Linguistics 567, Spring 2005, Jonathan Pool
Lab 4 (Esperanto)

Preparation

Subject optionality: Esperanto (as I interpret it) does not permit systematic (i.e. ubiquitous) omission of the sentential subject. Many verbs may be used without subjects, when there is no semantic subject, such as "malfruas" ('It's late') and "varmas" ('It's hot'), but a semantic subject, even if identifiable without ambiguity, may not be omitted.

Object optionality: Esperanto permits systematic omission of the sentential direct object, but no uniform inference about feature values of the omitted object is possible.

Object indefinite null instantiation: A default inference about an omitted object of an intrinsically transitive verb is that the object is indefinite. For some verb-context combinations, a definite focal object is inferred, instead.

Subject indefinite null instantiation: Esperanto does not permit systematic indefinite null instantiation of subjects. This seems to be inferrable from the first proposition above. The indefinite personal pronoun is used instead.

Unexpressed PP arguments: Some PP arguments in Esperanto are omitted, normally when they are part of the common information assumed shared by speaker and hearer. E.g., "la reĝo [de Anglujo]" ('the king [of England]).

Ditransitive Verbs

Sentences added to test suite for ditransitive verbs:

```
*tiu lernanto    igos      min freneza
;that student-NOM make-FUT me  crazy-SG-NOM
;'That student is going to drive me crazy.'
;Marked ungrammatical although actually grammatical, because grammar
;doesn't account for NP AP ditransitive verbs yet.

*tiu lernanto    igos      min frenezaj
;that student-NOM make-FUT me  crazy-PL-NOM

la   regxo    igos      min hundo
;the king-NOM make-FUT me  dog-NOM
;'The king is going to turn me into a dog.'
;'A dog will make me the king.'
;Ambiguous; 2 parses should exist.
```

```
igos     min hundo   la  regxo
;make-FUT me  dog-NOM the king-NOM
;'The king is going to turn me into a dog.'
;'A dog will make me the king.'
;Ambiguous; 2 parses should exist.

*la  regxo     igos     min hundon
;the king-NOM make-FUT me  dog-ACC

la  lernantoj    miaj      hundojn supozas     regxoj
;the students-NOM my-PL-NOM dogs-ACC assume-PRES kings-NOM
;'My students assume dogs are kings.'
;'Kings assume dogs are my students.'
;Ambiguous; 2 parses should exist.

*la  lernantoj    miaj      hundoj   supozas     regxoj
;the students-NOM my-PL-NOM dogs-NOM assume-PRES kings-NOM

*ni havas      hundon  regxo
;we have-PRES dog-ACC king-NOM

*mi kuiras     farunon    pano
;I  cook-PRES flour-ACC bread-NOM

mi igos      farunon    pano
;I make-FUT flour-ACC bread-NOM
;'I'll turn flour into bread.'
```

These sentences test whether the grammar accepts ditransitive verbs only with two NP complements, of which one is nominative and one is accusative, and whether it accepts sentences with such complements only when the verb is ditransitive. They also verify that the verb and its 3 arguments may appear in various (actually, any) order. Finally, they test that the grammar recognizes double-NP ditransitive sentences as ambiguous, because the subject and one complement are both nominative, and they can't be definitively distinguished by constituent order.

Ditransitive verbs more often take adjectival and prepositional complements. To account for adjectival complements we need to specify that they are nominative but agree with the NP complement in number.

Argument Optionality

Subjects are not optional, but complements generally are optional. Their omission is rare and awkward with some verbs, which might be interpreted as requiring expressed complements. I am doing that in the case of "havi" ('have') to assure that the grammar can account for any such verb that might be interpreted to exist.

It is arguable that no omitted complement of any particular verb can be guaranteed to have a uniform definiteness value. However, I have selected particular verbs to be treated as having omitted complements of uniform definiteness, to assure that the grammar can account for such verbs if they are deemed to exist. For example, I am analyzing "legi" ('read') as non-uniform, "kuiri" ('cook') as uniformly indefinite, and "supozi" ('assume') as uniformly definite.

To account for argument optionality, I defined these lexical types and made one of them the supertype of each transitive verb:

> trans-verb-lex: optional complement, nonuniform definiteness if omitted

> ini-comp-trans-verb-lex: optional complement, indefinite if omitted

> dni-comp-trans-verb-lex: optional complement, definite if omitted

> comp-trans-verb-lex: mandatory complement

To test for argument optionality, I used these sentences:

```
ili kuiras
;they perform-cooking-PRES
;'They are cooking.'

ili   kuiras    hundojn
;they cook-PRES dogs-ACC
;'They are cooking dogs.'

mi supozas
;I assume-PRES
;'I suppose.'

mi gxin    supozas
;I it-ACC assume-PRES
;'I suppose so.'

*gxin    supozas
;it-ACC assume-PRES

ni   havas regxon
;we have   king-ACC
;'We have a king.'

*ni havas
;we have
```

```
*ni havas regxo
;we have  king-NOM

*ni havas hundon  regxo
;we have   dog-ACC king-NOM
```

These sentences test whether the grammar assigns one structure to each grammatical sentence and none to each nongrammatical sentence. One of the verbs, "supozi" ('assume'), is also ditransitive, but ditransitive verbs in this language have only mandatory arguments, so the sentences with "supozi" also test that the grammar analyzes this verb only as an instance of dni-comp-trans-verb-lex when it has less than three arguments. Two of these sentences, namely the grammatical sentences with omitted complements, also test for the grammar's analysis of the definiteness value of the omitted complement. The feature structure of the V parent of the verb contains the correct "+" or "-" value of SYNSEM.LOCAL.CAT.VAL.COMPS.FIRST.LOCAL.CONT.HOOK.INDEX.DEF.

I have not discovered any discrepancies between the expected and the actual performance of the grammar.

Modification

Adjective Facts

Adjectives in Esperanto may modify nouns, including pronouns. Adjectives may precede or follow the nouns that they modify, and if a noun is modified by multiple adjectives some may precede it while others follow it. Generally, a determiner if present precedes all adjectives and the noun.

However, split NPs are found, as in:

```
kukon     mi neniam gustumis   tiom      dolĉan
cake-ACC I  never   taste-PAST that-much sweet-SG-ACC
I've never tasted cake this sweet.

ju pli vi strebos, des pli brilan vi akiros sukceson
the more you-NOM strive-FUT the more brilliant-SG-ACC you acquire-FUT
success-SG-ACC
The more you strive, the more brilliant success you'll achieve.
```

So are embedded determiners, as in:

```
la kalumnio portas sian langon la venenan rapide de poluso al poluso
the slander-NOM carry-PRES its-REFL-ACC tongue-ACC the
poisonous-SG-ACC rapidly from pole-SG-NOM to pole-SG-NOM
The poisonous tongue of slander darts from pole to pole.
```

I am temporarily defining these constructions as ungrammatical in the variety of Esperanto described by the grammar being developed.

Adjectives are inflected for number and case and agree with the nouns they modify or express predications about. Adjectival inflection is fully productive: Any stem not yet inflected for part of speech can be inflected as an adjective. This includes stems that can be used as words without inflection. For example, "en" ('in') is a preposition when uninflected, but when inflected adjectivally it becomes "ena" ('internal').

A closed set of lexemes can be analyzed as adjectives without inflection. These are the lexemes in the "correlative" table ending in "-a" and "-u", such as "tia" ('such'), "kiu" ('which'). These are entire morphemes, rather than adjectivally inflected stems, because what look like their stems aren't productively inflectable; for example, "kii" isn't a verb (however, stems are sometimes back-formed and then inflected).

Possessive pronouns obey the morphology of adjectives and may be used adjectivally, though when used prenominally they are usually interpreted as alternating with definite determiners.

Rare cases of multiple-part-of-speech inflection are attested. For example, the lexeme "viv" ('life/ live') has been found inflected as an imperative verb, "vivu", used as a quasi-lexeme with the meaning 'toast to one's health', and further inflected, e.g., as an adjective "vivua" ('relating to a toast'). Provisionally, I am choosing to treat such inflections as ungrammatical or to treat any such case as a case in which the initial inflection has evolved into an unanalyzed stem. The same solution can apply when quasi-stems in the correlative table become secondarily treated as ordinary lexemes.

Adverb Facts

A small fixed set of bare adverbs exists. In addition, adverbial inflection is fully productive in the same manner as adjectival inflection; thus, any stem, even a bare-adverb stem, can be inflected for adverbial part of speech. Any adverb that has been inflected for part of speech can is then inflected for nominative or accusative case.

Adverbs can modify lexical and phrasal nominal, pronominal, verbal, adjectival, prepositional, and adverbial constituents. The case of an adverb does not agree with the case of any other constituent.

Copular qualitative predications about clauses are made with adverbs. For example:

```
manĝi   po              ok    ovojn    tage    ŝajnas/estas      stulte
eat-INF at-the-rate-of eight eggs-ACC day-ADV seem-PRES/be-PRES stupid-ADV
Eating eight eggs a day seems/is stupid.
```

Adverbial word order is problematic. Because of the multiple parts of speech they can modify and the lack of agreement of adverbs, unless their locations determine their attachments their attachments can be highly ambiguous. Speakers usually use and interpret adverbs as attached to constituents that immediately follow them (see PAG, pp. 363-364). There appear to be some significant exceptions. One is that the emphatic uninflected adverb "mem" ('myself/yourself/ ...') appears immediately after the word or phrase it emphasizes. Another is that adverbs qualifying verbal nouns appear after them, as in "la vojaĝo teren" ('the trip earthward'). Another is that adverbial phrases, particularly heavy ones and ones in the accusative case, can appear sentence-finally when attached to a preceding VP. It might seem reasonable to enshrine the general rule and these exceptions as definitive. This would correctly make these sentences equivalent:

```
li parolas    tro laŭte
he speak-PRES too loud-ADV
He speaks too loud.

li tro laŭte    parolas
he too loud-ADV speak-PRES
He speaks too loud.
```

But it would incorrectly exclude as grammatical a third, likewise equivalent, sentence:

```
tro laŭte    li parolas
too loud-ADV he speak-PRES
He speaks too loud.
```

Solving this problem by liberating adverbial word order would create another problem by predicting ambiguity that in fact does not exist, as in:

```
nur li parolis
only he speak-PAST
Only he spoke.
```

which cannot be interpreted to mean 'He merely spoke', because, unlike "laŭte", "nur" is consistently attached to an immediately following non-sentential constituent.

Another exception is illustrated in the above example where the adverb in "po ok ovojn tage" ('eight eggs a day') is attached to its left.

Certain adverbs are interpreted as sentential modifiers in some contexts. Those adverbs are not always unambiguously sentential, however. For example, "kredeble" ('believably') is also used in the senses of 'apparently' and 'probably', so

```
kredeble        li parolis
believable-ADV he speak-PAST
```

or

```
li kredeble       parolis
he believable-ADV speak-PAST
```

could both mean either 'He spoke believably' or 'He probably spoke'. (And the first of these could also mean 'It's probably he that spoke'.)

Speakers wanting to avoid this ambiguity can place the adverb sentence-finally for VP attachment and sentence-initially to maximize the probability of sentential interpretation, or can assure a sentential interpretation by reformulating the adverb as a verb, as in

```
kredeblas        ke   li parolis
believable-PRES that he speak-PAST
It's probable that he spoke.
```

Suppose we develop the adverbial grammar incrementally. What generalization can we begin with to achieve the greatest coverage? The best candidate appears to be a generalization that every adverbial phrase is attached to some specifier-discharged constituent immediately to its right, with no constraint on the part of speech of that constituent's head, and with all adverbs belonging to a normal or bare inflectional subtype, with these subtypes not differing with respect to word order. I'll adopt that pattern as the first approximation.

Adjective Types

I have defined a type, all-adjective-lex, to contain the features and values shared by all adjectives. It has two subtypes: adjective-lex and bare-adjective-lex. The former covers adjectival lexemes that must be adjectivally inflected before being usable as words. The latter covers lexemes that have null inflection for the adjective part of speech. They differ in the value of their feature NEEDINF, which is appropriate for any subtype of word-or-lexrule and whose values tell what kind of inflection the lexeme is ready for. Ordinary adjectival lexemes (adjective-lex) are licensed for use as adjectives only after being inflected for part of speech, then number, and then case. The closed set of bare adjectives (bare-adjective-lex) skips part-of-speech inflection and must be inflected for number and then case.

Adverb Types

I have defined a type, all-adverb-lex, to contain the features and values shared by all adverbs. It has two subtypes: adverb-lex and bare-adverb-lex. The former covers adjectival lexemes that must be adverbially inflected before being usable as words. The latter covers lexemes that have null inflection for the adverb part of speech. They differ in the value of their feature NEEDINF. Ordinary adverbial lexemes (adverb-lex) are licensed for use as adverbs only after being inflected for part of speech, then case. The closed set of bare adverbs (bare-adverb-lex)

requires no inflection.

Intersectivity versus Scopality

On the suggestion of Christie Evans, I am surmising that the wide valence of adverbs described above is due in part or in whole to my inclusion of adverbs with scopal semantics in the test suite. Phrases such as "rapide biciklo" ('rapidly bicycle') may be impossible to interpret, and it may be useful to treat them as ungrammatical by defining adverb (and adjective) subtypes to distinguish scopal from intersective lexemes and limit the latter's valences more greatly than the former's.

Testing

Adjectives have been included in prior testing, because they were necessary for a grammar that would account for agreement.

Adverbs have not previously been included in the test suite. These items test for the grammar's predictions on sentences containing adverbs:

```
legas      nepre       junaj        lernantoj
;read-PRES definite-ADV young-PL-NOM students-NOM
;'Students that read are for sure the young ones.'
;'For sure it's young students that read.
;Ambiguous; 2 parses should exist.

*legas     nepr        junaj        lernantoj
;read-PRES definite-0 young-PL-NOM students-NOM

nepre         legas     junaj        lernantoj
;definite-ADV read-PRES young-PL-NOM students-NOM
;'It's absolutely true that young students read.'
;'Young students definitely read.'
;Ambiguous; 2 parses should exist.

nepre         junaj       lernantoj    legas
;definite-ADV young-PL-NOM students-NOM read-PRES
;'Students that read are for sure the young ones.'
;'For sure it's young students that read.
;'It's absolutely true that young students read.'
;Ambiguous; 3 parses should exist.

nepre         la junaj       lernantoj    legas
;definite-ADV the young-PL-NOM students-NOM read-PRES
;'For sure it's the young students that read.
;'It's absolutely true that (the) young students read.'
;Ambiguous; 2 parses should exist.
```

```
*legas      junaj       lernantoj     nepre
;read-PRES definite-0 young-PL-NOM students-NOM
```

Some limitations in the grammar's coverage vis-a-vis modification facts in Esperanto as actually used have been described above, and these are demonstrated in the results of testing against the revised test suite.

I have encountered, however, one unexpected failure of the grammar that I have not yet been able to debug. The grammar is incorrectly dealing with adverbial inflection. One manifestation of this defect is that uninflected ordinary adverbs, such as "nepr" (definite-0), do not cause the grammar to reject an otherwise grammatical sentence. Another is that the grammar is combining ordinary adverbs with their modified constituents before the adverbs are inflected for case. Case inflections are performed, but case-inflected adverbs are not used in the construction of the parental phrase.

I could not cure this defect simply by declaring uninflected adverbs to be INFLECTED -, similarly to adjectives, because this would only stop less than fully inflected adverbs from being used in parental phrases. It would not cause fully inflected adverbs to get used. My attempts to diagnose this defect have included inspecting the feature structures of the adj-head-int nodes and the nom-adv-lri or acc-adv-lri nodes that should be their non-head daughters, and also the unification of the fully inflected adverbial nodes with the values of ARGS.FIRST of the adj-head-int rule. I expected these to fail to unify and reveal a problem, because the rule isn't using the fully inflected adverb, but they did unify.