Linguistics 567, Spring 2005, Jonathan Pool
Lab 6 (Esperanto)

***Preparation***

*Redundancy*

In my prior report, I described the semantically identical multiple parses yielded by the grammar in the case of adverbs modifying transitive verbs. I showed that constraining adverbs' MOD values so as to prevent adverbs from modifying nodes of particular types couldn't stop all such redundancy without also causing under-acceptance.

In this section I describe how I eliminated the redundant parses arising from adverbial modification.

Emily Bender suggested investigating the effect of preventing the head-opt-comp rule's daughter from having a [ MODIFIED hasmod ] value. Following this idea for the phrase rules licensing the redundant parses, I added the following constraints:

> head-opt-comp-phrase: [ HEAD-DTR.SYNSEM.MODIFIED notmod ]
> head-comp-phrase: [ HEAD-DTR.SYNSEM.MODIFIED notmod ]

This decreased the redundant parses, but did not eliminate them. Some parses that couldn't be eliminated in this manner remained, as illustrated by:

> ```
> precize    tondas     harojn      ili
> precisely shear-PRES hair-PL-ACC they
> They cut hair precisely
> ```

This sentence previously had three semantically identical parses, and blocking the complementation of a modified verb decreased them to two. Those remaining were parses in which the adverb "precize" modified the verb phrase "tondas harojn" in one case and modified the sentence "tondas harojn ili" in the other case. The rule licensing both of these modifications was adj-head-int.

I investigated dealing with this surviving redundancy by requiring or prohibiting the presence of a subject in the head daughter of an adj-head-int phrase. I found that this would work, but would have adverse side-effects. Requiring a subject (i.e. constraining HEAD-DTR.SYNSEM.LOCAL.CAT.VAL.SUBJ to have the value "null") would eliminate the sole parse of

> ```
> ili precize tondas
> ```

in which the adverb modifies the verb before the subject is realized (adverbs attach only to their right). And prohibiting a subject (i.e. constraining HEAD-DTR.SYNSEM.LOCAL.CAT.VAL.SUBJ to have the value "cons") would eliminate a semantically unique (VP-modified) parse of

> ```
> precize ili tondas
> ```

leaving only a parse in which the adverb modifies the subject.

Finally, I succeeded in eliminating the remaining redundancies by barring the head daughter of an adj-head-int phrase from having a post-head subject descendant. To accomplish this, I first investigated the possibility of requiring the adj-head-int type to have a value of "-" on HEAD-DTR.SYNSEM.LOCAL.CAT.POSTHEAD. But the POSTHEAD value of the subj-head type is likewise "+", so this method would eliminate a semantically unique parse in sentences such as

> ```
> precize ili tondas
> ```

So, I resorted to defining a Boolean feature, "PREMODIFIABLE", that permits or prohibits

adverbial modification. I made it appropriate for the type "valence", constrained the "head-subj-phrase" type to have a value "-" on SYNSEM.LOCAL.CAT.VAL.PREMODIFIABLE, and constrained the "all-adverb-lex" type to make the first item of its SYNSEM.LOCAL.CAT.HEAD.MOD list have a "+" value on LOCAL.CAT.VAL.PREMODIFIABLE. Since the adj-head-int phrase requires these two values to be identical, adverbial modification is now prevented once a post-head subject is realized.

### *Imperatives*

Esperanto has a verbal suffix that expresses imperativity, exhortation, and volition. Occasionally it is used even to express states and actions considered but not realized, as in:

```
La  akvo      tro malprofundas ke   vi  plonĝu
the water-NOM too shallow-PRES that you dive-IMP
The water is too shallow for you to dive


Ili  permesis    ke   mi parolu
they permit-PAST that I  speak-IMP
They permitted me to speak
```

This -u suffix alternates with the suffixes expressing declarative tenses and the suffix expressing conditionality. For example, the noun "vintro" ('winter') and the verb stem "ven" ('come') can be combined in any of these ways (and equivalently in the noun-verb order):

```
venis vintro (Winter came)
venas vintro (Winter is coming)
venos vintro (Winter will come)
venus vintro (Winter would come)
venu vintro (Let winter come)
```

I am calling any verb with the -u suffix and "imperative" verb.

Imperative verbs are distributionally almost identical to declarative and conditional verbs, in that, like them, imperative verbs may have subjects of all persons and numbers, may be embedded in both declarative and interrogative complementizer clauses, and may be the verbs of wh-questions. Examples:

```
mi pensu
I  think-IMP
Let me think


cxu mi atendu
TF  I  wait-IMP
Shall I wait?


mi ne  scias     cxu mi atendu
I  not know-PRES TF  I  wait-IMP
I don't know whether (I ought) to wait


kial ili  tion      kredu
why  they that-N-ACC believe-IMP
Why should they believe that?


mi petas        ke   vi      rapidu
I  request-PRES that you-NOM fast-IMP
Quickly, please
```

However, when an imperative verb that would require a subject if declarative or conditional has no overt subject and is the verb of a declarative sentence, then the sentence is grammatical and the implied subject has the second person and an underspecified number. The number may be revealed if the verb has an adjectival predicative complement, as in:

```
estu    saĝa
be-IMP  wise-NOM-SG
```

```
Be wise (singular)

estu   saĝaj
be-IMP wise-NOM-PL
Be wise (plural)
```

The grammaticality of second-person imperative verbs with no overt subjects does not extend to other contexts and does not extend to non-imperative verbs in any context. Thus:

```
*kial tion       kredu
why   that-N-ACC believe-IMP

*mi petas        ke   rapidu
I    request-PRES that fast-IMP

*respondos baldaŭ
answer-FUT soon
```

Some verbs lexically license null subjects, and these may be inflected imperatively, as in:

```
mi preĝas    ke   pluvu
I  pray-PRES that rain-IMP
I pray that it rain

varmiĝu
become-hot-IMP
May it warm up
```

Imperative verbs of this class that are the verbs of declarative sentences and have no expressed subjects can produce ambiguities, arising from the two licenses. Thus, "varmiĝu", above, may also mean "(You) warm up". This kind of ambiguity is resolved in the event of a non-nominal predicative complement, since it will be adjectival if the subject is merely covert, but adverbial if null, as in:

```
estu   trankvila
be-IMP quiet
Be quiet

estu   trankvile
be-IMP quietly
Let it be quiet
```

### Interrogatives

Esperanto applies interrogativity to any constituent by preceding it with the question word "ĉu". This word may appear at the beginning of a verb clause. If it is to be attached to a subclausal constituent, that constituent appears immediately after it, though because of the free order of clausal constituents there are typically both clausal and subclausal interpretations of such an attachment. For example:

```
ĉu pluvas
TF rain-PRES
Is it raining?

ĉu min vi       kritikas
TF me  you-NOM criticize-PRES
Is it I that you're criticizing?
Are you criticizing me?

ĉu matene     vi      plej komforte    verkas
TF morning-ADV you-NOM most comfort-ADV write-as-compose-PRES
Is it in the morning that you're most comfortable writing?
Are you most comfortable writing in the morning?
```

When a speaker wants to disambiguate the attachment, the usual device is to ask a wh-question and follow it with an interrogative sentence fragment containing only the questioned

constituent, such as:

```
kiam vi      plej komforte    verkas;              ĉu matene
when you-NOM most comfort-ADV write-as-compose-PRES TF morning-ADV
When are you most comfortable writing; in the morning?
```

Any verb form may be subjected to interrogative treatment:

```
ĉu pluvas
TF rain-PRES
Is it raining?
```

```
se mi eksiĝus     ĉu vi      plendus
if I  ex-MED-COND TF you-NOM complain-COND
If I resigned, would you complain?
```

```
ĉu ni dancu
TF we dance-IMP
Shall we dance?
```

```
mi ne  scias      ĉu plori
I  not know-PRES TF weep-INF
I don't know whether to cry
```

It appears to me, then, that Esperanto syntax treats interrogativity as a higher-level attribute than imperativity. A clause can be interrogative or non-interrogative. In either case, its verb can be indicative (tensed), conditional, or imperative.

### *Clausal Embedding*

Esperanto exhibits embedded clauses introduced by both relative words and complementizers. This discussion is limited to the latter.

A complementizer-embedded clause's valence depends on the complementizer. If the complementizer is the interrogative one ("ĉu"), the clause has an interrogative character, which does not depend on whether its verb is indicative, conditional, or imperative. However, if the complementizer is the non-interrogative one ("ke"), the clause's character depends on the verb, with an imperative verb making the clause imperative and an indicative or conditional verb making the clause declarative.

This declarative, imperative, or interrogative character of the clause determines its ability to be selected as a complement by a matrix verb. Some verbs can and others cannot be complemented by a complementizer clause. Of those that can be, for some the complement clause must be interrogative, for others it must be non-interrogative, and for others it may be either. When it must be non-interrogative, in some cases it must further be imperative. In all of these cases, however, the embedded clause's verb may be imperative (even though that doesn't make the clause itself imperative when the clause's complementizer is interrogative).

For example, the verb "postul" ('demand') requires that a complementizer-clausal complement be imperative (thus with the "ke" complement and an imperative verb), as in:

```
mi postulas     ke   vi      silentu
I  demand-PRES that you-NOM silent-IMP
I demand that you be silent
```

The verb "esper" ('hope') requires that a complementizer-clausal complement be non-interrogative (with "ke") but not necessarily imperative, as in:

```
mi esperas   ke   infanoj      silentu    dum    koncertoj
I  hope-PRES that children-NOM silent-IMP during concerts-NOM
I hope children are expected to be silent during concerts
```

```
iuj        esperas   ke   ili  neniam mortos
some-N-PL hope-PRES that they never  die-FUT
Some hope that they'll never die
```

The verb "demand" ('ask to know') requires that a clausal complement be interrogative (with "ĉu"), optionally with an imperative verb, as in:

```
demandu ĉu vi      restu      hejme
ask-IMP TF you-NOM remain-IMP home-ADV
Ask whether you should stay home

mi demandos ĉu vi      rajtas      kunveturi
I  ask-FUT  TF you-NOM right-PRES with-travel-INF
I'll ask whether you can come along
```

Other verbs, including "sci" ('know') and "dub" ('doubt'), permit the full range of complementizer-clause complement types illustrated above.

### Syntactic Coverage

The current grammar covers both matrix and embedded clauses with the full range of indicative, conditional, and imperative verb forms, and covers embedded clauses with both interrogative and non-interrogative complementizers. Verbs correctly select for complement clause types, including the permitted combinations of complementizers and verb forms. The correct range of free word orders within clauses is recognized. As a result, bizarre and human-challenging center-embedded sentences are treated as grammatical, such as:

```
dubas       ke  dubas       ke  dubas       ke  dubas       mi mi mi mi
doubt-PRES that doubt-PRES that doubt-PRES that doubt-PRES I  I  I  I
I doubt whether I doubt whether I doubt whether I doubt
```

Realistically complex sentences combining verb forms, the forms of clausal embedding, and adjectival and adverbial modifiers are correctly parsed, such as:

```
postulis nur ili ke mi demandu cxu flavaj hundoj frenezaj timu ke iliaj
kadukaj estroj nepre tondos cxies harojn
;demand-PAST only they that I ask-IMP TF yellow-NOM-PL dog-NOM-PL
;crazy-NOM-PL fear-IMP that their-NOM-PL decrepit-NOM-PL masters-NOM
;definitely shear-FUT everybody's hairs-ACC
;'They alone demanded that I ask whether crazy yellow dogs should fear
;that their decrepit masters would definitely cut everybody's hair.'
```

Individual violations introduced into such sentences prevent them from being parsed.

The grammar correctly accepts covert-subject imperative sentences and correctly rejects attempts to embed such clauses.

Multiple parses that are semantically indistinguishable have been eliminated. The second-person pronoun "vi" has singular and plural interpretations, so any appearance of it in a sentence gives rise to dual parses. Possessive pronouns are treated as existing in determiner and adjective versions, so their use produces dual parses, whose MRSs differ but not in a way that seems entirely correct or useful, so I intend to seek a better treatment of these pronouns.

Various constructions not intended to be covered have not been covered. These include infinitive verbs, predicative complements, and categorial (noun-to-verb etc.) derivations.

### Syntactic Analysis

The main device used for the correct treatment of matrix and embedded clauses is a set of additional Boolean "cat" features:

IMPER marks a phrase as imperative or not.

QUESTION marks a phrase as interrogative or not.

SENTENCE marks a phrase as a semantically annotated clause or not.

ROOTMOM marks a phrase as mandatorily having the root node as its mother or not.

I have constrained the "headed-phrase" type to require it to copy its head daughter's values of IMPER and QUESTION to itself.

The imperative verb lexical rule gives a word an [ IMPER + ] value, and the other verb lexical rules give a word an [ IMPER - ] value. This value is copied to the root node by virtue of the just-mentioned constraint, except that the copying is stopped if the VP is a complement. In that case, the complementizer decides what its own IMPER value is. If it is the interrogative complementizer, it makes itself [ IMPER - ]. As for the "ke" complementizer, the grammar treats it as two separate lexemes with the same stem, a declarative complementizer and an imperative complementizer. The declarative "ke" requires that its complement be [ IMPER - ] and gives itself that value, too. The imperative "ke" requires that its complement be [ IMPER + ] and gives itself that value, too. This bifurcation seems to facilitate assigning two different MSG values to the complementizer depending on the IMPER value of its complement.

The QUESTION feature's value is set by complementizers. It is "+" in the interrogative and "-" in the declarative and imperative complementizers. The headed-phrase constraint copies this value to the CP mother and, if there is a CP grandmother, to it, too.

The grammar uses the SENTENCE feature to assure that a phrase is semantically annotated before being used as a complement or as a sentence. The grammar adds to the constraints on the "head-nexus-phrase" type by making it [ SENTENCE - ]. This value is inherited by the various subject-head and complement-head phrase types. The complementizers require their complement to be [ SENTENCE + ] and give themselves a [ SENTENCE - ] value. Four grammar rules, for declarative, imperative, command, and interrogative clauses, require their head daughters to be [ SENTENCE - ] (as phrases generally are) and give themselves a [ SENTENCE + ] value.

The ROOTMOM feature assures that imperative clauses with covert subjects appear only as sentences and not as embedded clauses. Phrases generally inherit [ ROOTMOM - ] from three supertypes, "basic-head-subj-phrase", "basic-head-comp-phrase", and "adj-head-int-phrase". The complementizer types require their complement clause to be [ ROOTMOM - ]. The "bare-imperative-phrase" type, which licenses covert-subject imperative phrases, makes itself [ ROOTMOM + ]. This prevents it from then being used as a complement by a complementizer. All clausal semantic types copy their daughters' ROOTMOM value to themselves. The "command" clause type requires itself (and thus its daughter) to be [ ROOTMOM + ], so this type is used for a covert-subject imperative sentence. The other clause types require themselves and their daughters to be [ ROOTMOM - ], so they are used for all other clauses.

Verb lexeme types use these features for complement selection. Those selecting imperative complementizer-clause complements require their complements to be [ SENTENCE - ], [ QUESTION - ], and [ IMPER + ]. Those selecting non-interrogative complements that may be declarative or imperative require their complements to be [ SENTENCE - ] and [ QUESTION - ] but don't specify IMPER. Those selecting interrogative complements require them to be [ SENTENCE + ], [ QUESTION + ], and [ IMPER - ]. The effect is to require interrogative complements to have semantic annotations added, reflecting the contribution of the interrogative complementizer, but not to require that of other complements, since "ke" doesn't require any additional semantic annotation.

This regime is compatible with the fact that complementizer phrases with the interrogative "ĉu", when semantically annotated, are qualified to be both stand-alone sentences and embedded clauses, while those with the non-interrogative "ke" may be used as complements but cannot become stand-alone sentences.

I looked for existing features that would have the same functionality but didn't find any that seemed practical. I considered trying to use constraints on semantic features, and in fact some of those do seem to have desired effects, but I wondered whether reliance on semantic constraints for syntactic modeling could cause future problems arising from the tendency for many syntactic forms to have the same semantic value and vice-versa. So I surmised that adding "cat" features would be the most robust strategy.

### *Semantic Analysis*

Lexical rules that inflect nouns and determiners for number constrain the lexemes' values of CONT.HOOK.INDEX.PNG.NUM accordingly. Lexical entries for pronouns do so to NUM and PER, with a lexical type doing this batchwise for the numerous third-person singular pronouns.

Several lexical and phrasal types constrain their CONT.MSG.PRED values, and verb types constrain those of their complements. However, these verb types also constrain their complements' CAT.IMPER, CAT.SENTENCE, CAT.QUESTION values, making it unclear, pending further investigation, whether the PRED constraints are superfluous.

Other semantic constraints are introduced in various type definitions, but I cannot explain their purposes and effects without further study.

The semantic contribution of the interrogative complementizer is not yet being correctly reflected in indexed MRS results. All proposition_m_rel elements above such a question's question_m_rel element are missing, and the question_m_rel's argument is an indirect rather than a direct reference to the embedded proposition.

Other than that problem, MRS results appear to be as expected for declarative and interrogative matrix clauses with any level of embedding of declarative clauses.

The grammar's definition of "complementizer-lex-item" does not include the recommended identification of the complementizer's MSG with its complement's MSG value. This is because this identification appeared to prevent the correct use of this type as a supertype for the interrogative complementizer type, and because the omission of this constraint didn't seem to cause any errors.

### *Testing*

My testing has been confined to single-sentence and batch parses within LKB. I intend to begin deploying the itsdb tool once I have corrected the semantics of the existing grammar.