

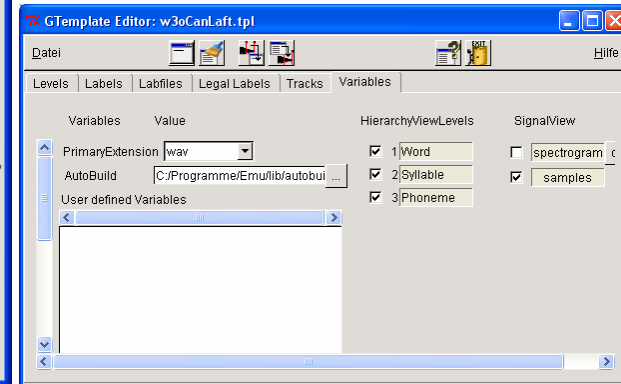
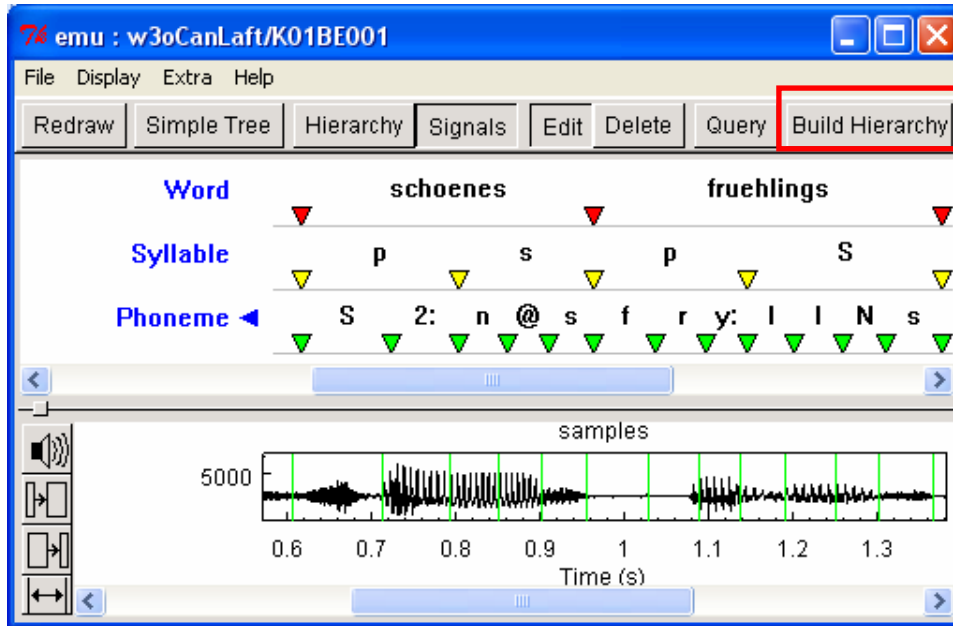
Automatische

Etikettierung in Sprachdatenbanken

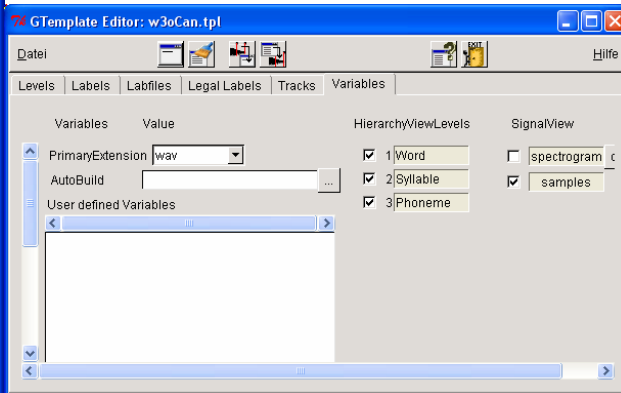
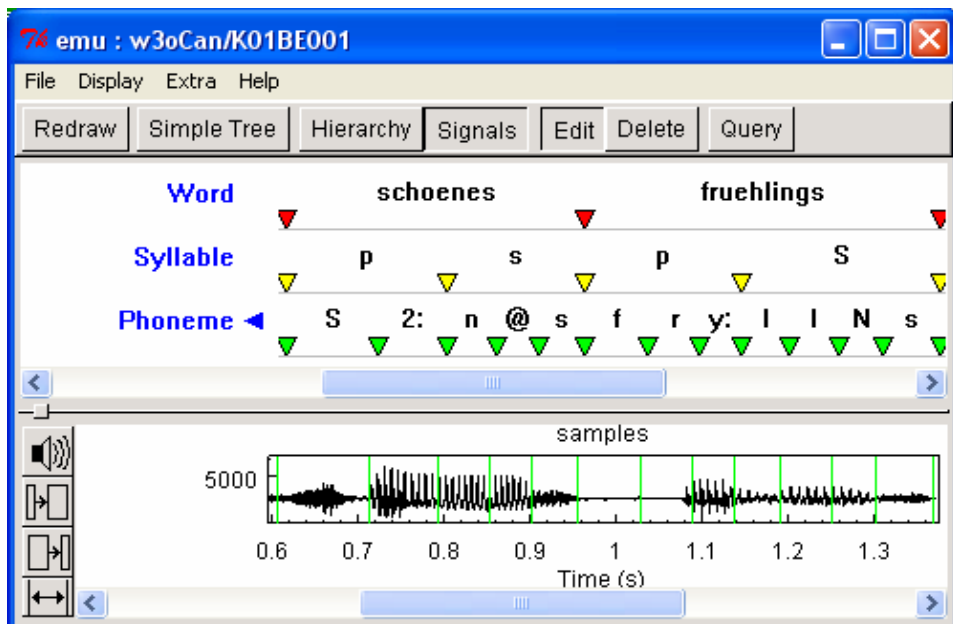
Tina John

Der Weg zur automatischen Etikettierung

automatisch

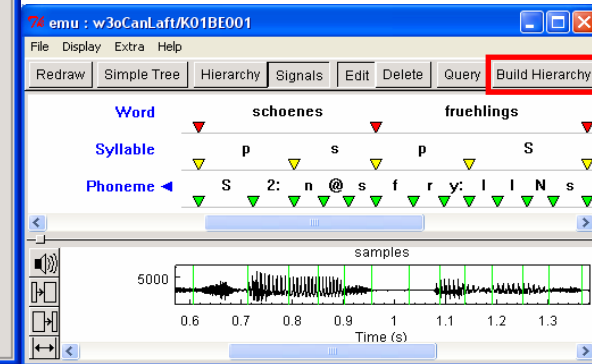
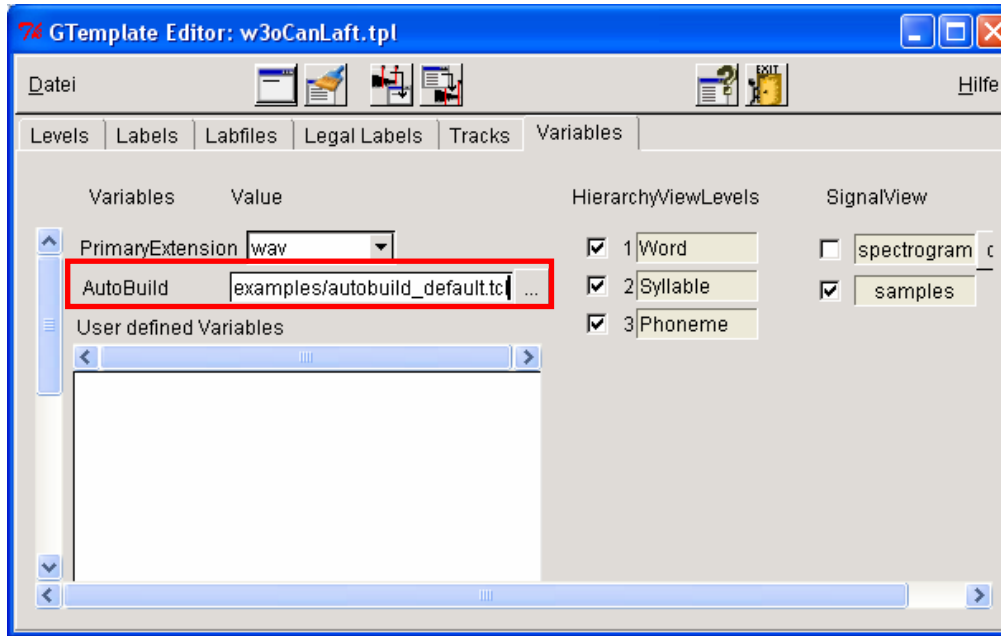


manuell

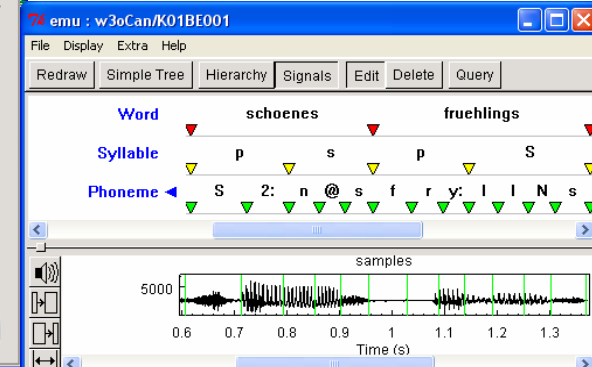
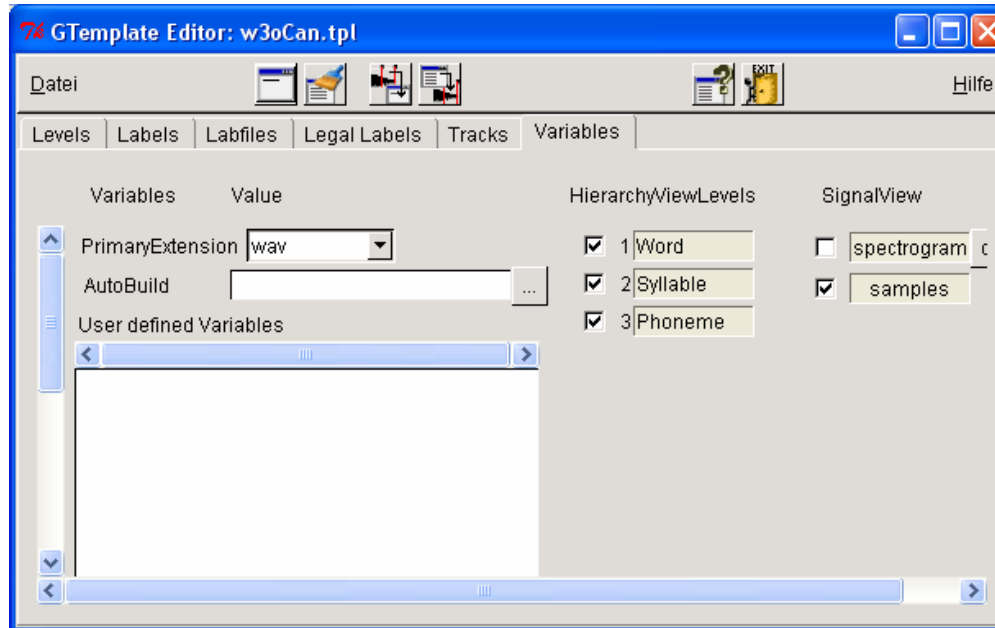


Der Weg zur automatischen Etikettierung

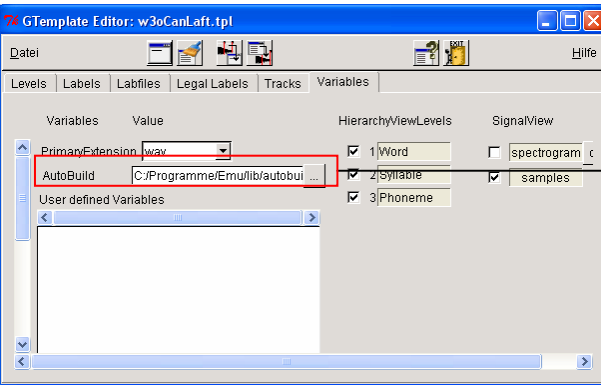
automatisch



manuell



Der Weg zur automatischen Etikettierung



Eine solche Datei kann selbst geschrieben werden und mit einem beliebigen Namen (ext: .tcl) an einem beliebigen Ort gespeichert werden.

Ein Beispiel einer solchen Datei befindet sich im Verzeichnis

*.../Emu/lib/autobuild/examples

*Windows: C:\Programme

Sun: /opt

```
autobuild_default.tcl - Editor
Datei Bearbeiten Format Ansicht ?

package require emu::autobuild

#####
## AutoBuildInit initialises any rulesets needed
## by AutoBuild
#####

proc AutoBuildInit {templ} {
}

#####
## AutoBuild builds the hierarchy using whatever
## means are appropriate
#####

proc AutoBuild {templ hier} {
    LinkAllFromTimes $templ $hier
}
```

Das AutoBuild-Script

Kommentare-nicht wichtig

Obligatorische Angaben-
das Grundgerüst

Vorbereitende Dinge-wird
von verschiedenen
Prozeduren verlangt

Aufruf einer Prozedur, die
die eigentliche Arbeit des
automatischen
Etikettierens macht-hier ist
es die Prozedur
LinkAllFromTimes, die alle
Segmente auf den
zeitgebundenen Ebenen
nach ihren Zeitpunkten
assoziiert (Links setzt)

```
autobuild_default.tcl - Editor
Datei Bearbeiten Format Ansicht ?

package require emu::autobuild

#####
## AutoBuildInit initialises any rulesets needed
## by AutoBuild
#####

proc AutoBuildInit {templ} {
}

#####
## AutoBuild builds the hierarchy using whatever
## means are appropriate
#####

proc AutoBuild {templ hier} {
    LinkAllFromTimes $templ $hier
}
```

Das AutoBuild-Script-Zeichensetzung

- _ Leerstelle
- ↵ Zeilenwechsel

```
autobuild_default.tcl - Editor
Datei Bearbeiten Format Ansicht ?

package_require_emu::autobuild↵

#####
## AutoBuildInit initialises any rulesets needed
## by AutoBuild
#####

proc_AutoBuildInit_{templ}_{↵
    ↵
} ↵

#####
## AutoBuild builds the hierarchy using whatever
## means are appropriate
#####

proc_AutoBuild_{templ_hier}_{↵
    ↵
    LinkAllFromTimes_$templ_$hier ↵
} ↵
```

LinkAllFromTimes

74 emu : w3oCanLaft/K01BE002

File Display Extra Help

Redraw Simple Tree Hierarchy Signals Edit Delete

Word * die sonne lacht *

Syllable * p p S p *

Phoneme * d h i: z O n @ l a x t h *

This screenshot shows the 'Hierarchy' view in the emu software. The hierarchy is flat, with the words 'die', 'sonne', and 'lach' listed at the top level. Below them, the syllables 'p p S p' and the phonemes 'd h i: z O n @ l a x t h' are listed, but they are not yet linked to the words above them.

74 emu : w3oCanLaft/K01BE002

File Display Extra Help

Redraw Simple Tree Hierarchy Signals Edit Delete Query Build Hierarchy

Word die sonne lacht

Syllable p p S p

Phoneme d h i: z O n @ l a x t h

5000

samples

0.8 0.9 1 1.1 1.2 1.3 1.4 1.5 1.6 1.7

Time (s)

This screenshot shows the 'Hierarchy' view with a hierarchical structure. The words 'die', 'sonne', and 'lach' are at the top. Below them, the syllables 'p p S p' are shown, and below that, the phonemes 'd h i: z O n @ l a x t h' are shown. A waveform is displayed below the hierarchy, with vertical green lines indicating the time boundaries for each syllable and phoneme. The x-axis is labeled 'Time (s)' and ranges from 0.8 to 1.7. The y-axis is labeled 'samples' and ranges from 0 to 5000.

AutoBuild Prozedur **LinkAllFromTimes**

74 emu : w3oCanLaft/K01BE002

File Display Extra Help

Redraw Simple Tree Hierarchy Signals Edit Delete Query Build Hierarchy

Word * die sonne lacht *

Syllable * p p S p *

Phoneme * d h i: z O n @ l a x t h *

This screenshot shows the 'Hierarchy' view with a fully linked hierarchical structure. The words 'die', 'sonne', and 'lach' are at the top. Below them, the syllables 'p p S p' are shown, and below that, the phonemes 'd h i: z O n @ l a x t h' are shown. The hierarchy is now fully linked, with lines connecting the words to their respective syllables and phonemes. The x-axis is labeled 'Time (s)' and ranges from 0.8 to 1.7. The y-axis is labeled 'samples' and ranges from 0 to 5000.

AutoBuild Prozeduren mal anders

Szenario: Nach der Vorlesung verlasse **ich** die Uni und frage mich wie es **jetzt** weitergeht

package require emu::autobuild

Prozeduren des Pakets:

- befinden sich in den Dateien im Verzeichnis ../Emu/lib/autobuild/tcl
- sind teilweise dokumentiert in ...\\Emu\\doc\\emu\\html\\ch08s02.html

Es können auch neue Prozeduren selbst programmiert werden.

proc AutoBuildInit {templ} {

}

proc AutoBuild {templ hier} {

LinkAllFromTimes \$templ \$hier

}

Paket von Möglichkeiten:

KaffeeTrinkenGehen {ich jetzt mitWem}

ZurUbGehen {ich jetzt}

MitDemBusNachHause {ich jetzt wannMgl}

erfahreWannNächsterBusFährt

Vorbereitung {ich}

Ausführung {ich jetzt}

ZurUbGehen ich jetzt

AutoBuild Prozeduren mal anders

Paket von Möglichkeiten:

KaffeeTrinkenGehen {ich jetzt mitWem}

ZurUbGehen {ich jetzt}

MitDemBusNachHause {ich jetzt wannMgl}

erfahreWannNächsterBusFährt

Vorbereitung {ich}

merkeDir wannMgl

wannMgl = erfahreWannNächsterBusFährt

Ausführung {ich jetzt}

merkeDir wannMgl

MitDemBusNachHause ich jetzt wannMgl

MapLevels

emu : w3MapLevel/K01BE001

File Display Extra Help

Redraw Simple Tree Hierarchy Signals Edit Delete Query Build Hierarchy

Pho h OY t h @ l s S 2: n @ s f r

Target O Y @

samples

5000

0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8

Time (s)

emu : w3MapLevel/K01BE001

File Display Extra Help

Redraw Simple Tree Hierarchy Signals Edit Delete Query Build Hierarchy

Can ◀

Pho h OY t h @ l s S 2: n @ s f r y

Target

rules.txt - Editor

Datei Bearbeiten Format Ansicht

?

t h -> t

p h -> p

k h -> k

AutoBuild Prozedur **MapLevels** Can Pho \$rules

emu : w3MapLevel/K01BE001

File Display Extra Help

Redraw Simple Tree Hierarchy Signals Edit Delete Query Build Hierarchy

Can h OY t @ l s S 2: n @ s f r y

Pho h OY t h @ l s S 2: n @ s f r y

Target ◀

MapLevels

```
package require emu::autobuild
```

Paket von Möglichkeiten:

KaffeeTrinkenGehen {ich jetzt mitWem}

ZurUbGehen {ich jetzt}

MitDemBusNachHause {ich jetzt wannMgl}

erfahreWannNächsterBusFährt

```
proc AutoBuildInit {templ} {
```

```
global rules
```

```
set rules [ReadLevelRules rules.txt]  
}
```

```
proc AutoBuild {templ hier} {
```

```
global rules
```

```
MapLevels $templ $hier Can Pho $rules
```

```
}
```

Vorbereitung {ich}

```
merkeDir wannMgl
```

```
wannMgl = erfahreWannNächsterBusFährt
```

Ausführung {ich jetzt}

```
merkeDir wannMgl
```

```
MitDemBusNachHause ich jetzt wannMgl
```

AutoBuild Kombinationen

```
package require emu::autobuild
```

Paket von Möglichkeiten:

KaffeeTrinkenGehen {ich jetzt mitWem}

ZurUbGehen {ich jetzt}

MitDemBusNachHause {ich jetzt wannMgl}

erfahreWannNächsterBusFährt

```
proc AutoBuildInit {templ} {
```

```
global rules
```

```
set rules [ReadLevelRules rules.txt]
```

```
}
```

```
proc AutoBuild {templ hier} {
```

```
global rules
```

```
LinkFromTimes $templ $hier
```

```
MapLevels $templ $hier Can Pho $rules
```

```
}
```

Vorbereitung {ich}

```
merkeDir wannMgl
```

```
wannMgl = erfahreWannNächsterBusFährt
```

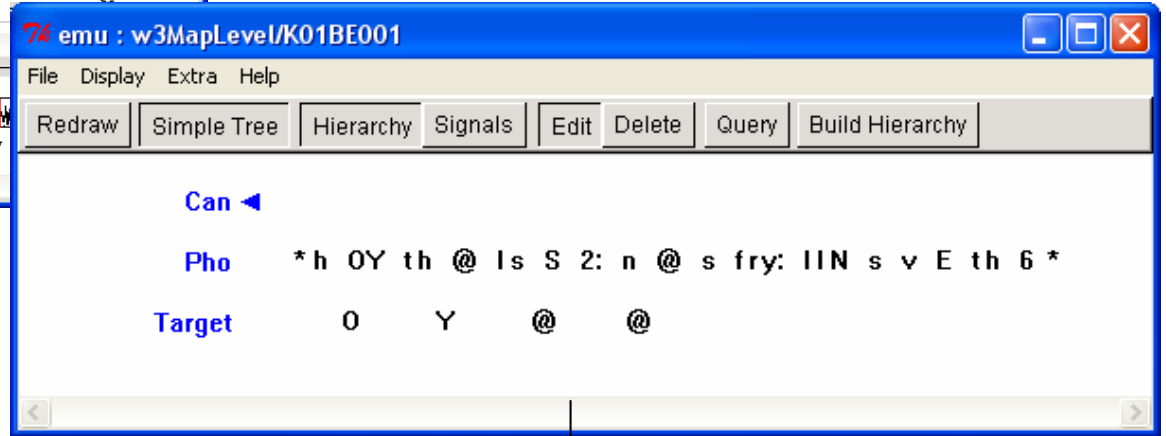
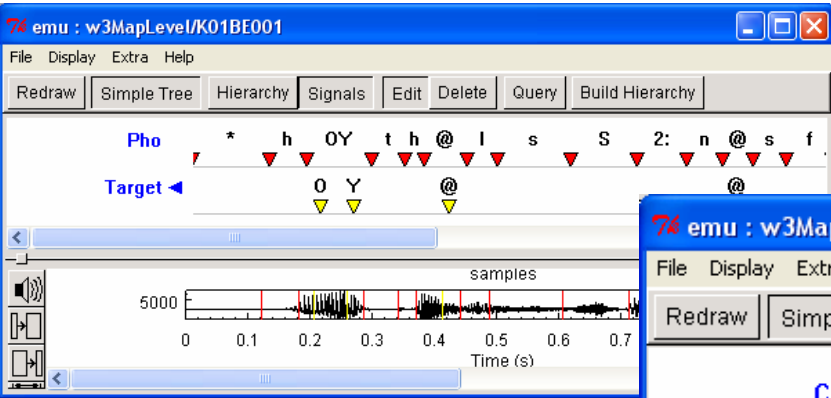
Ausführung {ich jetzt}

```
merkeDir wannMgl
```

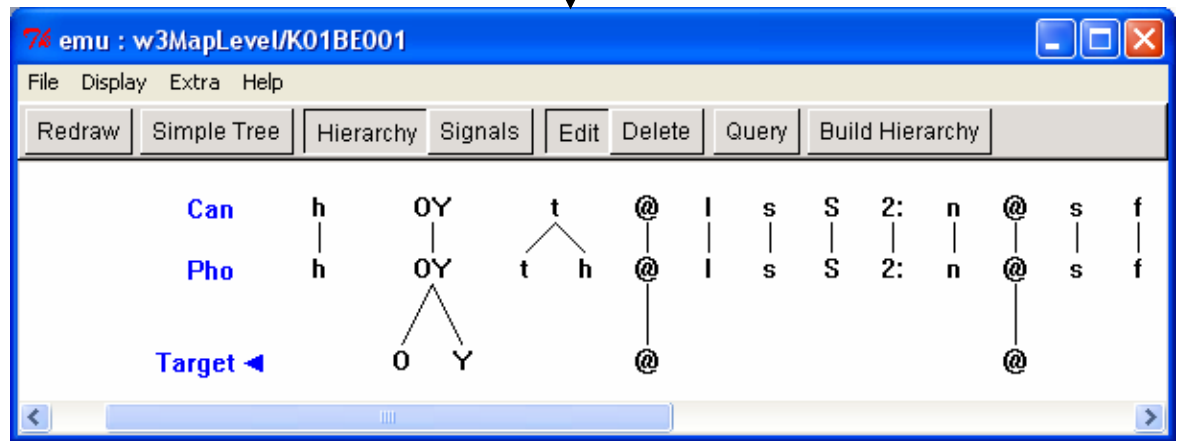
```
ZurUbGehen ich jetzt
```

```
MitDemBusNachHause ich jetzt wannMgl
```

LinkAllFromTimes + MapLevels



AutoBuild Prozedur **LinkAllFromTimes**
 AutoBuild Prozedur **MapLevels** Can Pho \$rules



Vorteile

und

Nachteile

Sehr schnell

Macht keine menschlichen Fehler
(vertippen, vergessen)

Kriterien sind immer konsistent

AutoBuild-Scripts müssen erst
geschrieben werden

Fehleranfällig, da sich die
Sprache nicht immer an die
Regeln hält

Kann nicht spontan auf
Ausnahmefälle reagieren

Ausnahmen können nachprogrammiert werden

Fehleranfälligkeit kann somit verringert werden

Zusammenfassung

Automatische Etikettierung

Das automatische Etikettieren ist eine **schnelle Methode** für den Aufbau einer Sprachdatenbank, dennoch sollten auch die **Nachteile** nicht ignoriert werden.

Das automatische Etikettieren funktioniert über **AutoBuild-Scripts**- selbsteditierte Textdateien.

Zusammenfassung

AutoBuild-Script

AutoBuild-Script enthalten die Zeilen:

package require emu::autobuild, ein Paket mit einer Auswahl an Prozeduren

proc AutoBuildInit {templ} {...}, hier werden vorbereitend verschiedene Dinge angegeben, die in der eigentlichen AutoBuild Prozedur benötigt werden.

proc AutoBuild {templ hier} {...}, die eigentliche AutoBuild Prozedur, in der Prozeduren aus dem emu::autobuild Paket bzw. selbst programmierte Prozeduren aufgerufen werden.

Zusammenfassung

Anwendung von AutoBuild-Scripts

Um ein AutoBuild-Script auf Äußerungen einer Sprachdatenbank anzuwenden, muss der **Pfad zum Script in der Template-Datei** unter **Variables-AutoBuild** angegeben werden.

(AutoBuild-Scripts können mit dem EMU-Tool AutoBuild-ein Wizard auch auf Sprachdatenbanken angewendet werden ohne sie in der Template-Datei anzugeben mit dem Nachteil: keine direkte Fehlerüberprüfung)