# The *GeLexi* MT Project

*Gábor Alberti and Judit Kleiber*

Linguistics Dept., University of Pécs
H-7624 Pécs, Hungary, Ifjúság útja 6.
gelexi@btk.pte.hu

**Abstract.** Our double Machine (-Aided) Translation project to be demonstrated here (sections 3-4) is based on the parsing project [5-9, 12] of our Pécs theoretical and computational linguistics research team, called *GeLexi* ('Generative LEXIcon'), whose (knowledge-based) parsers (section 2) rely on "totally lexicalist" [21, 2, 4, 7] representative mini-grammars of English and the (morphologically very rich) Hungarian language (section 1), and a developed version [3] of Kamp's DRT [19, 35, 20].

## 1. The Background Grammar

As for our team's "philosophical" background and attitude, we would like to verify that computational linguistics is worth returning from the nowadays wide-spread attitude characterized by "shallow parsing" (which is held to save expenses) [15] to the pure theoretical (generative) linguistic basis.[1] Our crucial argument relies on a double (parallel computational and linguistic) chance available in the early years of the new century: to use simultaneously, on one hand, a significantly greater number of huge patterns than earlier due to the immense increase in memory capacity [30], and to work out a formal grammar, on the other hand, showing the distribution of capacity advantageous in modern computer science (in harmony with the development mentioned above): "minimal processing – maximal database". This latter chance has something to do with the sweeping lexicalist turn [13, 14, 17, 28, 31] in generative linguistics, which used to be chiefly "process-oriented" (i.e. syntax-centered) in its first period [16]; the current attitude can be characterized by two mottoes of Joshi's [18], the father of mildly context-sensitive grammars [26]: "Complicate Locally, Simplify Globally", and "Grammar ≈ Lexicon".

What we propose as the ideal background grammar, in harmony with this favorable tendency, is a new sort of generative grammar, *GASG* ("Generative /Generalized Argument Structure Grammar", defined in [7] and demonstrated in a wide range of papers [1-9, 12]), which is more radically "lexicalist" [21] than any earlier one. It is a modified Unification Categorial Grammar [21, 36, 35], from which even the principal syntactic "weapon" of CGs, Function Application, has been omitted. What has remained is *lexical sign* [27] and the mere technique of *unification* as the engine of combining signs.

## 2. The Parser of *GeLexi*

Our GASG-parser, in accordance with the basic task of every generative grammar [16, 26], decides whether a sentence is *grammatical* (it provides, as a "by-product", a *morphophonological* analysis and a compilation of *grammatical relations*)[2], and it

---

[2] GASG is a *monostratal* grammar, with a formalized version of the Saussurean *sign* [27] in its core; the rich description of a lexical sign (say, out of a group of lexical signs selected from the Lexicon in order to combine them to form a sentence) serves a double purpose: it characterizes the *potential environment* of the given sign in possible grammatical sentences in order for the sign to find the morphologically (or in other ways) compatible elements and to avoid the incompatible ones in the course of forming a sentence, and the lexical description characterizes the sign itself in order for other signs to find (or not to find) it, on the basis of similar "environmental descriptions" belonging to the lexical characterizations of these other signs. In the model demonstrated in our earlier papers these richly structured lexical items were assigned to *words*, and the lexical description of morphologically complex words was claimed to be calculable in a multiple lexical inheritance network. Recently we [6] worked out a better method — "totally lexicalist morphology" — which suits the principle of total lexicalism directly: each single morpheme within words is to be assigned a lexical item. Another characteristic property of our grammar is that it is a generative grammar which, nevertheless, "builds" no PS trees: the crucial means of GASG substituting for phrase structure building mechanisms lies in *ranked* lexical requirements concerning *immediate precedence* between words/morphemes

supplies two kinds of (related) *semantic* representations: a DRS (discourse representation structure) [19] completed with information about its embedding in interpreters' information state also formulated as a DRS [3], and a network of *copredications,* primarily useful in translation (section 4) [8].

The Hungarian sentence in (1) serves as the first illustration:

(1) Énekel-tet-het-l-ek.
     sing-CAUSE-MAY-SG2-SG1
     'I may make you$_{sg}$ sing.'

Examples (2-6) below show our Prolog parser's full answer to the Goal gramm("Énekeltethetlek")., with the final conclusion at the end (6): yes. These points also reflect the main steps of the programme:

(2) LEXICAL ITEMS:

    énekel:
n(1,1,li(m("","énekel",""),labstem("sing",ph onfst(1,2,2,2),2,[["NOM"],["NOM","ACC"]])))
    tet:
n(1,2,li(m("t","A","t"),labder("cause",phonfsu (2,2,0.2,2),2, ac(-1,0,1))))
    het:
n(1,3,li(m("h","A","t"),labsuff("may", phonfsu(1,1,1,2),2,1)))
    l:
n(1,4,li(m("","l",""),labsuff("objperson2",phonfsu (3,2,1,1),2,2.5)))
    ek:
n(1,5,li(m("V","k",""),labsuff("sg1",phonfsu(1 ,1,2,3),2,3)))

(3) SYNTAX:
gr("suff","stem","free",1,2,1,1)
gr("suff","stem","free",1,3,1,1)
gr("suff","stem","free",1,4,1,1)
gr("suff","stem","free",1,5,1,1)

(4) COPREDICATIVE NETWORK:
copr("sing",1,1,"you-sg",1,4,1,1,"arg")
copr("cause",1,2,"I",1,5,1,1,"arg")
copr("cause",1,2,"sing",1,1,2,0,"arg")
copr("may",1,3,"cause",1,2,2,0,"arg")

(5) DISCOURSE SEMANTICS:
provref("fixpoint",[e(1,3,1)])
provref("new",[e(1,1,1)])
pred("sing",1,[e(1,1,1),r(0,1,2)])
provref("new",[e(1,2,1)])
provref("=",[e(1,2,1),e(1,1,1)])

pred("cause",1,[e(1,2,1),r(0,1,1),e(1,1,1)])
provref("new",[e(1,3,1)])
provref("<",[e(1,3,1),e(1,2,1)])
pred("may",1,[e(1,3,1),e(1,2,1)])

(6) yes

As a first step, the relevant *lexical items* are collected (2). The first two coordinates of function *n* express the correspondence between these lexical items and morphemes which can be identified in the input sentence. Suffix *-hAt* 'may', for instance, is the *third* morpheme of the *first* word.: n(1,3,...). The phonetic form of words is divided into three parts, e.g. m("h","A","t") and m("V","k",""), where capital letters, in accordance with the Prolog tradition, refer to variables, expressing underspecified parts of the given morphemes. The allomorphs denoted above are *-hat/het* and *-ek/ök/ok/ak/...;* thus letter strings het and ek are legitimate representatives of the given lexical items in the input sentence (see also Table 3 in (22) in section 4).

This kind of pattern matching is the first *filter* of our Prolog programme, which will be followed by another phonological filter checking *compatibility* between morphemes within words.[3] Then the number in the last but one coordinate position of function *n* encodes *category* of the containing word (noun / verb / etc.). In the case of stems, the last coordinate provides information on possible *argument structures*. *Énekel* 'sing' above, for instance, can be "surrounded" by one of the case frames included in the string occupying the position in question. In sentence (1) the intransitive frame is unified successfully because no "song" is mentioned; *énekel-tet* 'sing-cause',, however, is causative, but this fact is due to the causative morpheme whose intricate influence on case frames is encoded by the mysterious formula ac(-1,0,1) in (2) above. In the case of an affix, the last coordinate provides information on its distance from the stem: *-hAt* 'may', *-l* 'objperson2' and *-Vk* 'sg1' are required to be *immediately preceded* by the stem (see fn. 2 above) but the *rank* of this requirement is higher in the second

---

standing in different grammatical relations. This optimalistic [11] technique enables us to dispense with both "Merge" and "Move" [17]. See also section 5.

---

[3] In the characterization of a stem ('labstem'), after an (English) identifying label, four phonetic properties potentially relevant to (the vowels of) the environment (!) are provided: *frontness* (cf. *fiú-nak* 'boy-dat' vs. *őr-nek* 'guard-dat'), *roundness* (*űr-höz* 'space-all' vs. *hír-hez* 'news-all'), *vertical* position of tongue, "*lowering* property" peculiar to Hungarian (*vár-ak* 'castle-pl' vs. *cár-ok* 'czar-pl'). The label of an inflection ('labinfl'), practically a suffix in Hungarian, also shows four phonetic properties: whether it causes *lengthening* (cf. *apá-t* 'father-acc' vs. *apa-ként* 'father-form', *as father*), *shortening* (*kerek-et* 'wheel-acc' vs. *kerék-ként* 'wheel-form'), *epenthesis* (*fark-at* 'tail-acc' vs. *farok-ként* 'tail-form'), and *lowering*, again (*sors-ot* 'fate-acc' vs. *sors-ok-at* 'fate-pl-acc'). Derivative suffixes (e.g. *-tAt* 'cause' above) require other labels.

case than in the third case, and less high then in the first case (hence, eg. *énekeltet-l-ek-het* is, correctly, excluded).

Example (3) shows which lexical item stands in what kind of *grammatical relation* with which lexical item. The particular example demonstrates one kind of relation, typical of morphologically rich languages: where morphemes find each other within a word. Row (3.3) says, for instance, that the suffix referring to a second person object could find a suitable stem ('sing').[4]

As for the range of grammatical relations between separate words in a universal perspective [22], there are *unidirectional* (free) relations, e.g. an adjective "seeks" its noun, where the "seeking" lexical item may show certain properties (number, gender, case, definiteness) of the "sought" one (see (20f,h) in section 4 below), and *bidirectional* (regent – obj/subj/...) relations, e.g. an object and its regent ("predicator": in whose argument structure the former is) "seek" each other, where the argument may have a case-marking depending on the regent (*Acc* in Hungarian), and the regent may show certain properties (number, person, gender, definiteness) of the argument (person/def in Hungarian). Table 2 in (21) below will illustrate a systematic account.

Nominal arguments of regents (expressed in a separate word) are typically regarded as DPs in PS grammars (e.g. '[DP [D the] ... [N girl]]'; see [34: 213, (78)], which can be expressed in GASG by seeking two "pillars" in these cases: a determiner and a noun. The syntactic representation in (9) of the English equivalent in (7) below of the Hungarian sentence in (1) above illustrates these pillars: *make* needs a subject (rows 9.4-5), which needs a regent (rows 9.1-2) *vice versa,* and this subject should contain two (not necessarily different!) decisive elements (e.g. *the boys make...,* or *I make...,* or *Peter make(s)...*). If the argument is not a noun (e.g. row 9.6), only a single "pillar" is to be sought.

(7) I may make you sing.

(8) LEXICAL ITEMS:
    I:    n(1,1,li(m("","I",""),labsteme("I",1,
[["0","sg","1","NOM"]])))
    may:
n(2,1,li(m("","m","AY"),labsteme("may",2,
[["VERB"]])))
    make:
n(3,1,li(m("ma","K","e"),labsteme("cause",2,
[["NOM","VERB"]])))

    you:
n(4,1,li(m("","you",""),labsteme("you",1,
[["0","_","2","_"]])))
    sing:
n(5,1,li(m("s","I","ng"),labsteme("sing",2,
[["NOM"],["NOM","ACC"]])))

(9) SYNTAX:
    gr("noun","regent","subj",1,1,3,1)
    gr("det","regent","_",1,1,3,1)
    gr("regent","verb","arg",2,1,3,1)
    gr("regent","noun","subj",3,1,1,1)
    gr("regent","det","subj",3,1,1,1)
    gr("regent","verb","arg",3,1,5,1)
    gr("noun","regent","subj",4,1,5,1)
    gr("det","regent","_",4,1,5,1)
    gr("regent","noun","subj",5,1,4,1)
    gr("regent","det","subj",5,1,4,1)

(10)    COPREDICATIVE NETWORK:
    copr("may",2,1,"cause",3,1,1,0,"arg")
    copr("cause",3,1,"I",1,1,1,1,"arg")
    copr("cause",3,1,"I",1,1,1,0,"arg")
    copr("cause",3,1,"sing",5,1,2,0,"arg")
    copr("sing",5,1,"you",4,1,1,1,"arg")
    copr("sing",5,1,"you",4,1,1,0,"arg")

(11)    DISCOURSE SEMANTICS:
    provref("fixpoint",[e(2,1,1)])
    provref("old",[r(1,1,1)])
    pred("I",1,[r(1,1,1)])
    provref("new",[e(2,1,1)])
    pred("may",2,[e(2,1,1),e(3,1,1)])
    provref("new",[e(3,1,1)])
    provref("<",[e(2,1,1),e(3,1,1)])
    pred("cause",3,[e(3,1,1),r(1,1,1),e(5,1,1)])
    provref("old",[r(4,1,1)])
    pred("you",4,[r(4,1,1)])
    provref("new",[e(5,1,1)])
    pred("sing",5,[e(5,1,1),r(4,1,1)])
(12)    yes

Examples (4)/(10) show a special semantic network, "*copredicative* relations", which can be regarded as a preparatory phase towards DRS (but its independent usefulness will be argued for in section 4). Row (4.1) is a quite straightforward instance of *copredication*: it says that the lexical item belonging to the *first* morpheme of the *first* word in (1), labeled 'sing', stands in an *argument*-type copredicative relation with the lexical item belonging to the *fourth* morpheme of the same word, labeled 'you-sg'. Regarding the predicates belonging to these lexical items, this formula provides the information that *who is singing* is the hearer (*you*). The last two numbers in the formula (...1,1,...) refer to the fact that the *first* argument is concerned in the case

---

[4] It is the semantics (4-5), however, that is responsible for the expression of their „real" relation (i.e. *you sing*).

of both predicates. The formulae in rows (10.5-6) express the same, i.e. *you sing,* with the only difference that the relevant lexical items are the single morphemes of the fifth and the fourth words of the English sentence (7).

In row (4.3)/(10.4) a "zero-th" argument is referred to: it concerns the Davidsonian / eventuality argument of singing [25]: not the singer or the song but the action itself. In the case of *cause*, the second argument is concerned. All in all, the content of formula (4.3)/(10.4) is that *a singing event is caused.*

We would like to emphasize that in this intermediate level of copredicative network, before the last phase when  the programme constructs a DRS (see section 3) properly expressing the content of the input sentence, the major formal differences between the (one-word) Hungarian and the (five-word) English input essentially disappear.

## 3.  Machine-Aided Transl. Based on DRS

Relative to the copredicative network, the basic task concerns producing appropriate referent names. First let us consider the "English DRS" in (11) because it shows the more straightforward way of producing entity-type referents.

In rows (11.3) and (11.10) the subject's referent denoted by r(1,1,1) and the object's one denoted by r(4,1,1) have been created; the first two numbers refer to the pair of serial numbers belonging to the two morphemes serving as their source ('I', 'you'). Labels "old" in the formulas in (11.2) and (11.9) give information on the embedding of the DRS in (11) in an interpreter's information state IIS, which is formulated as a DRS as well [3]: these two "proto"-referents (r111, r4111) do *not* trigger the introduction of new referents in IIS. What are worth new referents, are the "possible" situation  (whose Davidsonian referent is e211; see rows (11.4-5)), the "causation" (e311, (11.6,8)) and the "singing" event (e511, (11.11-12)).

All in all, the DRS in (11) claims that r111, who coincides with the speaker (11.3), causes an event (11.8), which is one or more persons' singing (11.12), among whom the hearer can be found (11.10). Row (11.5) says, however, that the whole causing event is not sure, only probable ("it may happen that..."). Row (11.7) also expresses a piece of information relevant to the embedding of the DRS in IIS; the following formula has the same content in "Lifelong" DRT [3]: *world(e211) <world(e3111).* Its approximate meaning is that the world where "you" are singing in favor of "me" is a fictive world relative to, say, the real world; if the real world is the *fixpoint* of interpretation (11.1). That is, the

interpreter accepts the claim that something is possible ("it may happen") but (s)he does not accept that the singing event takes place indeed.

It can be checked that the "Hungarian DRS" in (5) essentially expresses the same content (they are isomorphic modulo referent numbering); the slight differences can be attributed to the fact that the Hungarian personal suffices enable us to get immediately to the "built-in" referents r011 and r012 (5.6, 5.3), denoting the speaker and the hearer, respectively.[5]

This isomorphic relation between the "Hungarian" and the "English" DRS suggests the idea of the application of our semantic parser in the area of machine-*aided* translation. It would take only a few hours (or days) to teach an intelligent (say) English-speaking person (who is no linguist or logician) how to interpret DRSs containing English names of predicators (see (5)); whilst it would take years to teach her even a basic level of Hungarian, Maltese, Estonian, Finnish, Slovenian and, say, other languages of current and would-be members of EU.

We are working on a universal semantic parser which is able to produce "English DRSs" as a response to weather reports, as input, in different (and mostly morphologically rich) languages of the would-be EU. The "English DRS" can be regarded as a text in a disambiguated, straightforwardly formalized, restricted variant of English, among whose favorable properties is the opportunity for explicitly referring to even *implicit* spatial and temporal entities due to the DRT-basis [19-20].[6] It is also easy to produce a uniform database from these formalized "translations" of the relevant weather reports (via manually unifying referents, as a first approximation to the problem), whose homogeneous Prolog-clause format is an excellent platform for database-query languages.

---

[5] The Hungarian sentence in (20a) below, which is a (less preferred, but still grammatical) variant of (1), would yield a DRS similar to the "English" one in (11), obviously due to the separate expression of the subject and the object. An interesting difference between (1) and (20a) is worth highlighting at this point:  'I' and 'you' can be regarded as predicates ("x is identical with the speaker / hearer"), whose sources in the *pro-*drop variant in (1) are the complex suffix *-lAk,* whilst in (20a) they come from the separate pronouns while *-lAk* is only an agreement marker. An "independent" argument for this approach is the observation  that *-lAk* as an agreement marker in (20a) is compatible with both singular and plural 'you' whereas *-lAk* as a predicator unambiguously refers to a singular 'you'. Thus (1) and (20a) are unambiguous (in the respect in question) with two different meanings, and (5) is ambiguous.

[6] These special referents are absent from the illustrations in (5) and (11) for the sake of simplicity.

# 4. Machine Translation Based on the Copredicative Network

Let us return to *copredicative networks* (e.g. (4), (10)) in order to discuss the more ambitious question of "pure" MT, because they contain more valuable information than DRSs ((5), (11)), as they "still defend" the original structure of the sentence in question.[7] One might think that, say, Hungarian and English differ so radically (cf. (1) and (7)) that we had better forget the structure of the input sentence and consider the abstract semantic content in the form of a DRS. Our comparative analysis in section 2, however, proves the opposite: Hungarian and English (already) show practically no difference on the level of the copredicative network. Hence, it is possible to follow a perfectly "conservative" method of translation where (bilateral) regent–argument relations will correspond to regent–argument relations[8], and unilateral "free adjunctions" will correspond to "free adjunctions".

Thus our method of translation relies on the mediating role of copredicative formulae considered one by one (4)/(10). These formulae already show no properties peculiar to the source language, which is a favorable fact. Nor does it show properties of the target language, however. The first problem to be coped with at this point concerns regent–argument relations, whose non-language-specific formulation should be substituted for *one* of the case frames of the given regent available in the target language. We claim [1] that the typical arguments of a verb has a practically absolute "causal order" but the strength of its agentive / patient-like nature shows (slight) language-specific differences. The five numbers in Table 1. (18) below from -2 to +2 mark five classes of argument roles from *strictly agentive* to *strictly patient-like* ones. The calculation of class-membership is a complex task, whose details are not entered into here (but see [1]); we let you know, however, a crucial property of members in class 0: these arguments can appear in *transitive* argument structure versions (out of ASV families demonstrated (partially) in (13-17) below) both as SUBJECT (nominatively marked in an accusative language) and as *OBJECT* (marked with Acc.).

(13)a.  PETER served *MARY DINNER*.[9]
b.  PETER served *DINNER* (to Mary).
c.  PETER served *MARY*.
d.  THE GUEST was served (with) *DINNER*.
e.  RATIONS were served *(to) the troops.

(14)a.  SMITH sold *THE FIRST CUSTOMER THE CAR*.
b.  SMITH sold *THE CAR* (to the first customer).
c.  *SMITH sold *THE FIRST CUSTOMER*. (!)
d.  THE FIRST CUSTOMER was sold *THE CAR*.
e.  THE CAR was sold *(to) the first customer.

(15)a.  PETER loaded *HAY* onto the wagon.
b.  PETER loaded *THE WAGON* with hay.

(16)a.  THE HORSES are walking.
b.  THE HORSES walked *FIVE MILES*.
c.  THEY walked *THE HORSES* for a while after the race.

(17)a.  PETER is bored with linguistics.
b.  LINGUISTICS bores *PETER*.
c.  MARY bores *PETER* with linguistics.

(18) **Table 1.** Agentive polarity pattern of argument structures in the case of a few typical English verb types

| POLARITY OF ARG'S ASV FAMILY | Agentive | | Patient-like | | |
|---|---|---|---|---|---|
| | -2 | -1 | 0 | +1 | +2 |
| X served Y Z | | X | Y | Z | |
| a. | | Nom | Acc | Acc | |
| b. | | Nom | Acc | | |
| c. | | | | | |
| d. | | | Nom | Acc | |
| e. | | | | Nom | |
| X sold Y Z | X | Y | | Z | |
| a. | N | A | | Acc | |
| b. | N | A | | Acc | |
| c. * | N | A̶ | | | |
| d. | | N | | Acc | |
| e. | | | | Nom | |
| X loaded Y onto Z | | X | | Y | Z |
| a. | | Nom | | Acc | |
| b. | | Nom | | | Acc |
| X, Y, Z | X | | Y | | Z |
| a. Y walked | | | Nom | | |
| b. Y walked Z | | | Nom | | Acc |
| c. X walked Y | Nom | | Acc | | |
| X bores Z with Y | X | Y | | Z | |
| a. | | | | Nom | |
| b. | | Nom | | Acc | |
| c. | Nom | | | Acc | |

After constructing, on the basis of the given polarity pattern, the *potential ASV family* of a given verb in the target language (and checking it in some corpus), we should select the appropriate AS version. What should know at this point, is the function of this selection in the particular target language. In English, but not in

---

[7] A simple DRS can be produced as the *symmetric–transitive closure* [26] of a network of copredications.
[8] The crucial point here is that it is irrelevant whether the source of this semantic relation is a formal relation between morphemes of the same word or between (morphemes of) separate words.
[9] The two bitransitive types exemplified in (13-14) are demonstrated in [23].

Hungarian [1], the selected subject will serve as the *topic* of the sentence. In Finnish argument selection depends on such factors as definiteness of arguments, aspect and mood (19a-c), and in Georgian tense and aspect are decisive factors (19d):

(19)a. Syö-t omena-n.  'You eat the apple.'
       eat-2SG apple-ACC
b. Syö-t omena-a        'You are eating some apple.'
   eat-2SG apple-PART
c. Syö omena!           'Eat the apple!'
   eat-2SG apple
d. Georgian *kill*:      ⟨Agent, Patient⟩ =
      present:           ⟨NOM, Dat⟩
      past:              ⟨ERG, NOM⟩
      present perfect:   ⟨Dat, NOM⟩

Suppose the MT programme could calculate the appropriate verbs (and other kinds of regents) with the appropriate case frames in the target language (and it could also check their existence in a corpus). Suppose, furthermore, it could also collect the appropriate adjectives and other kinds of adjuncts. Now the task is to complete this set of lexical items (out of the "generative lexicon" of the target language) with further lexical items whose characteristic property lies in the fact that they lack any background semantic content. As has been mentioned in section 2 with reference to [22], these "mysterious" elements are practically the language-specific *case* and *agreement* affixes. As is illustrated in (20) below and then summarized in Table 2. (21), their exact number and position (on the regent, the argument or the adjunctive element) can be determined in the function of the type of the copredicative relation and the target langue:

(20)a. Én énekel-tet-het-l-ek titek-et. (Hung.)
       I sing-CAUSE-MAY-2SG-1SG youₚₗ-ACC
       'I may make you sing.'
b. Mary sing-s.
c. Kedvel-em a ti barátságos kutyá-i-tok-at.(Hung.)
   like-1sg the youₚₗ friendly dog-PL-2PL-ACC
   'I like yourₚₗ friendly dogs.
d. Találkoz-ol Mari-val.   (Hung.)
   meet-2SG Mary- INST
   'Youₛ𝓰 meet Mary.'
e. A kutyá-k barátságos-ak.   (Hung.)
   the dog-PL friendly- PL
   '(The) dogs are friendly.'
f. Puhu-n viisa-i-sta suomalais-i-sta mieh-i-stä.
   speak-1SG clever-PL-ELAT Finnish- PL-ELAT man- PL-ELAT
   'I speak about (the) clever Finnish men.' (Finn.)
g. Mene-n syö-mä-än. (Finnish)
   go-1SG eat-INF-ILL
   'I am going to eat.'

h. Я чита-ю красив-ую книг-у.        (Russian)
   I read-1SG beautiful-SG+FEM+ACC book-ACC
   'I am reading a/the beautiful book.'
i. Я  чита-л-а  / чита -л.   (Russian)
   I read-past-fem / read-past(masc)
   'I (Judit / Gábor) was reading.'

(21) **Table 2.** Affixes appearing in (different-word) regent–argument and adjunction constructions across languages

| FEATURE CONSTRUCTION | PER-SON | DEF | NUM-BER | GEN-DER | CASE |
|---|---|---|---|---|---|
| H../Eng. Nom.: a,b | pred | - | pred | - | - |
| Hung. Acc.: a,c | pred | | - | - | arg |
| Hung. Possessive: c | pred | - | pred | - | -/arg |
| Hung., other case: d | - | - | - | - | arg |
| Hung. attr. adj.: c | - | - | - | - | - |
| Hung. pred. adj.: e | - | - | pred | - | - |
| Finnish attr. adj.: f | - | - | adj | - | adj |
| Finnish inf. arg.: g | - | - | - | - | arg |
| Russian Nom. pr.: h | pred | - | pred | - | - |
| Russian Nom. past:i | - | - | pred | pred | - |
| Russian attr. adj.: h | - | - | adj | adj | adj |

Table 3. (22) below illustrates a potential complication / „cost": the same lexical item, e.g. the agreement marker of the second person, is (to be) associated with several phonetic forms (-*Vl/Vsz/Vd/Vt/∅*), including the null form "in the neighborhood" of the suffix of the imperative mood (*üs-s* 'hit-imp': *you should hit me / us / indefinite third person(s)*[10]). The solution relies on a basic property of our totally lexicalist grammar: in the course of the "recognition" of each lexical item in a word, potential environments should be investigated thoroughly. In the case of *üss,* for instance, an empty sequence of letters can be accepted as a realization of the second person; and in the case of -*Vl,* too, it depends on the environment whether this suffix form refers to the person feature of the subject or the object.

The importance of the prediction of the precise number and position of the "meaningless" lexical items in the potential translation of a sentence depends on our method of producing the target sentence: we use our Prolog parser demonstrated in section 2 "inversely". Let us consider the simple Goal function below in (23a). Our programme's final answer to this question ("is it a grammatical English sentence?") will be *yes* (cf. (6)/(12)).

---

[10] '1/-3' in (22) refers to these choices whilst '+3' refers to a definite third person or more definite third persons.

(22) **Table 3.** Illustration of different realizations of a few Hungarian verbal suffixes

| stem | mood/ tense | obj person | subj person | number |
|---|---|---|---|---|
| *üt* hit |  | *-l* **2** | *-ek* 1 | - sg |
| *vesz* buy |  |  | *-ek* 1 | - sg |
| *vesz* buy |  |  | *-el* **2** | - sg |
| *üt* hit |  | - 1/-3 | *-sz* **2** | - sg |
| *üt* hit |  | - +3 | *-öd* **2** | - sg |
| *üt* hit |  | *-i* 1/-3 | *-t* **2** | *-ek* pl |
| *üs* hit | *-s* imp | - 1/-3 | *- / -él* **2** | - sg |
| *üs* hit | - imp | - +3 | *-d* **2** | - sg |
|  | *-s* imp | - +3 | *-ed* **2** | - sg |
| *üt* hit | *-n* cond | *-é* +3 | *-n* 1 | *-k* pl |
| *üt* hit | *-ött* past | *-é* +3 | - 3 | *-k* pl |

(23) a. gramme("Mary", "sing", "-s").
b.    gramme("Mary", X, "-s"), fail.
c.    gramme(X, "sing", Y), fail.

If the Goal contains variables, however, as in (23b-c) above, the parser will "generate" grammatical sentences by replacing the variables with appropriate lexical items: e.g. *Mary sings, Mary drinks* (23b), *Mary sings, Peter sings* (23c).[11]

Now suppose our "inverted parser" has managed to construct the proper *list* of words in the target language. The ultimate factor to be checked is the appropriate word order. This task can also be entrusted to our parser: it should check different *permutations* of the above mentioned list until it finds (the) one compatible with all the requirements associated with the given copredicative network in the target language. Let us return to (4) in order to illustrate this point. Rows of (24) below show the requirements (and their strength) concerning *immediate precedence* relations from row to row of (4). Finally, (25) shows the single solution satisfying all the requirements in (24), which is the English sentence in (7), the perfect translation of the Hungarian sentence in (1).

(24)   you << sing
   I  <  make
   make < sing
   may << make

(25)   I may make you sing.

## 5. Concluding Remarks

The theoretical and practical status of the would-be result of our MT project will be discussed in this concluding section.[12]

The relevant factors of our theoretical perspective can be elucidated in the framework of Nerbonne's introductory paper on computational semantics [24]. "There is a natural division", he says, "of theoretical labor between the disciplines of linguistics and computational linguistics, namely that linguistics is responsible for the description of language, and computational linguistics for the algorithms and architectures needed to compute with these. ... Computational linguistics is dependent on [general] linguistics for the characterization of the relations it computes."

As has been declared in section 1 and footnote 2, our research team's principal endeavor is of general linguistic nature: to "legitimize" a new sort of generative grammar (GASG), a "totally lexicalist" one, via verifying its *computational implementability*, among others, because a successful implementation is the best evidence for the exactness and consistency of a formal system. Practically we should create effective parsers carrying out the basic tasks of a generative grammar: to decide whether the given input consists of well-formed words, whether these words constitute a grammatical sentence, and whether this sentence can be assigned a well-formed semantic representation; and then we should go on to base Machine (-Aided) Translation, Information Extraction [8] and other tasks upon our parsers.

As dispensability with PS-tree building is the decisive trait of GASG, this grammar can be qualified as an (extreme) feature-based formalism, where the only engine of combining lexical signs is the mere technique of *unification* (of constraints that lexical signs exert on each other in potential sentences). Nerbonne [24] speaks about the advantage of the added freedom which constraint-based semantics allows as compared to strictly compositional treatments of syntax-semantics interaction as follows: "In feature-based formalisms, the structure shared among syntactic

---

[11] *Mary sang* is also among the solutions to (23c). The stem *sing* and the affix of past can be combined to yield *sang* due to the tripartite morpheme structure illustrated in (8) in section 2 (m("s", "I", "ng")).

[12] The existence of this section and partly its content are due to the anonymous reviewers' questions and comments concerning the first version of this paper. We are grateful to them.

and semantic values constitutes the syntax/semantics interface. Our earlier papers have explored several advantages of the constraint-based view of the syntax/semantics relation over standard views..."

Now his argumentation concerning processing is relevant to us: "It is natural to try to take advantage of such information ["selectional restrictions", e.g. between subject and predicate in the case of a sentence like *The chair decided on Mary*] early in processing — maximizing the efficiency benefits, and plausibly modelling human language users more faithfully in this respect as well. Taken to its extreme, this means that semantic processing must allow information flow along non-compositional lines. This does NOT reject the grammatical thesis that syntax-semantics dependence is ultimately compositional, only that processing is organized along the same lines." "The compositional view often extends to a natural and popular interpretation, that of bottom-up processing. This is a natural interpretation because compositional semantics specifies the semantics of phrases via functions on the semantics of their daughters. ... The fact that purely bottom-up processing is less than optimal is completely consistent with there being a compositional syntax-semantics interface."

Nerbonne argues for another sort of processing: "INCREMENTAL PROCESSING computes analyses while inputting strings one word at a time, in the order they are heard or read (the left-right order in text). ... In pure bottom-up processing, no parent node is processed until all of its daughters are. ... Since it is effectively the sentence node at which subject-verb agreement is enforced, including agreement of semantic selectional restrictions, there is no way of rejecting the senseless reading [*chair* ~ 'piece of furniture'] until the entire sentence is processed. Given the strong preference for right-branching structures in grammar, purely bottom-up processing can have no account of incremental understanding. ... In incremental schemes, it [the incompatibility of the furniture reading with the mental agency] must be enforced at the point at which the word *decide* is encountered. ... Purely bottom-up processing cannot be incremental [in this sense] (at least for standard grammars). ...Since the meaning of a phrase depends on its daughters' meanings (compositionality), there cannot be a (final) computation of phrasal meaning before all the daughters' meanings have been processed. One could imagine an argument for bottom-up evaluation proceeding from linguistic compositionality in this way. But several escape routes open before this conclusion is reached."

As is classified by Nerbonne [24], one "escape route" is Steedman's [33] solution, the Flexible Categorial Grammar, which "lies in the wholesale abandonment of standard assumptions about the constituent structure ... which Steedman takes to be rigorously left-branching in order to reconcile bottom-up and incremental processing. Since the syntax is completely left-branching, EVERY initial segment is a constituent." Well, "this solution is formally sound, but its linguistic assumptions are heterodox. ... It indicates that constituent structure — the primary explanatory device in syntax — is relatively insignificant."

Another "escape route" mentioned by Nerbonne [24] is that of Shieber and Johnson (1993) [32]. They claim that incremental interpretation follows not from grammatical structure (the flexible CG position), nor even from the control structure of particular algorithms, but rather from the asynchronous nature of understanding. According to this position, "it seems best not to explain incremental understanding be possible if grammatical structure is roughly as we know it (in fact, mostly right-branching). ... A popular choice for incremental parsing is (predictive) LEFT-CORNER (...) parsing. It is popular because it makes use of both bottom-up (lexical) and also top-down (grammatical) information."

After this long preparation, we claim that what our "totally lexicalist", non-PS-building (but) generative grammar offers is just the reconciliation of the advantages of the two "escape routes" from the above discussed conflict between "theoretical compositionality" and effective processing. As "constituents" provided by GASG are pairs of lexical signs[13] exerting constraints on each other, we should have recourse to neither "heterodox" (left-branching) syntactic constituents nor a processing strategy not following theoretical compositionality. Compositionality according to GASG means just the summation and unification of the proto-DRSs which are the contributions of the lexical signs formally checking and semantically exerting selectional restriction on each other.

Let us turn to the status of our projects [9] among those concerning the processing of Hungarian, famous for its rich morphology and free word order. The Proceedings of the (first) Conference on Hungarian Computational Linguistics [10], held in 2003, provides an exhaustive review on the projects. What seems to verify our mainly knowledge-based approach based on the above discussed combination

---

[13] These "constituents" differ from ones in the standard sense in that they do not make up a hierarchical structure with a top node, but they form an undered "network" of lexical requirements to be satisfied through unification.

of basic generative ideas with extreme lexicalism is that there is no working *semantic* parser "on the Hungarian market"[14], whilst, for instance, there are excellent morphological parsers.

One of the anonymous reviewers has raised the practical question why we are not using an existing Hungarian morphology for our project (such as the well-known one from Morphologic [29-30]). The answer directly comes from our often-mentioned theoretical-linguistic basis: as has been mentioned in footnote 2, it is a "totally lexicalist morphology" that suits our philosophy. In this approach, instead of inflected words, *morphemes* are assigned lexical signs (units of phonological, morphological, syntactic and (discourse-) semantic information). Thus our system dispenses with an autonomous morphological subsystem with whole words as input and words segmented into morphemes as output. That is, practically morphemes "seek" each other in the course of our parsing procedure; which does not mean, nevertheless, that the traditional grammatical level of *word* has disappeared from our grammar: a morpheme belonging to a lexical sign "seeks" certain lexical signs (i.e. morphemes) within the same word, and other ones definitely in different words in the sentence-internal "environment".

As for differences in the approach to Hungarian morphology between a version of the parsers of MorphoLogic described in [29] and our parser, it is worth comparing our analysis of, say, *ütnénk* 'we would hit that', demonstrated in Table 3 above in (22), to that mentioned on page 1041 in [29] (encoded as TFt1). In their practical approach, *-nénk* is a monolithic element with a complex meaning, whilst in our general background theory it is segmented into four parts following the separable elements of meaning (see Table 3). This theoretical viewpoint serves as a starting-point to the elaboration of an effective parser. What else is also to be considered is, for instance, the number of "zero" morphemes resulting from the rigorous theory; in this case we have decided on a segmentation into two parts (*-né-nk* 'would+3:OBJ-1PL:SUBJ). Another difference is that they store allomorphs [29] whilst we store underspecified morphemes (see (2) and (8) above).

Finally, after these concluding remarks on our project's status, we would like to call attention to the illustration in the Appendix of the English–Hungarian direction of translation.

---

[14] Our parsers are also only experimental versions operating on small fragments of English and Hungarian. Their exploitation in an MT project even in this early phase, nevertheless, is necessary for the permanent evaluation of practical results of our decisive "theoretical ambitions".

# References

[1] Alberti, G. 1997: *Argument Selection.* Peter Lang, Frankfurt am Main.

[2] Alberti, G. 1999: GASG: "The Grammar of Total Lexicalism". *Working Papers in the Theory of Grammar* 6/1. Theoretical Linguistics Programme, Budapest University and Research Institute for Linguistics, Hungarian Academy of Sciences.

[3] Alberti, G. 2000: Lifelong Discourse Representation Structures. *Gothenburg Papers in Computational Ling.* 00-5, Sweden, 13-20.

[4] Alberti, G. 2001: Indo-Germanic Word Order Phenomena in a Totally Lexicalist Grammar. *Sprachtheorie und germanistische Linguistik* 11.2., Debrecen, Hungary, 135-193.

[5] Alberti, G., K. Balogh, J. Kleiber 2002: GeLexi Project: Prolog Implementation of a Totally Lexicalist Grammar. *Proceedings of the Third and Fourth Tbilisi Symposium on Language, Logic and Computation,* ed. by de Jongh, Zeevat and Nilsenova. ILLC, Amsterdam, and Univ. of Tbilisi.

[6] Alberti, G., K. Balogh, J. Kleiber, A. Viszket 2002: Totally Lexicalist Morphology. Talk at the *Sixth Int. Conf. on the Structure of Hungarian.* Düsseldorf, Germany. To appear in *Approaches to Hungarian* 9, ed. by I. Kenesei and Ch. Piñón.

[7] Alberti, G., K. Balogh, J. Kleiber, A. Viszket 2003: Total Lexicalism and GASGrammars: *A Direct Way to Semantics.* Proceedings of CICLing2003, (Mexico City), ed. by A. Gelbukh. Lecture Notes in Computer Science N2588. Springer, 37-48.

[8] Alberti, G., J. Kleiber 2003: Extraction of Discourse-Semantic Information... In Cunningham, H., E. Paskaleva, K. Bontcheva, G. Angelova (eds.): *Information Extraction for Slavonic and Other Central and Eastern European Languages.* Borovets, Bulgaria. 63-69.

[9] Alberti, G., J. Kleiber and A. Viszket 2003: GeLexi projekt: GEneratív LEXIkonon alapuló mondatelemzés [GeLexi Project: Sentence Parsing Based on a GEnerative LEXIcon]. In [10], 79-85.

[10] Alexin, Z., D. Csendes and T. Gyimóthy 2003: *Magyar Számítógépes Nyelvészeti Konf.* [(First) Conference on Hungarian Computational Linguistics]. Univ. Scientiarum Szegediensis.

[11] Archangeli, D., T. D. Langendoen (eds.) 1997: *Optimality Theory: an overview*, Blackwell, Oxford.

[12] Balogh, K., J. Kleiber 2003: Computational Benefits of a Totally Lexicalist Grammar. In Matoušek, V., Mautner, P. (eds.): *Text, Speech and Dialogue*, Proceedings of TSD2003. Springer-Verlag, Berlin Heidelberg New York. 114-119.

[13] Borsley, R. D. 1996: *Modern Phrase Structure Grammar.* Blackwell, Oxford Cambridge.

[14] Bresnan, J. 2000: *Lexical Functional Syntax,* Blackwell, Oxford.

[15] Brill, E. 2003: Processing Natural Language without Natural Language Processing. In Gelbukh, A. (ed.): Proc. of CICLing2003 (Mexico City). LNCS N2588. Springer-Verlag. 360-369.

[16] Chomsky, N. 1957: *Syntactic Structures.* The Hague, Mouton.

[17] Chomsky, N. 1995: *The Minimalist Program.* MIT Press, Cambridge, Mass.

[18] Joshi, A. K. 2003: Starting with Complex Primitives Pays Off. In Gelbukh, A. (ed.): Proceedings of CICLing2003 (Mexico City). LNCS N2588. Springer-Verlag. 1-10.

[19] Kamp, H., U. Reyle 1993: *From Discourse to Logic.* Kluwer, Academic Publ.

[20] Kamp, H., J. van Genabith, U. Reyle 2004: Discourse Representation Structure. Draft to appear in *Handbook of Philosophical Logic.*

[21] Karttunen, L. 1986: Radical Lexicalism. Report No. CSLI 86-68. Stanford.

[22] Lehmann, Ch. 1988: On the Function of Agreement. In Barlow and Ferguson (eds.): *Agreement in Natural Languages. Approaches, Theories, Descriptions*. CSLI Stanford, 55–65.

[23] Levin, B., M. Rappaport 1986: The Formation of Adjectival Passives. *Linguistic Inquiry* 17, 623-662.

[24] Nerbonne, J. 1996: Computational Semantics – Linguistics and Processing. In Sh. Lappin ed.: *The Handbook of Contemporary Semantic Theory*. Blackwell.

[25] Parsons, T. 1995: Thematic Relations and Arguments. *Linguistic Inquiry* 26, 635–662.

[26] Partee, B. H., G. B. ter Meulen, R. P. Wall 1990: *Mathematical Methods in Linguistics*. Kluwer Academic Publ.

[27] Pollard, C., I. Sag 1987: Information-Based Syntax and Semantics, vol. I. CSLI Lecture Notes 13. Stanford.

[28] Pollard, C., I. Sag 1994: *Head-Driven Phrase Structure Grammar*. Stanford.

[29] Prószéky, G. 2000: A magyar morfológia számítógépes kezelése [Computational Processing of Hungarian Morphology]. In F. Kiefer ed.: *Strukturális magyar nyelvtan 3. Morfológia* [A structuralist grammar of Hungarian 3. Morphology]. Akadémiai, Budapest, 1021-1063.

[30] Prószéky, G. 2003: Megértéstámogatás és gépi fordítás: nyelvtechnológia a XXI. század elején [Text Understanding and Machine Translation: Language Technology in the Early Years of the 21st Century]. *Eighth National Neumann Congress*, Hungary.

[31] Schabes, Y., A. Abeillé, A. Joshi 1988: Parsing strategies with lexicalized grammars: Tree Adjoining grammars. *Proceedings of 12th COLING,* Budapest.

[32] Shieber, S. M. and M. Johnson 1993: Variations on incremental interpretation. *Journal of Psycholinguistic Research* 22, 287-318.

[33] Steedman, M. 1987: Combinatory grammars and human language processing. In J. Garfield ed.: *Modularity in Knowledge Representation and Natural-Language Understanding.* Cambridge MA: MIT Press. 187-205.

[34] Szabolcsi, A. 1994: The Noun Phrase. In *The Syntactic Structure of Hungarian, Syntax and Semantics* 27, ed. by K. É. Kiss and F. Kiefer. New York, Academic Press. 179-274.

[35] van Eijck, J., H. Kamp 1997: Representing discourse in context. In van Benthem, J., A. ter Meulen (eds.): *Handbook of Logic and Language.* Elsevier, Amsterdam & MIT Press, Cambridge, Mass.

[36] Zeevat, H. 1991: Aspects of Discourse Semantics and Unification Grammar. Ph.D., U. Amsterdam.

## Appendix

**Peter loves a clever Hungarian girl.**

LEXICAL ITEMS:

Peter: n(1,1,li(m("","Peter",""), labsteme("Peter",1,[["0"]])))

love: n(2,1,li(m("","love",""), labsteme("love",2,[["NOM","ACC"]])))
s: n(2,2,li(m("E","s",""), labsuffe("E/3",2,"infl")))

a: n(3,1,li(m("a","N",""), labsteme("a(n)",3,[])))

clever: n(4,1,li(m("","clever",""), labsteme("clever",4,[])))

Hungarian: n(5,1,li(m("","Hungarian","") , labsteme("Hungarian",4,[])))

girl: n(6,1,li(m("","girl",""), labsteme("girl",1,[])))

SYNTAX:
gr("noun","regent","subj",1,1,2,1)
gr("det","regent","_",1,1,2,1)
gr("regent","noun","subj",2,1,1,1)
gr("regent","det","subj",2,1,1,1)
gr("regent","noun","obj",2,1,6,1)
gr("regent","det","obj",2,1,3,1)
gr("det","noun","free",3,1,6,1)
gr("det","regent","_",3,1,2,1)
gr("adj","noun","free",4,1,6,1)
gr("adj","noun","free",5,1,6,1)
gr("noun","regent","obj",6,1,2,1)

regent-noun-subj: loves-Peter
regent-det-subj: loves-Peter
regent-noun-obj: loves-girl
regent-det-obj: loves-a
det-noun: a-girl
adj-noun: clever-girl
adj-noun: Hungarian-girl

COPREDICATIONS:
copr("love",2,1,"Peter",1,1,1,1,"arg")
copr("love",2,1,"Peter",1,1,1,0,"arg")
copr("love",2,1,"girl",6,1,2,1,"agr")
copr("love",2,1,"a(n)",3,1,2,0,"arg")
copr("a(n)",3,1,"girl",6,1,0,1,"free")
copr("clever",4,1,"girl",6,1,1,1,"free")
copr("Hungarian",5,1,"girl",6,1,1,1,"free")

SEMANTICS:
provref("fixpoint",[e(2,1,1)])
provref("old",[r(1,1,1)])
pred("Peter",1,[r(1,1,1)])
provref("new",[e(2,1,1)])
pred("love",2,[e(2,1,1),r(1,1,1),r(3,1,1)])
provref("new",[r(3,1,1)])
provref("<or=",[r(3,1,1),e(2,1,1)])
pred("clever",4,[r(3,1,1)])
pred("Hungarian",5,[r(3,1,1)])
pred("girl",6,[r(3,1,1)])

In Hungarian:
**Péter szeret egy okos magyar lányt.**

LEXICAL ITEMS (Hungarian):

li(m("","Péter",""),labstem("Peter", phonfst(1,2,0,2),1,[]))
li(m("","","",""),labsuff("NOM", phonfsu(3,3,3,3),1,4))

li(m("","szeret",""),labstem("love", phonfst(1,2,2,2),2,[["NOM","ACC"]]))
li(m("","","",""),labsuff("objpersonnot2", phonfsu(3,3,3,3),2,2.5))
li(m("","","",""),labsuff("sg3obj-def", phonfsu(3,3,3,3),2,3))

li(m("","egy",""),labstem("a(n)", phonfst(2,3,3,3),3,[]))

li(m("","okos",""),labstem("clever", phonfst(2,1,2,1),4,[]))

li(m("","magyar",""), labstem ("Hungarian",phonfst(2,2,1,1),4,[]))

li(m("","lány",""),labstem("girl", phonfst(2,2,3,2),1,[]))
li(m("V","t",""),labsuff("ACC", phonfsu(1,1,1,3),1,4))

SYNTAX:
gr("noun","regent","subj",1,1,2,1)
gr("det","regent","_",1,1,2,1)
gr("regent","noun","subj",2,1,1,1)
gr("regent","det","subj",2,1,1,1)
gr("regent","noun","obj",2,1,6,2)
gr("regent","det","obj",2,1,3,1)

gr("det","noun","free",3,1,6,1)
gr("det","regent","_",3,1,2,1)
gr("adj","noun","free",4,1,6,1)
gr("adj","noun","free",5,1,6,1)
gr("suff","stem","free",6,2,6,1)
gr("noun","regent","obj",6,2,2,1)

regent-noun-subj: szeret-Péter
regent-det-subj: szeret-Péter
regent-noun-obj: szeret-lányt
regent-det-obj: szeret-egy
det-noun: egy-lányt
adj-noun: okos-lányt
adj-noun: magyar-lányt

COPREDICATIONS:
copr("love",2,1,"Peter",1,1,1,1,"arg")
copr("love",2,1,"Peter",1,1,1,0,"arg")
copr("love",2,1,"girl",6,1,2,1,"arg")
copr("love",2,1,"a(n)",3,1,2,0,"arg")
copr("a(n)",3,1,"girl",6,1,0,1,"free")
copr("clever",4,1,"girl",6,1,1,1,"free")
copr("Hungarian",5,1,"girl",6,1,1,1,"free")

SEMANTICS:
provref("fixpoint",[e(2,1,1)])
provref("old",[r(1,1,1)])
pred("Peter",1,[r(1,1,1)])
provref("new",[e(2,1,1)])
pred("love",2,[e(2,1,1),r(1,1,1),r(3,1,1)])
provref("new",[r(3,1,1)])
provref("<or=",[r(3,1,1),e(2,1,1)])
pred("clever",4,[r(3,1,1)])
pred("Hungarian",5,[r(3,1,1)])
pred("girl",6,[r(3,1,1)])

yes

**Figure 1.** Translation from English to Hungarian