

## Einführung in die Grammatikentwicklung: Übung 6 (21. Oktober 2003)

### Ziele:

1. Umgestaltung der Grammatik, so daß ein Unterschied zwischen Wort und Lexem gemacht wird.
2. Hinzufügen von Flexion.

### Übungen:

Gestalten Sie die Grammatik so um, daß ein Unterscheid zwischen Wörtern und Lexemen gemacht wird und daß Flexionsregeln in die Grammatik aufgenommen werden. Auf diese Weise läßt sich Redundanz im Lexikon vermeiden. Wie im Lehrbuch von Sag und Wasow vorgeschlagen, werden wir alle Einträge im Lexikon zu Subtypen von Lexem machen, aber verlangen, daß nur Wörter (oder größere Einheiten, die aus Wörtern bestehen) als Töchter in den Grammatikregeln vorkommen können (so wie es jetzt in der Grammatik ist). Das bedeutet, daß aus jedem Lexikoneintrag vom Typ Lexem durch eine Lexikonregel ein Wort gemacht werden muß. Einige dieser Regeln haben einen morphologischen Effekt. Bei dieser Übung sind große Veränderungen an den Grammatikdateien nötig. Das ist leider eine Sache, die Grammatikentwickler im wirklichen Leben ziemlich oft machen müssen. Machen Sie die Änderungen schrittweise und fertigen Sie Backup-Kopien Ihrer Dateien an. Sie können das tun, indem Sie das Kommando `cp -r` in einem xterm-Fenster ausführen:

```
cp -r <mygrammar> <mygrammar>-bk1
```

1. Fügen Sie die Lexem/Wort-Unterscheidung in die Grammatik ein. Tun Sie dies, indem Sie die Typen *word* und *lexem* einführen, die beide von einem neuen Typ *lex-item* erben. Das Merkmal `ORTH` wird vom Typ *lex-item* eingeführt. Der Typ *word* ist der Obertyp aller Lexem-zu-Wort-Regeln. Die Strukturen, die Sie brauchen, sind die Folgenden:

```
lex-item := syn-struct &
[ ORTH string ].

lexeme := lex-item.

word := lex-item &
[ HEAD #1,
  SPR #2,
  COMPS #3,
  ARGS < lexeme &
    [ HEAD #1,
      SPR #2,
      COMPS #3 ]
  > ].
```

2. Die alten Typen *verb-word* usw. hatten Beschränkungen, die nach unserer neuen Sichtweise zu Lexemen gehören sollten. Fügen Sie Untertypen von *lexeme* für *verb-lxm*, *noun-lxm* usw. ein, die die Beschränkungen der alten Typen *verb-word*, *noun-word* usw. benutzen. Lassen Sie *det-lxm* und *prep-lxm* von *const-lxm* erben. *const-lxm* ist wie folgt definiert:

```
const-lxm := lexeme.
```

3. Fügen Sie die Typen *noun-form* und *verb-form* als Untertypen von *word* hinzu. Diese Typen sind Obertypen der Flexionsregeln: Ihr `ARGS`-Wert sollte so beschränkt werden, daß er dem passenden Lexemtyp entspricht.

4. Der alte Typ *verb-word* war ein Supertyp von Typen, die zwei verschiedene Arten von Information beschrieben haben, die in unserer gegenwärtigen Sichtweise inkonsistent sind. Ändern Sie Ihre Typdatei so, daß sie die Untertypen von *verb-word* aufteilen in solche, die für Lexeme passend sind und solche, die mit Flexio zu tun haben. Der Untertyp für transitive Verben wird dann z. B. ein Untertyp von *verb-lxm* sein und der Untertyp für dritte Person Plural wird ein Untertyp von *verb-form* sein. Entfernen Sie alle Typen, die sowohl von Flexionstypen als auch von Lexemtypen erben. Diese Typen werden durch die Flexionsregeln ersetzt.
5. Ändern Sie Ihre Datei *lexicon.tdl*:
  - (a) Entfernen Sie flektierte Formen von Einträgen, so daß pro Wort nur noch ein Lexemeintrag übrig bleibt, dessen ORTH-Wert dem Stamm entspricht. Das heißt, von *Hund*, *Hundes*, *Hunden* bleibt nur noch *Hund* übrig. Statt *gibt* und *gab* gibt es nur noch einen Eintrag für *geb*.
  - (b) Ändern Sie die Typen der verbleibenden Einträge, so daß alle entsprechend passende Untertypen von *lexeme* sind. Die Datei sollte nun Basisformen mit der Basisorthographie und einem einzigen Typ enthalten.
  - (c) Die Argumentumstellungslexikonregel wird zu einer Lexem-zu-Lexem-Regel. Verändern Sie die Regel in *lrules.tdl* entsprechend. Die Datei *rules.tdl* bleibt unverändert.
6. Sie sollten jetzt eine Grammatik haben, die ladbar ist, aber nichts analysiert, weil wir die Flexionsregeln noch nicht hinzugefügt haben. Überprüfen Sie, daß die Grammatik ohne Fehler geladen werden kann, bevor Sie zum nächsten Schritt gehen.
7. Sie haben jetzt eine Datei *inflr.tdl*, die die Flexionsregeln definiert. Machen Sie sich wegen der Zeilen, die mit % beginnen, keinen Kopf. Diese Zeilen sind Anweisungen für die orthographische Komponente. Sie müssen nur die Typen *sg-mas-neu-nom-dat-acc-noun*, *sg-mas-neu-gen-noun*, *sg-mas-neu-gen-noun*, *sg-fem-noun*, *pl-noun* und entsprechende Verb- und Adjektivtypen durch die entsprechenden von Ihnen definierten Typen ersetzen (ð, durch die Typen, die sie als Untertypen von *noun-form* und *verb-form* definiert haben).
8. Überprüfen Sie das Ergebniss, um sicherzustellen, daß die Abdeckung der Grammatik konstant geblieben ist.