Basic semantic error detection by sentence comparison
Sébastien L'haire, University of Geneva. Eurocall 2005,   Cracow

Motivations:
- Open questions rare in commercial CALL software
- Pattern matching techniques → poor feedback
- Anticipate every possible answer → complex formulae
- Correct answers does not fit into pattern → rejected
- Parsers detect ungrammatical answers
  - A grammatical sentence can be an incorrect answer
  - **Meaning must be also checked**

Pseudo-semantic structures (PSS)
- A formalism used for sentence generation and for automatic translation
- Predicate-argument structures
- Lexical information for noun, verbs, adjectives, adverbs
- Abstract information:
  - Thematic roles: agent, patient, experiencer
  - aspect, tense etc. for verbs
  - Determiner type, gender, number etc. for nouns
  - …

Sentence comparison technique

Learners' answers compared to model sentence stored into database. PSS (semantic structures) extracted for both sentences and compared.
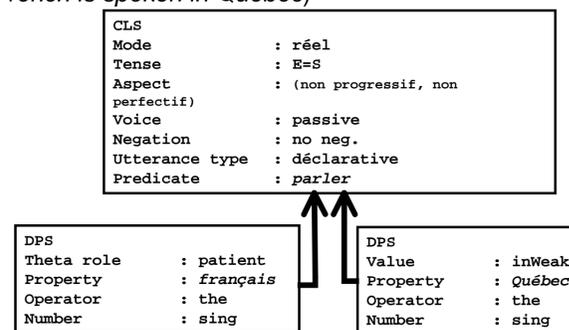
Question on image description: *de quelle couleur est la petite souris?* (*which colour is the small mouse*) Model answer: *la petite souris est grise* (*the small mouse is grey*). Learner answer: *elle est grise*. If pronominalisation allowed, sentence is considered correct.

Pronominalisation exercise: *as-tu vu les voitures?* (*did you see the cars?*). Model answer: *oui je les ai vues.* (*yes, I have seen-past-part-fem-pl them*). Learner answer: *je les ai vus* (I have seen-past-part-masc-pl) Grammatical sentence but considered incorrect.
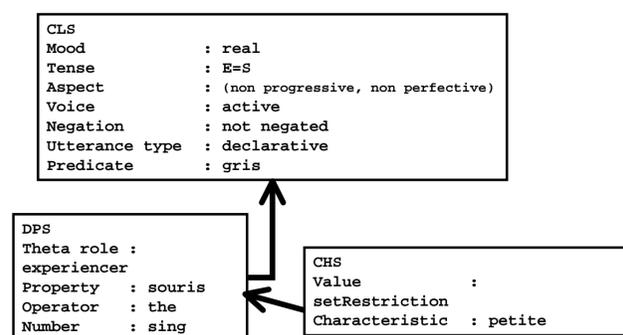
Question: *Qu'est-ce que Jean a acheté?* (*What did John buy?*). Model answer: *Jean a acheté un perroquet.* (*John has bought a parrot*),. Learner answer: *Jean a acheté un oiseau.* (*John has bought a bird*). Could be accepted if we use a semantic dictionary with hypernyms / hyponyms / synonyms relations.

Question: *quelle langue est parlée au Québec?* (*which language is spoken in Quebec?*). Expected answer: *Le français est parlé au Québec* (Cf. structure above). Learner: *On parle français au Québec.* Feedback: OK. | *On parle le français au Québec.* Feedback: OK, determiner added. | *Le français se parle au Québec:* Feedback: OK, different mode used.

Error detection by parsing:
- A parser assigns a structure to a sentence using combination rules and constraints on combination (e.g. subject-verb agreement, verb argument structures…)
- Constraint relaxation → builds sentences even if combination rules violated
- Other techniques: mal-rules, ad-hoc rules
- Better feedback for learners
- Interesting results achieved by parsers, but not good enough (the more complex the sentence, the worse the results)

Structure for: Le français est parlé au Québec
(*French is spoken in Quebec*)

```
CLS
Mode            : réel
Tense           : E=S
Aspect          : (non progressif, non
perfectif)
Voice           : passive
Negation        : no neg.
Utterance type  : déclarative
Predicate       : parler
```

```
DPS
Theta role    : patient
Property      : français
Operator      : the
Number        : sing
```

```
DPS
Value         : inWeak
Property      : Québec
Operator      : the
Number        : sing
```

Structure for: La petite souris est grise
(*The small mouse is grey*)

```
CLS
Mood            : real
Tense           : E=S
Aspect          : (non progressive, non perfective)
Voice           : active
Negation        : not negated
Utterance type  : declarative
Predicate       : gris
```

```
DPS
Theta role :
experiencer
Property   : souris
Operator   : the
Number     : sing
```

```
CHS
Value          :
setRestriction
Characteristic : petite
```

Similar technique
Dorr (1995) uses *Lexical Conceptual Structures*, which are abstract semantic structures:
Expected: John ran to the house
Structure: [Event GOLoc ([Thing JOHN], [Path TOLoc ([Position ATLoc ([Thing JOHN],[Property HOUSE])]), [Manner RUNNINGLY])]]
Learner answer: John went to the house
Structure: [Event GOLoc ([Thing JOHN], [Path TOLoc ([Position ATLoc ([Thing JOHN], [Property HOUSE])])])]
Concept of running is absent from learner's answer.

Construction	Model answer	Learner's sentence	Feedback
Bad use of clitic pronoun	Je les ai vues	Je les ai vus	Gender error on pronoun
Signification error	La grosse souris est grise.	La grosse souris est blanche.	Signification error on adjective
Global question (yes/no)	Est-ce que Jean a lu ce livre ?	Jean a-t-il lu ce livre.	No feedback : equivalent construction
Voice error	Le chat poursuit la souris.	La souris est poursuivie par le chat.	Voice error on verb
Relative clauses	Le voisin que Jean a vu	La voisine que Jean a vue	Signification error on noun
Cleft construction	C'est Pierre que Jean a vu	Jean a vu Pierre.	Error on cleft construction
Topicalization	Avec ses jumelles, Jean a observé cet homme	Jean a observé cet homme avec ses jumelles	Error on topicalization
Adjunction	Jean a regardé l'homme avec ses jumelles	Jean a observé l'homme.	Missing constituent
Partial questions	Quel livre est-ce que Jean lit ?	Quel livre Jean lit-il ?	No feedback : equivalent construction
Raising	Jean semble dormir	Il semble que Jean dort.	No feedback : equivalent construction