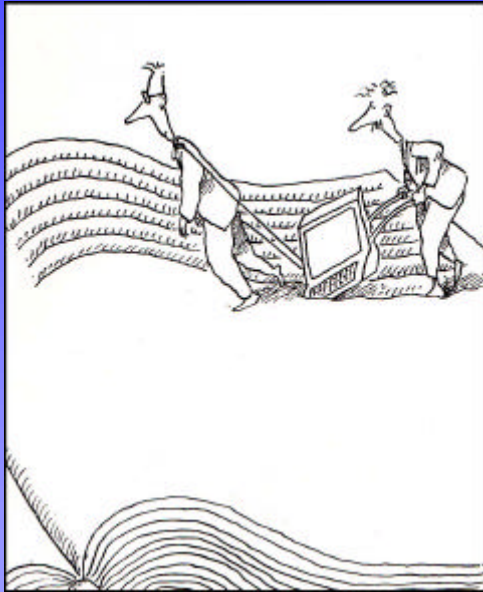


Textwissenschaftliches Colloquium



SLANG I

Zur Entstehung eines
komplexen
Textanalyse-Tools

AB *Computerunterstützte Textinterpretation*

Prof. Dr. Harald Schweizer

Einführung in SLANG I

1995 entstand eine erste Idee für ein Grammatikkonzept zur Beschreibung von Texten im Kontrast zu Schulgrammatiken:

LNNG

Lernende nicht normative Grammatik

Einführung in SLANG I

Aus LNNG wurde

SLANG

Search for a Learning non-normative Grammar

SLANG I vs. SLANG II

SLANG I

- Konstituierung des Textes
- Ausdruckssyntax

Ziel:

- Bereitstellung des Textes für die Analyse
- Bedeutungsfreie Untersuchung der Wortformen

SLANG II

- Semantik
- Pragmatik

Ziel:

- Bedeutungsanalyse
 - bzgl. Äußerungseinheiten
 - bzgl. gesamter Texte

Theorien

Theorie
SLANG I

besteht aus folgenden Theorien:

Theorie
Konstituierung

Theorie
Textkritik

Theorie
Segmentierung

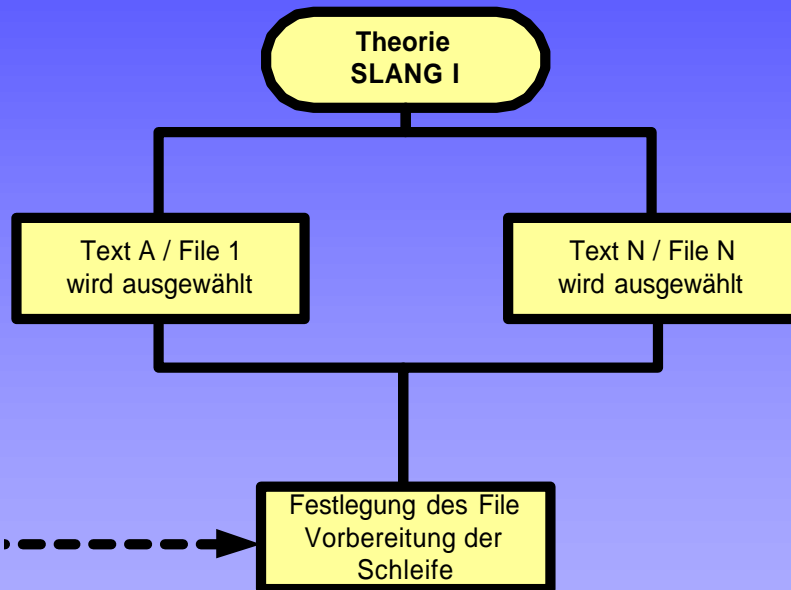
Theorie
Literarkritik

Theorie
SYNTAX

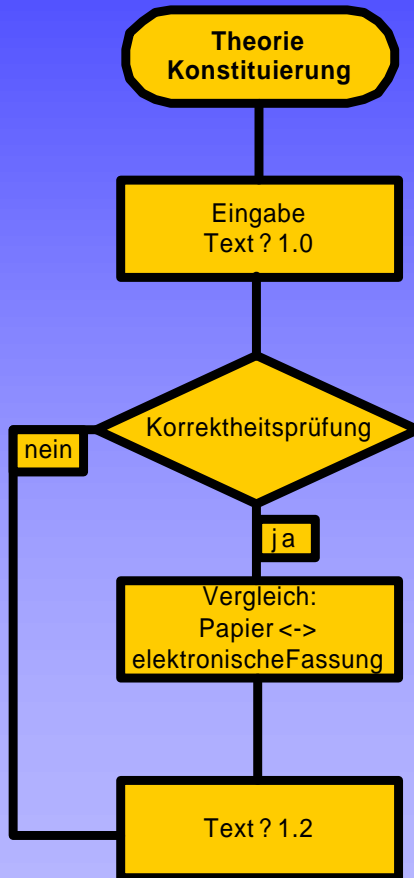
Theorie SLANG I

1. Schritt:

Auswahl der Texte



Theorie Konstituierung



2. Schritt:

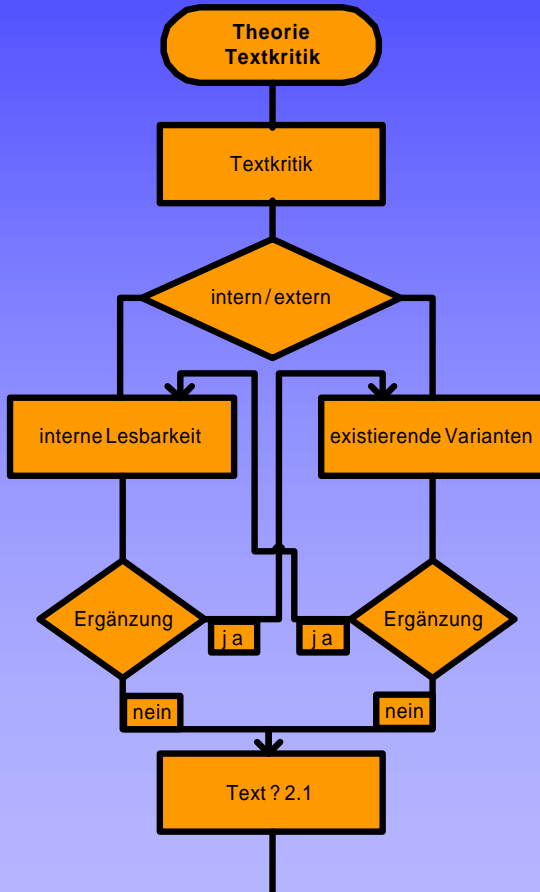
Text einlesen und digitale Form mit der analogen Form vergleichen

Ausgabe des neuen Textes

Programme:

- OCR-Software
- Manuelle Texteingabe
- Elektronische Vorlage

Theorie Textkritik



3. Schritt:

Textkritik,
ist der Text in sich lesbar?
Existieren Varianten?

Programme:

- Multi-Document-Editor
- Textvergleichssoftware

Beispiel existierende Varianten

Textkritik / GOETHES Faust

zu: vkonst3.1

aus:

SCHÖNE, A (ed.) Johann Wolfgang Goethe. Faust. Texte. Frankfurt/M 1994. (= Bd.I)

SCHÖNE, A Johann Wolfgang Goethe. Faust. Kommentare. Frankfurt/M 1994. (= Bd.II)

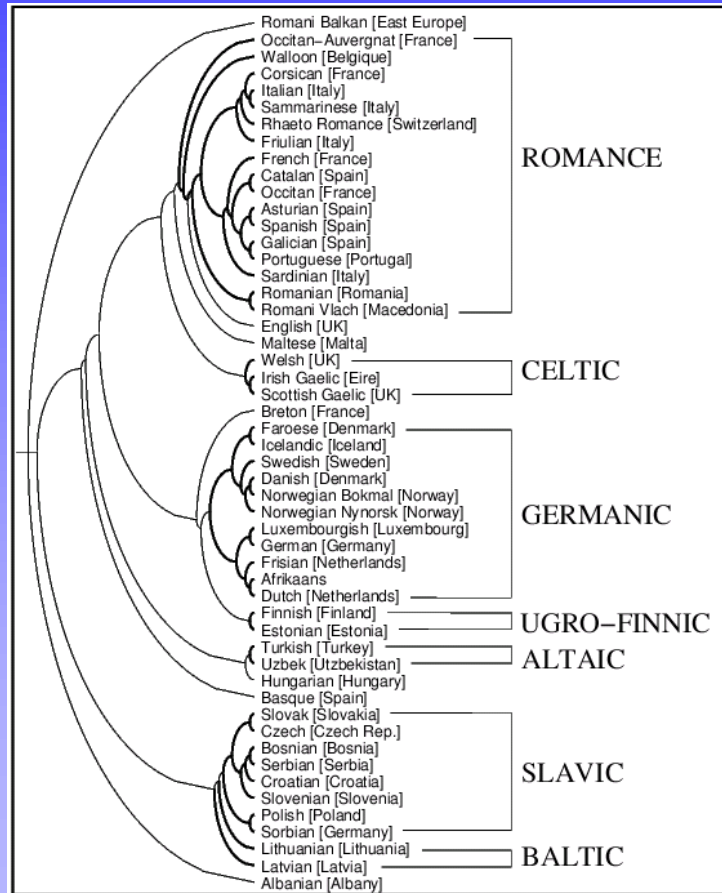
SCHÖNE, I,147 + II,321

Mein Busen drängt
sich nach ihm hin.

Ein glättend abschwächender Eingriff Goethescher Selbstzensur. In der *Frühen Fassung* hieß es, rücksichtslos direkt gegen den Verhaltenskodex des zeitgenössischen Publikums verstoßend:

Mein Schoos! Gott! drängt | Sich nach ihm hin - ein tonbeugender Aufschrei hob dort das Wort *Gott* aus der metrischen Senkung, so dass dieses für Margarete ganz ungeheure Eingeständnis als einzig dreihebiger Vers das zweihebig gefaßte Regelmaß ihrer Rede durchbrach.

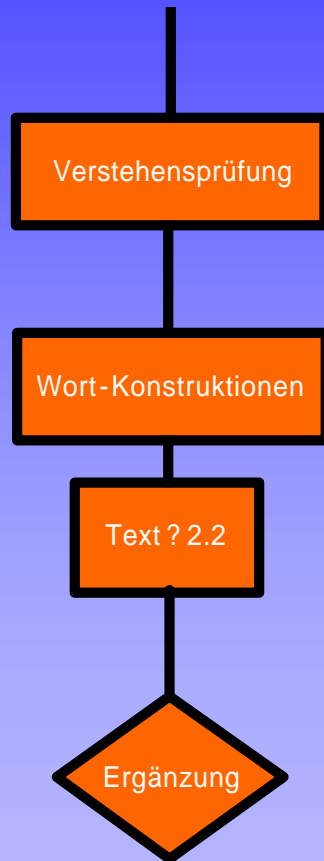
Beispiel Textkritik



Lempel/Ziv – Algorithmus:

Automatische Erzeugung
eines Textstammbaumes

Verstehensprüfung



4. Schritt:

Ungewohnte
Wortbedeutungen oder
Konstruktionen nachschlagen
und verstehen

Beispiel Verstehensprüfung

Textkritik / GOETHES Faust

zu: vkonst3.1

aus:

SCHÖNE, A (ed.) Johann Wolfgang Goethe. Faust. Texte. Frankfurt/M 1994. (= Bd.I)

SCHÖNE, A Johann Wolfgang Goethe. Faust. Kommentare. Frankfurt/M 1994. (= Bd.II)

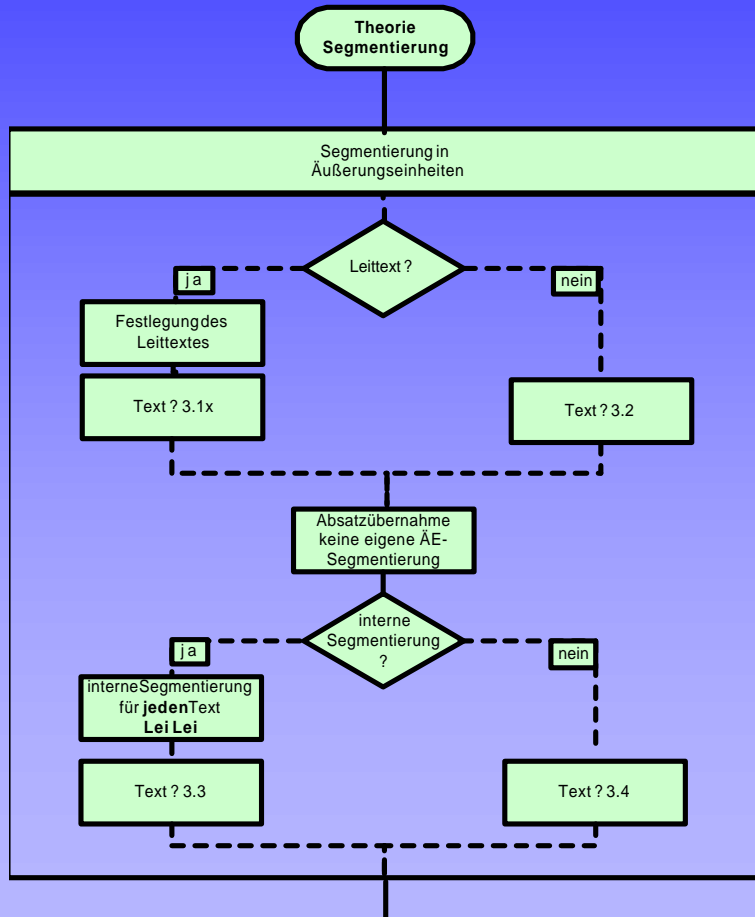
SCHÖNE, I,83 + II,269

Da wird der Geist euch wohl dressiert,
In spanische Stiefeln eingeschnürt,
Daß er bedächtiger so fortan
Hinschleiche die Gedankenbahn
Und nicht etwa, die Kreuz' und Quer
Irrelichteliere hin und her.

spanische Stiefeln] Die Schienbeine zusammenpressendes Marterinstrument beim zweiten Grad der Inquisitionstortur.

Irrlichteliere] Goethesche Wortbildung für eine unstete, sprunghaft - hüpfende Bewegung, wie das Irrlicht sie vortäuscht.

Theorie Segmentierung



5. Schritt:

- Äußerungseinheiten
Bei Bedarf:
- Vorbereitung morph-analoger
Übersetzung und Synopse

- Kriterien

<http://www-ct.informatik.uni-tuebingen.de/ct/aee.html>

Programm:

interne Segmentierung – Lei Lei

Beispiel Segmentierung

Die Hauptaufgabe von meinem Programm ist, einen Text in Äußerungseinheiten aufzuteilen. Ein eingelesener Text wird nach folgenden Kriterien aufgeteilt.

1.) **Starke Trenner, z.B. ;:~?!**

werden dann als Kriterium angenommen, wenn sie

- mit einem Leerzeichen eingelesen werden oder
- am Ende einer Zeichenfolge stehen

Beispiel Segmentierung

2.) Anführungszeichen und Klammer

- wenn ein Leerzeichen vor einem Anführungszeichen eingelesen wird, dann wird der Beginn einer speziellen Zeichenfolge angenommen.

Bsp: _"Anfang

- wenn ein Leerzeichen nach einem Anführungszeichen eingelesen wird, vor dem ein Trenner oder Buchstabe stehen können.

Bsp: Ende.“_

Dazwischen stehende Zeichenfolge, die aus einer ÄE oder mehreren ÄEen bestehen können, und wird weiter als normale Zeichenfolge analysiert und aufgeteilt.

Klammern, z.B. (){}[], usw, werden genau wie Anführungszeichen behandelt.

Beispiel Segmentierung

3.) schwacher Trenner, also Komma(,)

Wenn wir den Text durchlesen und ein Komma mit Leerzeichen gefunden haben, wissen wir noch nicht, ob danach eine neue ÄE beginnt.

Eine zusätzliche Prüfung ist notwendig:

a.) **Konjunktionale Nebensätze**, die z.B. mit "dass, damit, obwohl, weil" usw. nach dem Komma beginnen, werden abgetrennt.

Diese Konjunktionen können Sie löschen, ergänzen und verändern.

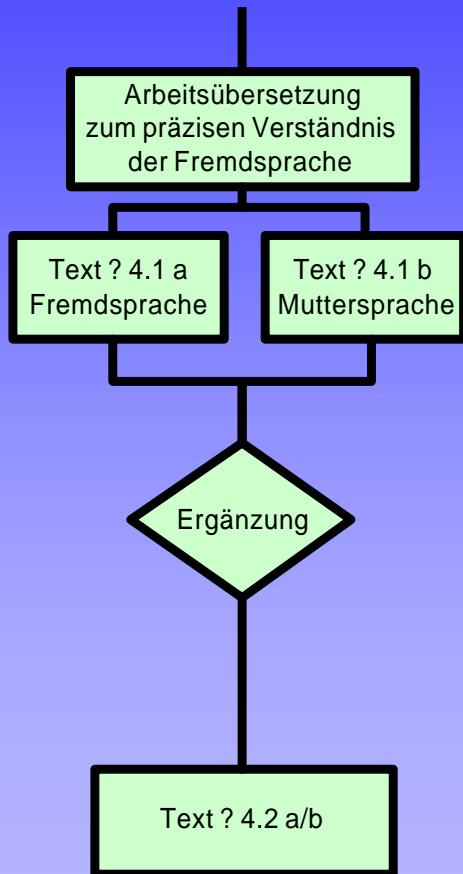
In meinem Programm habe ich eine Liste aufgestellt. Es enthält alle Konjunktionen. Damit prüfe ich das erste Wort des Satzes und sehe, ob es in der Liste enthalten ist. Dadurch wird beurteilt, ob der Satz ein konjunktionaler Nebensatz ist.

b.) **die andere Möglichkeit:** nach dem Komma folgt keine Konjunktion, es erfolgt keine Aufteilung. Darunter fallen z.B. Apposition u.ä.

Beispiel: David, der König, war ein Schuft.

Der Satz wird als eine ÄE angenommen, ohne Aufteilung.

Theorie Segmentierung



6. Schritt:

Weiterarbeiten mit Text und Arbeitsübersetzung (1. übersetzter Fremdsprachentext oder 2. „in eigene Worte“ übersetzter Text)

Programm:

- Editor mit Hilfesystem (Lexikon, etc.)

Beispiel Arbeitsübersetzung

Aus dem Forschungsprojekt Josefsgeschichte:

- 007::37,5b und+ er\erzählte[0027] !+ Brüdern+ seinen[0028]
008::37,6a und+ er\sprach[0029] zu+ihnen[0030]
009::37,6b {ÄE=>ÄE} hört[0031] doch[0032] den+ Traum [0033] den+selben[0034]
010::37,6c welchen[0035] geträumt-habe\ich[0036]
011::37,7a {ÄE=>ÄE} und+da[0037]
012::37,7b {ÄE=>ÄE} wir[0038] () Bindende[0039] Garben[0040] in+Mitten[0041] des+ Feldes[0042]
013::37,7c {ÄE=>ÄE} {ÄE=>ÄE} und+da[0043]
014::37,7d aufgestanden-ist[0044] Garbe+ meine[0045]
015::37,7e und+sogar[0046] stehenblieb\sie[0047]
016::37,7f {ÄE=>ÄE} und+da[0048]
017::37,7g (im-Kreis-)stellten-sich[0049] Garben+ eure[0050]
018::37,7h und+ sie\verneigten-sich[0051] vor+ Garbe+ meiner[0052]
019::37,8a und+ sprachen[0053] zu+ihm[0054] Brüder+ seine[0055]
020::37,8b {ÄE=>ÄE} ?+ als-König-Regieren[0056] du-möchtest/als-König-regieren[0057]
über+uns[0058]
021::37,8c (?) -oder[0059] Herrschen[0060] du-willst/herrschen[0061] über+uns[0062]
022::37,11a und+ eifersüchtig-waren[0063] auf+ihn[0064] Brüder+ seine[0065]

Beispiel Ausgangstext

Aus dem Forschungsprojekt Josefsgeschichte:

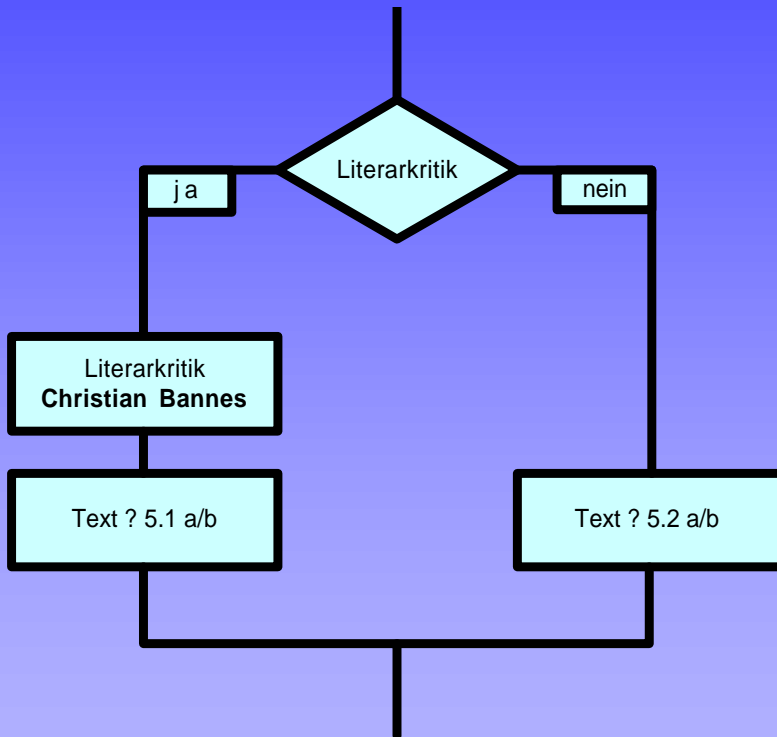
zu: vkonst3.1 aus: 2 Sam 14

וַיֵּדַע יוֹאָב בֶּן זֵרוּיָהּ	wa=ye [n] da ^c YW ^y B bãn ŞRYH	1a	und=er/erfuhr JOAB, Sohn (des) -ZERUJA
כִּי לִב הַמֶּלֶךְ עַל אַבְשָׁלוֹם	ki (y) leb ha=mäläk	1b	dass das-Herz des=Kö- nigs ()
וַיִּשְׁלַח יוֹאָב תְּלוּמָה	^c al ^y BŠLWM wa=yišlah YW ^y B	2a	über ABSCHALOM und=er/sandte JOAB
וַיִּקַּח מִשָּׁמַיִם אִשָּׁה חַכְמָה	wa=yi [l] qah mi [n] =šãm ^y iššã (h) ḥkãmã (h)	2b	und=er/nahm von=dort eine-Frau eine- weise

Theorie Literarkritik

7. Schritt:

Literarkritik =
Textarchäologie



Programm:

Literarkritik – C. Bannes

Beispiel Literarkritik

1.) Sammeln von Leseauffälligkeiten (LA):

alles was möglicherweise für den Leseprozess störend ist wird notiert. Jede LA benötigt zwei Pole.

2.) Minimale Leseinheiten (MLE):

Textbereiche, in denen keine Leseauffälligkeiten mit beiden Polen vorliegen

3.) Diskussion der LA / Erarbeitung der Teiltexpte:

Text wird von vorn gelesen, indem das Verhältnis der ersten usw. MLE zur jeweilig folgenden überprüft wird.

Gewichten der in 1) gesammelten LA, Bewertung nach Qualität und Quantität

Beispiel Literarkritik

4.) Verhältnisbestimmung der Teiltexthe (TT):

Unmittelbar benachbarte TT schließen sich aus. Man nimmt immer vom Textanfang her weiter entfernte TT und überprüft sie nach Kriterien von Schritt 1),

- wirken zwischen ihnen ebenfalls LA, die einen textlichen Anschluss verhindern?

- entsteht durch aneinanderkoppeln ein problemlos lesbarer, störungsfreier Text?

5) Gegenkontrolle:

Sind alle unter 1) gesammelten LA aufgegriffen, wurde eine zufriedenstellende Erklärung gefunden?

Schema Literarkritik

1.) Beobachtung *möglicher* Verstehensprobleme

Text in Äußerungseinheiten

2.) Minimale Leseinheiten

AAAABBBBCCCCCCCCCCCCCCCCDDEEEEEEEEEEEEEEEEEFFFGHHHHHHIIIIIIJJJKK

3.) Teiltexte

AAAABBBB

CCCCCCCCCCCC

DDEEEEEEEEEEEEEEFF

GHHHH

IIIIII...

4.) Schichten

AAAABBBB

DDEEEEEEEEEEEEEEFF

CCCCCCCCCCCC

GHHHH

IIIIII...

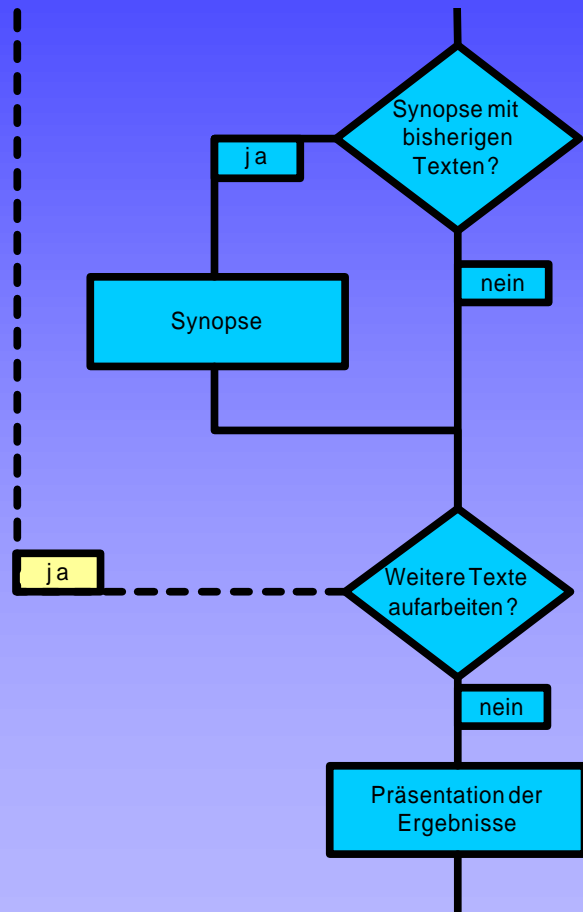
5.) Gegenkontrolle

Sind alle Beobachtungen aus 1.) berücksichtigt?

Ergebnis Literarkritik

Äußerungseinheiten Zählung pro ÄE Endtext, inkl. Überarbeitungen Transkription A	innerhalb ÄEen absolute Zählung ursprüngliche Textschicht Transkription B	Zählung pro ÄE an Endtext orientiert ursprüngliche Textschicht Arbeitsübersetzung morph-analog	KEY absolute Zählung Endtext
WF:01 YV.C.P	(1)iu+s+^p [יְסוּפִּי]	W:1JOSEF	11
WF:02 B*.N			12
WF:03 W.B.J			13
WF:04 J.S.R.H			14
WF:05 W.N.H			15
WF:06 H.Y.H	(2)h+y+h [הַעֲהֵ]	W:6war	16
WF:07 R.J.H	(3)r+y+h [רַעֲהֵ]	W:7ein-Hüter	17
WF:08 %.T	(4)a+t [אֵת]	W:8mit	18
WF:09 %.X.YV	(5)a+x+iu [אֵתְּיָם]	W:9Brüdern+seinen	19
WF:10 B*.3*.			20

Synopse



8. Schritt:

Synopse mit bisherigen Texten,
sollen weitere Texte bearbeitet werden -> Schleife zum Anfang

Ausdrucksanalyse

9. Schritt:

Ende, wenn keine
Syntaxanalyse durchgeführt
werden soll.

