

a framework for teaching
multimodal interface
construction

Keerthi Rajendran & Simon Lynch

University of Teesside

aims

UG & PG modules & projects to include MMI / MMD

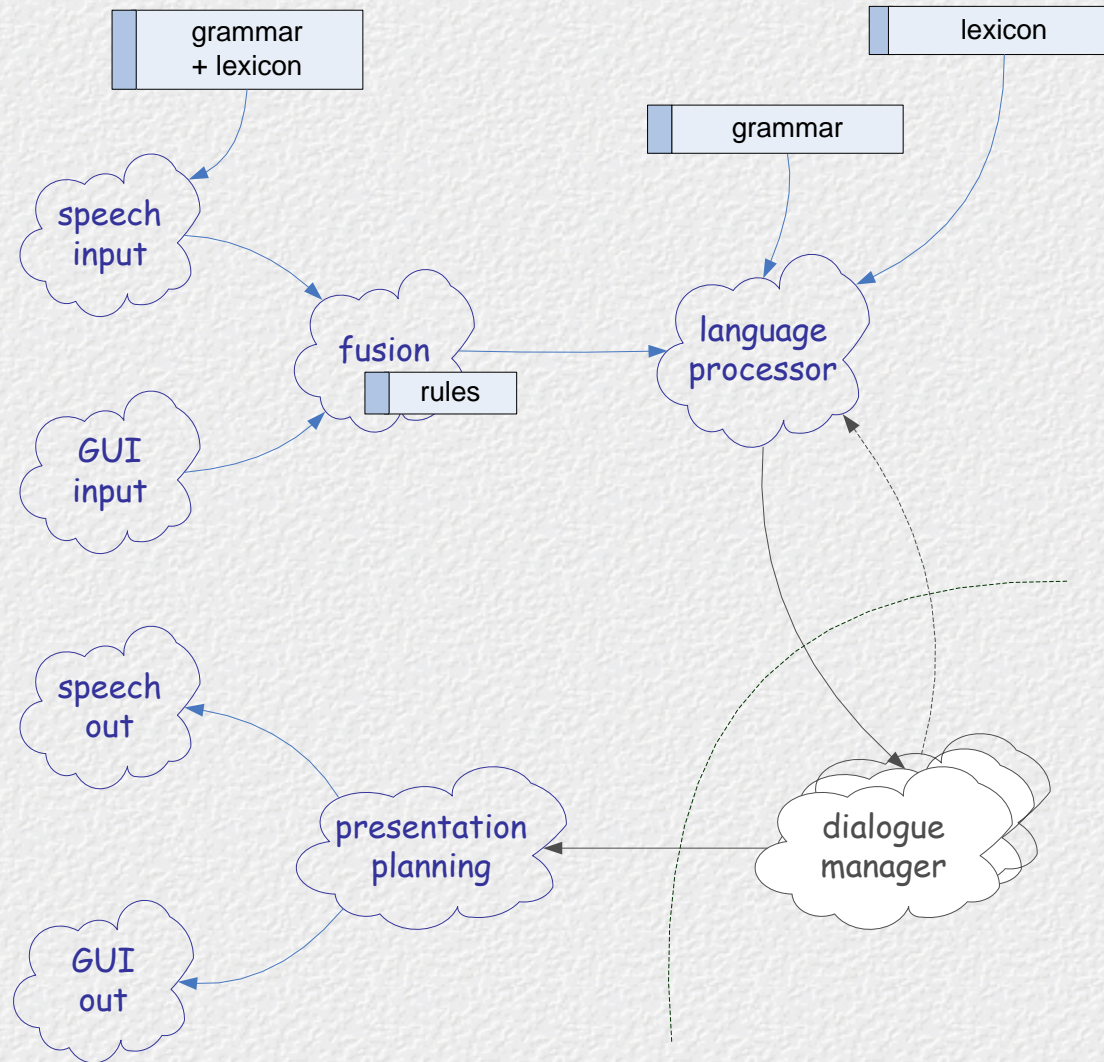
but MMI / MMD...

primarily research

educational use only to improve T&L in general

- not to advance teaching MMI / MMD

example architecture



difficulties

technologies (speech, etc)

- more difficult, more error prone

issues like...

- fusion, fission, parallelism
- dialog, conversation tracking

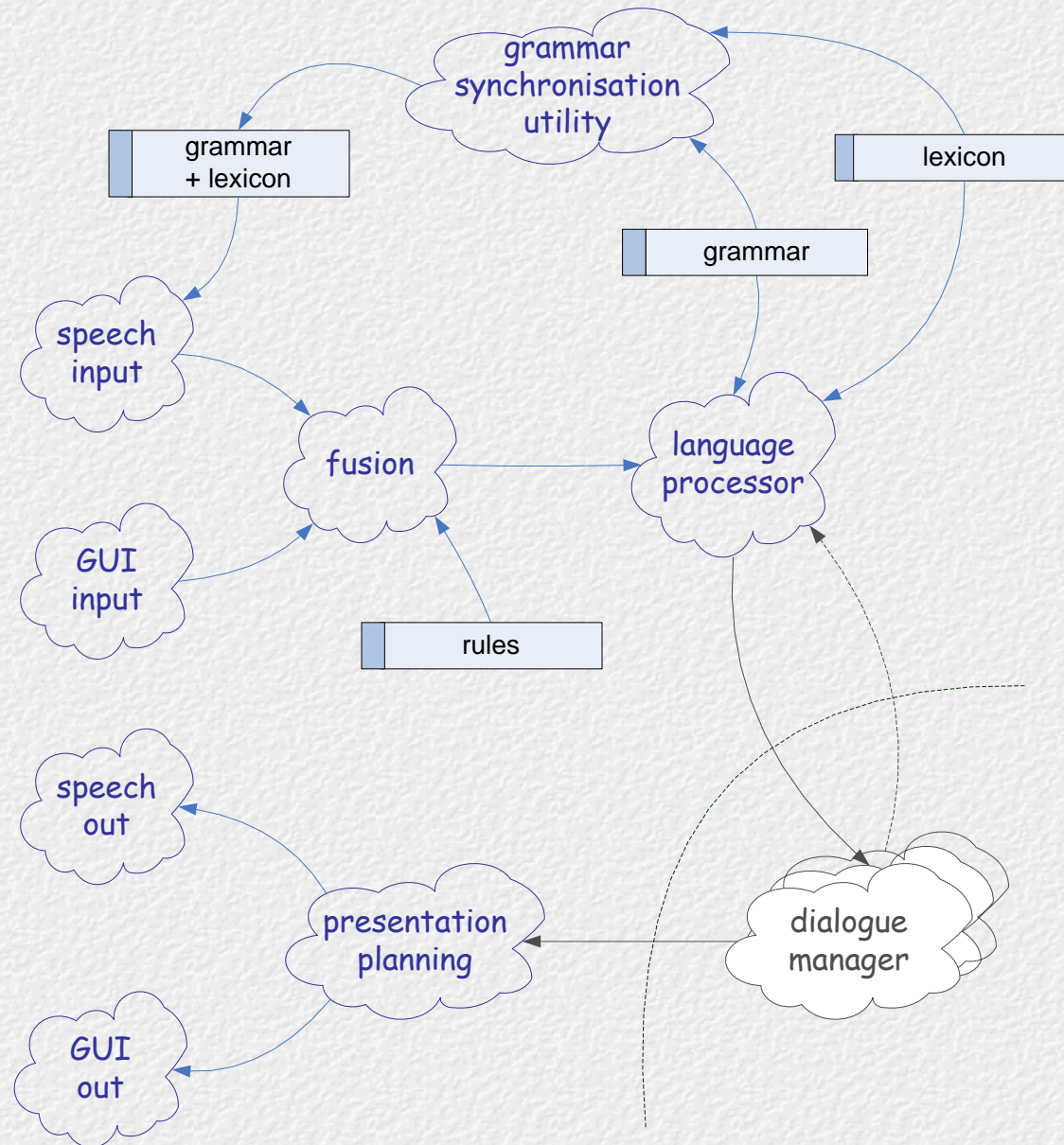
need for base technologies (eg: MAS)

complex system architectures

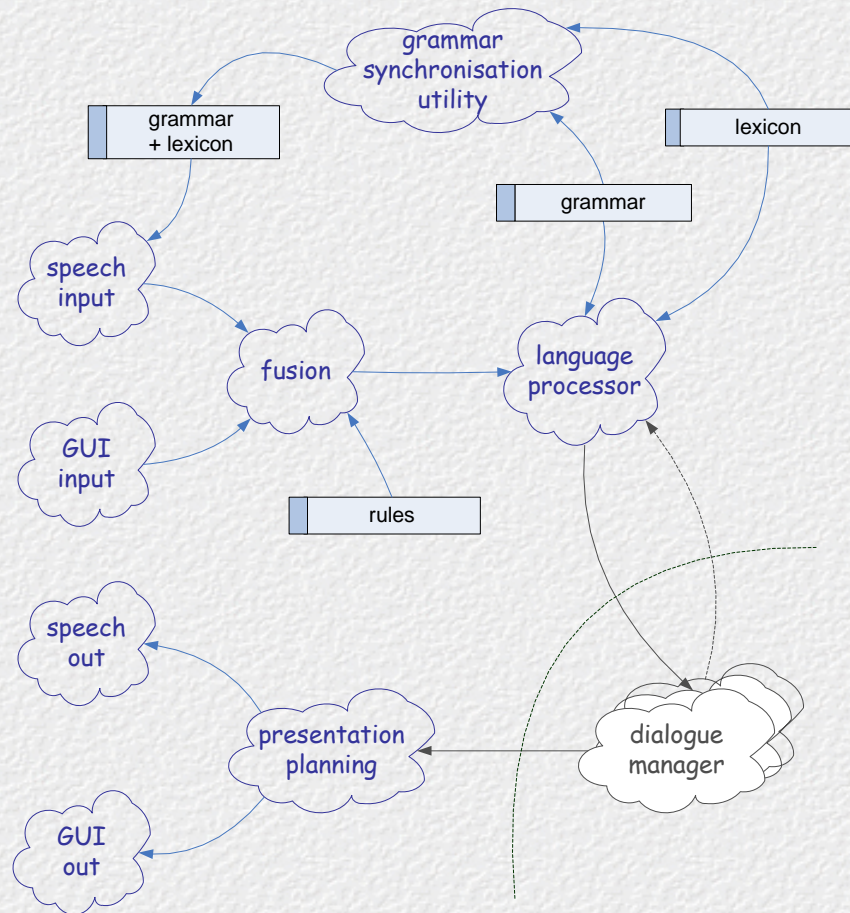
approaches

1. make the course theoretical but...
 - no student projects
 - limited objectives (no rules, experimentation)
2. focus on small number of specific MM components but...
 - no wider context
 - no design patterns
3. script / configure example MM systems but...
 - need support systems / framework

a framework architecture

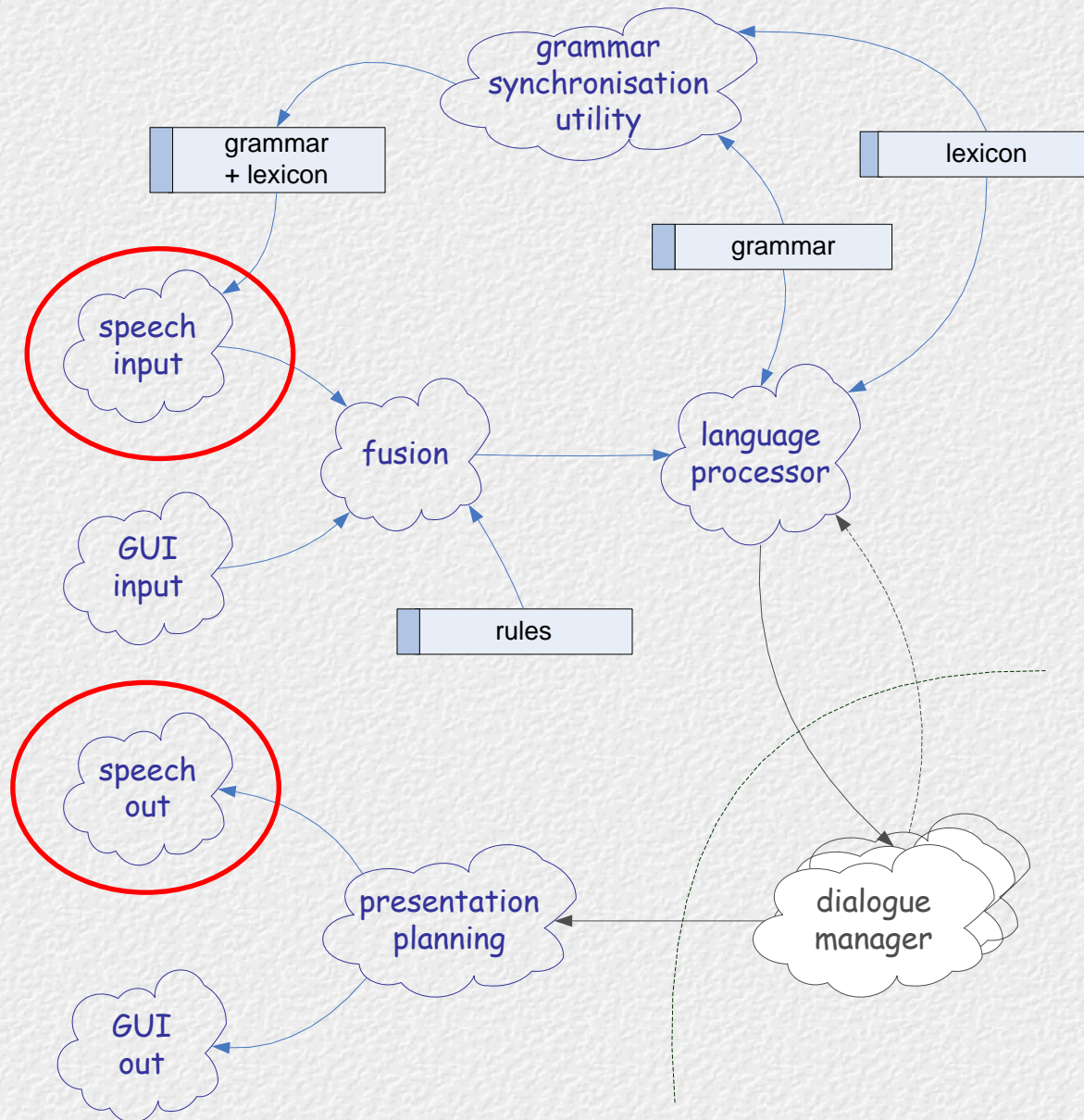


a framework architecture

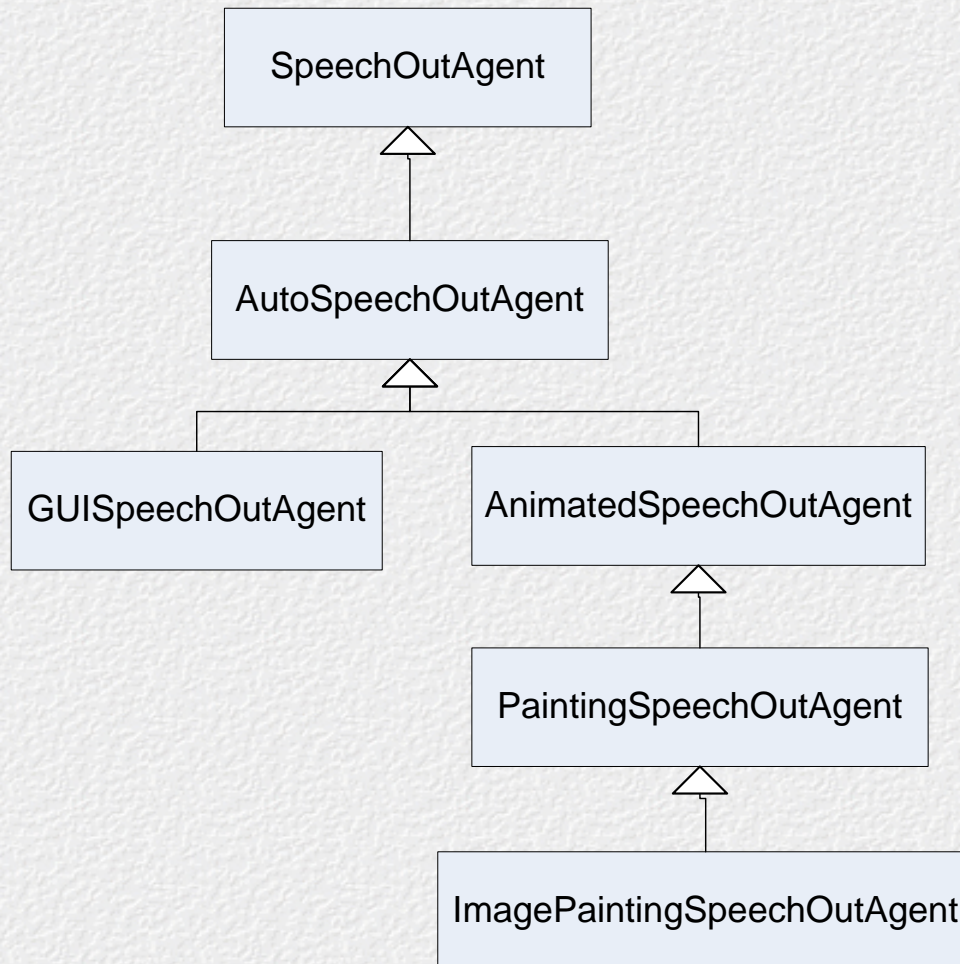


- a MAS architecture
- base technology: *Boris*
see www.agent-domain.org
- flexible & extendible
- components are agents
- agents pre-written
- generic/scriptable
- need driven by student experience

a framework architecture



speech output



Java example...

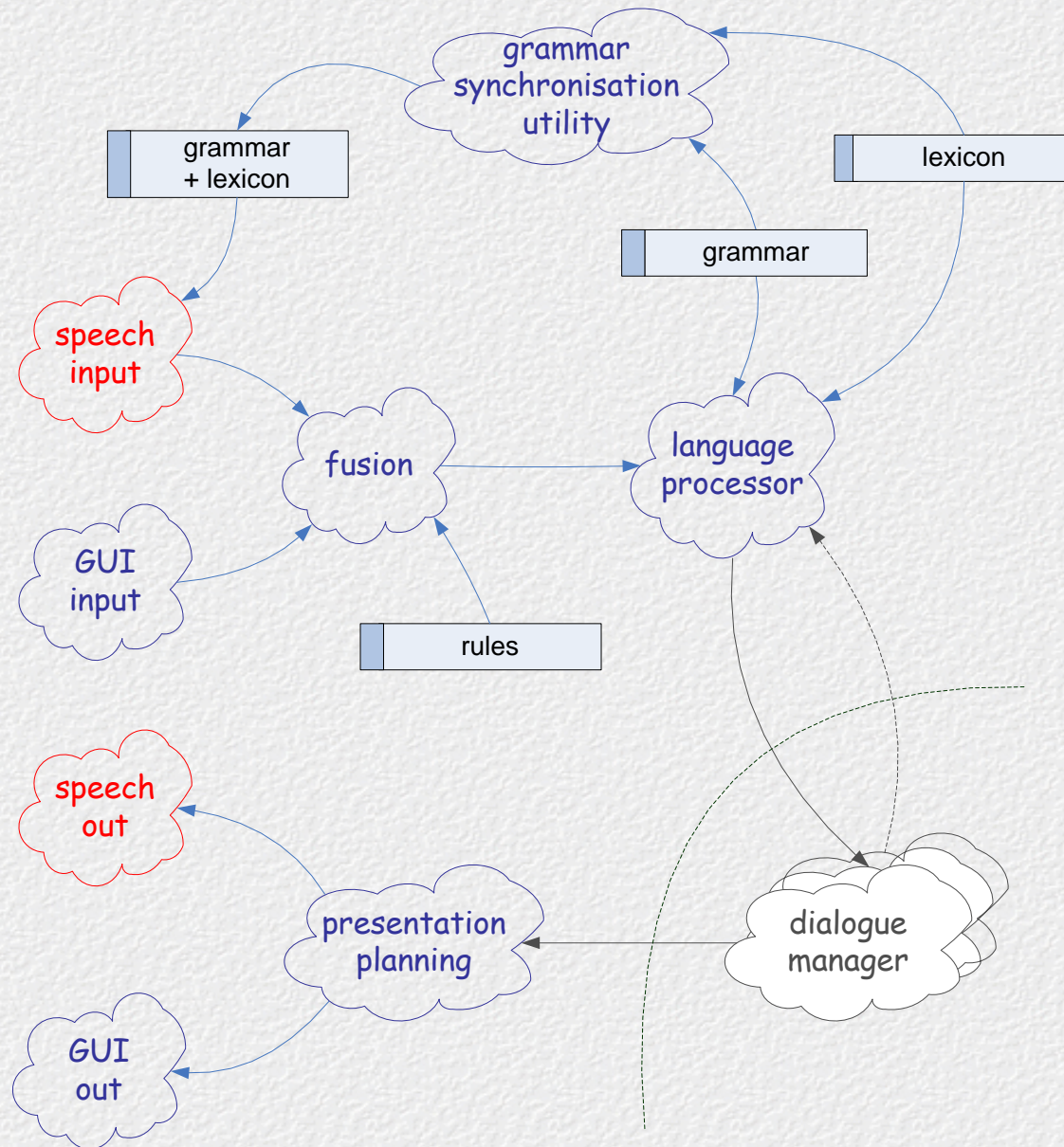
```
SpeechOutAgent s =  
    new SpeechOutAgent();
```

```
s.speak( "hello" );
```

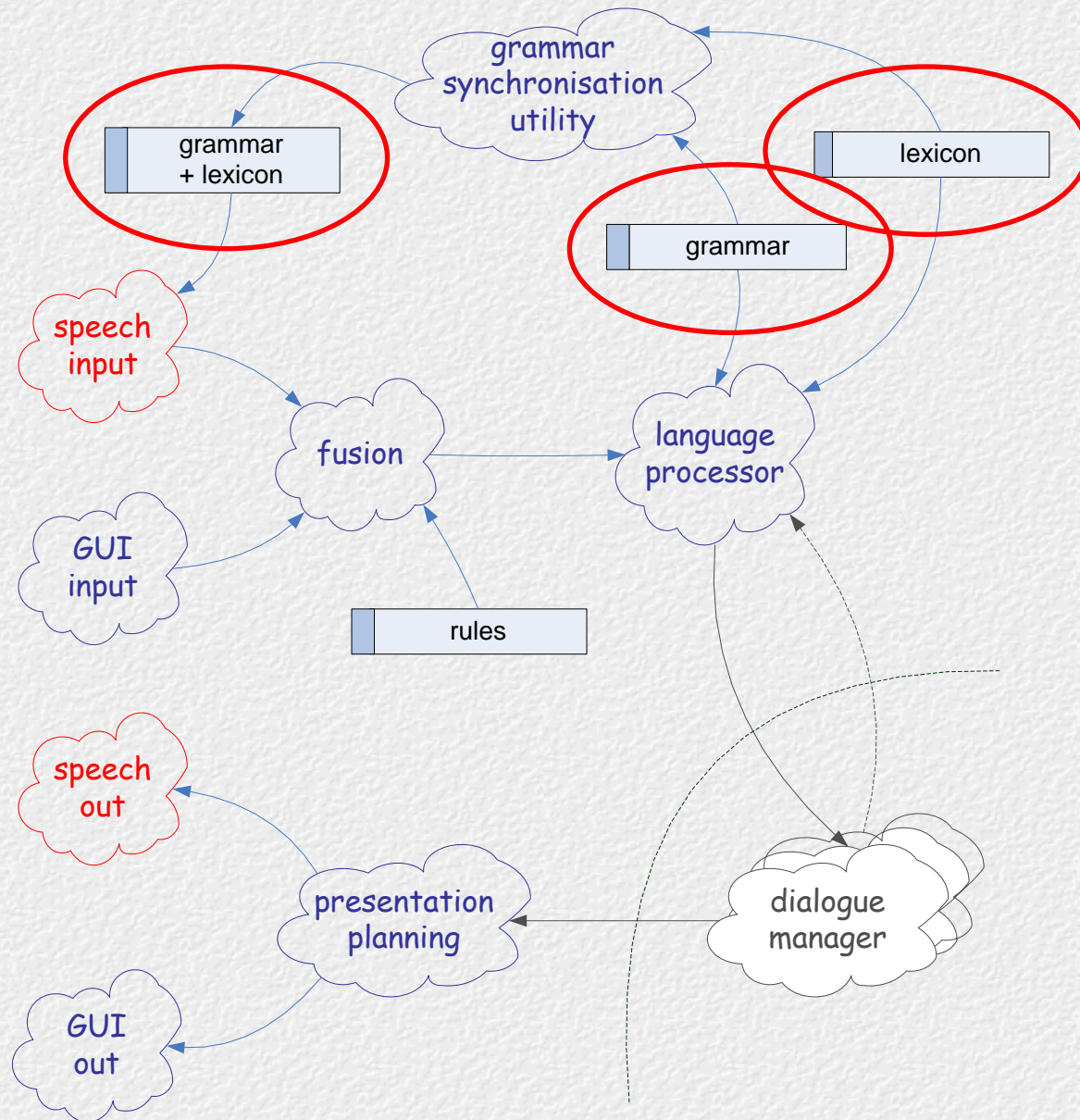
optional...

```
s.setVoiceModel( v );
```

a framework architecture



a framework architecture



grammars

<color> =

red | blue | pink | green;

<object> =

the <color> box | prism | it;

public <sentence> =

pick up <object> [and <continue>];

public <continue> =

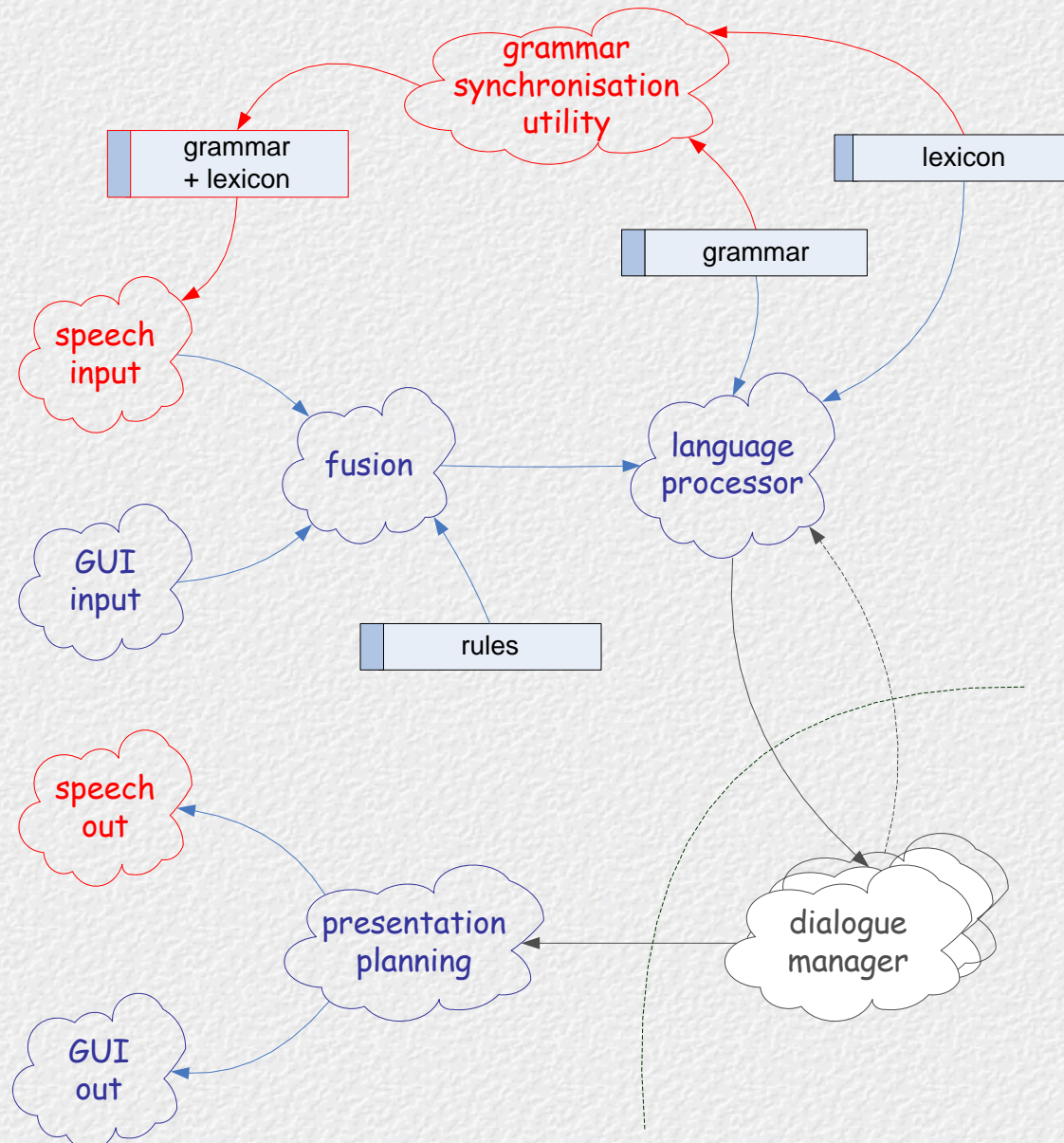
put <object> on the <object>;

grammars

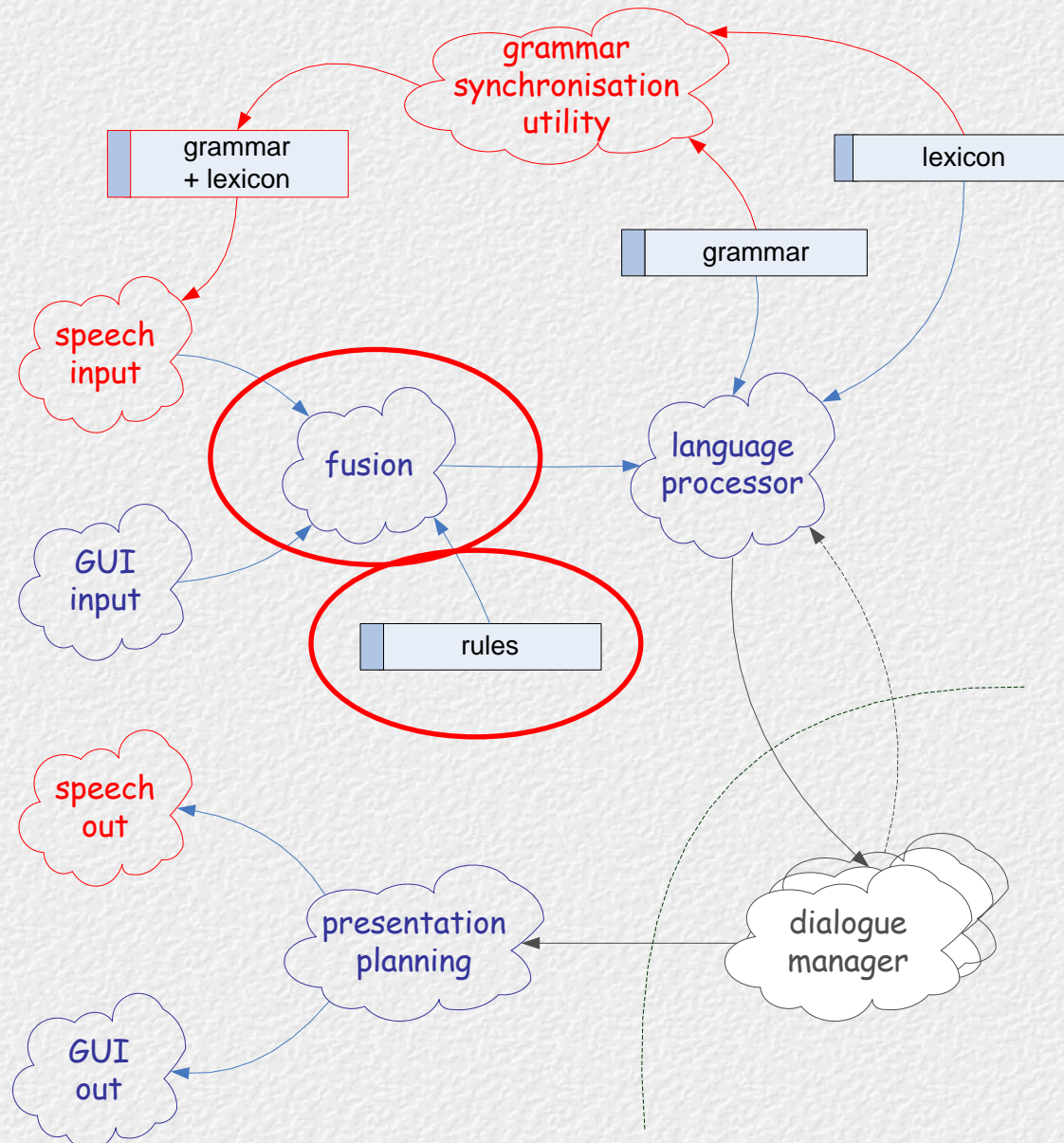
lexicon (a det (quantifier any))
 (block noun (manipulable object...) ...etc...)

grammar (s1 (sentence -> noun-phrase verb-phrase)
 (actor . noun-phrase)
 (action . verb-phrase.action)
 (object . verb-phrase.object))
(np (noun-phrase -> det noun)
 (det . noun))
(vp (verb-phrase -> verb noun-phrase)
 (action . verb))

a framework architecture



a framework architecture



fusing

- complex process (various constraints)
- observed various problems for students with fusion (also semantic analysis & dialog)
- student control groups understand problem but unable to build a solution

fusing example

input

```
(speech-in (that 1 5 20))  
(gui-in g0 type $obj)  
(gui-in g0 time 12)  
(gui-in g0 target blue-block)
```

output required

```
(fused ($obj blue-block 1 5 20))
```

fusing rules

that+click: matching preconditions

(speech-in (that ?edge ?start ?end))

(gui-in ?click type \$obj)

(gui-in ?click target ?target)

(gui-in ?click time ?t)

guards (<= ?start ?t ?end)

add (fused (\$obj ?target ?edge ?start ?end))

delete (speech-in (that ?edge ?start ?end))

(gui-in ?click ==)

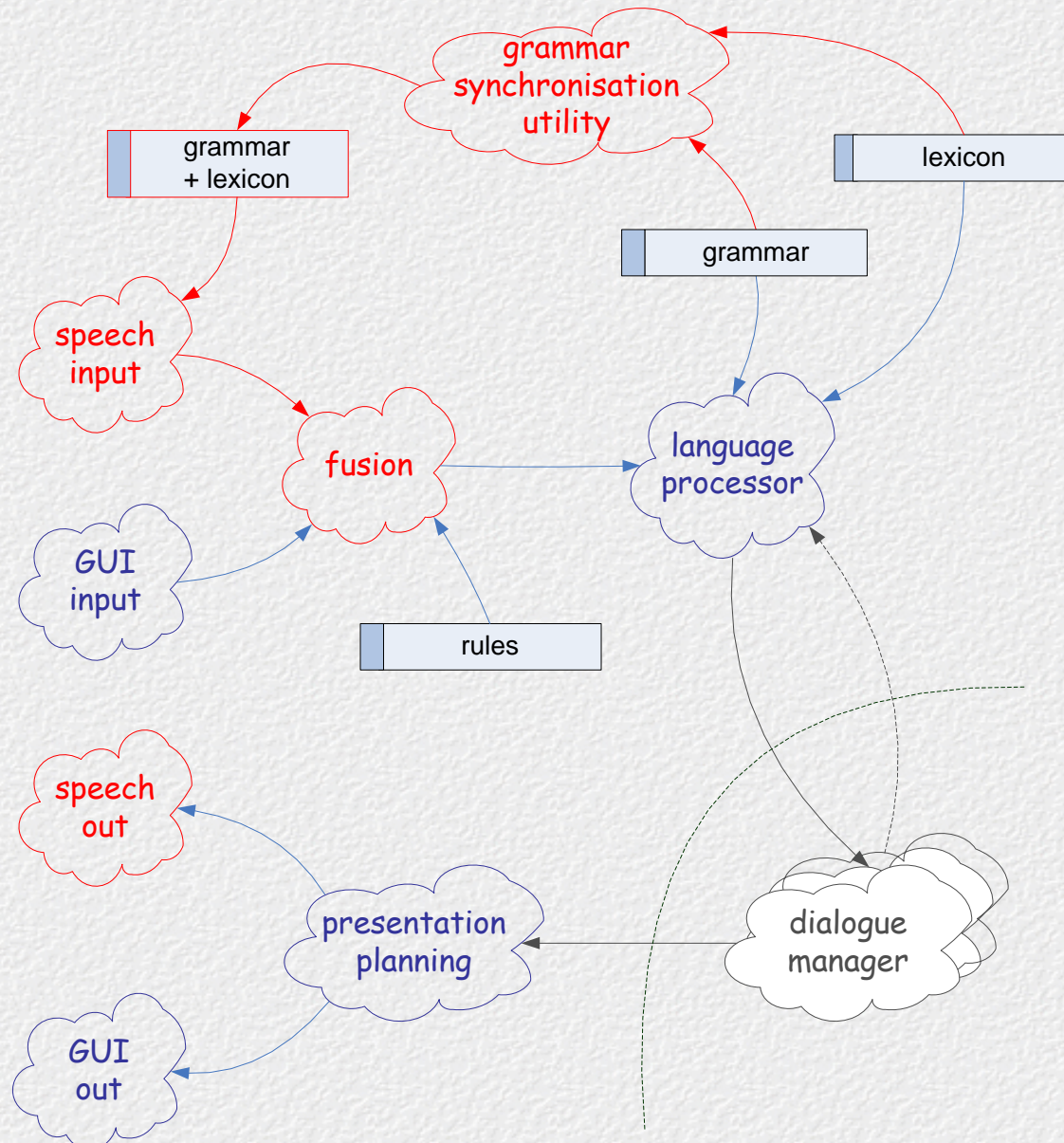
fusing rules

in brief...

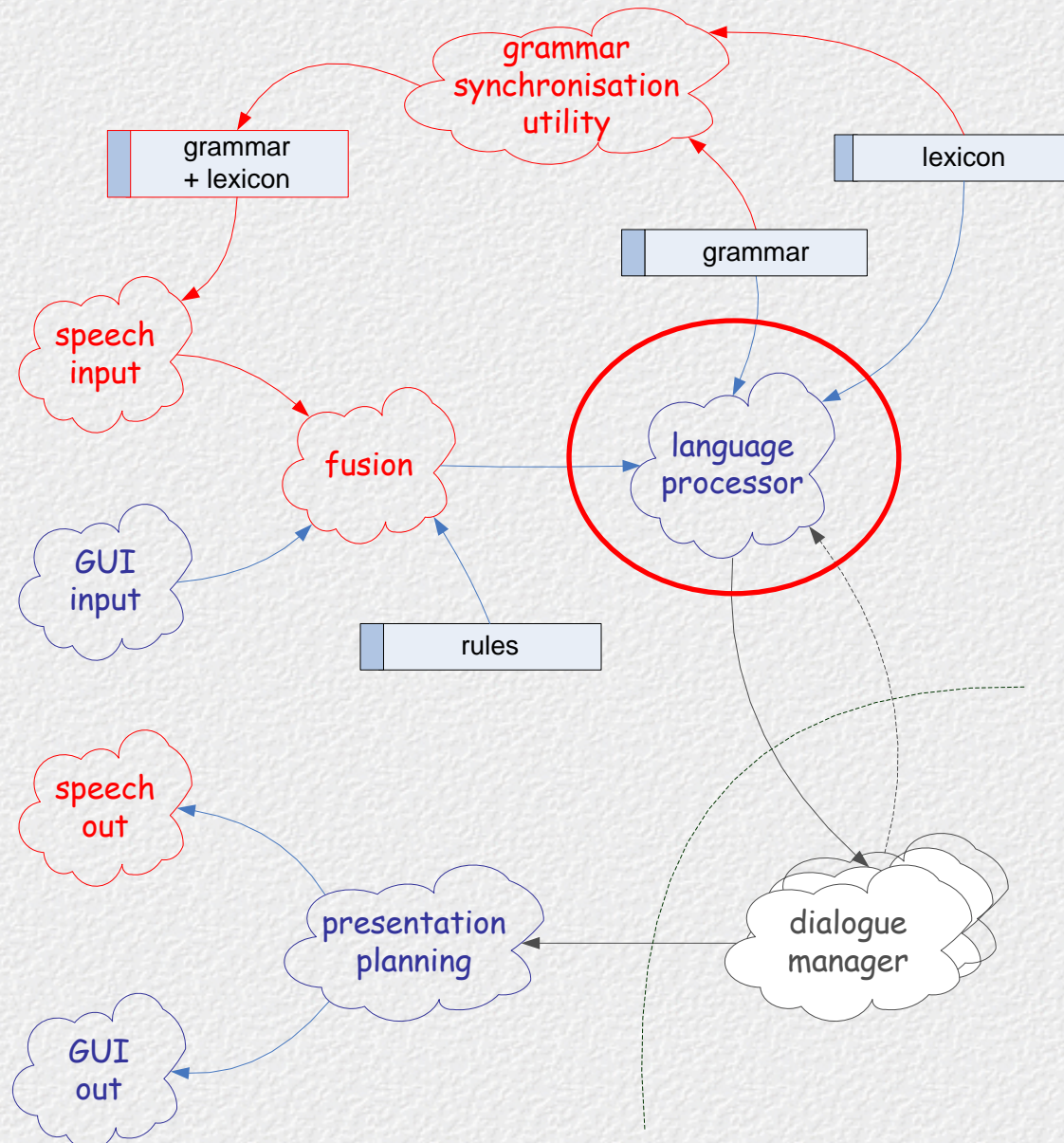
```
(rule that+click
  (fuse (speech-in that) (gui-in $obj)
    => (fused (gui-in type) (gui-in target))))
```

```
(rule there+click
  (fuse (speech-in there) (gui-in $place)
    => (fused (gui-in type) (gui-in target))))
```

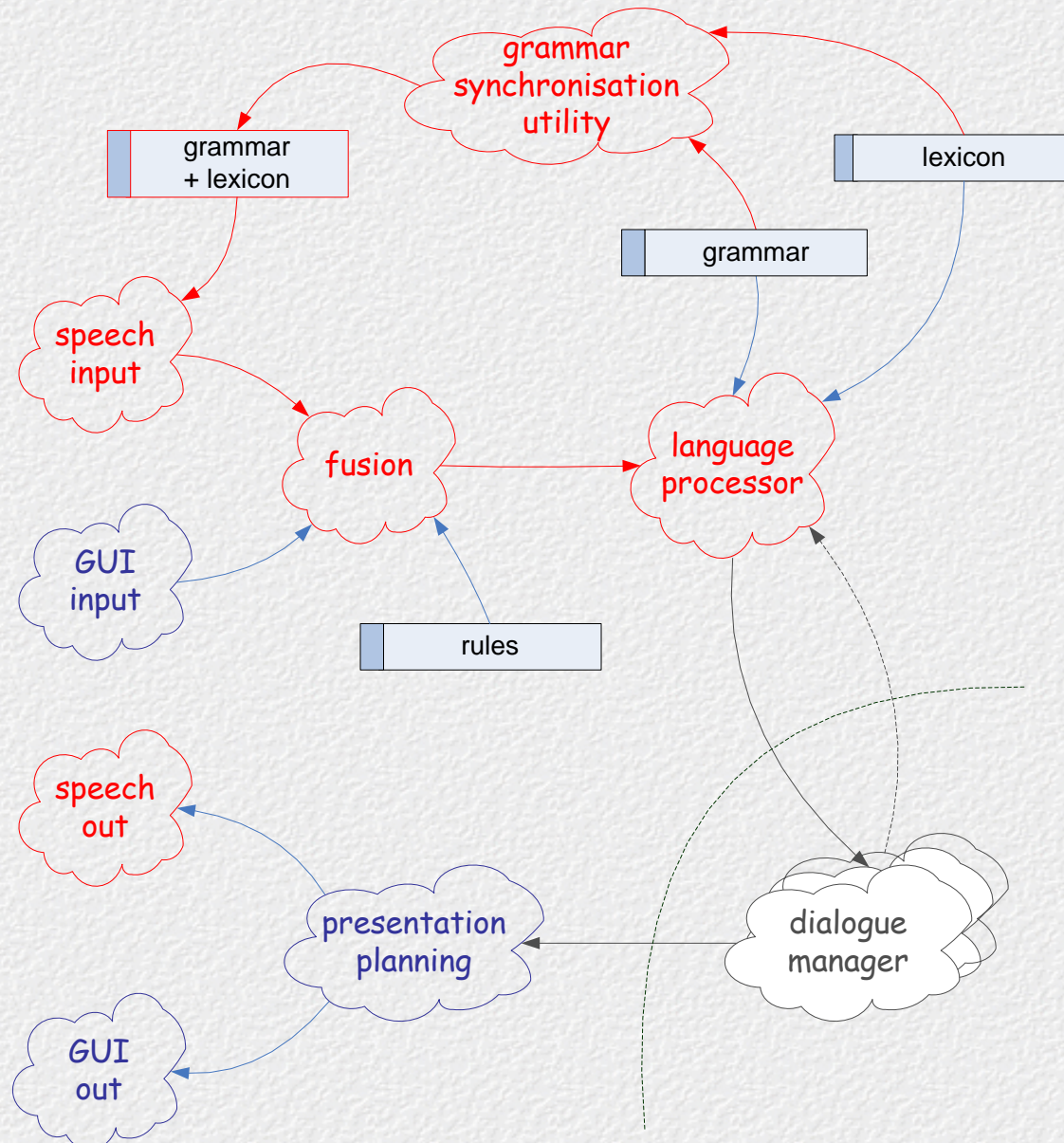
a framework architecture



a framework architecture



a framework architecture



results to date...

3 groups: 3-4 students each
level 2 completed
12 week internship

1. no tool support - v.limited results
2. MAS platform - some speech I/O
hacked fusion
no MM integration
3. MM framework - prototype completed

