



Textlesbarkeit für Alpha-Levels

Annotationsrichtlinien für Lesetexte

Zarah Weiss und Theresa Geppert

Universität Tübingen, DIE Bonn

Version 1.2, September 2018

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



dIE

Inhaltsverzeichnis

1 Übersicht über Alpha-Kriterien	3
1.1 Kriterien für Alpha-Level 3	3
1.2 Kriterien für Alpha-Level 4	4
1.3 Kriterien für Alpha-Level 5	5
1.4 Kriterien für Alpha-Level 6	6
2 Level-Übergreifende Erläuterung der einzelnen Kriterien	7
2.1 Textlänge	7
2.2 Satzlänge	10
2.3 Satzstruktur	12
2.4 Zeitstruktur	14
2.5 Wortlänge	16
3 Glossar	19
4 Frequently Asked Questions	23

Vorwort

Die folgenden Richtlinien zur Annotation der Textlesbarkeit für funktionale Analphabeten dienen der Zuordnung von Texten zu den Alpha-Levels A3 bis A6 und basieren auf den Lesbarkeitskriterien, die in Kretschmann & Wieken (2010) und Gausche et al. (2014) ausgewiesen werden. Die Alpha-Level A1 und A2 werden nicht abgedeckt, da diese unterhalb der Satzebene ansetzen und somit nicht auf das Medium Text angewandt werden können.

Die Richtlinien wurden im Rahmen des KANSAS Projektes in Kooperation der ICALL Research Group der Universität Tübingen und des Deutschen Instituts für Erwachsenenbildung, Bonn angefertigt (<https://www.die-bonn.de/kansas>). Sie wurden erstellt, um eine Gold Standard Annotation von Textlesbarkeit auf einer Untermenge des KANSAS Lese-Korpus zu erzielen und um darauf aufbauend eine automatische Klassifikation von Textlesbarkeit mittels computerlinguistisch informierter Algorithmen zu ermöglichen. Über den Kontext des KANSAS Projektes hinaus können diese Richtlinien jedoch auch als wertvolle Grundlage für weitere wissenschaftliche Untersuchungen sowie für die Evaluation und Replikation von Lesbarkeitsannotationen von Texten für funktionale Analphabeten und Literarisierungskontexte dienen.

Die Richtlinien wurden in enger Orientierung an Kretschmann & Wieken (2010) und Gausche et al. (2014) erstellt, den beiden zentralen Quellen, welche die Spezifikation von Kompetenzniveaus funktionaler Analphabeten sowie die Erstellung entsprechender Curricula und förderdiagnostischer Arbeitsmaterialien fokussieren. Die aus diesen Darstellungen und den dazugehörigen Referenzmaterialien entnommenen Kriterien wurden im Rahmen der vorliegenden Richtlinien unter der Maßgabe höchster definitorischer Klarheit operationalisiert. Kriterien, die sich nicht eindeutig operationalisieren ließen, wurden nicht übernommen. Dies betrifft zum Beispiel auf Sprachnutzung und Wortfrequenz basierende Kriterien wie gebräuchliche Wörter, eingedeutschte Fremdwörter, und geläufige Anglizismen (vgl. Gausche et al. 2014, S. 39). Des Weiteren wurden alle zu Grunde gelegten linguistischen Einheiten explizit für die Zwecke dieser Richtlinien definiert. Sofern dies möglich war, orientieren sich die Definitionen an den Begriffsbestimmungen von Kretschmann & Wieken (2010) und Gausche et al. (2014). In Fällen, in denen die Ausführungen von Kretschmann & Wieken (2010) und Gausche et al. (2014) keine ausdrücklichen Definitionen enthielten, wurden eigene Definitionen festgelegt. Abgesehen vom Kriterium der definitorischen Klarheit wurde bei der Erstellung der Richtlinien und der Definition der zu Grunde liegenden linguistischen Einheiten darauf geachtet, dass für die Anwendung der Richtlinien eine möglichst geringe linguistische Expertise benötigt und eine zeiteffiziente Anwendung der Richtlinien im Annotationsprozess gewährleistet wird.

Danksagung

An dieser Stelle möchten wir all jenen danken, die an der Erstellung dieser Richtlinien beteiligt waren. Besonderer Dank gilt Frau Prof. Dr. Angelika Wöllstein, deren Anmerkungen und Vorschläge für die erfolgreiche Operationalisierung dieser Annotationsrichtlinien unabdingbar waren. Weiterhin gilt unser spezieller Dank unseren beiden Annotatorinnen, Valentina Mayer und Nadja Ramsaier, die mit ihrer Arbeit die Validierung dieser Richtlinien erst möglich gemacht haben und uns mit ihren Rückmeldungen und Vorschlägen dabei unterstützt haben, diese beständig zu verbessern. Nicht zuletzt danken wir auch Herrn Prof. Dr. Detmar Meurers und Herrn PD Dr. Hannes Schröter für ihre Unterstützung während der Erstellung der Richtlinien.

I Übersicht über Alpha-Kriterien

I.1 Kriterien für Alpha-Level 3

Definition

Auf Alpha-Level 3 können bereits einzelne Wörter sowie Alltagssprachliche Satzgefüge sinnerfassend gelesen werden (vgl. Gausche et al. 2014, S. 96f; Kretschmann & Wieken 2010, S. 237). Ziel ist es, die Leseflüssigkeit sowie die Fähigkeit zu steigern, einen zusammenhängenden Text aus mehreren Sätzen erlesen und verstehen zu können (vgl. Gausche et al. 2014, S. 97). Um dies zu gewährleisten, soll die Strategie des Lautierens seitens der Lesenden zunehmend durch die orthographische Strategie ersetzt werden. Größere Segmente, mehrgliedrige Einheiten und orthographische Besonderheiten können dann schneller erlesen werden (vgl. Gausche et al. 2014, S. 39). Damit Wort-für-Wort-Lesungen zunehmend durch flüssiges Lesen ersetzt werden können, ist es daher besonders wichtig, dass die einzelnen Sätze auf diesem Alpha-Level nicht zu viel sprachliches Material enthalten und eine gewisse sprachliche Komplexität nicht überschreiten.

Kriterien

1. **Textlänge:** Der Text besteht aus maximal 5 Sätzen.
2. **Satzlänge:** Jeder Satz hat jeweils maximal 10 Wörter.
→ **Ausnahme:** Pro 3 Sätze darf 1 Satz aus bis zu 15 Wörtern bestehen, falls sonst alle Kriterien ohne Ausnahme erfüllt werden.
3. **Satzstruktur:** Maximal 50% der Sätze enthalten einen Nebensatz. Kein Satz enthält mehr als einen Nebensatz. Kein Satz enthält wörtliche Rede.
4. **Zeitstruktur:** Der Text ist durchgängig im Präsens geschrieben.
5. **Wortlänge:** Jedes Wort hat maximal 3 Silben.
→ **Ausnahme:** Pro 20 Wörter darf 1 Wort-Typ bis zu 5 Silben aufweisen, falls ansonsten alle Kriterien ohne Ausnahmen erfüllt werden.

Beispieltext

Leinwand

Ein Maler malt sein Bild auf der Leinwand. Das ist ein Stoff, der sehr fest ist. Der Stoff ist aus Leinen.
Oder: Das ist eine große weiße Fläche vorne im Kino. Darauf wird der Film gezeigt.

Leicht abgewandelt aus dem MDR-Wörterbuch in Leichter Sprache, Eintrag für „Leinwand“.¹

¹<https://www.mdr.de/nachrichten-leicht/woerterbuch/glossar-leinwand100.html>, zuletzt aufgerufen am 16.04.18.

1.2 Kriterien für Alpha-Level 4

Definition

Auf Alpha-Level 4 können kurze, schriftsprachlich konstituierte Texte gelesen werden. Einzelne Textinformationen können kombiniert und zueinander in Beziehung gesetzt werden, sodass die Gesamtaussage des Textes verstanden und die Inhalte reflektiert und bewertet werden können (vgl. Gausche et al. 2014, S. 131). Insbesondere durch die Lektüre von Sachtexten soll die Fähigkeit gesteigert werden, einen Text zu erlesen, die Inhalte zu verstehen und im Leseprozess neues Wissen zu erwerben. Um das Textverständnis Lernender zu sichern, sind die Texte auf diesem Alpha-Level kurz gehalten und weisen eine einfache Erzählstruktur auf (vgl. Gausche et al. 2014, S. 131).

Kriterien

1. **Textlänge:** Der Text besteht aus maximal 10 Sätzen.
2. **Satzlänge:** Jeder Satz hat jeweils maximal 10 Wörtern.
→ **Ausnahme:** Pro 3 Sätze darf 1 Satz aus bis zu 15 Wörtern bestehen, falls sonst alle Kriterien ohne Ausnahme erfüllt werden.
3. **Satzstruktur:** Maximal 50% der Sätze enthalten einen Nebensatz. Kein Satz enthält mehr als einen Nebensatz. Kein Satz enthält wörtliche Rede.
4. **Zeitstruktur:** Der Text enthält nur die Zeitformen: Präsens, Perfekt und Präteritum.
5. **Wortlänge:** Jedes Wort hat maximal 5 Silben.
→ **Ausnahme:** Pro 20 Wörter darf 1 Wort-Typ bis zu 7 Silben aufweisen, falls ansonsten alle Kriterien ohne Ausnahmen erfüllt werden.

Beispieltext

Gleich treffe ich meine Freundin Marie. Sie **war** heute in der Schule ganz traurig. Gestern **ist** ihre Katze weggelaufen. Die Katze **heißt** Minka. Sie **ist** erst ein halbes Jahr alt. Nun **macht** sich Marie große Sorgen um die Katze. Marie **hofft**, dass der Katze nichts **passiert ist**. Oder dass sie sich **verlaufen hat**. Vielleicht **ist** die Katze mittlerweile schon wieder **zurückgekommen**. Ich **drücke** Marie die Daumen!

Text von Theresa Geppert.

1.3 Kriterien für Alpha-Level 5

Definition

Bei Alpha-Level 5 handelt es sich um eine zusätzliche Ausdifferenzierung der Lesekompetenzfassung durch Kretschmann & Wieken (2010). Auf dieser Ebene sind Lernende so weit literarisiert, dass sie längere, mittelschwere Texte sinnverstehend erfassen können (vgl. Kretschmann & Wieken 2010, S. 239). Aus mittelschweren Texten, die über erläuternde Bilder verfügen, können zudem direkt sowie indirekt enthaltene Textinformationen entnommen werden.

Kriterien

1. **Textlänge:** Der Text besteht aus maximal 15 Sätzen.
2. **Satzlänge:** Jeder Satz hat jeweils maximal 12 Wörtern.
→ **Ausnahme:** Pro 3 Sätze darf 1 Satz aus bis zu 15 Wörtern bestehen, falls sonst alle Kriterien ohne Ausnahmen erfüllt werden.
3. **Satzstruktur:** Maximal 50% der Sätze enthalten einen Nebensatz. Kein Satz enthält mehr als einen Nebensatz. Kein Satz enthält wörtliche Rede.
4. **Zeitstruktur:** Der Text enthält nur die Zeitformen: Präsens, Perfekt, Präteritum und Futur I.
5. **Wortlänge:** Die Wortlänge in Silben ist beliebig.

Beispieltext

Ein **Asylbewerber** ist eine Person. Ein Asylbewerber sucht Schutz in einem anderen Land. Asylbewerber wollen in einem sicheren Land leben. So lange bis es in ihrem eigenen Land wieder sicher ist. [Ein Asylbewerber stellt einen Asylantrag in einem neuen Land, **in dem er dann leben wird.**] [Nachdem er den Antrag **gestellt hat, wird entschieden**, ob er in dem Land bleiben darf.] Das **Asylantragsverfahren** dauert oft sehr lange. In Deutschland gibt es verschiedene Gesetze und Regeln. [Durch die Gesetze **wird auch bestimmt**, ob ein Asylbewerber im Land bleiben darf oder nicht.] Zum Beispiel durch **Paragraph** Nummer 16a im Grundgesetz. Dort steht: Politisch Verfolgte genießen **Asylrecht**. Alle Anträge **werden** für jeden Einzelfall **geprüft**. Danach **wird entschieden**, ob ein Asylbewerber im Land bleiben darf.

Leicht abgewandelt aus dem Hurraki Wikipedia Eintrag für „Asylbewerber“.²

²www.hurraki.de/wiki/Asylbewerber, zuletzt aufgerufen am 16.04.18.

1.4 Kriterien für Alpha-Level 6

Definition

Bei Alpha-Level 6 handelt es sich um eine zusätzliche Ausdifferenzierung der Lesekompetenzerfassung durch Kretschmann & Wieken (2010). Auf dieser Ebene sind Lernende zunehmend darin geübt, mittel-schwere, zusammenhängende Texte mit ansteigender Länge zu lesen und die Inhalte sowie den gesamten Textzusammenhang ohne zusätzliche Visualisierung der Textinformationen zu verstehen (vgl. Kretschmann & Wieken 2010, S. 240).

Kriterien

1. **Textlänge:** Der Text besteht aus maximal 20 Sätzen.
2. **Satzlänge:** Jeder Satz hat jeweils maximal 12 Wörtern.
→ **Ausnahme:** Pro 3 Sätze darf 1 Satz aus bis zu 15 Wörtern bestehen, falls sonst alle Kriterien ohne Ausnahmen erfüllt werden.
3. **Satzstruktur:** Der Text kann mehr als einen Nebensatz und wörtliche Rede enthalten.
4. **Zeitstruktur:** Der Text enthält nur die Zeitformen: Präsens, Perfekt, Präteritum und Futur I.
5. **Wortlänge:** Die Wortlänge in Silben ist beliebig.

Texte, die in einem oder mehreren Punkten diese sprachlichen Kriterien überschreiten, liegen außerhalb der Alpha-Level Kategorisierung.

Beispieltext

Die Mietpreisbremse ist ein Gesetz. **Darin steht, dass eine Wohnung nur ein bestimmtes Geld kosten darf, damit die Miete nicht zu teuer ist.** Die Miete ist das Geld, das für die Wohnung bezahlt werden muss. Die Mietpreisbremse ist eine Abmachung, damit jemand nicht zu viel Miete zahlen muss. Wie viel Geld das ist, wird ausgerechnet. Durch den Mietspiegel wird das ausgerechnet. Die Miete darf nicht viel mehr kosten als bei den Nachbarn. Genauer: die Miete darf nicht mehr als 10 Prozent teurer sein. Die Abmachung für das Gesetz gibt es seit dem 5. März 2015. Wenn der Bundespräsident das Gesetz unterschreibt, ist das Gesetz ab dem 1. Juli 2015 gültig. Das heißt: ab diesem Tag muss man sich an das Gesetz halten. Der Vorschlag für die Mietpreisbremse kam von der Partei SPD. **Sie gilt nicht überall, sondern nur, wo Wohnungen fehlen und viele Menschen eine Wohnung suchen.** Wo die Mietpreisbremse gültig ist, bestimmen die Landesminister. Bei neu gebauten Wohnungen gilt die Mietpreisbremse auch nicht. **„Die Mieten sind immer noch zu teuer!“**, beklagen sich trotzdem viele Mieter.

Leicht abgewandelt aus dem Hurraki Wikipedia Eintrag für „Mietpreisbremse“.³

³www.hurraki.de/wiki/Mietpreisbremse, zuletzt aufgerufen am 16.04.18.

2 Level-Übergreifende Erläuterung der einzelnen Kriterien

2.1 Textlänge

Kriterienübersicht

Alpha 3 Der Text besteht aus maximal 5 Sätzen.

Alpha 4 Der Text besteht aus maximal 10 Sätzen.

Alpha 5 Der Text besteht aus maximal 15 Sätzen.

Alpha 6 Der Text besteht aus maximal 20 Sätzen.

Während innerhalb des DVV-Rahmencurriculums *Lesen* (Gausche et al. 2014, S. 38ff) keine konkrete Empfehlung zur Limitierung der Textlänge für Alpha-Level 3 erfolgt, wird die Anzahl der Sätze pro Text in den *Kann-Beschreibungen* von Kretschmann & Wieken (2010, S. 237) auf zwei Sätze beschränkt. Innerhalb der Beispielmateriale für Alpha-Level 3, die Gausche et al. (2014, S. 98ff) bereitstellen, sowie in jenen, die im Projekt *lea* zur Diagnose der Alpha-Levels für die Dimension *Lesen* entwickelt wurden, sind jedoch ebenfalls Texte mit einer größeren Gesamtanzahl von Sätzen enthalten.⁴ Für Alpha-Level 3 soll daher an dieser Stelle von einem erhöhten Satzlimit von maximal 5 Sätzen ausgegangen werden.

Für Alpha-Level 4 wird von Kretschmann & Wieken (2010, S. 238) eine maximale Anzahl von 8 Sätzen pro Text empfohlen. Gleichmaßen wie bei Alpha-Level 3 umfassen die innerhalb des Projekts *lea* entwickelten Aufgaben sowie die im DVV-Rahmencurriculum (Gausche et al. 2014) enthaltenen Aufgabenblätter für Alpha-Level 4 Texte, die diese Satzbegrenzung überschreiten. Aus diesem Grund soll für Alpha-Level 4 von einem leicht erhöhten Satzlimit von 10 Sätzen ausgegangen werden.

Weiterhin wird basierend auf den Beschreibungen von Kretschmann & Wieken (2010, S. 239f) für Alpha-Level 5 eine maximale Textlänge von 15 Sätzen sowie von 20 Sätzen für Alpha-Level 6 angenommen.

Linguistische Einheiten

Die primäre linguistische Einheit, die für das Textlängenkriterium relevant ist, ist der **Satz**. Sätze werden entweder durch Textgrenzen beendet oder durch satzbeendende Zeichen.

- Satzbeendende Zeichen sind: . ? !
- Nicht als satz-beendend gelten: , ; ... : - /

Auch Satzäquivalente fallen unter diesen Satzbegriff. Eine ausführlichere Darstellung des hier angenommenen Satzbegriffes findet sich in Abschnitt 3, Seite 20.

Ebenfalls findet sich dort (Abschnitt 3, Seite 20) die Definition der zweiten für dieses Kriterium relevanten linguistischen Einheit **Text**. Wir definieren Texte als lineare Aneinanderreihung von zusammenhängend zu lesenden Texten. Text-Überschriften werden nicht als Teil des Textes gewertet, da sie für das unmittelbare Verständnis des Textes in der Regel irrelevant sind. Überschriften sind daher für die Erhebung der Textlänge nicht relevant.

⁴<http://blogs.epb.uni-hamburg.de/lea/aufgaben-lesen/>

Beispiele

Die folgenden Beispiele illustrieren die Beschränkung der Textlänge exemplarisch für Alpha-Level 3 und 4 anhand ausgewählter Ausschnitte aus den in Abschnitt 1 eingeführten Beispieltextrn für diese Levels. Zur besseren Erkennbarkeit wurden die Sätze in den Textbeispielen durch eckige Klammern voneinander abgetrennt und fortlaufend nummeriert. In Abschnitt 1 sind auch für die anderen Alpha-Levels Beispieltextrn aufgeführt, die das entsprechende Textlängenkriterium veranschaulichen.

- (1) a. [Ein Maler malt sein Bild auf der Leinwand.]₁ [Das ist ein Stoff, der sehr fest ist]₂ [Der Stoff ist aus Leinen.]₃ [Oder: Das ist eine große weiße Fläche vorne im Kino.]₄ [Darauf wird der Film gezeigt.]₅
- b. [Gleich treffe ich meine Freundin Marie.]₁ [Sie war heute in der Schule ganz traurig.]₂ [Gestern ist ihre Katze weggelaufen.]₃ [Die Katze heißt Minka.]₄ [Sie ist erst ein halbes Jahr alt.]₅ [Nun macht sich Marie große Sorgen um die Katze.]₆ [Marie hofft, dass der Katze nichts passiert ist.]₇ [Oder dass sie sich verlaufen hat.]₈ [Vielleicht ist die Katze mittlerweile schon wieder zurückgekommen.]₉ [Ich drücke Marie die Daumen!]₁₀

Beispiel 1a zeigt einen Text, der insgesamt aus 5 Sätzen besteht. Er liegt somit innerhalb der Satzlimitierung für Alpha-Level 3, erfüllt auch alle zusätzlichen Kriterien und kann somit diesem Alpha-Level zugeordnet werden. Der Text in Beispiel 1b enthält hingegen insgesamt 10 Sätze und überschreitet somit die Textlängenbegrenzung für Alpha-Level 3. Da er das Textlängenkriterium von Alpha-Level 4 sowie alle weiteren Kriterien erfüllt, stellt er ein Beispiel für Alpha-Level 4 dar.

Sonderfälle

Sonderfall 1. Bei Leichter Sprache wird Interpunktion nicht zwangsläufig entsprechend der gängigen Sprachnorm verwendet. Insbesondere werden Punkte statt Kommas verwendet und Nebensätze oder Auflistungen wie eigenständige Sätze behandelt. Ähnliches gilt auch für Doppelpunkte, nach denen groß weitergeschrieben wird, um den Beginn eines neuen Satzes zu markieren. Da diese von den Regeln der Standardsprache abweichende Nutzung jedoch nicht systematisch durchgeführt wird, ist es nicht möglich, diesen spezifischen Gebrauch in unsere Regeln aufzunehmen. Bei der Identifikation von Sätzen in Texten in Leichter Sprache kann es daher zu Annotationen kommen, die dem eigenen Satzverständnis widersprechen. Beispiel 2 zeigt dies anhand eines Beispieltextrns.

- (2) [Das sind Menschen aus anderen Ländern.]₁ [Zum Beispiel aus dem Land Syrien.]₂ [Sie kommen nach Deutschland: Weil sie in ihrem Land nicht leben können.]₃ [Zum Beispiel: Weil dort Krieg ist.]₄ [Und viele Menschen getötet werden.]₅

Dieser Text besteht, basierend auf unseren Richtlinien, aus insgesamt 5 Sätzen. Gemäß der standard-sprachlichen Norm sollte Satz 2 allerdings nicht von Satz 1 abgetrennt werden. Außerdem sollte in Satz 3 der mit *weil* eingeleitete Nebensatz durch ein Komma und nicht durch Doppelpunkte abgetrennt werden. Auch Satz 4 und 5 gehören syntaktisch zu dem Hauptsatz in Satz 3. Aufgrund der Interpunktion werden die Satzgrenzen jedoch wie im Beispiel angegeben gesetzt.

Sonderfall 2. Um die Gesamtanzahl der Sätze eines Textes zu ermitteln und basierend darauf Rückschlüsse auf das zu Grunde liegende Alpha-Level zu ziehen, werden lediglich die Sätze innerhalb des Fließtextes, nicht jedoch die Überschriften gezählt.

- (3) **[Was sind eigentlich Journalisten?]** [Sie berichten über viele Sachen aus der ganzen Welt.]₁ [Sie schreiben Berichte.]₂ [Sie machen Filme.]₃ [Oder sie sprechen im Radio darüber.]₄ [Und sie sind bei den Ereignissen selbst vor Ort.]₅⁵

⁵Leicht abgewandelt aus dem mdr Wörter-Buch-Eintrag in Leichter Sprache zu „Journalist“, <https://www.mdr.de/nachrichten-leicht/woerterbuch/glossar-journalisten-100.html>, zuletzt aufgerufen am 17.04.18.

Beispieltext 3 besteht aus der Überschrift „*Was sind eigentlich Journalisten?*“ sowie aus 5 Sätzen. Insgesamt enthält der Text somit eine Gesamtanzahl von 5 Sätzen. Da auch alle weiteren Kriterien erfüllt sind, ist der Text Alpha-Level 3 zuzuordnen. In Fällen, in denen nicht klar erkenntlich ist, ob es sich um eine Überschrift handelt oder nicht, wird ein Textbestandteil nicht als Überschrift gewertet.

2.2 Satzlänge

Kriterienübersicht

Alpha 3 Jeder Satz hat jeweils maximal 10 Wörter.

→ **Ausnahme:** Pro 3 Sätze darf 1 Satz aus bis zu 15 Wörtern bestehen, falls sonst alle Kriterien ohne Ausnahme erfüllt werden.

Alpha 4 Satzlängenkriterium und Ausnahmenregelung für Alpha 4 sind identisch zu Alpha 3.

Alpha 5 Die Sätze in dem Text enthalten jeweils maximal 12 Wörter.

→ **Ausnahme:** Pro 3 Sätze darf 1 Satz aus bis zu 15 Wörtern bestehen, falls sonst alle Kriterien ohne Ausnahme erfüllt werden.

Alpha 6 Satzlängenkriterium und Ausnahmenregelung für Alpha 6 sind identisch zu Alpha 5.

Diese Begrenzungen wurden aus den Empfehlungen des DVV-Rahmencurriculums *Lesen* (Gausche et al. 2014, S. 97) abgeleitet, in welchem die Anzahl der Wörter pro Satz für Alpha-Level 3 auf 10 Wörter beschränkt wird. Da für Alpha-Level 4 keine abweichenden Spezifizierungen erfolgen (vgl. Gausche et al. 2014, S. 134ff), wird die Regel aus Alpha-Level 3 übernommen. Für Alpha-Level 5 und 6 wird basierend auf den Beschreibungen von Kretschmann & Wieken (2010, S. 239f) eine maximale Satzlänge von 12 Wörtern angenommen.

Linguistische Einheiten

Die zentrale linguistische Einheit zur Erhebung dieses Kriterium ist das **Wort**. Wörter werden als kontinuierliche Zeichenketten aus Buchstaben, Nummern oder Sonderzeichen definiert, die durch Leerzeichen getrennt werden, und eigenständige, bedeutungstragende sprachliche Einheiten sind. Eine ausführlichere Darstellung des Wortbegriffes findet sich in Abschnitt 3, Seite 21.

Ein zweite wichtige linguistische Einheit zur Erhebung dieses Kriterium ist der **Satz**. Sätze werden entweder durch Textgrenzen beendet oder durch satzbeendende Zeichen.

- Satzbeendende Zeichen sind: **. ? !**
- Nicht als satz-beendend gelten: **, ; ... : - /**

Auch Satzäquivalente fallen unter diesen Satzbegriff. Eine ausführlichere Darstellung des hier angenommenen Satzbegriffes findet sich in Abschnitt 3, Seite 20.

Das Kriterium zur Satzlänge findet ausschließlich innerhalb der **Text**grenzen Anwendung. Überschriften werden nicht als Teil des Textes gewertet, da sie für das unmittelbare Verständnis des Textes in der Regel irrelevant sind. Überschriften sind daher bei der Erhebung der Satzlänge nicht relevant.

Beispiele

Die folgenden Beispiele in 4 illustrieren die Beschränkung der Satzlänge über Alpha-Levels hinweg anhand ausgewählter Ausschnitte aus den in Abschnitt 1 eingeführten Beispieltextrn und aus einem Eintrag aus dem MDR Lexikon in Leichter Sprache. Zur besseren Erkennbarkeit wurden die Wörter in den Beispielen 4a und 4b durch eckige Klammern separiert und fortlaufend nummeriert. In Beispiel 4c wurde zur Bewahrung der Übersichtlichkeit hierauf verzichtet.

- (4) a. [Ein]₁ [Maler]₂ [malt]₃ [sein]₄ [Bild]₅ [auf]₆ [einer]₇ [Leinwand]₈.
 b. [So]₁ [lange]₂ [bis]₃ [es]₄ [in]₅ [ihrem]₆ [eigenen]₇ [Land]₈ [wieder]₉ [sicher]₁₀ [ist]₁₁.
 c. Der Vorteil ist: Mit einem Auto kommen die Menschen schnell von einem Ort zu einem anderen Ort.⁶

Beispiel 4a zeigt den ersten Satz aus dem Beispieltext für Alpha-Level 3. Der Satz besteht aus 8 Wörtern und liegt damit innerhalb der Limiterung für Alpha 3 oder Alpha 4. Beispiel 4b hingegen zeigt den vierten Satz aus dem Beispieltext für Alpha 5. Dieser überschreitet mit 11 Wörtern das Limit für Alpha-Level 3 und 4. Der dazugehörige Text liegt daher über diesen Alpha-Levels, sofern nicht die Ausnahmeregelung greift, siehe unten. Er liegt jedoch innerhalb der Grenzen für Alpha-Level 5 und 6. Beispiel 4c schließlich zeigt einen Satz, der 17 Wörter beinhaltet und damit außerhalb der Kriterien für die Alpha-Level 3, 4, 5 und 6 sowie außerhalb der durch die Ausnahmeregelung gegebenen Toleranzgrenze liegt. Ein Text mit diesem oder einem ähnlich langen Satz liegt damit automatisch außerhalb der Alpha-Kriterien.

Ausnahmeregelung

Das Kriterium zur Satzlänge beinhaltet eine Ausnahme, die Anwendung findet, wenn alle anderen Alpha-Level Kriterien ausnahmslos erfüllt werden. Diese Ausnahme speist sich aus praktischen Erfahrungen mit Alpha-Texten – beispielsweise aus Kretschmann & Wieken (2010) und Gausche et al. (2014) – und erlaubt, dass einer von drei Sätzen in einem Text aus bis zu 15 Wörtern bestehen darf. Beispiel 5 verdeutlicht dies. Zur besseren Übersicht sind hier Sätze durch eckige Klammern voneinander abgetrennt und durchgehend nummeriert.

- (5) [Ein Dirigent ist der Chef von einem Orchester.]₁ [Ein Dirigent steht immer vor seinem Orchester.]₂
 [Und der Dirigent zeigt den Musikern in seinem Orchester zum Beispiel: So müsst ihr spielen.]₃
 [Oder: Jetzt müsst ihr eine Pause machen.]₄⁷

Der Text besteht insgesamt aus vier Sätzen. Mit einer Gesamtanzahl von 15 Wörtern überschreitet lediglich Satz 4 die auf Alpha-Level 3 maximal zulässige Satzlänge von 10 Wörtern. Der Text kann daher Alpha-Level 3 zugeordnet werden, sofern alle weiteren Kriterien für Alpha-Level 3 ohne weitere Ausnahmen erfüllt werden. Die Ausnahme-Regel kann äquivalent für andere Alpha-Level angewandt werden.

Sonderfälle

Derzeit sind keine Sonderfälle in diesem Kontext bekannt.

⁶Aus dem mdr Wörterbuch-Eintrag in Leichter Sprache zu „Vorteil“, <https://www.mdr.de/nachrichten-leicht/woerterbuch/glossar-journalisten-100.html>, zuletzt aufgerufen am 17.04.18.

⁷Aus dem ndr Wörterbuch-Eintrag in Leichter Sprache zu „Dirigent“, https://www.ndr.de/fernsehen/service/leichte_sprache/Dirigent,dirigentleichtesprache100.html, zuletzt aufgerufen am 17.04.18.

2.3 Satzstruktur

Kriterien-Übersicht

Alpha 3 Maximal 50% der Sätze enthalten einen Nebensatz. Kein Satz enthält mehr als einen Nebensatz. Kein Satz enthält wörtliche Rede.

Alpha 4 Das Satzstrukturkriterium für Alpha 4 ist identisch zu Alpha 3.

Alpha 5 Das Satzstrukturkriterium für Alpha 5 ist identisch zu Alpha 3.

Alpha 6 Der Text kann mehr als einen Nebensatz und wörtliche Rede enthalten.

Den Empfehlungen von Kretschmann & Wieken (2010, S. 237ff) zu Folge können die in den Texten für Alpha-Level 3, 4 und 5 enthaltenen Sätze sowohl Haupt- als auch Nebensatzkonstruktionen enthalten. Aus dem DVV-Rahmencurriculum *Lesen* (Gausche et al. 2014) können weitere Hinweise zur Beschränkung der Komplexität der Satzstruktur für die einzelnen Alpha-Levels abgeleitet werden. Da es auf Alpha-Level 3 zunächst um das Erlesen von Sätzen bzw. auf Alpha-Level 4 und 5 um das Lesen von Texten mit einer einfachen Erzählstruktur geht (vgl. Gausche et al. 2014, S. 96ff; Kretschmann & Wieken 2010, S. 239), ist innerhalb der Texte von einfachen Satzkonstruktionen auszugehen. Für Alpha-Level 3, 4 und 5 wird daher angenommen, dass maximal jeder zweite Satz innerhalb des Textes einen Nebensatz enthält. Für Alpha-Level 6 legen Kretschmann & Wieken (2010, S.240) fest, dass ebenfalls „Schachtelsätze“ sowie wörtliche Rede innerhalb der Texte zulässig sind. Im Umkehrschluss leiten wir hieraus ab, dass diese Konstruktions-typen für niedrigere Alpha-Level nicht angemessen sind.

Linguistische Einheiten

Die zentralen linguistischen Einheiten zur Anwendung dieses Kriteriums sind die folgenden: **Direkte Rede** (auch: wörtliche Rede) ist eine direkte Redewiedergabe, die durch Anführungszeichen kenntlich gemacht wird und bei der ein Fremdsprecher eindeutig kenntlich gemacht wird, siehe auch Abschnitt 3, Seite 19.

Ein **Nebensatz** lässt sich daran erkennen, dass das **finite Verb** am Ende der Satzkonstruktion steht.⁸ Nebensätze können nicht alleine stehen, sondern hängen von einem anderen Nebensatz oder einem Hauptsatz ab. Weitere Informationen zur Nebensatzdefinition finden sich in Abschnitt 3, Seite 19.

Ein **Schachtelsatz** ist ein Satz mit mehrfach untergeordneten Nebensätzen oder mit mehreren nebengeordneten Nebensätzen, siehe Abschnitt 3, Seite 20.

Ein **Satz** wird entweder durch Textgrenzen beendet oder durch satzbeendende Zeichen.

- Satzbeendende Zeichen sind: **. ? !**
- Nicht als satz-beendend gelten: **, ; ... : – /**

Auch Satzäquivalente fallen unter diesen Satzbegriff. Eine ausführlichere Darstellung des hier angenommenen Satzbegriffes findet sich in Abschnitt 3, Seite 20.

Das Kriterium zur Satzstruktur findet ausschließlich innerhalb der **Textgrenzen** Anwendung. Überschriften werden nicht als Teil des Textes gewertet, da sie für das unmittelbare Verständnis des Textes in der Regel irrelevant sind. Die Satzstruktur von Überschriften ist daher für die Erhebung des Alpha-Levels eines Textes nicht relevant.

⁸Finite Verben sind Verben, die mit dem Subjekt kongruieren, siehe Abschnitt 3, Seite 19.

Beispiele

Die Beispiele in 6 illustrieren die Beschränkung der Nebensatzkonstruktionen über die verschiedenen Alpha-Levels hinweg anhand ausgewählter Ausschnitte aus den in Abschnitt 1 eingeführten Beispieltex-ten. Zur besseren Erkennbarkeit sind die Nebensätze und direkte Rede fett markiert. Die Schachtelung von Nebensätzen wird – wo erforderlich – überdies mit eckigen Klammern verdeutlicht. Auf eine separate Kennzeichnung von Satz- und Wortgrenzen wurde der Übersichtlichkeit halber verzichtet.

- (6) a. Gleich treffe ich meine Freundin Marie. Sie war heute in der Schule ganz traurig. Gestern ist ihre Katze weggelaufen. Die Katze heißt Minka. Sie ist erst ein halbes Jahr alt. Nun macht sich Marie große Sorgen um die Katze. Marie hofft, **dass der Katze nichts passiert ist**. Oder dass sie sich verlaufen hat. Vielleicht ist die Katze mittlerweile schon wieder zurückgekommen. Ich drücke Marie die Daumen!
- b. Die Mietpreisbremse ist ein Gesetz. Darin steht, **[dass eine Wohnung nur ein bestimmtes Geld kosten darf, [damit die Miete nicht zu teuer ist]]**. Die Miete ist das Geld, **das für die Wohnung bezahlt werden muss**. Die Mietpreisbremse ist eine Abmachung, **damit jemand nicht zu viel Miete zahlen muss. Wie viel Geld das ist**, wird ausgerechnet. Durch den Miet-
spiegel wird das ausgerechnet. Die Miete darf nicht viel mehr kosten als bei den Nachbarn. Genauer: die Miete darf nicht mehr als 10 Prozent teurer sein. Die Abmachung für das Gesetz gibt es seit dem 5. März 2015. **Wenn der Bundespräsident das Gesetz unterschreibt**, ist das Gesetz ab dem 1. Juli 2015 gültig. Das heißt: ab diesem Tag muss man sich an das Gesetz halten. Der Vorschlag für die Mietpreisbremse kam von der Partei SPD. [...] Bei neu gebauten Wohnungen gilt die Mietpreisbremse auch nicht. **„Die Mieten sind immer noch zu teuer!“**, beklagen sich trotzdem viele Mieter.

Beispiel 6a zeigt einen Text, in dem 1 Nebensatz auf insgesamt 10 Sätze kommt. Dieser Text kann daher einem niedrigeren Alpha-Level als Alpha-Level 6 zugeordnet werden. Da der Text aufgrund seiner Länge nicht für Alpha-Level 3 geeignet ist, sonst jedoch alle Kriterien für Alpha-Level 4 erfüllt, kann er Alpha-Level 4 zugeordnet werden.

Beispiel 6b zeigt einen Textausschnitt, in dem 6 von 13 Sätzen einen Nebensatz enthalten, also weniger als 50%. Allerdings beinhaltet der Textausschnitt einen Satz, in dem ein Nebensatz einem anderen Nebensatz untergeordnet ist, wie durch die eckigen Klammern kenntlich gemacht wird. Eine solche Schachtelung ist nicht mit den Kriterien für die Alpha-Level 3, 4 und 5 vereinbar. Beispiel 6b liegt daher über den Alpha-Levels 3, 4 und 5 und ist Alpha-Level 6 zuzuordnen.

Sonderfälle

Derzeit sind keine Sonderfälle in diesem Kontext bekannt.

2.4 Zeitstruktur

Kriterien-Übersicht

Alpha 3 Der Text ist durchgängig im Präsens geschrieben.

Alpha 4 Der Text enthält nur die Zeitformen: Präsens, Perfekt und Präteritum.

Alpha 5 Der Text enthält nur die Zeitformen: Präsens, Perfekt, Präteritum und Futur I.

Alpha 6 Das Zeitstrukturkriterium für Alpha 6 ist identisch zu Alpha 5.

Die Kriterien zur Zeitstruktur wurden aus den Empfehlungen von Kretschmann & Wieken (2010, S. 237ff) abgeleitet. Für Alpha-Level 4 werden Präsens, Perfekt und Präteritum als zulässige Zeitformen ausgewiesen (vgl. Kretschmann & Wieken 2010, S. 238). Da für Alpha-Level 3 keine äquivalente Spezifizierung zulässiger Zeitformen vorliegt, wird davon ausgegangen, dass Alpha-Level 3 Texte nur das Präsens verwenden dürfen, während auf Alpha-Level 4 zusätzlich auch Präteritum und Perfekt verwendet werden können. Auf Alpha-Level 5 und 6 wird Futur I als zusätzliches Tempus zugelassen (vgl. Kretschmann & Wieken 2010, S. 239). Plusquamperfekt und Futur II sind für kein Alpha-Level zulässig.

Linguistische Einheiten

Die für die Anwendung dieses Kriterium notwendigen linguistischen Begriffe sind die Zeitformen des Deutschen. **Präsens** ist die synthetisch gebildete Gegenwartsform, siehe Abschnitt 3, Seite 20.⁹

Präteritum ist die synthetisch gebildete Vergangenheitsform, siehe Abschnitt 3, Seite 20.¹⁰

Perfekt ist die periphrastisch gebildete vollendete Vergangenheitsform, siehe Abschnitt 3, Seite 19.¹¹

Futur I ist die periphrastisch gebildete Zukunftsform, siehe Abschnitt 3, Seite 19.¹²

Futur II ist die periphrastisch gebildete Form der abgeschlossenen Zukunft, siehe Abschnitt 3, Seite 19.¹³

Plusquamperfekt ist die periphrastisch gebildete Form der Vorvergangenheit, siehe Abschnitt 3, Seite 19.¹⁴

Das Kriterium zur Zeitstruktur findet ausschließlich innerhalb der **Text**grenzen Anwendung. Überschriften werden nicht als Teil des Textes gewertet, da sie für das unmittelbare Verständnis des Textes in der Regel irrelevant sind. Die Zeitstruktur von Überschriften ist daher für die Erhebung des Alpha-Levels eines Textes nicht relevant.

Beispiele

Die Beispiele in 7 illustrieren die Beschränkung der temporalen Komplexität über die verschiedenen Alpha-Levels hinweg anhand ausgewählter Ausschnitte aus den in Abschnitt 1 eingeführten Beispieltexten. Zur besseren Erkennbarkeit des vorliegenden Tempus wurden alle zum Ausdruck des Tempus relevanten Verben fett markiert.

⁹Siehe auch <https://grammis.ids-mannheim.de/kontrastive-grammatik/4163>

¹⁰Siehe auch <https://grammis.ids-mannheim.de/kontrastive-grammatik/4164>

¹¹Siehe auch <https://grammis.ids-mannheim.de/kontrastive-grammatik/4166>

¹²Siehe auch <https://grammis.ids-mannheim.de/kontrastive-grammatik/4165>

¹³Siehe auch <https://grammis.ids-mannheim.de/kontrastive-grammatik/4169>

¹⁴Siehe auch <https://grammis.ids-mannheim.de/kontrastive-grammatik/4168>

- (7) a. Leinwand
Ein Maler **malt** sein Bild auf der Leinwand. Das **ist** ein Stoff, der sehr fest **ist**. [...]
- b. Gleich **treffe** ich meine Freundin Marie. Sie **war** heute in der Schule ganz traurig. Gestern **ist** ihre Katze **weggelaufen**. [...] Marie **hofft**, dass der Katze nichts **passiert ist**. Oder dass sie sich **verlaufen hat**. [...]

Beispieltext 7a enthält ausschließlich die Zeitform Präsens und erfüllt damit das Zeitstruktur-Kriterium für Alpha-Level 3 Texte. Da er auch alle weiteren Kriterien für Alpha-Level 3 erfüllt, kann er diesem Level zugeordnet werden.

Beispieltext 7b enthält mehrere Zeitformen: Präsens (treffe, hofft), Präteritum (war) und Perfekt (ist ... weg-gelaufen, passiert ist, verlaufen hat). Die zeitliche Struktur des Textes ist daher zu komplex für Alpha-Level 3 Texte. Da der Text kein Futur I enthält und auch kein Plusquamperfekt oder Futur II, eignet sich die zeitliche Struktur für einen Alpha-Level 4 Text.

Sonderfälle

Derzeit sind keine Sonderfälle in diesem Kontext bekannt.

2.5 Wortlänge

Kriterien-Übersicht

Alpha 3 Jedes Wort hat maximal 3 Silben.

→ **Ausnahme:** Pro 20 Wörter darf 1 Wort-Typ bis zu 5 Silben aufweisen, falls ansonsten alle Kriterien ohne Ausnahme erfüllt werden.

Alpha 4 Jedes Wort hat maximal 5 Silben.

→ **Ausnahme:** Pro 20 Wörter darf 1 Wort-Typ bis zu 7 Silben aufweisen, falls ansonsten alle Kriterien ohne Ausnahme erfüllt werden.

Alpha 5 Die Wortlänge in Silben ist beliebig.

Alpha 6 Das Wortlängenkriterium für Alpha 6 ist identisch zu Alpha 5.

Die Kriterien zur Wortlänge entstammen den Empfehlungen des DVV-Rahmencurriculums *Lesen* (Gausche et al. 2014, S. 97) sowie den *Kann*-Beschreibungen von Kretschmann & Wieken (2010, S. 237). Für Alpha-Level 3 wird in diesen die zulässige Wortlänge auf maximal drei Silben limitiert. Für Alpha-Level 4 wird die Wortlängenbegrenzung auf maximal 5 Silben pro Wort erweitert. Für Alpha-Level 5 und 6 ist die zulässige Wortlänge unbeschränkt (vgl. Kretschmann & Wieken 2010, S. 238ff).

Linguistische Einheiten

Die zentrale linguistische Einheit für die Anwendung dieses Kriteriums ist die **Silbe**. Eine Schreibsilbe besteht notwendigerweise aus mindestens einem Vokalgraphem (**a, e, i, o, u, ä, ö, ü, ie, au, äu, eu, oi, ai, ui, aa, ee, uu, oo, ...**) und gegebenenfalls aus einem oder mehreren Konsonantengraphemen, siehe auch Abschnitt 3, Seite 20. Da die genaue Position der Silbengrenze für die Bestimmung der Anzahl der Silben nicht relevant ist, genügt es die Anzahl der Schreibsilben eines Wortes mit der Anzahl der Vokalgrapheme gleichzusetzen.

Bei der Ausnahmeregelung bezüglich der Wortlänge wird zudem der Begriff **Wort-Typ** relevant. (Wort-)Typen sind die einzigartigen Wörter in einem Text.¹⁵ Während identische, wort-wertige Zeichenfolgen als zwei verschiedene Wörter zählen, zählen sie als der selbe Wort-Typ, siehe auch Abschnitt 3, Seite 22. Dies wird im Folgenden ebenfalls an einem Beispiel illustriert (siehe Beispiel 9).

Die letzte wichtige linguistische Einheit zur Erfassung dieses Kriteriums ist das **Wort**. Wörter werden als kontinuierliche Zeichenketten aus Buchstaben, Ziffern oder Sonderzeichen definiert, die durch Leerzeichen getrennt werden und eigenständige, bedeutungstragende sprachliche Einheiten sind. Eine ausführlichere Darstellung des Wortbegriffes findet sich in Abschnitt 3, Seite 21.

Das Kriterium zur Wortlänge findet ausschließlich innerhalb der **Text**grenzen Anwendung. Überschriften werden nicht als Teil des Textes gewertet, da sie für das unmittelbare Verständnis des Textes in der Regel irrelevant sind. Die Länge von Wörtern in Überschriften ist daher für die Erhebung des Alpha-Levels eines Textes nicht relevant.

Beispiele

Die folgenden Beispiele in 8 illustrieren die Beschränkung der Wortlänge über die verschiedenen Alpha-Levels hinweg anhand ausgewählter Ausschnitte aus den in Abschnitt 1 eingeführten Beispieltexten. Zur

¹⁵Siehe für eine ausführlichere Erläuterung zu Typen und Tokens: https://de.wikipedia.org/wiki/Token_und_Type

besseren Identifikation der einzelnen Silben werden diese in allen folgenden Beispielen mit Bindestrichen voneinander getrennt.

(8) a. **Leinwand**

Ein Ma-ler malt sein Bild auf der Lein-wand. Das ist ein Stoff, der sehr fest ist. Der Stoff ist aus Lei-nen. Oder: Das ist eine gro-ße wei-ße Flä-che vor-ne im Ki-no. Dar-auf wird der Film ge-zeigt.

- b. Gleich tref-fe ich mei-ne Freun-din Ma-rie. Sie war heu-te in der Schu-le ganz trau-rig. Gestern ist ih-re Kat-ze **weg-ge-lau-fen**. [...] Viel-leicht ist die Kat-ze **mitt-ler-wei-le** schon wie-der **zu-rück-ge-kom-men**. Ich drü-cke Ma-rie die Dau-men!

Beispiel 8a zeigt einen Text, in dem kein Wort mehr als drei Silben enthält. Er entspricht damit dem Wortlängenkriterium für Alpha-Level 3. Da auch alle weiteren Kriterien erfüllt sind, kann der Beispieltext Alpha-Level 3 zugeordnet werden. Im Gegensatz hierzu zeigt Beispiel 8b einen Textausschnitt, in dem 3 von 32 Wörtern aus 4 bzw. 5 Silben bestehen (fett markiert). Die Silbenbegrenzungsregel für Alpha-Level 3 wird somit überschritten. Der Text kann jedoch Alpha-Level 4 zugeordnet werden.

Ausnahmeregelung

Das Kriterium zur Wortlänge beinhaltet eine Ausnahme, die Anwendung findet, wenn alle anderen Alpha-Level Kriterien ausnahmslos erfüllt werden. Die Ausnahmeregelung erlaubt für Alpha-Level 3 und 4 alle 20 Wörter die Überschreitung der zulässigen Maximallänge um 2 Silben bei einem Wort-Typ. Die Regelung wird auf Wort-Typen und nicht auf Worte angewandt, da davon ausgegangen wird, dass die mehrfache Nennung des gleichen, zu langen Wort-Types das Leseverständnis nicht zusätzlich erschwert. Für Alpha-Level 3 Texte bedeutet dies, dass 1 Wort-Typ alle 20 Wörter aus bis zu 5 Silben bestehen darf. Bei Alpha-Level 4 Texten führt die Ausnahmeregelung dazu, dass 1 Wort-Typ alle 20 Wörter aus bis zu 7 Silben bestehen darf. Für die Alpha-Level 5 und 6 ist aufgrund der unbegrenzten Wortlänge keine Ausnahmeregelung notwendig. Die Ausnahme wurde aus den Materialien abgeleitet, die im Projekt *lea*. zur Diagnose der Alpha-Levels *Lesen* entwickelt wurden.¹⁶ In Teilen überschreiten diese ebenfalls die zulässige Maximalwortlänge der ihnen zugewiesenen Alpha-Level.

Beispiel 9 verdeutlicht die Anwendung der Ausnahmeregelung an einem leicht erweiterten Ausschnitt aus dem Text „Kochrezept“, der zur *lea*-Diagnose für die Dimension *Lesen* auf Alpha-Level 3 entwickelt wurde.¹⁷ Zur besseren Identifikation der Silbengrenzen werden Silben erneut durch Bindestriche getrennt. Wörter, welche die für Alpha-Level 3 zulässige Gesamtlänge überschreiten, werden zusätzlich fett markiert.

- (9) Re-zept für Sah-ne So-ße. Die-se Zu-ta-ten in ei-nen **mit-tel-gro-ßen** Topf ge-ben: [. . .] Al-les bei we-nig Hit-ze fünf bis acht Mi-nu-ten ko-chen las-sen, da-bei ab und zu um-rüh-ren. [Den **mit-tel-gro-ßen** Topf vom Herd neh-men.] Fer-tig!

Der Textausschnitt besteht aus insgesamt 33 Wörtern. Hiervon überschreitet der zwei Mal vorkommende Wort-Typ *mittelgroßen* mit einer Länge von 4 Silben als einziger das Wortlängenlimit von 3 Silben. Der Text erfüllt außerdem alle anderen Kriterien für Alpha-Level 3. Daher kann die Ausnahmeregelung angewandt und der Text Alpha-Level 3 zugeordnet werden. Die Anwendung der Regel auf beide Instanzen des Wort-Typs gilt als einmalige Anwendung der Ausnahmeregel. Die Ausnahme-Regel kann äquivalent für Alpha-Level 4 angewandt werden.

Sonderfälle

Sonderfall 1. Bei Begriffsdefinitionen wird die Wortlänge des erklärten Begriffs nicht berücksichtigt. Beispiel 10 zeigt einen Sonderfall, für den das Wortlängenkriterium nicht zutrifft. Es handelt sich um einen

¹⁶<http://blogs.epb.uni-hamburg.de/lea/aufgaben-lesen/>

¹⁷<http://blogs.epb.uni-hamburg.de/lea/files/2011/02/Kochrezept.pdf>, S. 266, zuletzt aufgerufen am 16.04.18. Auslassungen und nicht zum Originaltext gehörige Ergänzungen befinden sich in eckigen Klammern.

Glossar-Eintrag, dessen inhaltliches Ziel die verständliche Erklärung des Wortes *Journalisten* ist. Derartige begriffliche Erläuterungen sind ein häufiger Bestandteil von Texten, die sich an funktionale Analphabeten richten.

- (10) Was sind ei-gent-lich **Jour-na-lis-ten**? Sie be-rich-ten ü-ber vie-le Sach-en aus der gan-zen Welt. Sie schrei-ben Be-rich-te.¹⁸

Mit insgesamt 4 Silben überschreitet *Journalisten* die zulässige Silbenlänge für Alpha-Level 3 Texte. Die darauf folgende Begriffserklärung ist jedoch im Rahmen der für einen Alpha-Level 3 Text zulässigen sprachlichen Gestaltung. Da der Text zudem dazu dient, einen Begriff zu erläutern, der mutmaßlich außerhalb des Wortschatzes des Rezipienten liegt, ist es nicht zielführend, diesen Text aufgrund des erläuterten Begriffes auszuschließen. Da das erklärte Wort hinreichend durch die Erklärung substituiert wird, unterliegt es nicht der Wortlängenbeschränkung. Dieser Sonderfall gilt jedoch nur für Texte, deren einzige oder primäre Funktion die Erklärung oder Definition des fraglichen Begriffes ist. Das betrifft in der Regel insbesondere Glossen und Wörterbucheinträge. Eine Erläuterung eines Begriffes innerhalb eines Textes, dessen Funktion über die Erklärung dieses Begriffes hinaus geht, reicht nicht aus, um das Wortlängenkriterium außer Kraft zu setzen.

Die Ausnahme gilt für alle Vorkommnisse des Definitionsgegenstands in all seine Flexionsformen, also zum Beispiel auch für **Journalist**, wenn der Definitionsgegenstand **Journalisten** ist. Die Ausnahme erstreckt sich jedoch nicht auf Derivationsformen des Definitionsgegenstands (z.B. **journalistisch**) oder Komposita, die den Definitionsgegenstand enthalten (z.B. **Journalistentagung**).

Sonderfall 2. Für das Kriterium der Wortlänge zählen nur die tatsächlich ausgeschriebenen Wortbestandteile. Für Ellipsen wie **Industrie-** in Beispiel 11 heißt dies, dass der ausgelassene Bestandteil nicht für die Wortlänge mitgezählt wird.

- (11) **Industrie-** und Handels-Kammer

Das Beispiel besteht also aus drei Wörtern, die jeweils 3, 1 und 4 Silben haben

¹⁸Leicht abgewandelt aus dem mdr Wörter-Buch-Eintrag in Leichter Sprache zu „Journalist“, <https://www.mdr.de/nachrichten-leicht/woerterbuch/glossar-journalisten-100.html>, zuletzt aufgerufen am 17.04.18.

3 Glossar

Direkte Rede Bei direkter (oder wörtlicher) Rede handelt es sich um eine direkte Redewiedergabe, die durch Anführungszeichen kenntlich gemacht wird. Direkte Rede kann von einem vorangehenden oder nachstehenden Einleitungssatz begleitet werden, wie „sagte sie“, sie kann jedoch auch alleine stehen. In jedem Fall muss ein Fremdsprecher eindeutig kenntlich gemacht werden.

Finites Verb Verbform, die die grammatischen Merkmale Person (1, 2, 3), Numerus (Singular, Plural), Tempus (Präsens, Präteritum), Modus (Indikativ, Konjunktiv, Imperativ) und Genus Verbi (Aktiv/Passiv) ausdrückt.

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Definition finiter Verben.

Futur I Beschreibt als Zukunftsform zur Lesezeit / Jetzt-Zeit des Textes nachzeitige Ereignisse. Es wird gebildet aus einer flektierten Form des Verbs *werden* im Präsens sowie dem Vollverb im Infinitiv. Für Details, siehe <https://grammis.ids-mannheim.de/kontrastive-grammatik/4165>

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Definition von Futur 1.

Futur II Beschreibt als abgeschlossene Zukunftsform ein vorzeitiges Ereignis in Relation zu einer nach der Lesezeit / Jetzt-Zeit des Textes liegenden Zeit. Es wird gebildet aus einer flektierten Form des Verbs *werden* im Präteritum sowie dem Vollverb im Infinitiv. Für Details, siehe <https://grammis.ids-mannheim.de/kontrastive-grammatik/4169>

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Definition von Futur 2.

Hauptsatz Alleinstehender Satz oder übergeordneter Satz in einer Satzstruktur mit mehreren Sätzen. Hauptsätze können sinnhaft alleinstehend gelesen werden und weisen keine Verbletzt-Stellung auf. Das finite Verb befindet sich also nicht an letzter Stelle im Satz. Meistens befindet sich das finite Verb an erster oder zweiter Stelle.

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Hauptsatz-Definition.

Kompositum Eine Wortzusammensetzung bestehend aus mindestens zwei voneinander unabhängig bestehenden Wörtern.

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Definition von Komposita.

Nebensatz Nebensätze können nicht alleine stehen, sondern hängen von einem anderen Neben- oder einem Hauptsatz ab. Nebensätze haben in der Regel eine Verbletzt-Stellung. Das finite Verb befindet sich also am Ende der Satzkonstruktion.

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Nebensatz-Definition.

Perfekt Zeitform der vollendeten Gegenwart. Beschreibt als Vergangenheitsform zur Lesezeit / Jetzt-Zeit des Textes vorzeitige Ereignisse. Perfekt wird aus einer flektierten Form der Verben *sein* oder *haben* im Präsens sowie dem Vollverb im Partizip II gebildet. Für Details, siehe <https://grammis.ids-mannheim.de/kontrastive-grammatik/4166>

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Perfekt-Definition.

Plusquamperfekt Beschreibt als Vorvergangenheit ein nachzeitiges Ereignis in Relation zu einer nach der Lesezeit / Jetzt-Zeit des Textes liegenden Zeit. Es wird gebildet aus einer flektierten Form der Verben *sein* oder *haben* im Präteritum sowie dem Vollverb im Partizip II. Für Details, siehe <https://grammis.ids-mannheim.de/kontrastive-grammatik/4168>

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Definition von Plusquamperfekt.

Präsens Beschreibt als unmarkiertes Tempus in der Regel die Gegenwartszeit, kann jedoch auch andere Zeitbezüge ausdrücken. Das Präsens wird aus einem finiten Verb mit Präsensflexion gebildet. Für Details, siehe <https://grammis.ids-mannheim.de/kontrastive-grammatik/4163>

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Präsens-Definition.

Präteritum Beschreibt als Vergangenheitsform zur Text-Zeit vorzeitige Ereignisse. Das Präteritum wird aus einem finiten Verb mit Präteritumsflexion gebildet. Für Details, siehe <https://grammis.ids-mannheim.de/kontrastive-grammatik/4164>

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Präteritum-Definition.

Satz Eine Folge von Wörtern und anderen Zeichenfolgen, die durch Textgrenzen oder satzbeendende Zeichen getrennt wird.

- Als satzbeendende Zeichen gelten: **. ? !**
- Nicht als satz-beendend gelten die Zeichen: **, ; ... :- /**

Bei Kretschmann & Wieken (2010) findet sich keine Definition von Sätzen. Gausche et al. (2014, S. 188) gehen davon aus, dass ein Satz erst „[m]it der Ergänzung von Subjekt und Prädikat zu einer vollständig in sich geschlossenen Aussage entsteht [...] [und] immer auf einen Punkt [endet].“ Diese stark einschränkende Auffassung von Sätzen schließt durch Frage- und Ausrufezeichen abgeschlossene Äußerungseinheiten ebenso aus wie Satzäquivalente (ohne Subjekt-Prädikat-Struktur). Der Einfluss dieser Konstruktionen auf das Leseverständnis würde mit der Verwendung dieses Satzbegriffes nicht bei der Erfassung der Alpha-Level berücksichtigt werden. Um dies zu vermeiden, schließen wir – entgegen der Definition von Gausche et al. (2014) – alle durch Punkte, Ausrufezeichen, und Fragezeichen abgeschlossenen Wort- und Zeichenfolgen in unseren Satzbegriff mit ein.

Diese Definition beinhaltet auch Satzäquivalente ohne (vollständige) Subjekt-Prädikat-Struktur. Ob Satzäquivalente als Sätze gelten sollen, ist eine andauernde Forschungsfrage, ebenso wie eine nicht auf satzbeendende Interpunktion festgeschriebene Definition von Sätzen (vgl. Schmidt 2016). Da es für unsere Annotation unabdingbar ist, reproduzierbare Kriterien festzulegen, werden Sätze ohne Ausnahme über das Kriterium der gängigen satzfinalen Interpunktion (!?) sowie unabhängig von ihren weiteren linguistischen Charakteristika definiert. Ausnahmen von dieser Regel, beispielsweise für Doppelpunkte, die Hauptsätze abtrennen und nach denen groß geschrieben wird, wurden als nicht praktikabel verworfen, auch wenn Doppelpunkte in leichter Sprache in der Praxis oft wie satzbeendende Zeichen behandelt werden (siehe Sonderfälle für das Textlängenkriterium, Abschnitt 2.1, Seite 8). Da in leichter Sprache jedoch kein systematischer Umgang mit Interpunktion festgelegt wird, konnte keine zuverlässige Ausnahme hierfür formuliert werden.

Schachtelsatz Laut Duden ist ein Schachtelsatz ein „langer, kompliziert gebauter Satz mit mehrfach untergeordneten Nebensätzen“.¹⁹ Wir operationalisieren dies als Sätze mit entweder mindestens einem Nebensatz, der einem anderen Nebensatz untergeordnet ist, oder mit mehreren einander nebengeordneten Nebensätzen.

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Definition von Schachtelsätzen.

Silbe Wir verstehen Silben hier als Schreibsilben. Diese bestehen notwendigerweise aus mindestens einem Vokalgraphem (**a, e, i, o, u, ä, ö, ü, ie, au, äu, eu, oi, ai, ui, aa, ee, uu, oo, ...**) und gegebenenfalls aus einem oder mehreren Konsonantengraphemen.

Weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) findet sich eine Konkretisierung des Silbenbegriffes.

¹⁹ www.duden.de/rechtschreibung/Schachtelsatz, zuletzt aufgerufen am 03.07.18

Text Eine maximale, lineare Sequenz von Sätzen, die sich aufeinander beziehen und aufeinander aufbauend eine zusammenhängende, strukturierte Bedeutungseinheit formen.

Wir folgen hiermit der Definition von Gausche et al. (2014), die „kontinuierliche Texte“ definieren als

„lineare Folge von Sätzen. Formal sind Texte in Kapitel, Kapitel wiederum in Absätze und Absätze in einzelne Sätze gegliedert. Die Organisation eines (langen) Textes wird durch Absätze, Einrückungen und ggf. eine hierarchische, durch Überschriften kenntlich gemachte Gliederung des Textes sichtbar.“ (Gausche et al. 2014, S. 52)

Wir schließen Überschriften jedoch aus unserem Textbegriff aus, da sie in der Regel alleinstehend zu verstehen sind und sich der Rest des Textes nicht auf diese bezieht. Kretschmann & Wieken (2010) geben keine Textdefinition an.

Wort Eine durch Leerzeichen getrennte, kontinuierliche Zeichenkette aus Buchstaben, Ziffern, oder Sonderzeichen. Wir verwenden den Wortbegriff daher im Sinne des in der Linguistik üblichen Begriffes *Token*. Wörter bilden eigenständige sprachliche Einheiten und haben eine sprachliche Bedeutung.

Der Wortbegriff wird weder bei Kretschmann & Wieken (2010) noch bei Gausche et al. (2014) explizit definiert. Kretschmann & Wieken (2010, S. 241) gehen bei ihrer Bestimmung von Wortschwierigkeit über Konsonanten-Vokal-Abfolgen implizit davon aus, dass Wörter durch ununterbrochene Buchstabenfolgen geformt werden. Mit dieser annähernden Definition ist jedoch nicht geklärt, wie Zeichenfolgen zu bewerten sind, die Ziffern oder Sonderzeichen enthalten, z.B. **Auto-Bahn** oder **R2D2**. Wir konkretisieren unsere Wortdefinition daher wie folgt:

Folgende Zeichenketten werden überdies für die Zwecke der Richtlinien als eigenständige Wörter gewertet:

1. Abkürzungen mit und ohne Interpunktion (**KFZ** oder **z.B.**)
2. Wortpartikel (**ab-** in **ab- und zuführen**)
3. Ausgeschriebene Zahlen (**Fünfundvierzig**)
4. Ausgeschriebene Währungsbezeichnungen (**Euro**)
5. Ausgeschriebene Bestandteile von Datumsangaben (**Februar** in **15. Februar 2018**)

Als ein einziges Wort gelten

1. Durch Bindestrich getrennte Wörter (**Schiff-Fahrt**, **Aus-bau**)
2. Wörter mit Zusammenschreibung mehrerer Flexionsformen (**Le(i)hen**, **freundlich/er**, **Angestellte_r**)
3. Wörter mit Kennzeichnungen nicht-binärer Geschlechtsmarkierung (**Mitarbeiter***, **MitarbeiterX**)

Als zwei Wörter gewertet werden

1. Ausgeschriebene Wörter ohne Leerzeichentrennung (**Lehrer/Lehrerin**)

Nicht als Wort gewertet werden:

1. URLs (www.die-bonn.de/kansas/)
2. Formeln ($\mu = \sum n_i \frac{1}{n}$)
3. Ziffernfolgen mit oder ohne Sonderzeichen (**31**, **345,30** oder **15.1123.1467.11**)
4. Numerische Datumsangaben (**15.02.2018**)
5. Numerische Währungsangaben (**€ 11,99**, **12.68\$**) oder Maßangaben (**30°C**, **15cm**)
6. Interpunktions- und Satzzeichen (., ,)
7. Smileys (☺, ;))

Wort-Typ Wort-Typen sind die einzigartigen Wörter in einem Text. Während identische wort-wertige Zeichenfolgen als zwei verschiedene Wörter zählen, zählen sie als derselbe Wort-Typ. Sie stehen im Kontrast zu den Vorkommnissen von Wörtern, die in diesem Kontext auch als (Wort-)Token bezeichnet werden. Im Rahmen dieser Richtlinien bezeichnen wir diese als "Wort". Folgender Satz illustriert dies: "Weil wir gerne laufen, laufen wir den ganzen Tag."Dieser besteht aus insgesamt 9 Wörtern (oder Tokens), jedoch nur aus 7 (Wort-)Typen, nämlich **Weil, wir, gerne, laufen, den, ganzen, Tag**. Für Details, siehe https://de.wikipedia.org/wiki/Token_und_Type

4 Frequently Asked Questions

Dieser Abschnitt beinhaltet Fragen und Antworten, die im Verlauf des Annotationsexperiments aufkamen, mit dem diese Richtlinien verifiziert wurden. Das Annotationsexperiment wurde auf Texten für funktionale Analphabeten, Kinder oder Deutsch L2 Lerner verifiziert, die automatisch von Internetseiten gecrawlt und bereinigt wurden. Einige der Fragen und Entscheidungen beziehen sich spezifisch auf die besonderen Erfordernisse dieser Texte, in denen beispielsweise Überschriften nicht klar erkennbar sind, Tippfehler vorhanden sind, oder Fehler durch die automatische Textextraktion von den Internetseiten eingeführt wurden. Es kann daher der Fall sein, dass sie sich nicht oder nur teilweise auf andere Datenkontexte anwenden lassen. Aus diesem Grund wurden die hier dokumentierten Entscheidungen nicht in die Richtlinien selbst aufgenommen.

1. **Wenn ich zwei unabhängige Sätze eindeutig identifizieren kann, das satzbeendende Zeichen zwischen ihnen jedoch fehlt / vergessen wurde, zähle ich dies als einen oder als zwei Sätze?**

Kriterien: Textlänge, Satzlänge

Beispiel: Einen Kuchen kann man **essen** **Er** passt gut zum Tee.

Antwort: Auch in diesen Fällen gilt die Satzdefinition, die es erfordert, dass Sätze entweder durch satzbeendende Zeichen oder durch Textgrenzen abgegrenzt werden. Beim Beispiel handelt es sich nach unseren Richtlinien also um einen Satz mit 10 Wörtern.

2. **Wenn zwei Wörter aufgrund eines Tippfehlers zusammengeschrieben wurden zählen diese als ein Wort oder als zwei Wörter? Was ist mit Fällen, in denen ein Wort versehentlich durch ein Leerzeichen getrennt wurde?**

Kriterien: Satzlänge, Wortlänge

Beispiel: Am **erstenTag** lernen sich alle kennen. Wir **fin den** das wichtig.

Antwort: Auch in diesen Fällen gilt die formale Wortdefinition, die für diese Richtlinien angenommen wird. Das heißt, **erstenTag** zählt als ein Wort mit drei Silben und **fin den** als zwei Wörter mit je einer Silbe.

3. **Soll ich Inhalte ignorieren, die eindeutig nicht in den Text passen, zum Beispiel eine Liste mit Schlagworten, die bezuglos im Text steht?**

Kriterien: alle

Beispiel: Das heißt: niemand weiß wie der Autor heißt. Und wer die Texte wirklich geschrieben hat. Auf dem Bild ist ein Ghostwriter zu sehen **Autor Schriftsteller Ghost Writer Geisterschreiber Auftragschreiber Ghostwriting Phantomschreiber Ghost-Writer** Ein Ghostwriter schreibt Texte.

Antwort: Schlagwörter und anderes sprachliches Material, das nicht in den Text zu passen scheint, wird wie andere Textinhalte als Teil des Textes analysiert. Für das Beispiel heißt dies, dass der Satz von **Auf dem Bild ...** bis **Ein Ghostwriter schreibt Texte.** geht und aus 21 Wörtern besteht.

4. **Wie viele Sätze zähle ich, wenn in direkter Rede mehrere Sätze durch satzbeendende Zeichen abgetrennt sind, die direkte Rede als Ganzes jedoch in einen Matrixsatz eingebettet ist?**

Kriterien: Textlänge, Satzlänge

Beispiel: „Die Mieten sind zu hoch. Und sie steigen immer weiter. Wir können das nicht bezahlen“, beklagt sich Frau Wesel.

Antwort: In diesem Fall zählen alle Sätze in der direkten Rede als einzelne Sätze. Der Matrixsatz **beklagt sich Frau Wesel** wird an den letzten Satz in der direkten Rede **Wir können das nicht bezahlen** angeschlossen. Im Beispiel finden sich also 3 Sätze, die ersten zwei mit je 5 Wörtern und der letzte mit neun Wörtern.

5. **Was mache ich, wenn ich nicht sicher bin, ob ein Begriff der Definitionsgegenstand eines Textes ist, i.e. wenn ich nicht sicher bin, ob Sonderfall 1 des Wortlängenkriteriums gilt?**

Kriterien: Wortlänge

Beispiel: Ein **Direktor** ist ein Chef. Ein Schuldirektor ist der Chef einer Schule. Jede Schule hat einen Schuldirektor. Kinder müssen in die Schule gehen.

Antwort: Wenn nicht vollkommen klar ist, ob oder welcher Begriff der Definitionsgegenstand eines Textes ist, dann findet die Sonderregel keine Anwendung. In diesem Fall wird also kein Begriff als Definitionsgegenstand des Textes behandelt.

6. **Kann ein Text mehr als einen Begriffgegenstand haben, i.e. kann Sonderfall 1 des Wortlängenkriteriums auf mehr als ein Wort im Text zutreffen?**

Kriterien: Wortlänge

Beispiel: Ein **Direktor** ist ein Chef. Ein **Schuldirektor** ist der Chef einer **Schule**. Jede Schule hat einen Schuldirektor. Kinder müssen in die Schule gehen.

Antwort: Jeder Text hat maximal einen Definitionsgegenstand. Wenn mehrere Begriffe als Definitionsgegenstand in Frage kommen, dann findet die Sonderregel keine Anwendung. In diesem Fall wird also kein Begriff als Definitionsgegenstand des Textes behandelt.

7. **Wenn ein Satz eine Nebensatzstellung aufweist, jedoch als unabhängiger Satz geschrieben wurde, zählt dieser als eigenständiger Satz? Und zählt er als Nebensatzkonstruktion im Sinne des Satzstrukturkriteriums?**

Kriterien: Textlänge, Satzlänge, Satzstruktur

Beispiel: Ich gehe früh ins Bett. Darum bin ich morgens immer ausgeruht.

Antwort: Sätze, die syntaktisch Nebensätze sind, aber als eigenständige Sätze geschrieben wurden, werden als separate Sätze gezählt und zählen nicht als Nebensätze im Sinne der Satzstruktur. Das Beispiel beinhaltet also zwei Sätze, aber keine komplexe Satzstruktur.

Literatur

Gausche, S., Haase, A. & Zimper, D. (2014), *Lesen. DVV-Rahmencurriculum*, 1. edn, Deutscher Volkshochschul-Verband e.V., Bonn.

Kretschmann, R. & Wieken, P. (2010), *Lesen. Alpha Levels*, lea., Hamburg.

Schmidt, K. (2016), 'Der graphematische Satz', *Zeitschrift für germanistische Linguistik* **44**(2), 215–256.