   
 $H_3: S \rightarrow a, S \rightarrow aA_1, A_1 \rightarrow a, A_1 \rightarrow aA_2, A_2 \rightarrow a$   
... etc.
    - (b)  $H_0: S \rightarrow a, S \rightarrow aS$
  2. Man kann sich weiter überlegen, dass für zwei Grammatikklassen  $\mathcal{A}, \mathcal{B}$  gilt: Wenn  $\mathcal{A} \subset \mathcal{B}$ , dann  $L(\mathcal{A}) \subseteq L(\mathcal{B})$ .
  3. Da  $\mathcal{H}$  Teilmenge von  $\mathcal{C}$  ist und  $\mathcal{H}$  nicht mengenlernbar ist, folgt dann aus dem Obermengentheorem, dass  $\mathcal{C}$  nicht mengenlernbar ist.
  4. Es folgt aus dem Lemma, dass  $\mathcal{C}$  nicht textlernbar ist.

## Lernbarkeit von Transformationsgrammatiken

- Behauptung: Die Klasse der Transformationsgrammatiken (TfGen)  $\mathcal{T}$  ist nicht textlernbar.
- Beweis:
  1. Eine degenerierte Teilmenge von  $\mathcal{T}$  enthält alle TfGen, die nur die Basiskomponente umfassen.
  2. Die Basiskomponente ist eine ktf Grammatik.
  3. Es gibt eine solche degenerierte TfG für jede ktf Grammatik.
  4. Damit ist die Klasse der ktf Grammatiken  $\mathcal{C}$  eine echte Teilmenge von  $\mathcal{T}$ .
  5.  $\mathcal{C}$  ist nicht textlernbar (siehe letzter Beweis).
  6. Dann folgt aus dem Obermengentheorem, dass  $\mathcal{T}$  auch nicht textlernbar ist.
- Behauptung (ohne Beweis) : Die Klasse der TfGen  $\mathcal{T}_B$  über einer fixen Basiskomponente  $B$  ist nicht mengenlernbar.

## Lernbarkeit von Transformationsgrammatiken 2

- Problem:
  1. Die Klasse  $\mathcal{T}$  der TfGen ist nicht mengen- oder textlernbar.
  2. TfGen sind für Wexler & Culicover (1980) aber das Mittel der Wahl, um eine beschreibungsadäquate linguistische Theorie zu erreichen.
  3. Wie kann man dann linguistische Theorie und Lernbarkeit in Einklang miteinander bringen?

“At present there is reason to suspect that transformational grammars cannot be sufficiently limited to allow both descriptive adequacy and set-learnability.”

(Wexler & Culicover 1980, 53)

- Lösung: Der Lernbarkeitsbegriff muss verändert werden, so dass TfGen lernbar werden.

## Modifikation der Lernprozedur

- Vorschlag 1: Die Lernprozedur (eine Funktion) muss verbessert werden, so dass  $\mathcal{T}$  oder eine Teilmenge von  $\mathcal{T}$  lernbar wird.
- Problem:
  1. Viele der TfGen, die Linguisten vorschlagen würden, sind nicht einmal mengenlernbar.
  2. Das heißt, dass keine Funktion **existiert**, die die Abbildung von **endlichem** Input auf Zielgrammatik leisten kann.
  3. Wenn aber eine solche Funktion noch nicht einmal existiert, dann kann sie auch nicht verbessert werden.

“Thus there is no hope that any kind of increase in the power of procedures will yield learnability according to our definitions.”

(Wexler & Culicover 1980, 55)

## Modifikation des Input

- Vorschlag 2: Der Input wird angereichert, so dass Lernbarkeit möglich wird.
- Frage: Wie soll der Input angereichert werden?
- Bisher:
  1. Der Lerner kann davon ausgehen, dass jeder Satz, den er hört, zu seiner Zielsprache gehört.
  2. Wenn er einen Satz  $S$  nicht hört, kann er **nicht** davon ausgehen, dass  $S$  ungrammatisch ist.
  3. Es kann ja sein, dass er ihn im nächsten Moment hören wird.
- Neue Idee: Der Lerner erhält positive und **negative** Evidenz: Zu jedem Satz wird Information geliefert, ob er grammatisch ist, oder nicht.
- Gold (1967) nennt das **Informantenpräsentation**. Er zeigt:  $\mathcal{C}$  (Klasse der ktf Grammatiken) ist informantenlernbar (aber wie gesehen nicht textlernbar).

## Negative Evidenz

- Frage: Wird das Kind systematisch korrigiert?
- Brown & Hanlon (1970), Slobin (1972): Nein.
- Anekdotische Evidenz dafür, dass selbst wenn das Kind korrigiert wird, es daraus keinen Nutzen für den Spracherwerb zieht (McNeill 1966, 69):
  1. Kind: Nobody don't like me.
  2. Mutter: No, say "nobody likes me".
  3. Kind: Nobody don't like me.
  4. (. . . Dialog wiederholt sich acht mal . . . )
  5. Mutter: No, now listen carefully; say "*nobody likes me*".
  6. Kind: Oh! Nobody don't likes me.

## Baby Talk

- Andere Idee: Das Kind erhält die Daten in bestimmter Weise präsentiert. Dies soll das Lernen erleichtern. Man nennt das auch die Hypothese vom **Baby Talk** (BT).
- Wexler & Culicover (1980, 66ff) diskutieren dann folgenden Fragen:
  1. Inwiefern könnte das beim Lernen helfen?
  2. Was sind die tatsächlichen Fakten?

“First, there is no theoretical justification for the claim that simplified input aids language learning. Second, the best empirical evidence is that input to children in general does not have the special characteristics anyway.”

(Wexler & Culicover 1980, 66)

## Baby Talk 2

- Erinnerung:  $\mathcal{H}$  war nicht textlernbar. Angenommen  $L(\mathcal{H})$ s Sätze sind dem Lerner wie folgt gegeben:
  1. Sätze werden in aufsteigender Reihenfolge der Länge nach präsentiert (kurze zuerst).
  2. Für die endlichen Sprachen  $L(H_{i>0})$  gilt: Wenn der längste Satz erreicht ist, fängt man wieder beim kürzesten an.
- Konsequenz: Es gibt ein Verfahren, das  $\mathcal{H}$  textlernt.
  1. Vermute Grammatik  $H_0$  bis der erste Satz ( $a$ ) wiederholt wird.
  2. Wenn der Satz vor dem wiederholten Satz  $a^i$  war, dann vermute Grammatik  $H_i$ .

“It is highly unlikely that much information about linguistic structure is coded into the input sequence in natural language in the same way as in this artificial example . . .” (Wexler & Culicover 1980, 70)

## Baby Talk 3

- Alternative: Das Kind erhält zunächst nur “simple” Daten, später kompliziertere.
- Probleme:
  1. Ärmerer Input verlangt **mehr** Beschränkungen für die Grammatik, stärkt also die Nativismusthese.
  2. Was ist die Metrik für “einfach”?
  3. Woher sollten Eltern wissen, wie sie den Input präsentieren sollen oder was “einfach” ist? Solches Wissen müsste instinktiv sein.

“ . . . we do not mean to say that special characteristics of mother’s speech have *no* effect on *any* aspect of the growth of language in the child . . . We simply are saying that what is known about BT in no way can eliminate the need for an innate structural component for language learning.”

(Wexler & Culicover 1980, 75)

## Starke vs. schwache Lernbarkeit

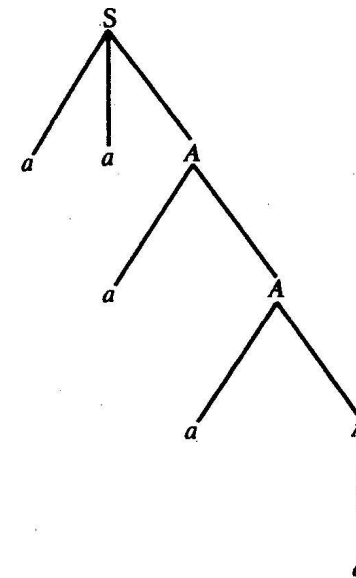
- Mengen- und Textlernbarkeit verlangen bisher nur, dass die Grammatik, die der Lerner findet, schwach äquivalent zur Zielgrammatik ist.
- Das heißt, ultimative Lernergrammatik und Zielgrammatik generieren dieselbe Menge an Ketten. Damit ist die erreichte Lernergrammatik aber nicht beschreibungsadäquat.
- Eine Klasse von Grammatiken ist **schwach lernbar**, wenn die gefundene Grammatik schwach äquivalent zur Zielgrammatik ist.
- Eine Klasse von Grammatiken ist **stark lernbar**, wenn die gefundene Grammatik stark äquivalent zur Zielgrammatik ist.

## Starke vs. schwache Lernbarkeit 2

- Erinnerung:  $\mathcal{G} =$ 
  1.  $L(G_1) = \{a, aa, aaa, \dots\}$
  2.  $L(G_2) = \{aa, aaa, aaaa, \dots\}$
  3.  $L(G_3) = \{aaa, aaaa, aaaaa, \dots\}$
  4. ...
  5.  $L(G_i) = \{a^i, a^{i+1}, \dots\}$
  6. ...
- Ziel- $G_3$  könnte wie folgt aussehen:
  1.  $S \rightarrow aaA$
  2.  $A \rightarrow aA$
  3.  $A \rightarrow a$
- Angenommen Lerner- $G_3$  bestünde aus folgenden Regeln:
  1.  $S \rightarrow Aaa$
  2.  $A \rightarrow Aa$
  3.  $A \rightarrow a$

## Starke vs. schwache Lernbarkeit 3

- Der Baum für  $aaaaa (\in L(G_3))$ , den Ziel- $G_3$  generiert ist (vereinfacht:  $[aa[aaa]]$ ):



- Lerner- $G_3$  weist  $aaaaa$  aber  $[[aaa]aa]$  zu, ist also nicht stark äquivalent zu Ziel- $G_3$ .  $G_3$  (und damit  $\mathcal{G}$ ) wäre nicht stark lernbar.

## Tiefenstruktur und Input

- Ziel: Eine Lernbarkeitstheorie, die die starke Lernbarkeit von TfGn sicherstellt.
- Behauptung: Kinder erhalten Information über die Bedeutung mancher Sätze, obwohl sie ihnen keine vollständige Struktur zuweisen können.
- Diese Information erhalten Sie aus der Situation, in der die Sätze geäußert werden.
- Idee: Diese Information reichert den Input an und soll dadurch das Lernbarkeitsproblem lösen.

## Tiefenstruktur und Input 2

- Annahme (Chomsky 1965): Die Bedeutung eines Satzes ist in der TS des Satzes kodiert.
- Idee: Kinder können aus der Bedeutung eines Satzes  $S$  die TS von  $S$  konstruieren, auch wenn sie die OS von  $S$  noch nicht beherrschen.
- Das Kind erhält als Input eine genügend große Zahl an Paaren  $\langle b, s \rangle$ , wobei
  1.  $b$  eine TS ist, die das Kind sich rekonstruiert,
  2.  $s$  die Kette der OS ist, die durch Transformationen aus  $b$  abgeleitet wurde.

## Literatur

- Brown, R. & C. Hanlon (1970): Derivational complexity and the order of acquisition of child speech. *In: J. Hayes, ed., Cognition and the Development of Language*. Wiley, New York.
- Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Massachusetts.
- Gold, Mark E. (1967): 'Language Identification in the Limit', *Information and Control* **10**, 447–474.
- McNeill, D. (1966): Developmental Psycholinguistics. *In: F. Smith & G. A. Miller, eds, The Genesis of Language*. MIT Press, Cambridge, Massachusetts.
- Slobin, D. (1972): 'Children and language: they learn the same way all around the world', *Psychology Today* **6**, 71–82.
- Wexler, Ken & Peter Culicover (1980): *Formal Principles of Language Acquisition*. MIT Press, Cambridge, Massachusetts.