

## Syntax – Begriffe

Regeln, mit denen »richtige« (== syntaktisch korrekte) Ergebnisse erstellt werden, werden als Grammatik oder Syntax bezeichnet. Die Syntax von (Programmier-)Sprachen wird mit drei verschiedenen Strukturelementen beschrieben.

Bezeichnung	Bedeutung	Beispiel
terminale Symbole	finden sich im Ergebnis als hingeschriebene Symbole wieder	L
nicht terminale Symbole	Zwischenelemente für die Erstellung (oder Analyse)	identifizier
Metasymbole	Bildungsstruktur (Sequenz, Alternative, Wiederholung)	

## BNF – EBNF – ISO-EBNF

Um 1960 wurde eine Darstellung für die Syntax von Programmiersprachen entwickelt, die leicht gelesen und verstanden werden kann. Diese Darstellung wurde nach den Entwicklern als Backus-Naur Form (BNF) bezeichnet. Ein Beispiel aus der Syntax von ALGOL 60 zeigt die Regel, um für diese Sprache ein korrektes nicht terminales Symbol „identifizier“ zu erzeugen. BNF-Darstellung:

```
<identifizier> ::= <letter> | <identifizier> <letter> | <identifizier> <digit>
```

Die BNF wurde zur Extended BNF (EBNF) erweitert (indem weitere Metasymbole hinzugenommen wurden) und 1996 genormt: ISO-EBNF – Standard Extended BNF. EBNF-Darstellung:

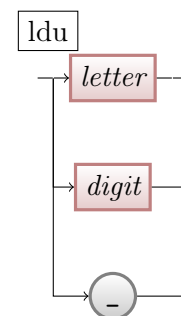
```
identifizier = letter { letter | digit }.
```

Die Regeln besagen (in Worten): Ein »identifizier« wird gebildet durch genau ein nicht terminales Element »letter« gefolgt von beliebig vielen Elementen, die »letter« oder »digit« sein dürfen.

## Syntax-Diagramm als alternative Darstellungsform zu [E]BNF

Mit Hilfe sogenannter *Syntax*-Diagramme<sup>a</sup> werden die Metasymbole so dargestellt, dass sie nicht als Zeichen in den Regeln auftreten, sondern in eine grafische Darstellung eingebettet sind. Die Durchlaufrichtung erfolgt von links nach rechts bzw. in Pfeilrichtung. Knotenpunkte sind Verzweigungen im Sinne von Alternativen oder Zusammenführungen von Wegen.

Im nebenstehenden Beispiel ist das Nonterminal »ldu« aus der Syntax der Programmiersprache Python als Syntax-Diagramm abgebildet. Einträge in den ovalen Textboxen bezeichnen terminale Symbole, während in rechteckigen Textboxen nichtterminale Symbole stehen.



<sup>a</sup>durch ihre Form auch teilweise Railroad (deutsch: Schienenweg) genannt

