

Linguistics 567, Spring 2005, Jonathan Pool  
Lab 7 (Esperanto)

**Preparation**

*Semantic Corrections*

My prior report described errors in indexed MRS lists when sentences included interrogative complementizers. Further testing revealed that MRS errors also occurred with embedded clauses containing singular noun complements of transitive verbs and with any nouns modified by adjectives.

I added some items to the test suite to be sure that it would include sentences of the kinds that had produced MRS errors.

On Emily Bender's advice and in accord with comments in the matrix.tdl file, I investigated the nodes of each sentence yielding erroneous MRSs and discovered HCONS and RELS lists with indeterminate lengths. When their lengths were made determinate, and when new parsing errors arising from these corrections were corrected, the MRS errors were eliminated, the test suite's sentences got the correct numbers of parses, and illustrative sentences' parse trees were as expected.

**Sentential Negation**

*Esperanto Facts*

Sentential negation in Esperanto typically involves the adverb "ne", which, like other adverbs, precedes the modified constituent. This adverb may modify a constituent of any head type, including nouns, pronouns, adjectives, determiners, adverbs, verbs, complementizers, and prepositions. Its modified constituent follows it. When it is negating a sentence, it does so by attaching to an immediately following verb, VP, or sentence. If any argument of the verb separates it from the verb, "ne" is interpreted as modifying the intervening argument instead of the verb. But, if there is another intervening modifier of the same verb, the attachment of "ne" is open to multiple interpretations. Example of a negation that will **not** be interpreted as sentential:

ne espero vin savos  
not hope-NOM you-ACC save-FUT  
What saves you won't be hope

Example of an ambiguous negation:

mi ne rapide respondos tian demandon  
I not fast-ADV answer-FUT such-SG-ACC question-ACC  
I won't answer such a question quickly, if at all  
I'll answer such a question, but not quickly

Example of an unambiguously sentential negation:

ni ne sukcesis kapti la muson  
we not succeed-PAST catch-INF the mouse-ACC  
We didn't succeed in catching the mouse

Since compounding is generally productive in Esperanto, speakers can attach "ne" to a negated word as a prefix, too. Doing so is one method for preventing ambiguities like the one illustrated above. If "ne rapide" became "nerapide", constituent negation would be unambiguous. Sentential negation by "ne"-prefixation is so rare as to be arguably ungrammatical, but for a class of verbs whose opposites are equivalent to their negations, like "sukces" ('succeed'), speakers achieve the equivalent of sentential negation by prefixing the verb with the inversional prefix "mal" ("malsukcesis" = "ne sukcesis").

The negating adverb, and more generally scopal adverbs, are generally interpreted as generous, i.e. as attaching to the minimal rightward constituent. For example:

```
ne tio doloras min
not that hurt-PRES me
What hurts me isn't that

ne min tio doloras
not me that hurt-PRES
I'm not the one whom it hurts
```

The most common order for neutral sentential negation depends on the constituents' lightnesses, but in any case involves the negator "ne" immediately preceding the verb, as in:

```
tio ne doloras min
that not hurt-PRES me
It doesn't hurt me

tio min ne doloras
that me not hurt-PRES
It doesn't hurt me
```

### *Implementation*

To cover sentential negation, I tried to capture the facts about "ne" by defining a type of scopal adverb and treating "ne", along with several already-defined adverbs, as an instance of that type. This left adverbs partitioned into intersective and scopal, and within each type sub-partitioned into bare and inflected.

The scopal adverb types are defined as attaching to the modified constituents by the "adj-head-scop-phrase" rule.

### *Adverb Strategy*

Although my purpose was to account for sentential negation, trying to do so in a way that improves adverbial coverage made it necessary to address some general adverb issues.

As a first approximation, scopal adverbs can modify any head type, while intersectal adverbs can modify only verbs, adjectives, and adverbs.

Thus, creating separate adverb types permits the grammar to correct its over-acceptance of sentences containing scopal adverbs. For example, the previous grammar treated this sentence as ambiguous by permitting the adverb to modify the subject pronoun, as well as the entire sentence, in order not to prevent other adverbs, such as "nur" ('only'), from being able to modify pronouns:

facile mi manĝas vitron  
easily I eat-PRES glass-ACC  
I easily eat glass

By separating scopal adverbs, we can restrict the modification valence of this adverb and correctly treat this sentence as unambiguous.

The split of adverbial types can also be expected to help in the correction of the non-acceptance of sentences in which adverbs modify constituents on their left. This type split should help here because such sentences appear to be uniformly unacceptable when the adverb is scopal, as in:

\*mi rompis la vitron preskaŭ  
I break-PAST the glass-ACC almost  
  
mi rompis la vitron hieraŭ  
I break-PAST the glass-ACC yesterday  
I broke the glass yesterday

It isn't clear, however, that all Esperanto adverbs can be felicitously classified as intersectal versus scopal, or that this classification will suffice to constrain their distributions. For example, "tre" ('very') and "tro" ('too, excessively') seem to be intersectal insofar as they are arguably unable to modify nominal constituents, but they seem to have more in common semantically with the adverbs that can modify nouns, such as "preskaŭ" ('almost'), "nur" ('only'), and "ne".

#### *Adverb Tactics*

To make the intersectal-scopal partition of adverbs work, I limited the use of the PREMODIFIABLE feature to intersectal adverbs and added another Boolean "valence" feature PREQUALIFIABLE for scopal adverbs. The new feature has a negative value for two types that are not subject to modification by scopal adverbs. One of these is "subj-head-phrase"; in this way, a sentence such as

nur mi manĝas vitron  
only I eat-PRES glass  
I, alone, eat glass

is to be parsed only one way, with "nur" modifying just "mi", in contrast to the sentence above where "facile" is the adverb. The same holds for "comp-head-phrase".

#### *Intersective Adverb Results*

I did not expect the grammar's performance on intersective adverbs to change, but in testing I did discover some parsing redundancy that I had previously overlooked. This can be illustrated with the family of sentences containing these words:

ili = they  
hejme = home-ADV  
kuiris = cook-PAST  
panon = bread-ACC

The adverb, being intersective, can modify only a verbal constituent in any of the sentences in this family. Every word order in which the adverb precedes the verb is grammatical. There are

twelve such orders. Of these, ten yield unique parses, while two yield two parses each. These are

```
hejme kuiris ili panon
hejme panon kuiris ili
```

In each case, the semantic representations of the two parses are identical, so the parses are redundant.

I have not discovered a parsimonious and robust way to prune the two excess parses. Each parse that could be considered redundant is similar to the unique parse of another word order, with the only differences appearing two levels below the node where a constraint might rule out one of the parses. A constraint that would decrease redundancy while doing no harm would need to express a rule such as "head-comp may not have a head daughter whose head daughter is modified by an intersectal adverb". If feature copying were introduced to make such a rule possible, it isn't obvious that the rule would work and suffice with more complex sentences.

### *Scopal Adverb Results*

The results of the current implementation with respect to scopal adverbs were a mixture of success and failure. Success was achieved in the rejection of incorrect interpretations previously accepted as in examples such as "nur mi manĝas vitron" above. The grammar no longer interprets "nur" as possibly modifying "mi manĝas vitron". However, the grammar incorrectly rejects the correct interpretation in which "nur" modifies "mi". More generally, the grammar does not accept the attachment of scopal adverbs to nominal constituents. Thus, although both of the sentences below are grammatical, only the first is parsed by the grammar:

```
ili      kuiris  nur nigran      panon
they-NOM cook-PAST only black-SG-ACC bread-SG-ACC
They cooked only black bread
```

```
ili      kuiris  nur panon
they-NOM cook-PAST only bread-SG-ACC
They cooked only bread
```

And the grammar parses the first sentence in only one way, with "nur" modifying "nigran", instead of accepting the possibility that it modifies "nigran panon".

The reason for this error appears to be that a scopal adverb and its modified constituent combine in an "adj-head-scop-phrase" constituent, whose index type is "event" (even if its head daughter's index type is "ref-ind", as with a nominal head daughter), but the "head-comp" type requires its head daughter's complement's index type to be "ref-ind". This is an unresolved issue with the current grammar, constituting a degradation compared with the prior grammar.

Although for now nominal negation is failing to be parsed, the grammar correctly parses cases of sentential negation (as well as adjectival, adverbial, and complementizer-clause negation).

Some word orders yield two redundant parses. These are:

```
ne doloras min tio
ne doloras tio min
```

Since adverbial revisions are anticipated, I am not currently seeking to prune these.

The MRS lists representing tested cases of sentential negation have conformed to expectations.

### **Potentiality**

#### *Esperanto Facts*

The typical mechanism whereby speakers of Esperanto express the potential of an agent to exhibit an act or state is the "modal auxiliary verb" "pov" ('be able'). Its complement is an infinitive clause with no expressed subject. For example:

hundoj povas kuiri librojn  
dogs-NOM can cook-INF books-ACC  
Dogs can cook books

The generally free order of sentential constituents extends to this verb and its complement, and its complement may be split. Thus, four-word sentences containing

mi = I  
povas = can  
manĝi = eat-INF  
vitron = glass-ACC

can appear grammatically in any of the 24 possible permutations.

In at least some ways, "pov" appears to operate like a raising verb, although comments in the most authoritative traditional grammar (PAG, p. 159) suggest that with sentient agents "pov" may express agential perceptions and/or aptitudes and not protect truth-functional equivalence under passivization.

#### *Implementation*

Notwithstanding the above caveat, I have implemented potentiality by defining a class of raising verbs and an infinitive lexical type.

The infinitive suffix is one of a set of mutually exclusive verb suffixes that express tense and mood. To account for the infinitive form, I revised the treatment of imperative verbs by eliminating the Boolean IMPER feature and defining a multi-valued VFORM feature appropriate for the "head-min" type. The value of VFORM is "vform", and its subtypes form a hierarchy. At the top are "imperative", "infinitive", and "finite". The "finite" type has subtypes "conditional" and "indicative". The "indicative" type has subtypes "past", "present", and "future". These types are selected so that one of them will be appropriate for any constraint on the forms of verbs. I then revised the type definitions where IMPER occurred, so as to make them use VFORM instead. I also added and instantiated a lexical rule for infinitive verbs.

The new "raising-verb-lex" type, instantiated by the verb "pov" ('be able'), is defined as follows:

```
raising-verb-lex := verb-lex & trans-first-arg-raising-lex-item &  
  [ SYNSEM.LOCAL.CAT.VAL  
    [ SUBJ.FIRST #sj,  
      COMPS  
        [ FIRST #comp &
```

```

[ LOCAL.CAT
  [ HEAD verb &
    [ VFORM infinitive ],
    VAL
    [ SUBJ.FIRST #sj,
      COMPS null ] ] ],
  REST null ] ],
  ARG-ST.REST.FIRST #comp ].

```

It requires its complement to be an infinitive VP and requires its complement's subject to be identical to its own subject.

### Results

This implementation appears to have operated correctly, with one significant exception.

The parses and the indexed MRS representations appear to be as expected. The MRS for

```

mi povas manĝi vitron
I can-PRES eat-INF glass-SG-ACC
I can eat glass

```

is

```

<h1,e2:SEMSORT:TENSE:ASPECT:MOOD,
 {h3:pronoun_n(x4:SEMSORT:FIRST:SG:BOOL),
 h5:pronoun_q(x4, h6, h7),
 h8:_can_v(e2, h9),
 h9:_eat_v(e10:SEMSORT:TENSE:ASPECT:MOOD, x4, x11:SEMSORT:THIRD:SG:BOOL),
 h12:_glass_n(x11),
 h13:indef_q(x11, h14, h15),
 h1:proposition_m(h16)},
 {h6 qeq h3,
 h14 qeq h12,
 h16 qeq h8}>

```

The raising-verb construction appears to be robust in the grammar. For example, we can insert adverbs, such as

```

mi povas vitron nur manĝi
I can-PRES glass-SG-ACC only eat-INF
The only thing I can do to glass is eat it

```

and do multiple levels of embedding, such as

```

la hundo freneza povis facile supozi ke la nigra kato tre dubas ĉu
hundoj kadukaj povus katojn manĝi
the dog-SG-NOM crazy-SG-NOM can-PAST easily assume-INF that the
black-SG-NOM cat-SG-NOM very doubt-PRES whether dogs-NOM decrepit-PL-NOM
can-COND cats-ACC eat-INF
The crazy dog could easily assume that the black cat highly doubted
whether decrepit dogs could eat cats

```

The grammar is also correctly intolerant of ungrammaticalities, such as changing the present "dubas" to the infinitive "dubi" in the last sentence above.

The most serious defect of the grammar with respect to "pov" appears to be that it rejects some acceptable word orders. Examples of wrongly rejected orders:

```
vitron  mi povas  mangxi
glass-ACC I can-PRES eat-INF
I can eat glass
```

```
mi vitron  povas  mangxi
I glass-ACC can-PRES eat-INF
I can eat glass
```

This problem resembles the grammar's failure to accept split NP constituents, as in:

```
vitron  mi mangxas  nur verdan
glass-ACC I eat-PRES only green-SG-ACC
Green is the only kind of glass I eat
```

### ***Interactions***

Sentences combining potentiality with sentential negation are correctly parsed, and their MRS representations are appropriate, subject to the limitations described above. This includes sentences such as

```
mi ne povas  ne mangxi vitron
I not can-PRES not eat-INF glass-ACC
I can't avoid eating glass
```

Some alternative sentences differing only in the attachment of negation are semantically different in ways that are subtle. An example:

```
mi ne esperas  ke ili  mortos
I not hope-PRES that they-NOM die-FUT
I don't hope they will die
```

```
mi esperas  ke ili  ne mortos
I hope-PRES that they-NOM not die-FUT
I hope they won't die
```

```
mi esperas  ne ke ili  mortos
I hope-PRES not that they-NOM die-FUT
My hope is not that they will die
```

The MRS representations of these examples appear to reflect intuitively their differences.

### ***Generation***

The grammar appears to support correct generation from at least some sentences with potential and sententially negative semantics.

The sentence "mi povas mangxi vitron" ('I can eat glass') yields one parse, whose semantics generate 48 sentences. This number is the product of 12 (the number of accepted word orders) and 4 (the number of verb forms with propositional semantics, i.e. the number of finite verb forms: past, present, future, and conditional).

The sentence "tio ne doloras min" ('It doesn't hurt me') yields one parse, whose semantics generate 288 sentences. This is the product of 8 (the number of pronouns sharing the third-person-singular semantics of "tio" in the current grammar), 4 (the number of finite verb forms), 6 (the number of accepted word orders), and 1.5 (the mean number of generated copies per unique sentence). Sentences beginning with "min ne" and sentences beginning with "ne" are all generated twice each. The number of unique generated sentences is 192. The duplicate generations reflect semantically duplicate parses that the grammar produces for those sentences, in which the negative adverb "ne" alternatively modifies just the verb or modifies the phrase containing the verb and its postverbal argument(s).

Sentences combining potentiality and negation also generate correctly. Examples tested:

```
mi ne povas dormi
I not can-PRES sleep-INF
I can't sleep
```

```
mi povas ne dormi
I can-PRES not sleep-INF
I'm able not to sleep
```

```
mi ne povas ne dormi
I not can-PRES not sleep-INF
I can't avoid sleeping
```

Generating from this last sentence required increasing the LKB edge limit above 4000.

### *To do*

Correct failure of grammar to allow scopal adverbs to modify nominal constituents.

Investigate conditional permission to intersectal adverbs to modify constituents on their left.

Investigate Emily Bender's advice to amend "verb-lex" to identify subject's INDEX with verb's own XARG, which should eliminate the need to identify the #sj values in raising-verb-lex.

Begin testing with itsdb.