Black-Box/Glass-Box Evaluation in Shiraz

Rémi Zajac, Steve Helmreich and Karine Megerdoomian
Computing Research Laboratory, New Mexico State University
{zajac,shelmrei,karine}@crl.nmsu.edu

The Shiraz project included an evaluation component: two 'glass-box' evaluations have been performed during the project as well as a black-box evaluation at the end of the project. The evaluations were based on the use of a bilingual tagged test corpus of 3000 sentences. Evaluation tools were developed in order to automate the evaluation process. The glass-box evaluations included the evaluation of components of the MT system, and in particular the Persian morphological analyzer, the dictionary and the parser. The evaluation of the translations themselves (black-box evaluations) were performed manually on a subset of the test corpus. This paper outlines the problems encountered in trying to use these evaluations for development and testing purposes as well as traditional 'off-line' evaluations.

1 Introduction

The Shiraz machine translation system is a transfer-based MT prototype that translates Persian text into English. The project began in October 1997 and the final version was delivered on August 1999 (http://crl.nmsu.edu/shiraz/). The system uses typed feature structures and an underlying unification-based formalism to describe Persian linguistic phenomena. It is able to run on Unix as well as Windows machines. The Shiraz system uses an electronic bilingual Persian to English dictionary consisting of approximately 50,000 terms, a complete morphological analyzer and a syntactic parser. The system components were tested on a bilingual tagged corpus developed from a large Persian corpus of on-line material (approximately 10MB). The machine translation system is mainly targeted at translating news material.

The evaluations were identified as important tasks in the project. In order to minimize duplication of work, one constraint set in the project's statement-of-work was the dual use of these evaluations for external purposes (to give the sponsors an image of the system at a given point in time) as well as for internal testing. In this paper, we present the lessons of these evaluations with regard to the use of evaluation for testing purposes. Two glass-box evaluations were meant as a tool for developers (to help them identify recurring problems) and also as a tool for the sponsor, to measure the internal quality of the software and to measure progress. These two goals proved somewhat conflicting and although the evaluation tools attempted to provide answers for both kinds of users, the results were not really satisfactory to either.

This experience led us to develop more specific tools that answer more specific needs, and in particular testing tools. In particular, we consider issues such as the construction of the test corpus, the choice of evaluation criteria and the automation of the evaluation process. The parallel test corpus was initially prepared by collecting on-line news articles from Persian news web sites and extracting a set of sentences varied along syntactic and domain dimensions. The translation of these test items was done manually by a Persian native speaker. We present the initial guidelines for corpus preparation, the problems encountered in building and using this test corpus, and in retrospect, what could have been a better procedure. Black-box evaluations were performed (manually) during the course of the project to assess the global quality of the system and identify problems that were not taken into account during the glass-box evaluations. Given the constraints on resources, this could be done only on a subset of the test corpus. We outline an automated evaluation procedure that tests some of the properties of a appropriate translation.

2 Construction of the test suite

The sentences of the test suite are representative of contemporary journalistic prose; they were all collected from the on-line Iranian newspaper Hamshahri (http://www.hamshahri.org/). All sentences were manually translated at CRL. The Persian sentences were manually tagged for part of speech and bracketed to indicate Noun Phrases and Prepositional Phrases. The English sentences have been tagged automatically for POS and phrases using UPenn's SuperTagger (Doran et als. 94). The sentences below represent an example of a tagged Persian sentence (in the

Shiraz transliteration) and its corresponding English translation. The example in (1) represents a sample of a tagged sentence from the corpus. Noun Phrases are tagged as *np*, Preposition Phrases as *pp*. Relative clauses that have been separated from their heads by an intervening verb are linked to the head NP by an index (e.g., *np1* in this example). A gloss is provided in (2).

(1) [flsTyny|n<n>]np rvz sh ^snbh<day> nyz<av>[zd v xvrd<n> xvnyny<ad>[b|<pre> nyrvh|y<n> |sr|iyl<pn>]pp]np1 d|^stnd<v>[kh<rel> [Ty<pre> anh|<per>]pp, [[yk<num> flsTyny<n> [bh n|m<pre> n|dr s@yd<pn>]pp]np 24<dig> s|lh<ad>]np k^sth<v1> v<con>[yk<num> nfr<n> dygr<ad>]np zxmy<ad> ^sd<v>.]np1

(2) [Palestinians<n>]np day Tuesday<day> also<av>[clash<n> bloody<ad>[with<pre> forces<n> Israel<pn>]pp]np1 had<v>[that<rel> [during<pre>them<per>]pp, [[one<num>Palestinian<n> [to name<pre> Nader Said<pn>]pp]np 24<dig> year old<ad>]np dead<v1>and<con>[one<num> person<n>other<ad>]np injured<ad> became<v>.]np1

The corpus was collected and translated at the very beginning of the project and no strict guidelines were provided to the translators. The only guidelines provided for selecting the sentences were that syntactic constructions were to be as varied as possible, and the sentences were to be extracted from articles in various domains (politics, economy, religion, education, entertainment, etc.). In the course of using the translations as a benchmark for evaluating the automatic translations (black-box evaluations), we realized that manual translations were, as should have been expected, not 'literal translations' but sometimes paraphrases that would add new material or remove some material from the source sentence. When possible, these translations were corrected to follow stricter guidelines, for example:

- Produce a translation that is as close as possible to the source text
- All words should be translated (that is, all content words) and no extra word should be added in the translation
- Change the word order only if the English is not understandable (meaning that the order of complements and adjuncts, for example, should stay the same unless there is a compelling reason to change it).

These guidelines produced translations that were still good translations of the original and were much closer to the translations produced by the MT system, thereby facilitating the comparison for evaluation purposes. For example, in the translation below, the phrase *the unemployment figures* does not correspond with the grammatical structure of the Persian phrase it is translating. "byk|r" *[bikAr]* is a noun referring to a jobless person. However, in the translation provided, the word *unemployment* is used. In the revision, the appropriate equivalent of the word is used, keeping the grammatical relation of the word with the other components in that phrase. Also, in the Persian text, the word *tsly bx^s [tasalli bakhsh]* is used, which exists in the Shiraz dictionary as an adjective meaning *comforting*. In translation, the word is paraphrased: *has had a calming effect on.* In order to keep the translation closer to the Persian text, and to use the valid equivalent already existing in the dictionary, in revision, the word *comforting* is used which conveys the exact meaning:

> **Persian**: b| vJvdykh<con> [|yn<det> t@d|d<n> byk|r<n>]np [dr<pre> |yn<det> m|h<n> [|z<pre> s|l<n>]pp]pp bys|bqh<ad> |st<n>, [k|h^s<n> an<per>]np t| Hdy<av> tsly bx^s<ad> dvlt<n> |st<v>.
> **Translation:** Despite the fact that *the unemployment figures* of this month are unprecedented, the reduction in numbers *has had a calming effect on* the government.
> **Revised:** Despite the fact that this number of the unemployed of this month is unprecedented, the reduction in numbers to some extent is comforting to the government.

In the following translation, *kmyth n_Z|rt br tvlyd |@Z|y |vpk [komite-ye nezArat bar tolide a'zAye opek]*, referring to an OPEC committee, is translated word for word. After conducting an on-line research in the OPEC's official web page (http://www.opec.org/faqs.htm#a10e) the actual name of the committee, the Ministerial Monitoring Sub-Committee was found and used in the revision. Also, in the same translation, the name of the Persian month, *Ordibehesht [ordibehesht]*, was translated into *May* followed by the Persian date of the day. That is, 16th and 17th of Ordibehesht was translated as *May 16 and 17*. Since these months do not actually correspond, 16th and 17th of Ordibehesht was used in the revision.

Persian: qr|r |st<lv> [kmyth n_Z|rt br tvlyd |@Z|y |vpk<n>]np[dr<pre> rvz <n>16<dig> v<con> 17<dig> |rdybh^st<mon>]pp [dr<pre> |Sfh|n<pn>]pp t^skyl Jlsh bdhd<lv>.
Translation: *The Supervisory Committee on production of OPEC members* is due to meet in Esfehan on *May 16 and 17*.
Revised: OPEC's Ministerial Monitoring Sub-Committee is due to hold a meeting in Esfehan on 16th and 17th of Ordibehesht.

In the next translation, *riys hv|pym|yy k^svry |rdn [raise havApeimAyiye keshvariye ordon],* meaning "the head of aviation of Jordan," is translated as *the countrys air minister. riys,* meaning *head* or *chairman* is translated as *minister,* and *hv|pym|yy k^svry* meaning *aviation* as *air.* No translation is available for |*rdn* meaning *Jordan.* The translation hardly conveys the intended meaning of the Persian text. In the revision, to prevent ambiguity, the head of aviation of Jordan is used, every word of which is present in the Shiraz dictionary. Furthermore, in the translation for the same sentence, the critical words *brn|mh prv|z[barnAmeye parvAz],* meaning *flight schedule* and *@rf hv|nvrdy [orfe havAnavardi],* meaning *aviation convention* are not translated at all. The translation provided for the part of the sentence, **[[brn|mh prv|z<n>]np kh<rel> [@rf hv| nvrdy<n>]np** , meaning ...flight schedule that is regarded as aviation convention is missed; the translated phrase *its flight path did not count as entering Jordans air,* does not seem appropriate. In the revision, these problems are fixed. Note that the revised translation is actually conveying the ambiguity present in the Persian sentence. Although understandable, the Persian sentence does not make it clear whether, permission for passing the sky of Jordan is regarded as an aviation convention or asking for the flight schedule or both. Only the readers' knowledge would help judge the case. The revision is meant to convey the exact meaning, maintaining the original ambiguity, rather than interpreting it.

Persian: |m|<con>[dr<pre> |rdn<pn>]pp [[stv|n<ti> J|sm zy|d<pn>]np[riys<n> hv|pym|yy k^svry<n> |rdn<pn>]np]np gft<v>: [hv|pym|y<n> @r|qy<ad>]np drxv|st<lv1>[|J|zh<n> @bvr<n> [|z<pre> asm|n<n> |rdn<pn>]pp]np v<con>[[brn|mh prv|z<n>]np kh<rel> [@rf hv| nvrdy<n>]np mHsvb my~^svd<lv>]np, nkrdh |st<lv>
Translation: But in Jordan Lieutenant Jasim Ziad, the *countrys air minister,* said: the Iraqi plane did not request permission to enter the Jordanian airspace *and its flight path did not count as entering Jordans air.*
Revised: But in Jordan, Lieutenant Jasim Ziad, the head of aviation of Jordan said: the Iraqi plane did not ask for permission for passing the sky of Jordan and flight schedule that is regarded as aviation convention.

3 Glass-Box Evaluations

The bilingual tagged corpus was used in the Shiraz project for testing the results of the morphological analyzer, the dictionary entries and the syntactic parser. For testing the output of morphological analysis and dictionary lookup, the Glass-Box Evaluation system compared the results with the hand-produced annotations. The results were grouped as:

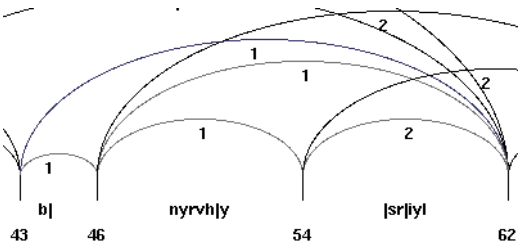
- Non-Matching Structures, if the part-of-speech marked in the corpus did not match the POS obtained from morphological analysis and dictionary look-up or, for syntax, if the syntactic tree does not match the parenthesized phrase.
- Non-Matching Spans, if the range of the annotation in the corpus is different from the range produced by the system. This is the case, for instance, if a compound is not recognized and only the individual parts are analyzed by the system,
- Matches, if both the span and the structure of the corpus annotation match the span and structure produced by the system.
- No matches when a corpus annotation is not comparable to any result produced by the system.

As each stage of the analysis may produce multiple outputs, the results are computed over all outputs. That is, if morphological analysis for example produces three outputs, two of which are wrong, and one right, that would count as one match and two non-matches. As a result, in our glass-box evaluations, the precision is often low.

Matches In the example below, the Glassbox module has matched the Preposition Phrase given in curly brackets *{b| nyrvh|y |sr|iyl}* (= with forces Israel) with the tag *pp* in the corpus. The numbers in the brackets preceding the phrasal entry, ([43,61] in this example) mark the annotation span. The numbers in the parentheses indicate the number of matches, the number of non-matches, and the number of overlapping spans or analyses, respectively. In

this case, there is one match and no non-matches. There are 6 overlapping analyses. These are not indicating a problem, however, since the Glassbox system considers any analysis such as the ones for each entry as a sub-analysis. The graph of the syntactic parses for this Preposition Phrase is shown in Figure 1 with the corresponding number of analyses.

[43,61] {b| nyrvh|y |sr|iyl} (1,0,6)
Matches
1. per.Rule.PrepositionPhrase



Certain elements are matched more than once, without any apparent difference. This is due to different morphological analyses that are not distinguishable for the Glassbox module. In the example below, the word *t_Z|hr|t [tazAhorAt]* is a Noun in both matches. In the first case, however, it is interpreted as the irregular plural meaning “demonstrations” and in the second case as the noun *t_Z|hr* “pretension” with a plural morpheme.

[[0,17] {^sdt t_Z|hr|t dyrvz} (2,0,14)
Matches
1. per.Rule.NounPhrase
2. per.Rule.NounPhrase

Non-matching structure The first word of example (1) results in one match and one non-match as shown below. This is due to lexical ambiguity of the entry, since *flsTyny* (Palestinian) can be either a noun or an adjective. Hence, the system produces two parses, one NounPhrase (which has been matched against the corpus *np*) and one AdjectivePhrase (which results in a non-matching structure against the corpus tag)[1]
.

[0,9] {flsTyny|n} (1,1,0)
Matches
1. per.Rule.NounPhrase

Non-matching structures
1.
per.Rule.NO
per.Rule.AdjectivePhrase

Type mismatch The following example shows a Preposition Phrase which matches the top-level phrasal type but with a mismatch at the word level (followed by a second perfect match). The first match has found a Preposition Phrase which matches the phrasal tag in the corpus, but the POS tags have resulted in a mismatch as shown. The system has produced a Noun for the word *flsTyny* (Palestinian), whereas the corpus specifies an Adjective. The second match does not have a type mismatch because both corpus and system have agreed on the POS of the

1. per.Rule.NO is used by the Glassbox to refer to the *np* tag which can be mapped into either a NounPhrase or an ObjectPhrase (a NP + object marker).

entry.

[58,81] {|z Jmlh yk plys flsTyny} (2,0,29)
Matches
1. per.Rule.PrepositionPhrase
Bolero:
Type mismatch for {flsTyny}: (system) per.Type.Noun, (corpus) per.Type.Adjective

2. per.Rule.PrepositionPhrase

Such type mismatches are expected because of the lexical ambiguities present in Persian words. Words carrying more that one part-of-speech are common and the lack of short vowels also adds to such ambiguities. Another mismatch takes place when a compound has been detected in either the corpus or the system, but not in both. The compound *hiyt vzyr|n* (cabinet) has been recognized by the system parser, but in the corpus, it has been tagged as two separate nouns. This mismatch is presented by the Glassbox report as "Corpus remaining: per.Type.Noun", indicating that the corpus contains an additional Noun POS. Note in this case that there are three possible matches for this Noun Phrase. One of the main problems in ambiguity encountered in the Glassbox report is due to the recognition of compounds in the system. When a compound is recognized by the Compound Lookup component in the machine translation system, the inflectional information is underspecified. This results in unification of several rules at the syntactic level, giving rise to ambiguous structures.

[28,44] {hyit vzyr|n k^svr} (3,0,23)
Matches
1. per.Rule.NounPhrase
Bolero:
Corpus remaining: per.Type.Noun

2. per.Rule.NounPhrase
Bolero:
Corpus remaining: per.Type.Noun

3. per.Rule.NounPhrase
Bolero:
Corpus remaining: per.Type.Noun

The results of the Glass-Box Evaluation component were used to correct and edit any mistakes in the dictionary, in the stemmer or in the morphological component. The Glass-Box tool produces a report as a simple text format which contains some additional corpus-wide statistics. It also produces the evaluation report as a set of browsable HTML pages which allows for browsing of the report by test item (sentence) or by tag category (e.g. results for all prepositional phrases). The result allows evaluation of the coverage of the various components in the system (only the dictionary, the morphological analyzer and the parser were evaluated using this tool).

It was initially envisaged that the linguist would also use the glass-box as a testing tool in the course of dictionary and grammar development. It turned out that the results of the glass-box were used essentially to correct and add lexical entries. Perhaps because the grammar had a relatively small coverage (in term of number of constructions covered), the linguists did not use the Glass-Box for the actual testing of the morphological and syntactic grammars. Rather, a specific testing tool was developed for morphology using morphological test suites (see e.g. Dauphin & Lux 96, Klein et als. 98) and a graphical debugging tool for syntax.

4 Black-box evaluation

Translation quality was evaluated as the project progressed by simply looking at the resulting translation, and for non-Persian speakers, by comparing the translation to the manually produced one. No specific evaluation criteria were used, except the direct comparison with the manually produced translation. As mentioned above, this comparison was made difficult since the original translations were produced without any guidelines. In the end, a

subset of these were corrected and used for further evaluation. These problems suggest a way of automating the evaluation of the quality of translations. A machine translation system will produce a target translation where:

1. All lexical equivalents for content words, single words or multi-word expressions, are contained in the system's dictionaries and no unknown lexical material can be introduced (except for the case of unknown words which appear as is –or in the case of the Shiraz system, in Latin transliteration– in the output).
2. Functions words are generated during transfer and/or generation and could be introduced based either on lexical equivalences (e.g., conjunctions) and/or on pure syntactic constraints (e.g., pronouns).
3. The word order will differ and in the best of the case, follow a normative target grammar. Typically, changes in the word order correspond to the smallest set of changes that is necessary to produce a grammatical output.[1]

This suggests a rather strict set of corresponding guidelines for producing the target side of the test suite:

1. The translation should be basically a literal translation: an MT system will not produce a radically different kind of translation.
2. All words or expressions in the translation must be specified in a reference bilingual dictionary (which can be a paper dictionary for example). Expressions could even include idioms and frozen metaphors if they are listed in the dictionary.
3. The word order should follow closely the source word order except in those cases were this would produce an ungrammatical sentence.
4. There should be no arbitrary changes in the syntactic structure of phrases and clauses unless using the same structure would results in an ungrammatical sentences or in a major change in meaning. For example, a passive sentence should remain in passive unless it is impossible to produce a passive target sentence for any equivalent verb, or if the use of passive implies a major shift in meaning. Similarly, relative clauses should be preferably translated by relative clauses, etc.

If the target side of the test suite were to be produced manually following these guidelines (which could be built into a checking software), we can imagine an automatic evaluation procedure which would measure the distance between the two translations. This procedure would be similar to an alignment procedure used to align bilingual corpora: the alignment is done modulo lexical equivalent specified in a dictionary (which could be used to measure accuracy in word-sense selection), modulo function words, and modulo some word order changes.

5 Conclusion

The Shiraz project initially included the double use of glass-box and black-box evaluations as internal testing and development tools as well as an evaluation tool for the project's sponsors. These evaluations were helpful to external consumers in order to have a picture of the coverage and the quality of the system at various times during the development of the system, and especially when the system did not yet produce any translations. However, these evaluations were not useful for the development team. Instead, the same results have been achieved more efficiently by developing special purpose tools that could be integrated more easily in the debugging loop. The testing tools that were developed included a specific component for morphology testing that used a systematically constructed morphological test suite (pairs of input/output in the form of inflected words and citation form plus morphosyntactic features), and a browser to inspect internal data structures at some points during processing that was used for debugging the syntactic grammar. A special purpose tool would also have been more efficient in the case of dictionary testing and debugging.

Our evaluations were intended for external consumption (that is, by our funders). Attempts to use the results of these evaluations (particularly the glass-box evaluations) for internal consumption, that is, to guide the development and to single out problematic areas or areas where coverage was lacking, were not successful. We suggest that it is better to separate the task of testing for feedback and improvement from the task of external evaluation.

1. This is not necessarily the case though, especially for systems with a sophisticated syntactic transfer components, or for systems using a semantic representation (semantic transfer or KBMT).

References

D. Arnold, R.L. Humphreys and L. Sadler (eds.). 1994. Special Issue on the Evaluation of Machine Translation. *Machine Translation* 8(1-2).

Dauphin E., Lux V. 1996. “Corpus-Based annotated Test Set for Machine Translation Evaluation by an Industrial User”. In *Proceedings of the 16th International Conference on Computational Linguistics, COLING-96*, August 5-9, Center for Sprogteknologi, Copenhagen, Denmark, pp.1061-1065.

Doran, Christy, Dania Egedi, Beth Ann Hockey, B. Srinivas. 1994. “Status of the XTAG System”. In Proceedings of TAG+3. See also http://www.cis.upenn.edu/~xtag.

O’Connell, T., O’Mara, F. and White, J. 1994a. “The ARPA MT evaluation methodologies: Evolution, lessons and further approaches”. *Proceedings of the First Conference of the Association for Machine Translation in the Americas*, Columbia, U.S.A.

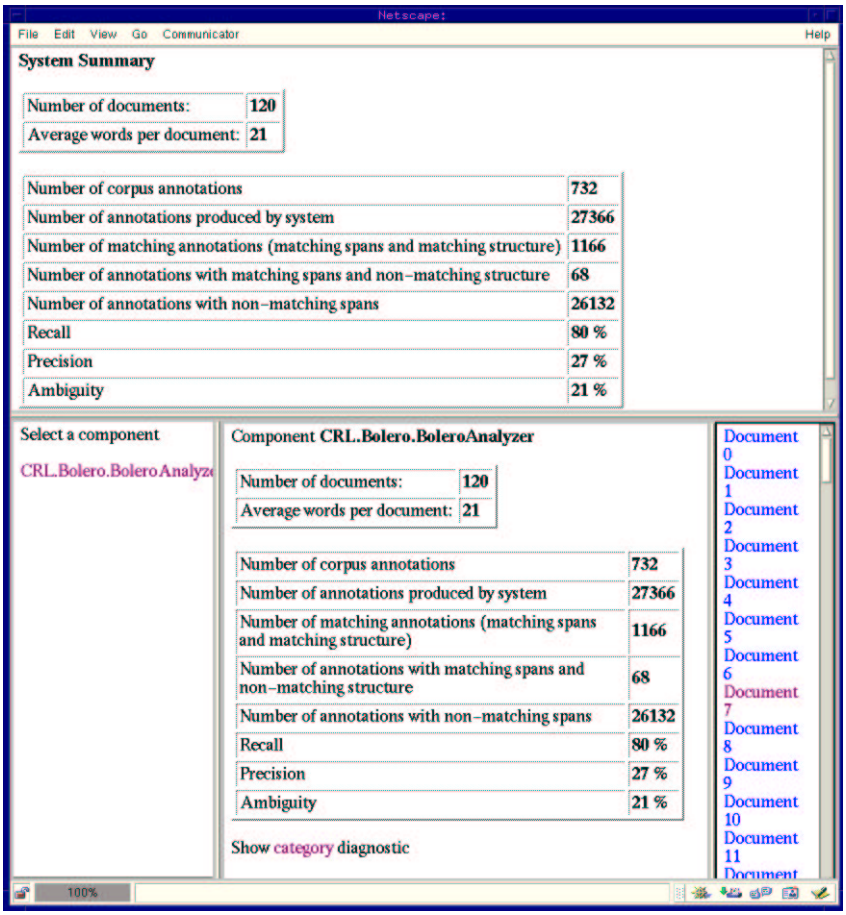
O’Connell, Theresa, Francis E. O’Mara, Kathryn B. Taylor. 1994b. “Sensitivity, Portability and Economy in the ARPA Machine Translation Evaluation Methodology”. Technical Report, PRC Inc. Available at http://ursula.georgetown.edu/mt_web/Meth.htm.

Klein, Judith, Sabine Lehmann, Klaus Netter, Tillman Wegst. 1998. “DIET in the context of MT evaluation”. In *Proceedings of KONVENS-98*.

Wagner, Simone. 1998. “Small Scale Evaluation Methods”. In *Proceedings of the Workshop on Evaluation of the Linguistic Performance of Machine Translation Systems, KONVENS-98*. Bonn. pp93-105.

Appendix: HTML Interface for Glass-Box evaluations

The main frame:



Display by test item (sentence):

Document 72

mq|mh|y ^cyny my gvynd : |yn dv nfr kh b| mrdm ^cyn m@|^srt nd|^stnd , zb|n ^cyny nmy d|nstnd v tnh| ^shrt d|dh ^sdh bvd kh b| n|r|Zy|n ^cyny tm|s d|^stnd ^cgvnh my tv|nnd bdvn _zkr dlyl v m|x_z ^cnyn |d@|yy r| m TrH knnd .

[0,12] {mq|mh|y ^cyny} (3,0,8)

Matches

1. *per.Rule.NounPhrase*
2. *per.Rule.NounPhrase*

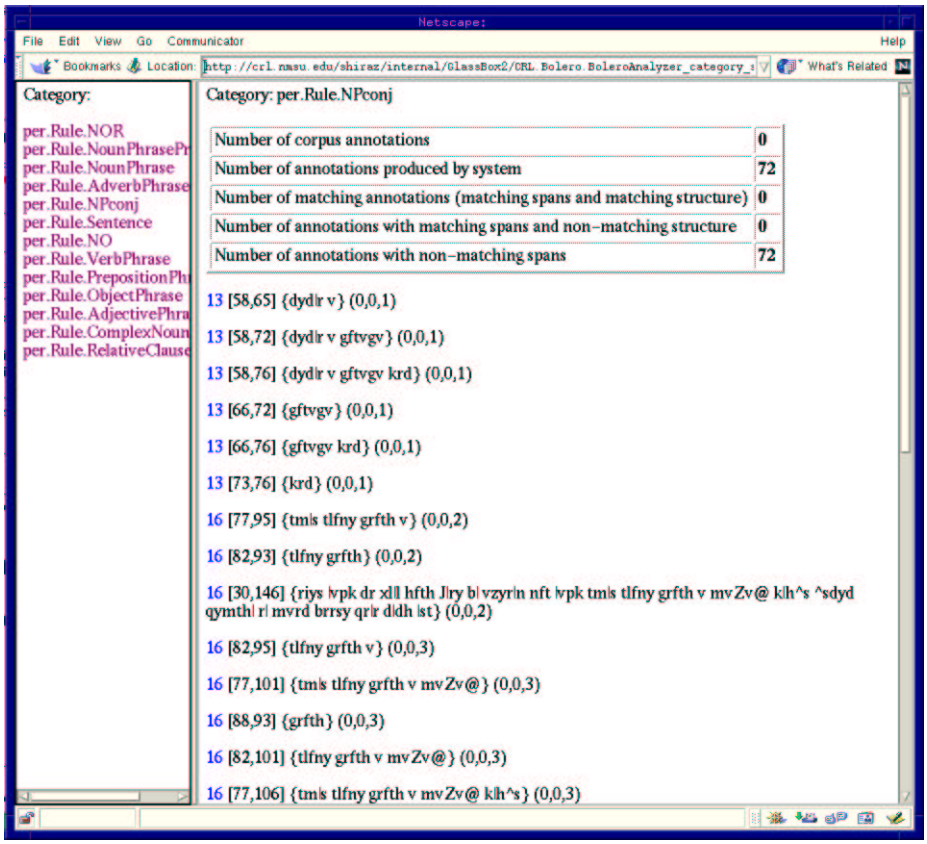
   Bolero

   `Type mismatch for {^cyny}: (system) per.Type.Noun, (corpus) per.Type.Adjective`

3. *per.Rule.NounPhrase*

```
1. NounPhrase[
     aTree : "[[[Noun [[Adjective]A']AP]N']NPo]NP"
     posList : <:
       Noun
       Adjective :>
     head : Entry[
       form : Form[
         morph : NounMorphology[
           lex : NominalLexical[
             pos : #111=Noun
             number : Number]
           infl : #79=NominalInflection[
             ezafe : EzTrue
             clitic : Clitic[
               number : Number
               person : Person]
             indefiniteEnclitic : False
             number : Plural]]
         orth.exp : "Ù?Ù‚Ø§Ù?"]
```

Display by category:



Netscape:

File Edit View Go Communicator Help

Bookmarks Location: http://crl.nmsu.edu/shiraz/internal/GlassBox2/CRL.Bolero.BoleroAnalyzer_category_s What's Related

Category:

per.Rule.NOR
per.Rule.NounPhrasePr
per.Rule.NounPhrase
per.Rule.AdverbPhrase
per.Rule.NPconj
per.Rule.Sentence
per.Rule.NO
per.Rule.VerbPhrase
per.Rule.PrepositionPh
per.Rule.ObjectPhrase
per.Rule.AdjectivePhra
per.Rule.ComplexNoun
per.Rule.RelativeClause

Category: per.Rule.NPconj

Number of corpus annotations	0
Number of annotations produced by system	72
Number of matching annotations (matching spans and matching structure)	0
Number of annotations with matching spans and non-matching structure	0
Number of annotations with non-matching spans	72

13 [58,65] {dyd|r v} (0,0,1)

13 [58,72] {dyd|r v gftvgv} (0,0,1)

13 [58,76] {dyd|r v gftvgv krd} (0,0,1)

13 [66,72] {gftvgv} (0,0,1)

13 [66,76] {gftvgv krd} (0,0,1)

13 [73,76] {krd} (0,0,1)

16 [77,95] {tm|s tlfny grfth v} (0,0,2)

16 [82,93] {tlfny grfth} (0,0,2)

16 [30,146] {riys |vpk dr xl|l hfth J|ry b| vzyr|n nft |vpk tm|s tlfny grfth v mvZv@ k|h^s ^sdyd qymth| r| mvrd brrsy qr|r d|dh |st} (0,0,2)

16 [82,95] {tlfny grfth v} (0,0,3)

16 [77,101] {tm|s tlfny grfth v mvZv@} (0,0,3)

16 [88,93] {grfth} (0,0,3)

16 [82,101] {tlfny grfth v mvZv@} (0,0,3)

16 [77,106] {tm|s tlfny grfth v mvZv@ k|h^s} (0,0,3)