A neural network sentence generator: Results from Māori, Slovak and English

Alistair Knott[1], Martin Takac[1,2]

[1]Department of Computer Science, University of Otago
[2]Centre for Cognitive Science, Comenius University, Slovakia

Our high-level research question

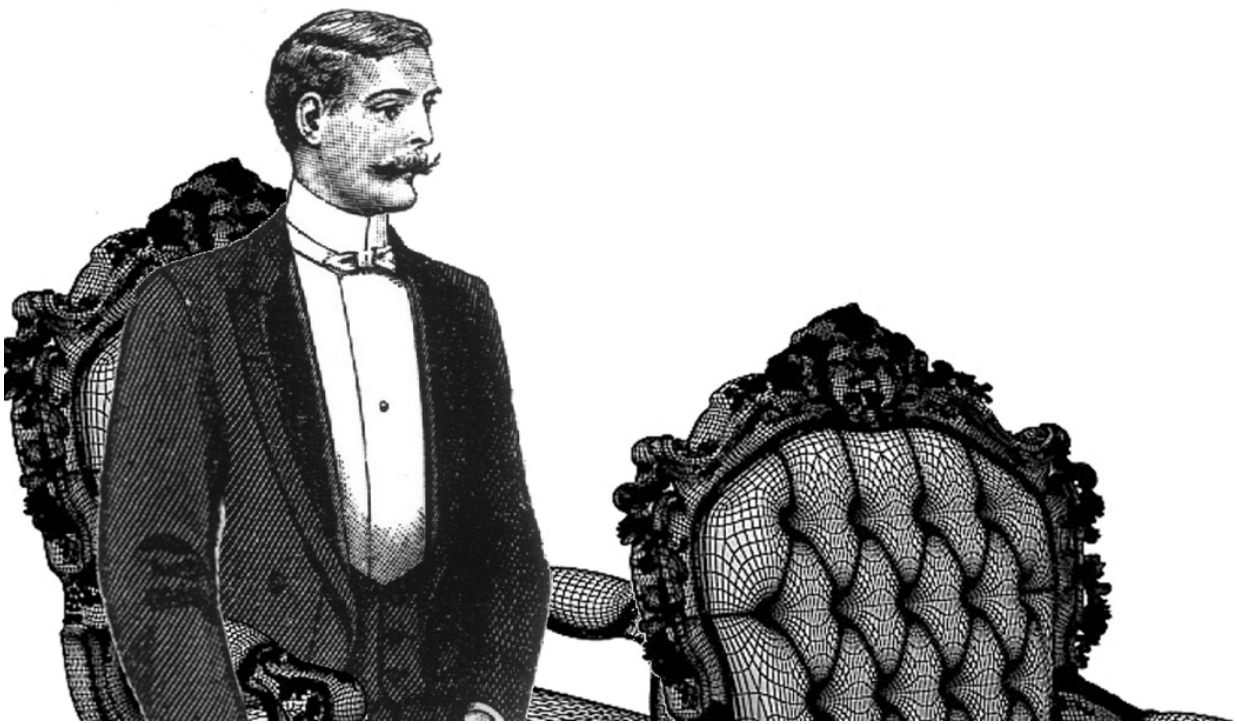
Our high-level research question

How can we talk about the things we see and do in the world?

Our high-level research question

How does language interface with the sensorimotor (SM) system?

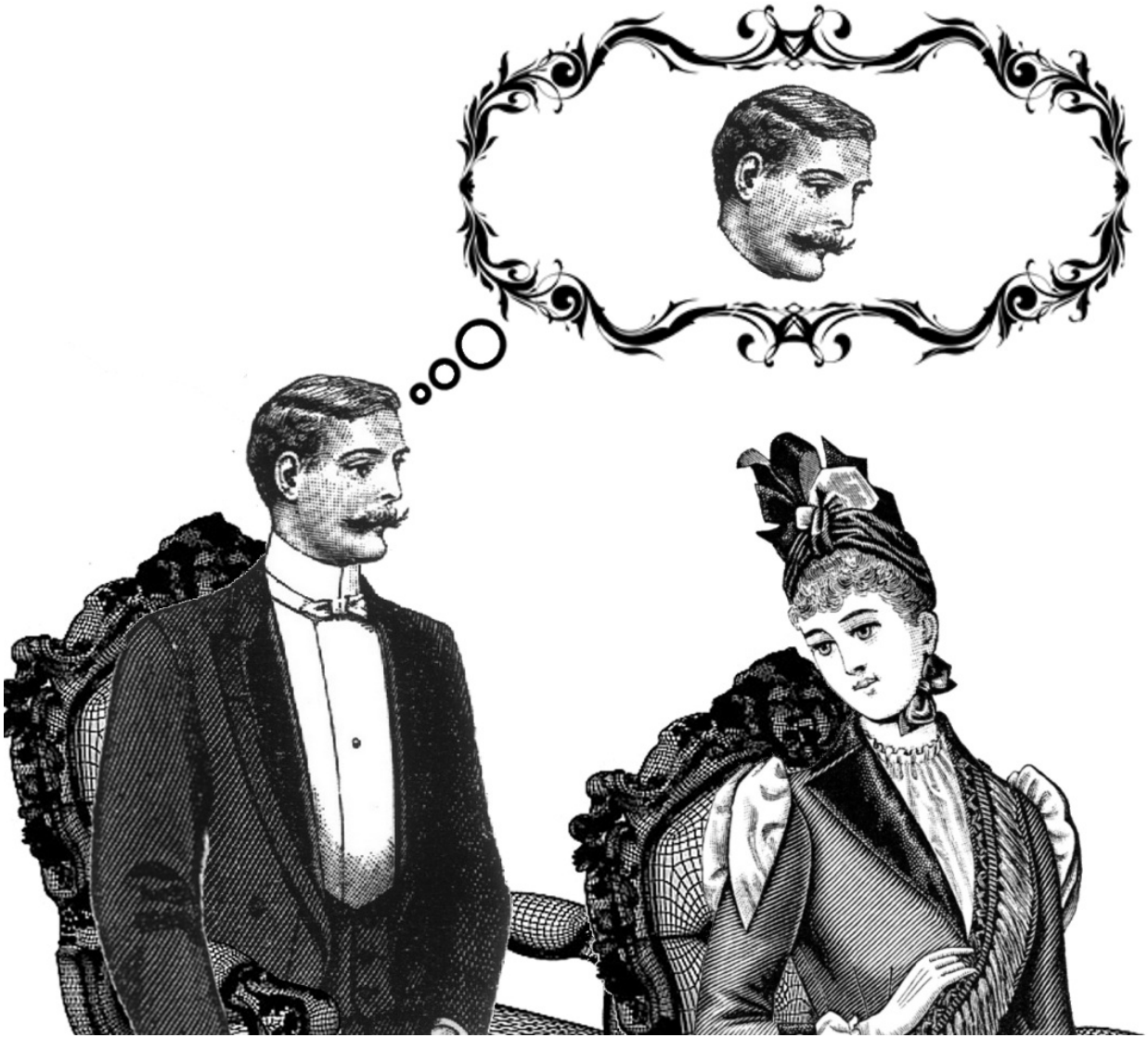
An example SM process: AGENT grabs TARGET



An example SM process: AGENT grabs TARGET

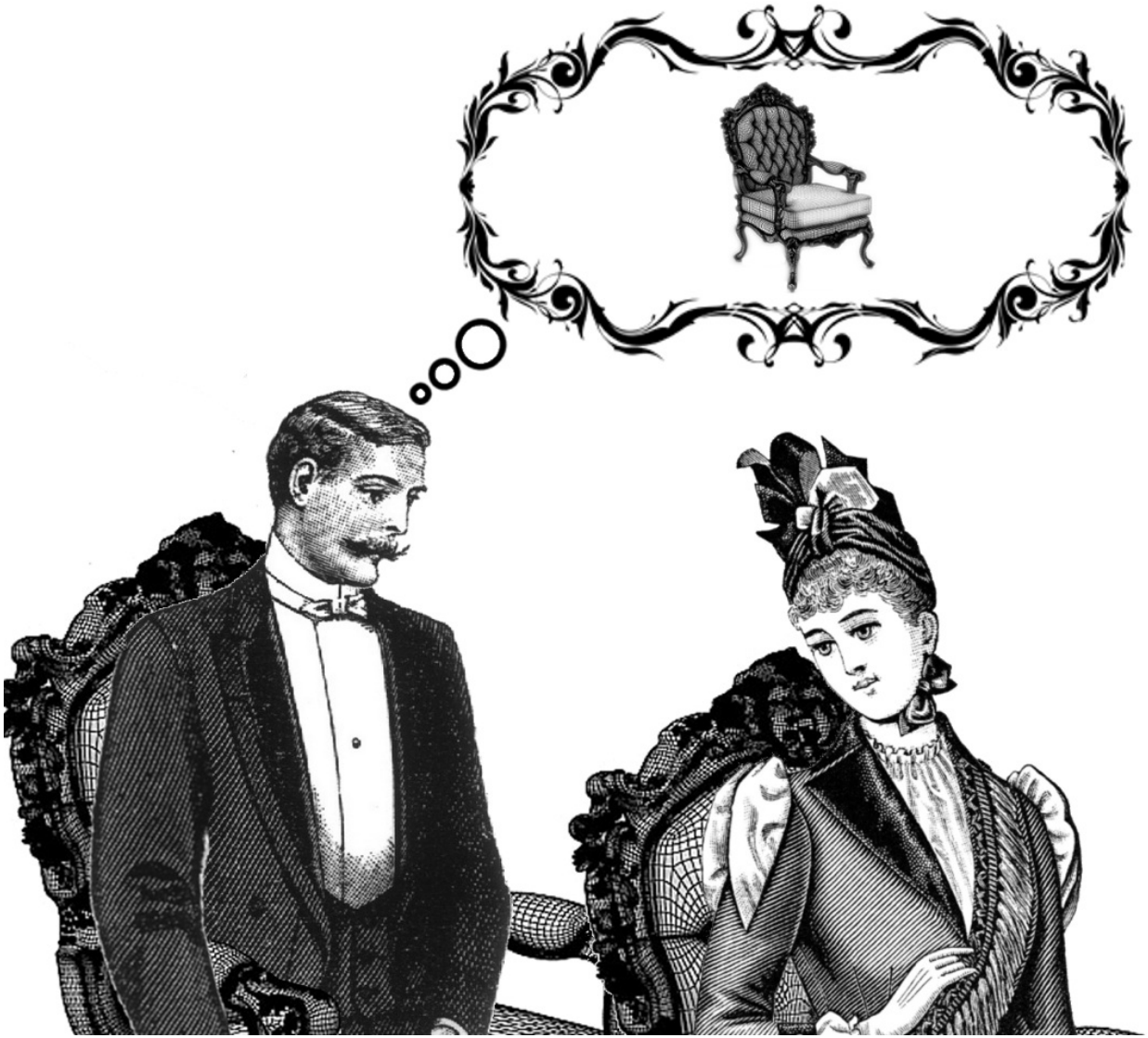
An example SM process: AGENT grabs TARGET

An example SM process: AGENT grabs TARGET



An example SM process: AGENT grabs TARGET

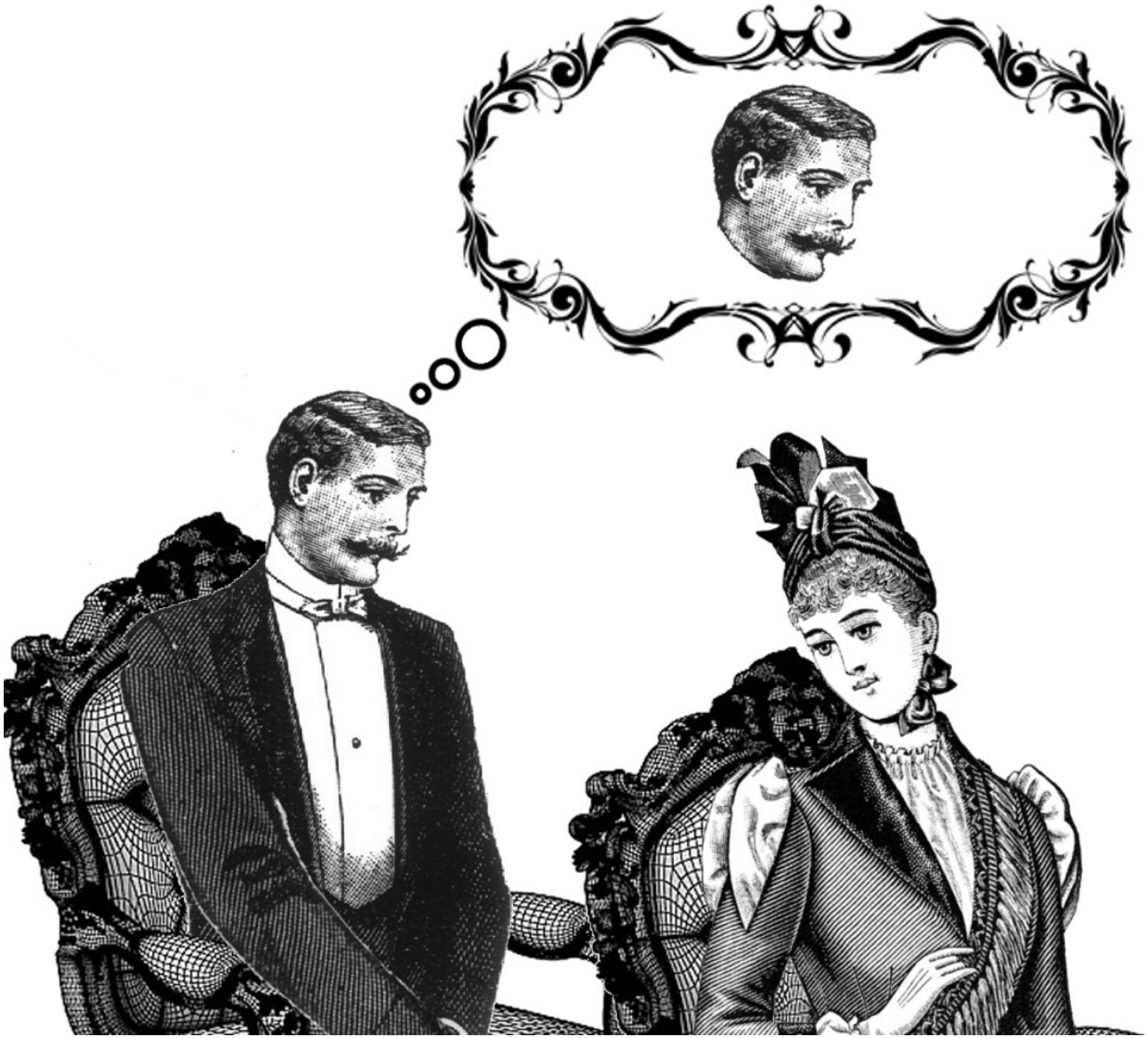
An example SM process: AGENT grabs TARGET



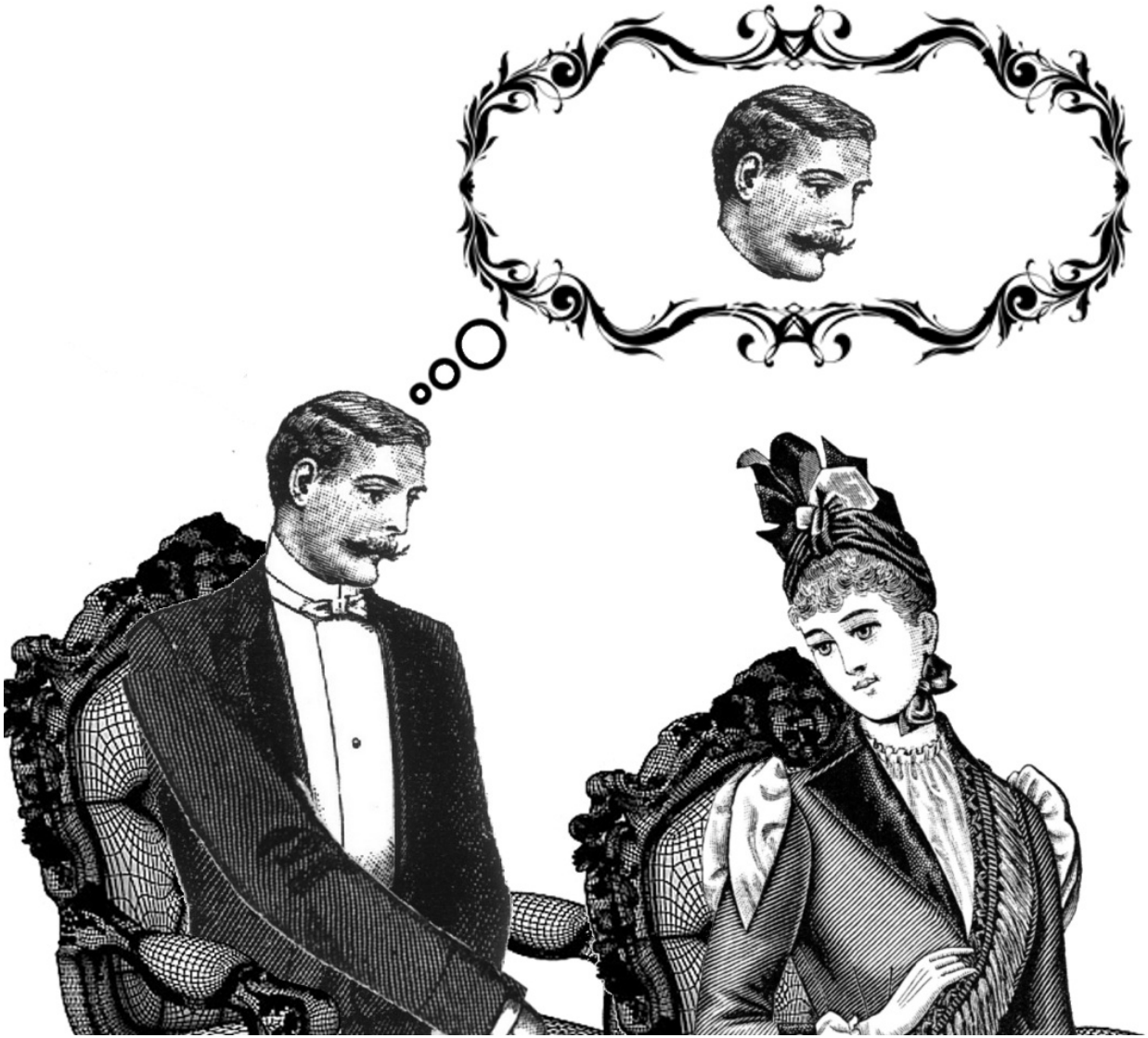
An example SM process: AGENT grabs TARGET

An example SM process: AGENT grabs TARGET

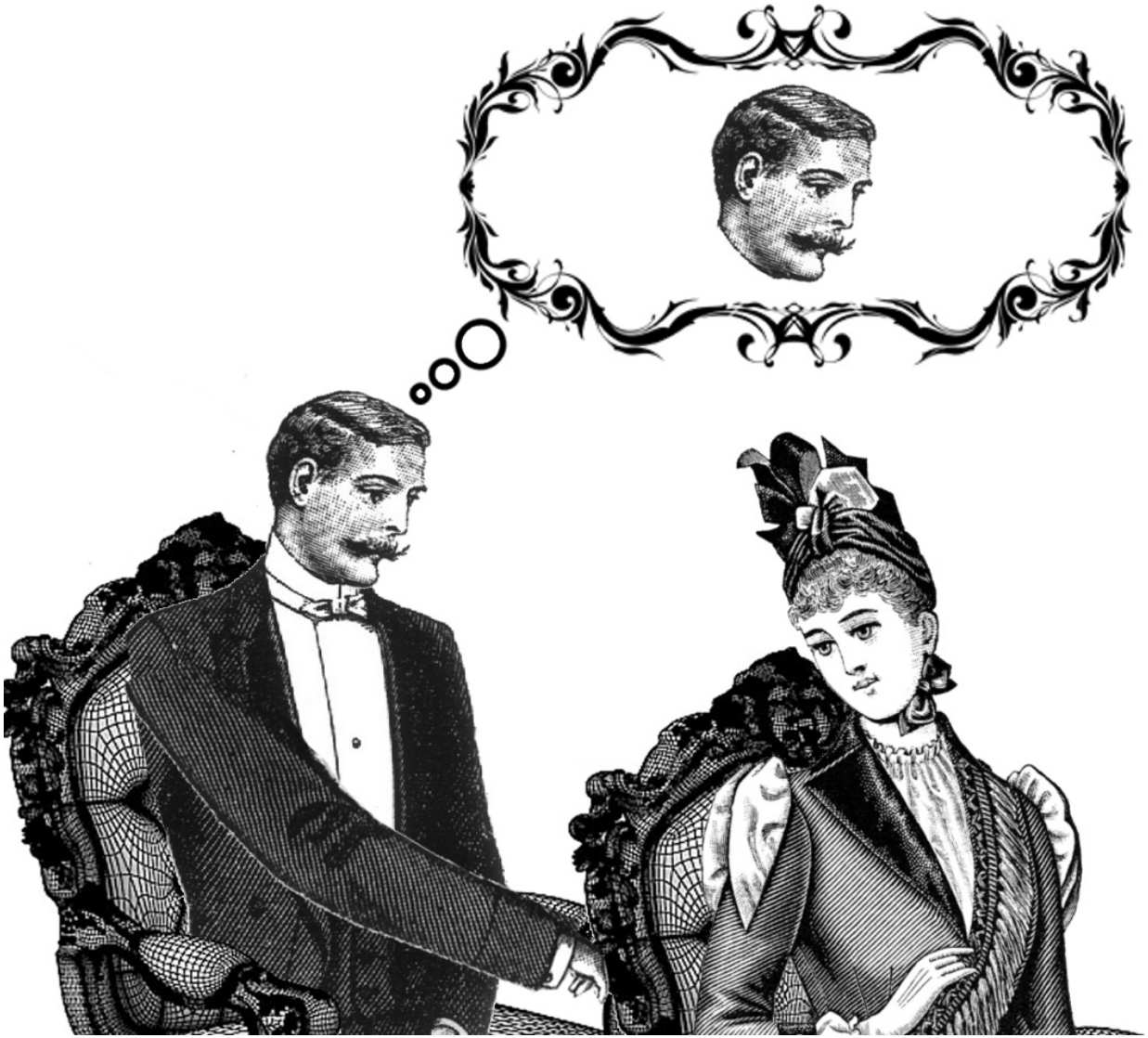
An example SM process: AGENT grabs TARGET



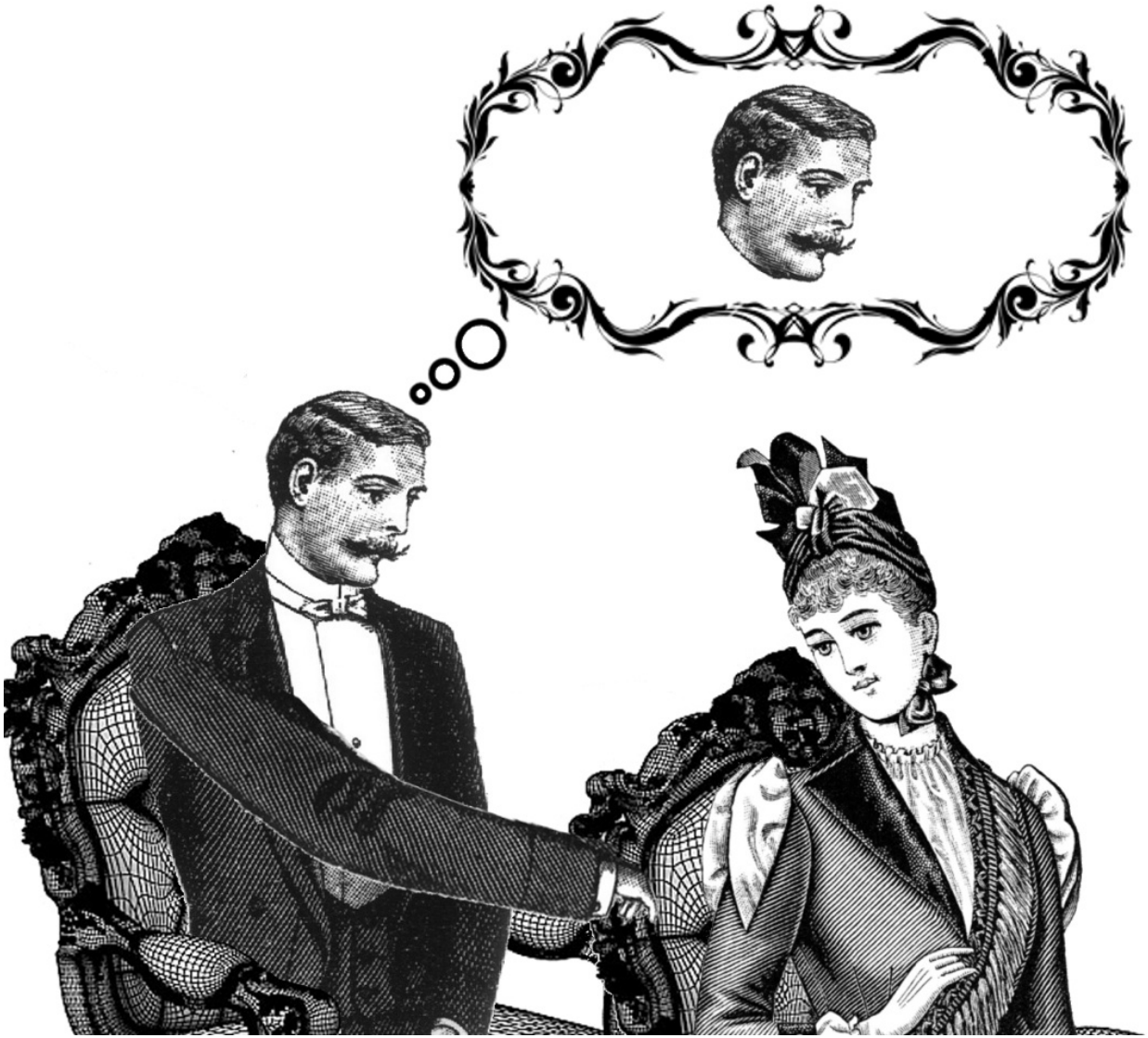
An example SM process: AGENT grabs TARGET



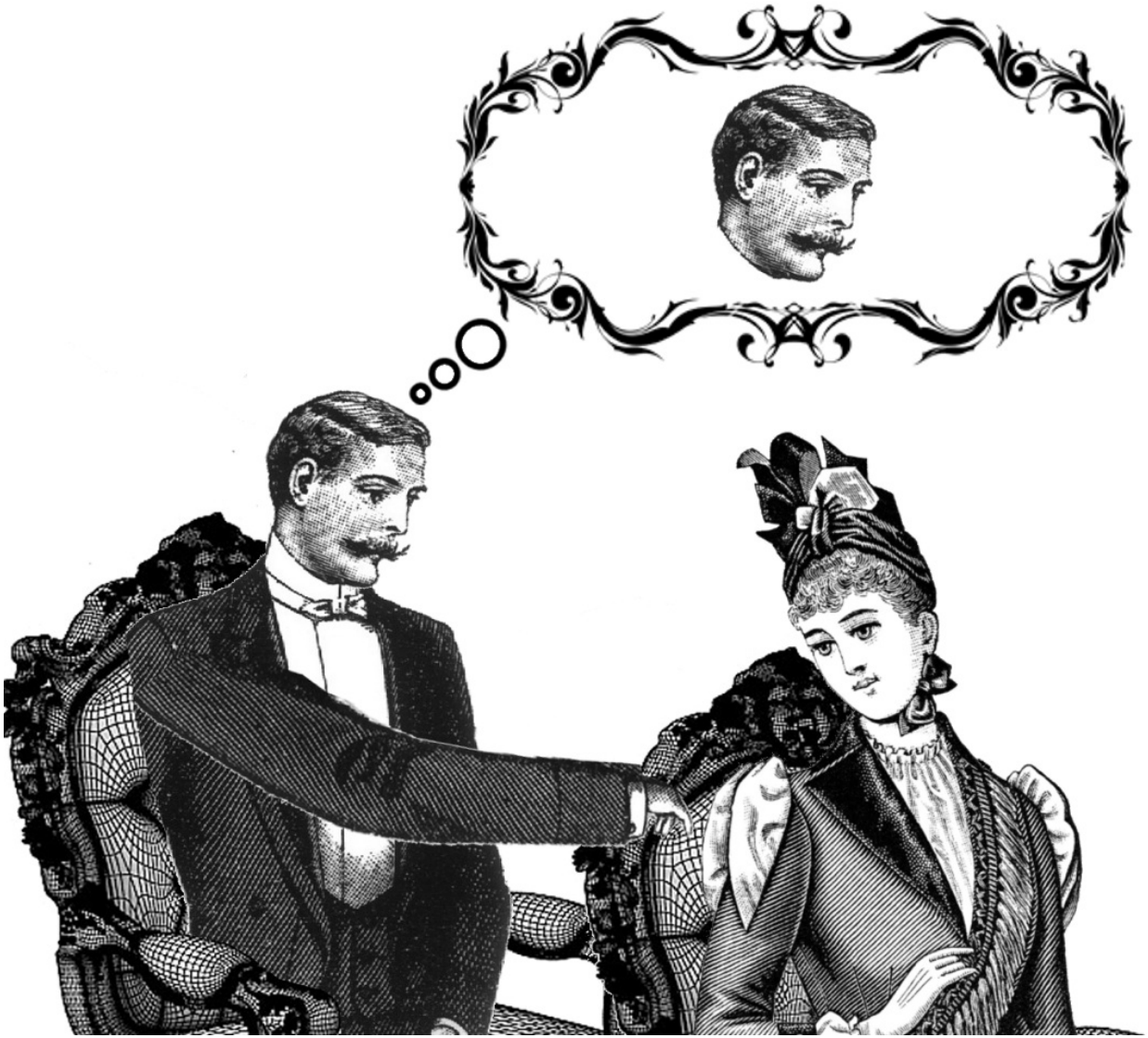
An example SM process: AGENT grabs TARGET



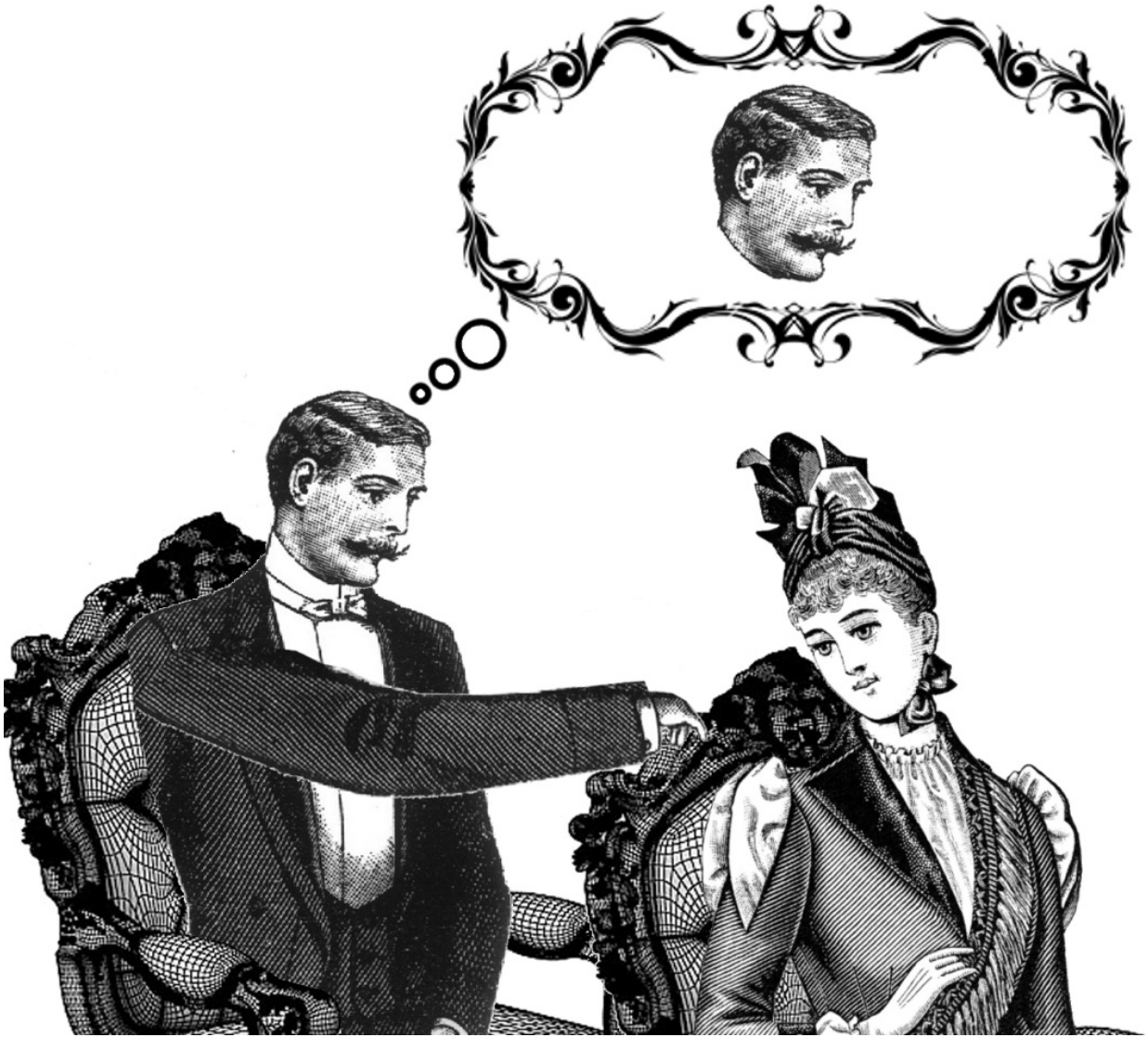
An example SM process: AGENT grabs TARGET



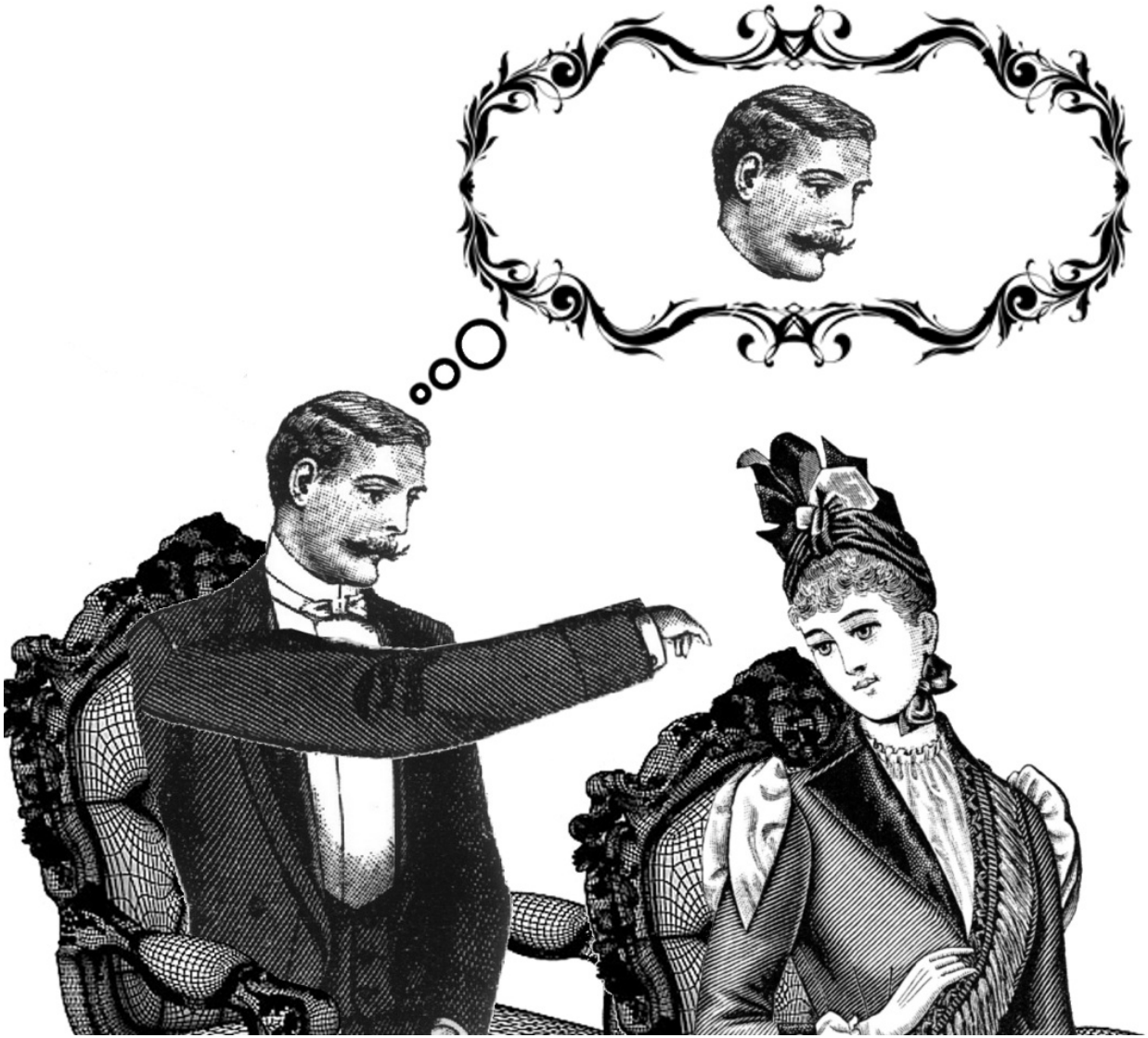
An example SM process: AGENT grabs TARGET



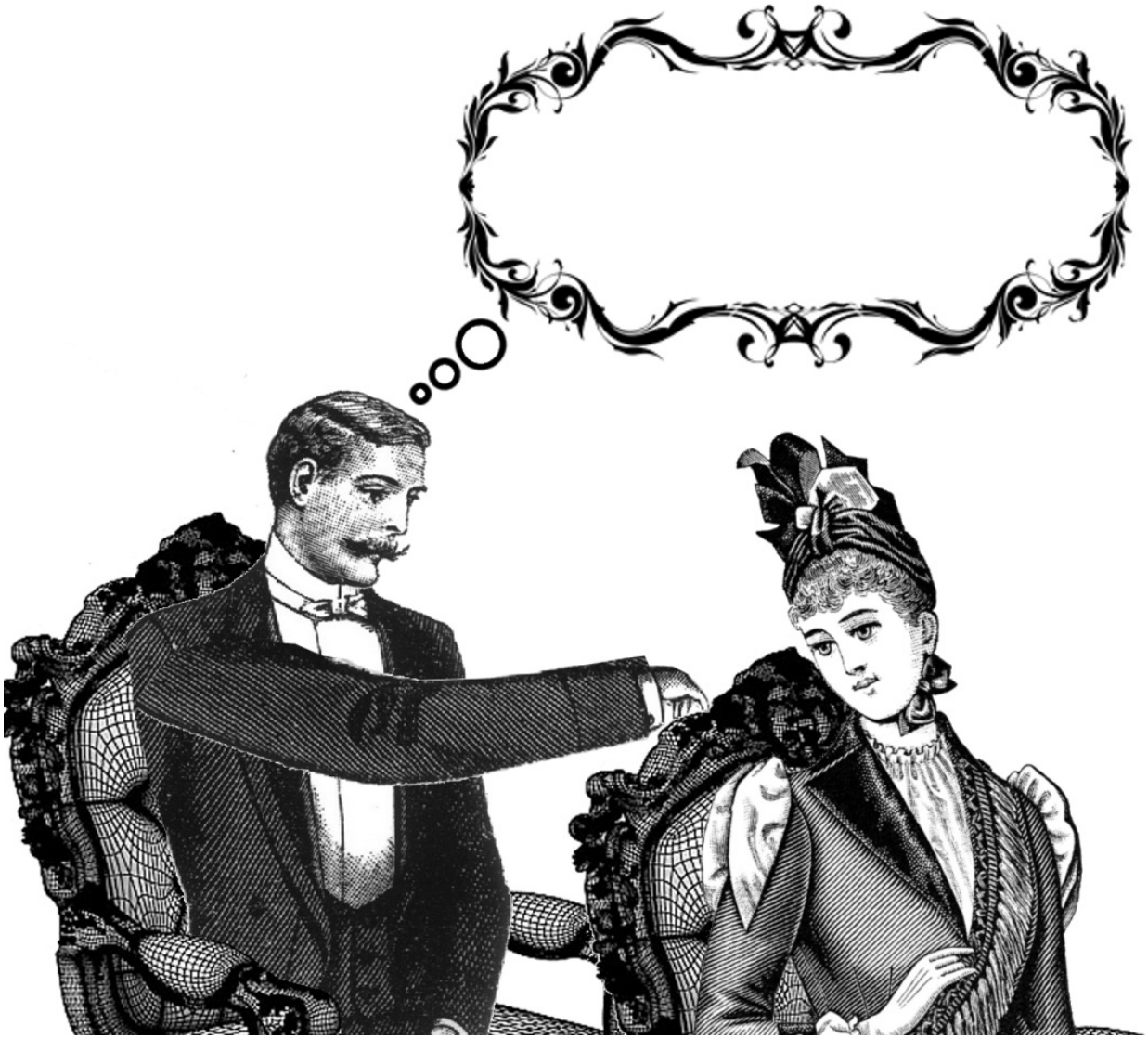
An example SM process: AGENT grabs TARGET



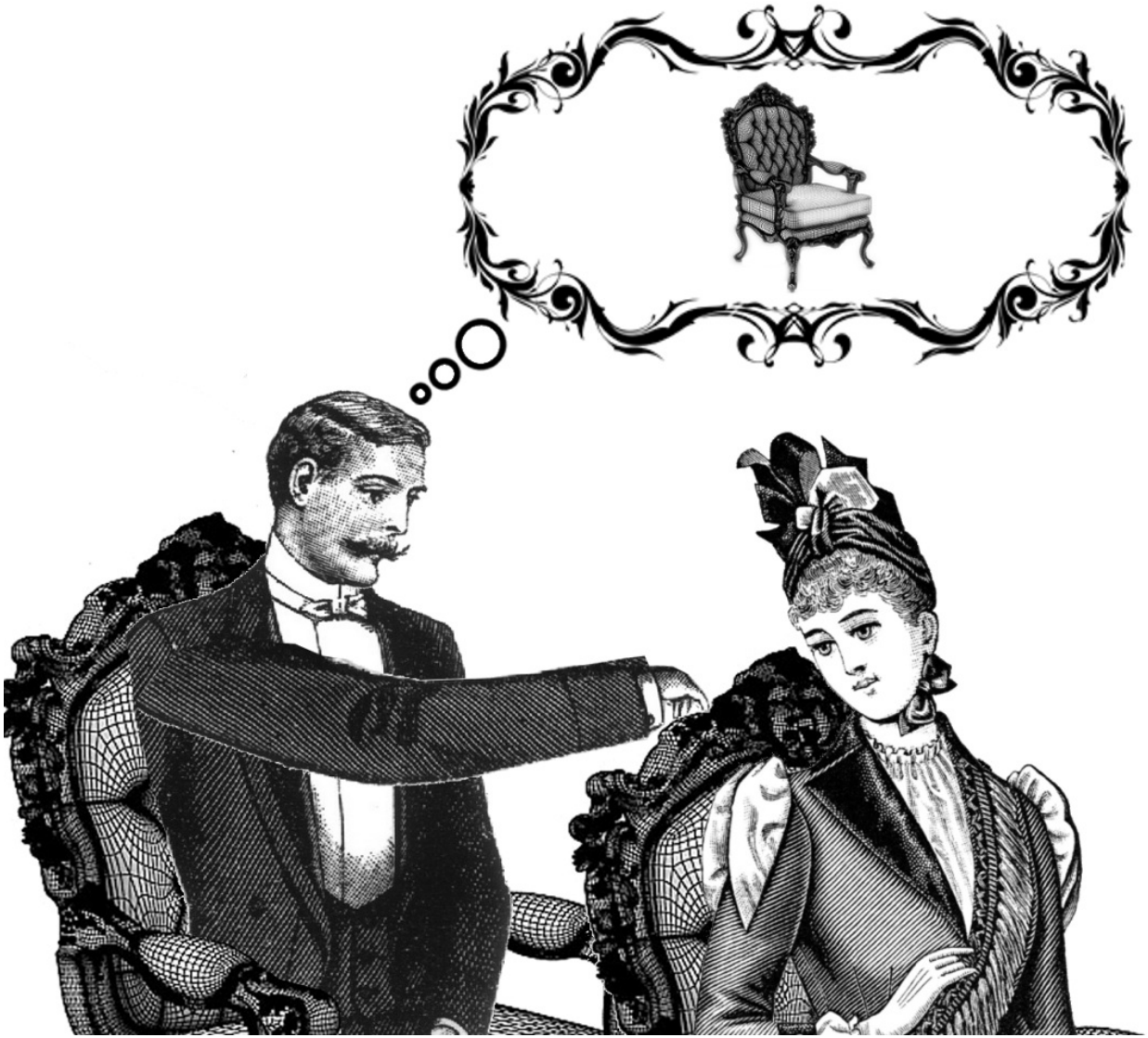
An example SM process: AGENT grabs TARGET



An example SM process: AGENT grabs TARGET



An example SM process: AGENT grabs TARGET



A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1		

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2		

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2	*attend_target*	

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2	*attend_target*	*target*

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2	*attend_target*	*target*
c_3		

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2	*attend_target*	*target*
c_3	*action*	

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2	*attend_target*	*target*
c_3	*action*	*agent*

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2	*attend_target*	*target*
c_3	*action*	*agent*
c_4		

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_1	*attend_agent*	*agent*
c_2	*attend_target*	*target*
c_3	*action*	*agent*
c_4		*target*

A model of SM processing for transitive actions

In our model, experiencing any concrete episode involves a *sequence* of SM operations, interleaved with reafferent sensory signals.

E.g. for a transitive episode ⟨AGENT ACTION TARGET⟩:

Context	SM operation	Reafferent signal
c_0	*select_time*	*time*
c_1	*attend_agent*	*agent*
c_2	*attend_target*	*target*
c_3	*action*	*agent*
c_4		*target*

A model of the interface to language

A model of the interface to language

We propose that experienced episodes are stored in semantic working memory (WM) as *prepared SM routines*, that can be actively rehearsed (executed or simulated).

A model of the interface to language

We propose that experienced episodes are stored in semantic working memory (WM) as *prepared SM routines*, that can be actively rehearsed (executed or simulated).

The planned SM operations are active *in parallel* in the planning area of the brain. So during rehearsal there's a mixture of 'sustained' and 'transient' SM signals.

A model of the interface to language

We propose that experienced episodes are stored in semantic working memory (WM) as *prepared SM routines*, that can be actively rehearsed (executed or simulated).

The planned SM operations are active *in parallel* in the planning area of the brain. So during rehearsal there's a mixture of 'sustained' and 'transient' SM signals.

In our model, an agent produces a sentence by rehearsing a planned SM routine, in a special mode where SM signals can activate words.

A rehearsed transitive episode

Here are the brain areas involved in experiencing a transitive episode.

	Transient signals		
		SM operation	Reafferent signal
	c_0	*select_time*	*time*
	c_1	*attend_agent*	*agent*
	c_2	*attend_target*	*target*
	c_3	*action*	*agent*
	c_4		*target*

A rehearsed transitive episode

Here are the brain areas involved in *rehearsing* this experience.

Sustained signals	Transient signals		
		SM operation	Reafferent signal
$plan_{select\_time/attend\_agent/attend\_target/action}$	c_0	*select_time*	*time*
$plan_{select\_time/attend\_agent/attend\_target/action}$	c_1	*attend_agent*	*agent*
$plan_{select\_time/attend\_agent/attend\_target/action}$	c_2	*attend_target*	*target*
$plan_{select\_time/attend\_agent/attend\_target/action}$	c_3	*action*	*agent*
	c_4		*target*

A rehearsed transitive episode

In our model, only some of these areas interface with phonology.

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			*agent*
select_time / *attend_agent* / *attend_target* / *action*			*target*
select_time / *attend_agent* / *attend_target* / *action*			*agent*
			target

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
			agent

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
			target

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
			agent

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			

A rehearsed transitive episode

Rehearsal produces a *sequence* of SM signals:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
			target

A rehearsed transitive episode

There are two opportunities to pronounce the agent:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
			agent
			agent

A rehearsed transitive episode

There are two opportunities to pronounce the target:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
			target
			target

A rehearsed transitive episode

And four opportunities to pronounce the planned SM routine.

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			

A rehearsed transitive episode

An infant has to learn which opportunities to take.

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			*agent*
select_time / *attend_agent* / *attend_target* / *action*			*target*
select_time / *attend_agent* / *attend_target* / *action*			*agent*
			target

A rehearsed transitive episode

Our hypothesis about the language-SM interface:

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			*agent*
select_time / *attend_agent* / *attend_target* / *action*			*target*
select_time / *attend_agent* / *attend_target* / *action*			*agent*
			target

A rehearsed transitive episode

Our hypothesis about the language-SM interface:

- Syntactic heads in a clause are read from sustained signals. . .

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			*agent*
select_time / *attend_agent* / *attend_target* / *action*			*target*
select_time / *attend_agent* / *attend_target* / *action*			*agent*
			target

A rehearsed transitive episode

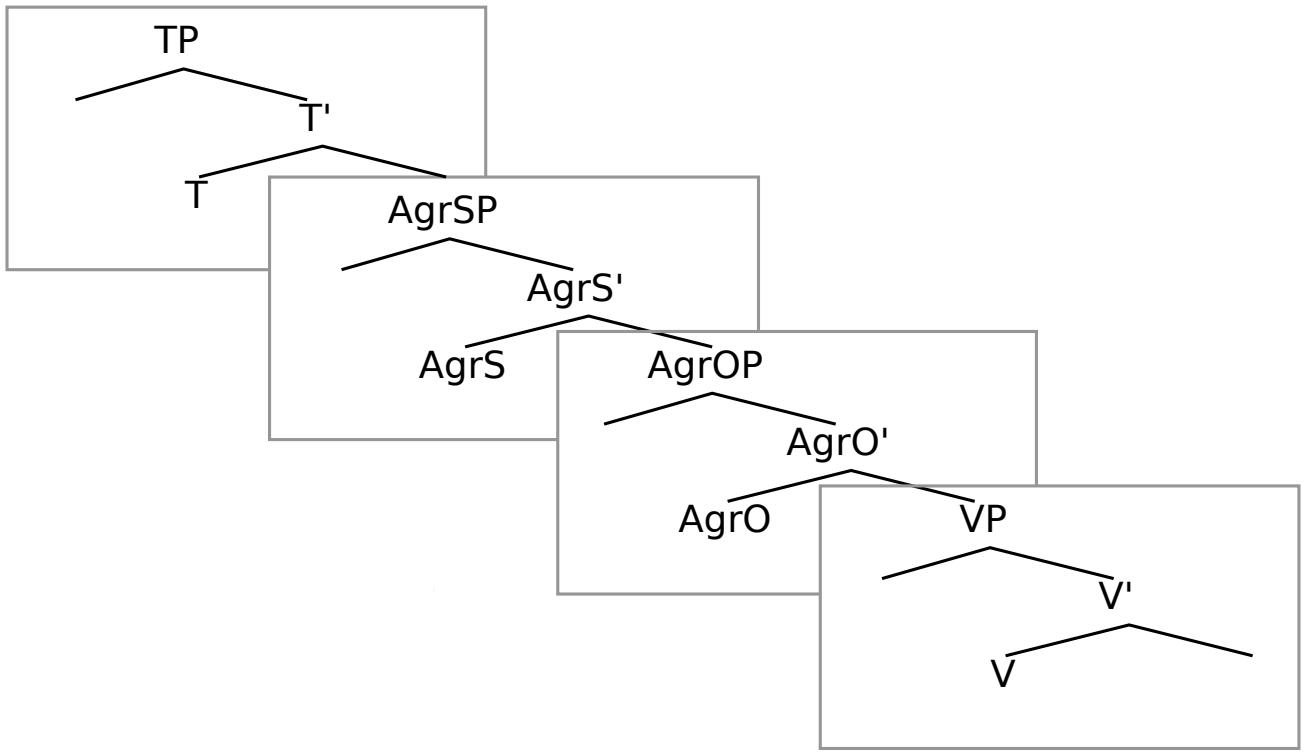
Our hypothesis about the language-SM interface:

- Syntactic heads in a clause are read from sustained signals. . .
- Specifiers—in particular DPs—are read from transient signals.

Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			*agent*
select_time / *attend_agent* / *attend_target* / *action*			*target*
select_time / *attend_agent* / *attend_target* / *action*			*agent*
			target

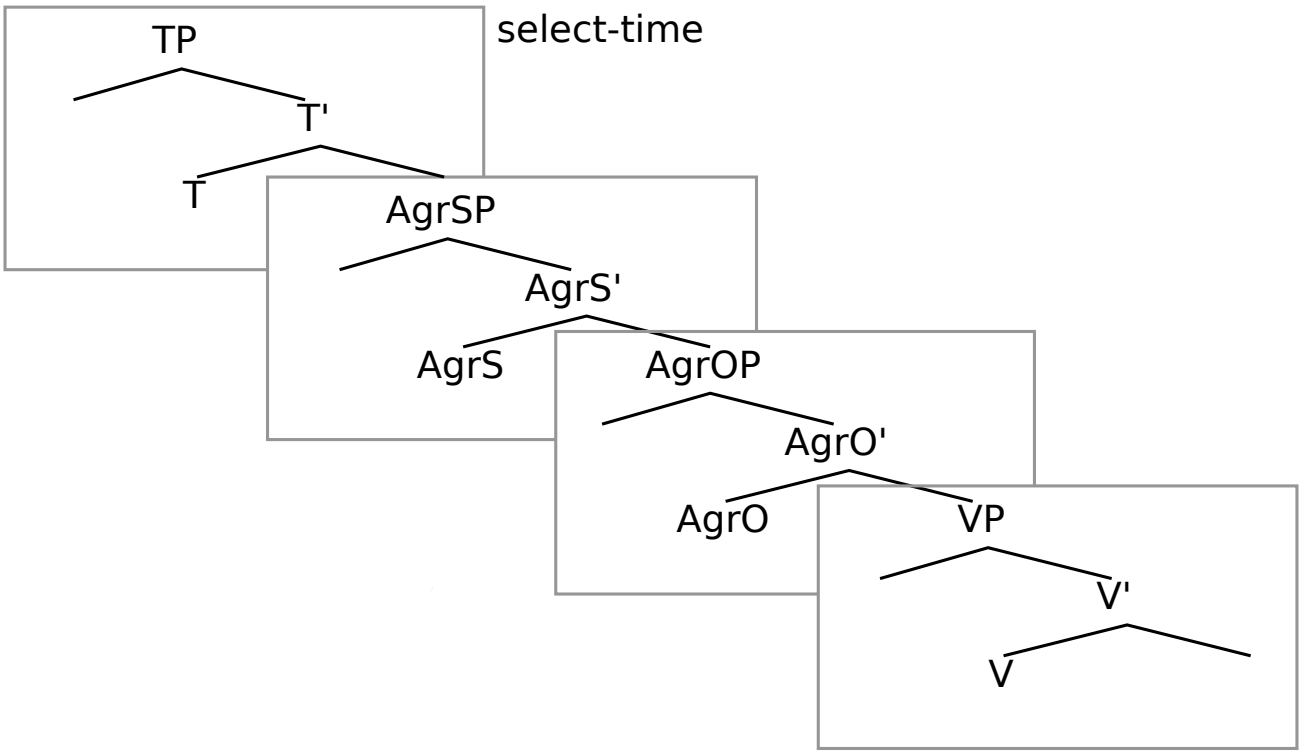
In syntactic terms

On this hypothesis, we can interpret Chomsykan ‘logical form’ (LF) structures in SM terms.



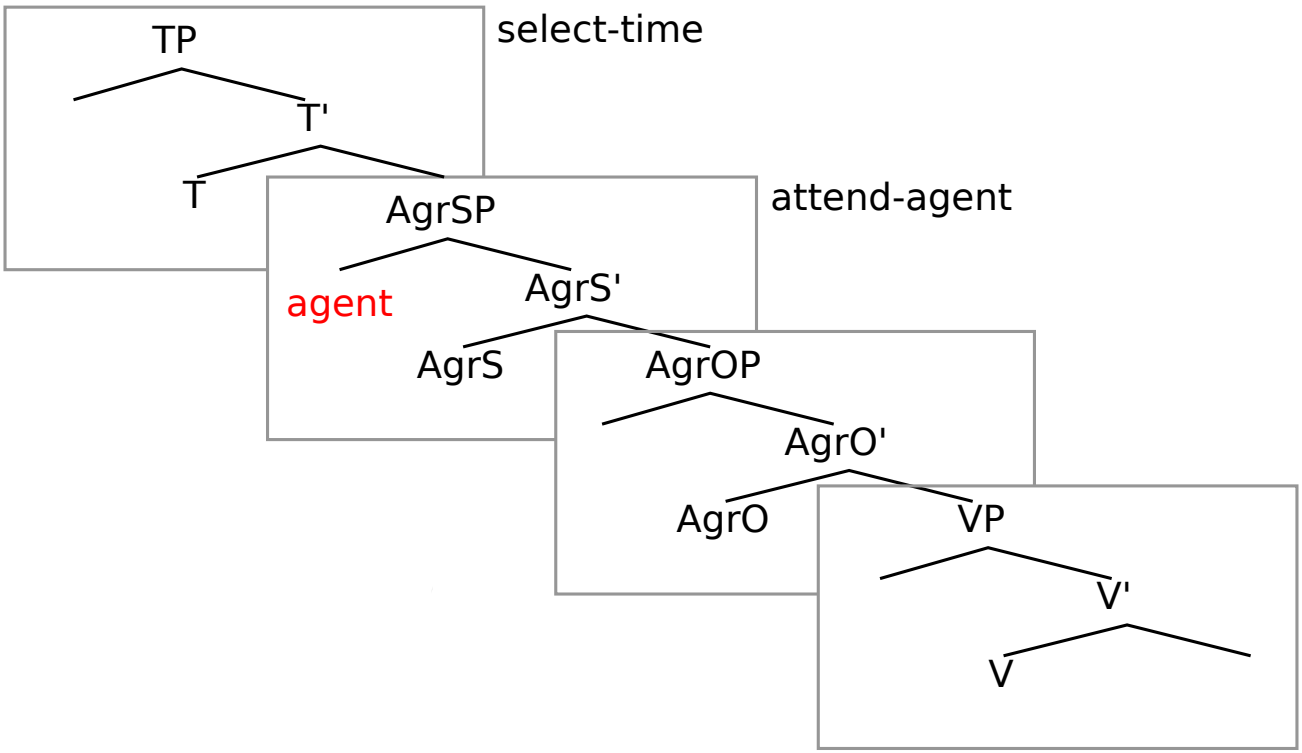
In syntactic terms

TP describes the operation of selecting a time (past or present)



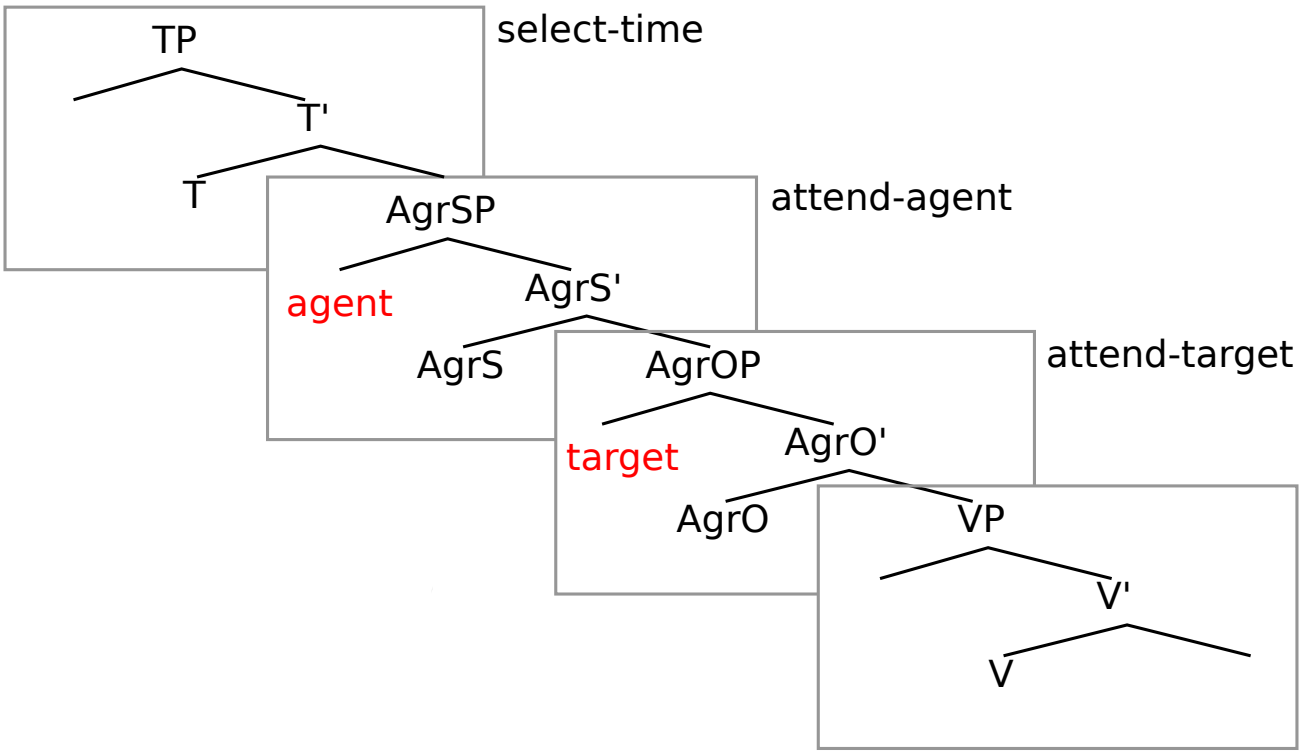
In syntactic terms

AgrSP describes the operation of attending to the agent



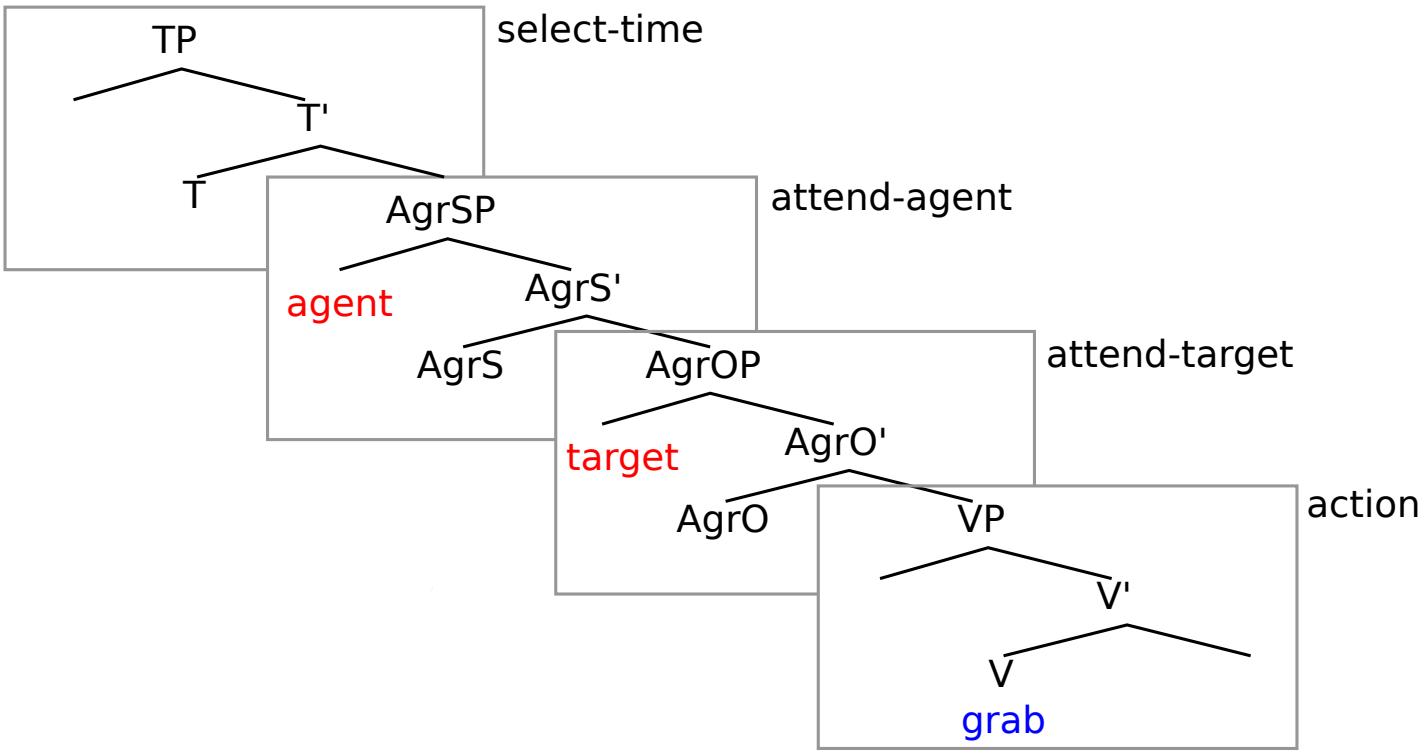
In syntactic terms

AgrOP describes the operation of attending to the target



In syntactic terms

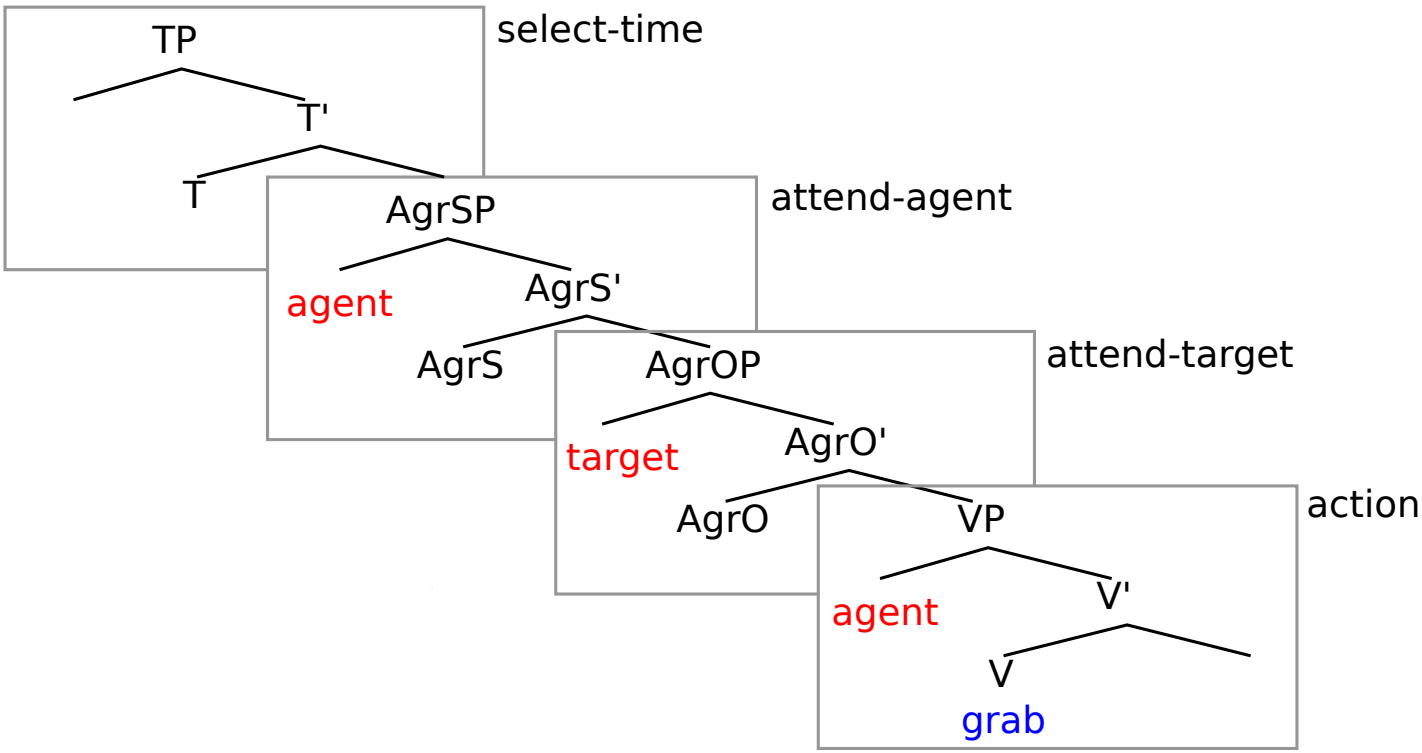
VP describes the motor operation. . .



In syntactic terms

VP describes the motor operation. . .

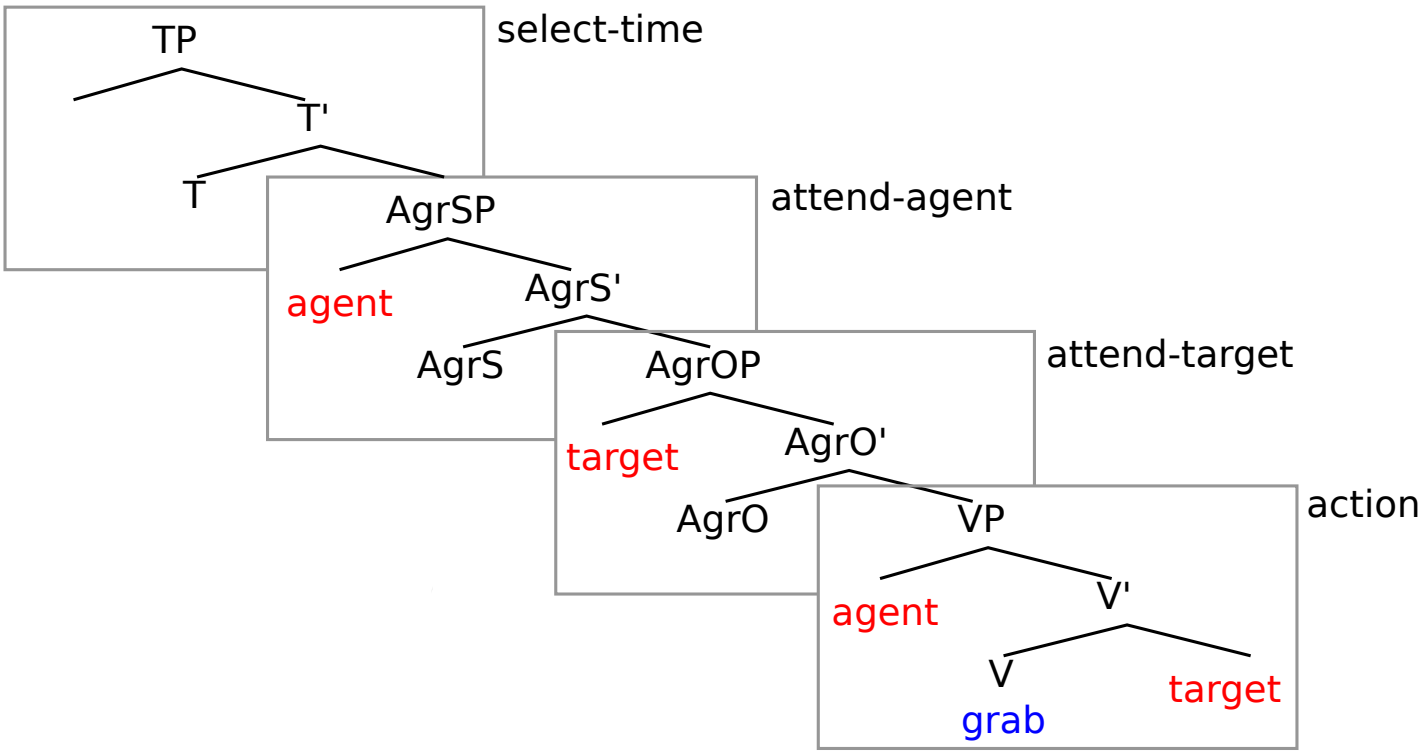
- with reattention to the agent. . .



In syntactic terms

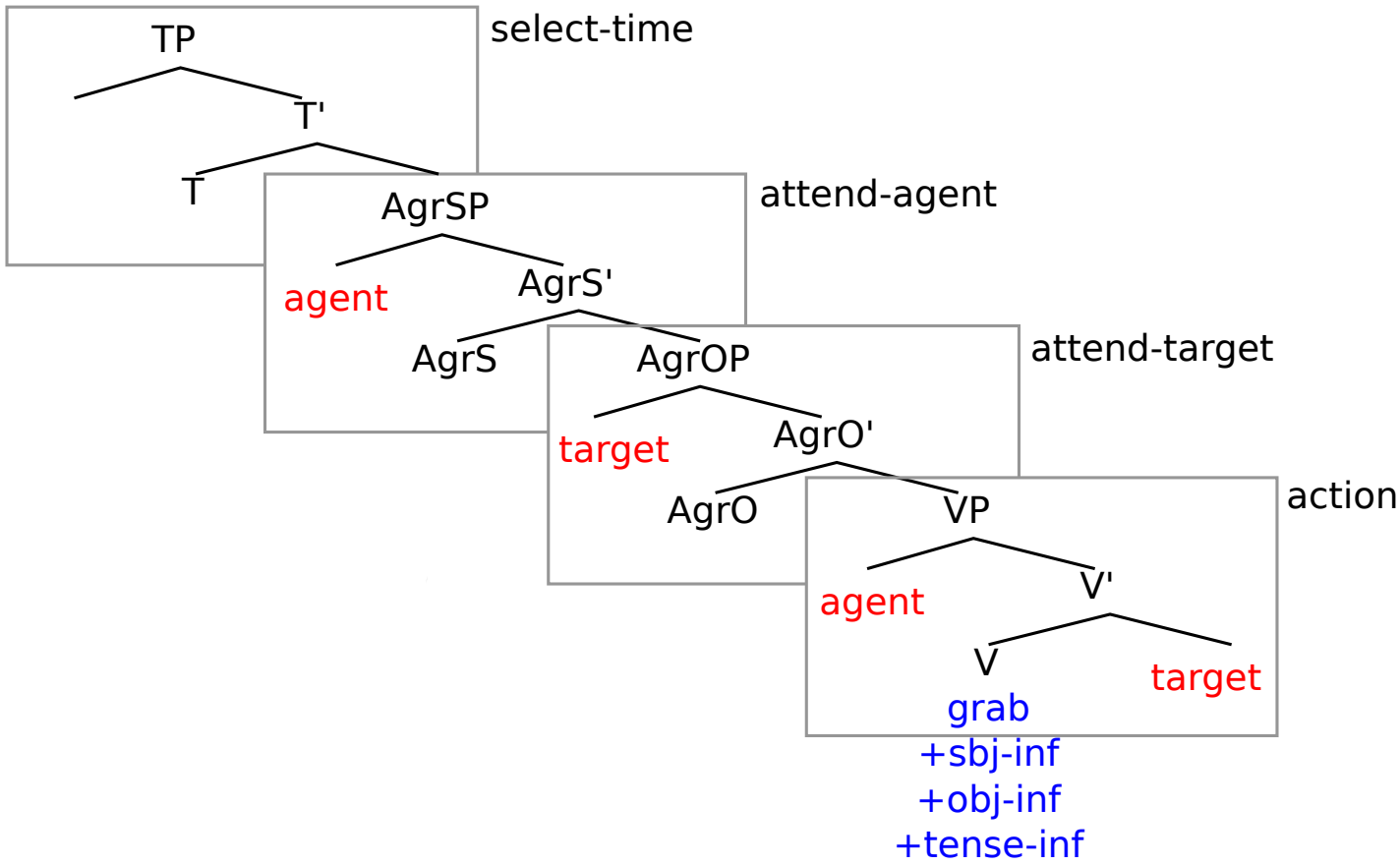
VP describes the motor operation. . .

- with reattention to the agent. . . and to the target.



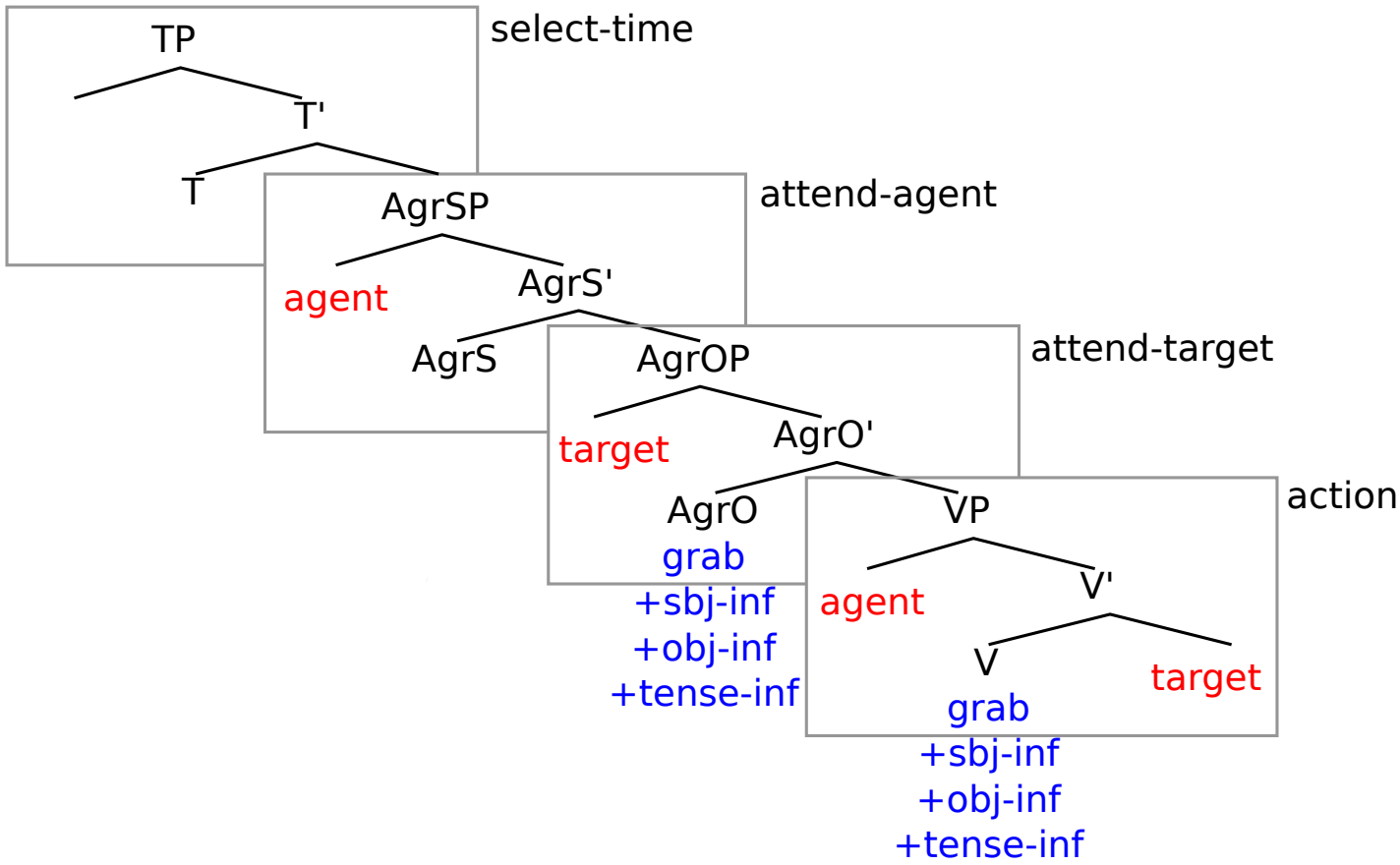
In syntactic terms

The verb can carry inflections signalling agent/target properties, and also tense. . .



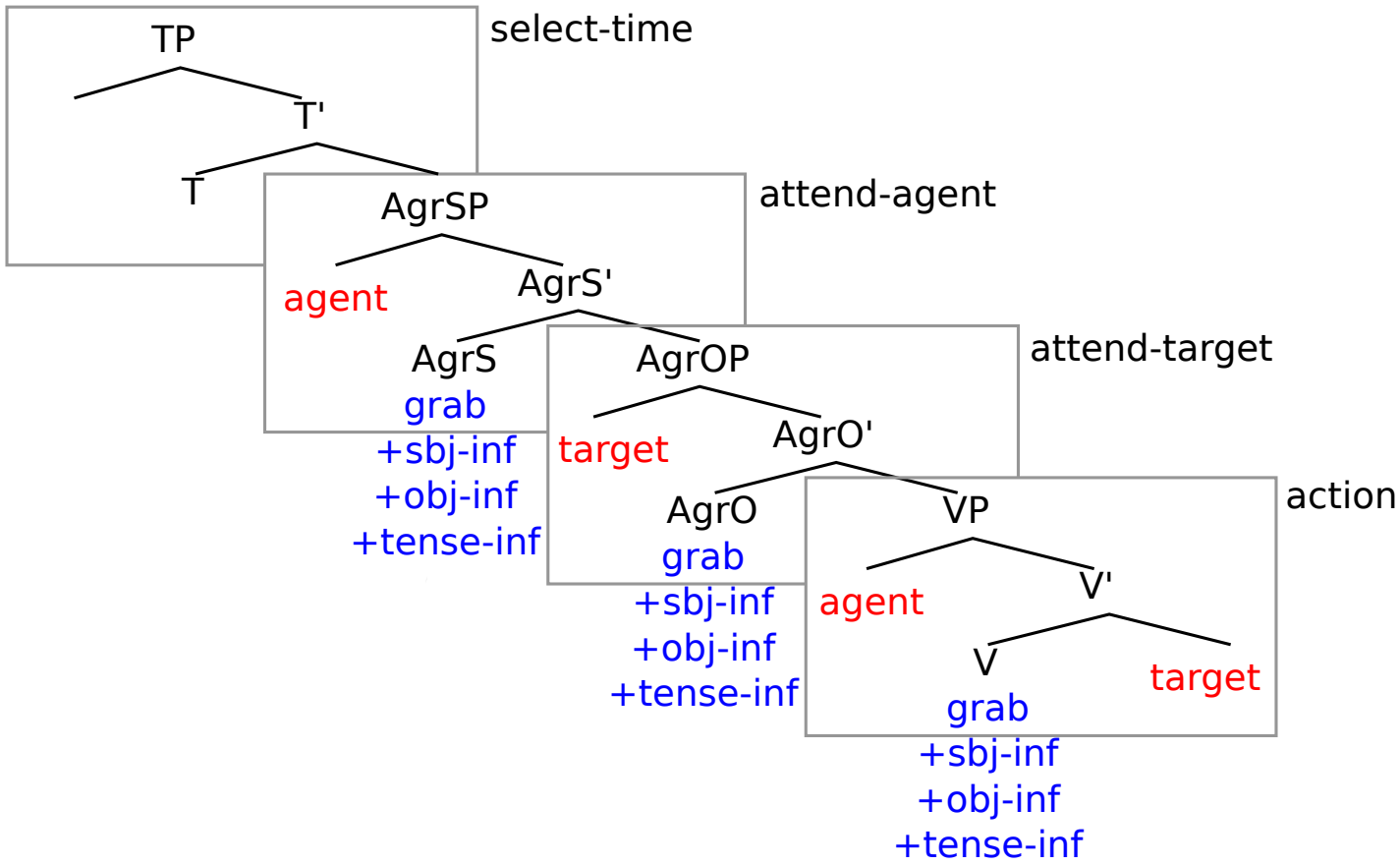
In syntactic terms

If heads are read from the planning medium, we can read out information about *each* SM operation at *each* head position.



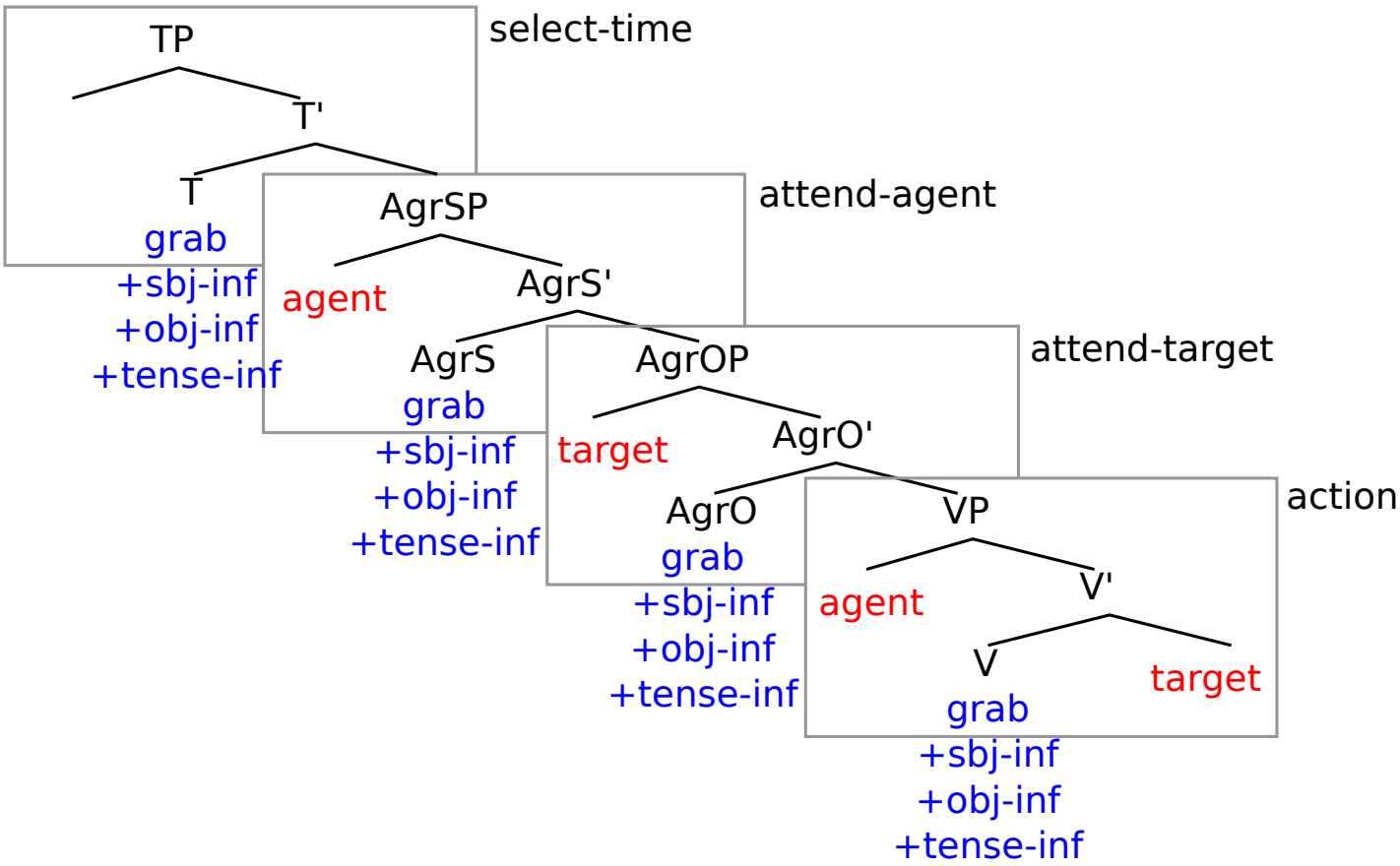
In syntactic terms

If heads are read from the planning medium, we can read out information about *each* SM operation at *each* head position.

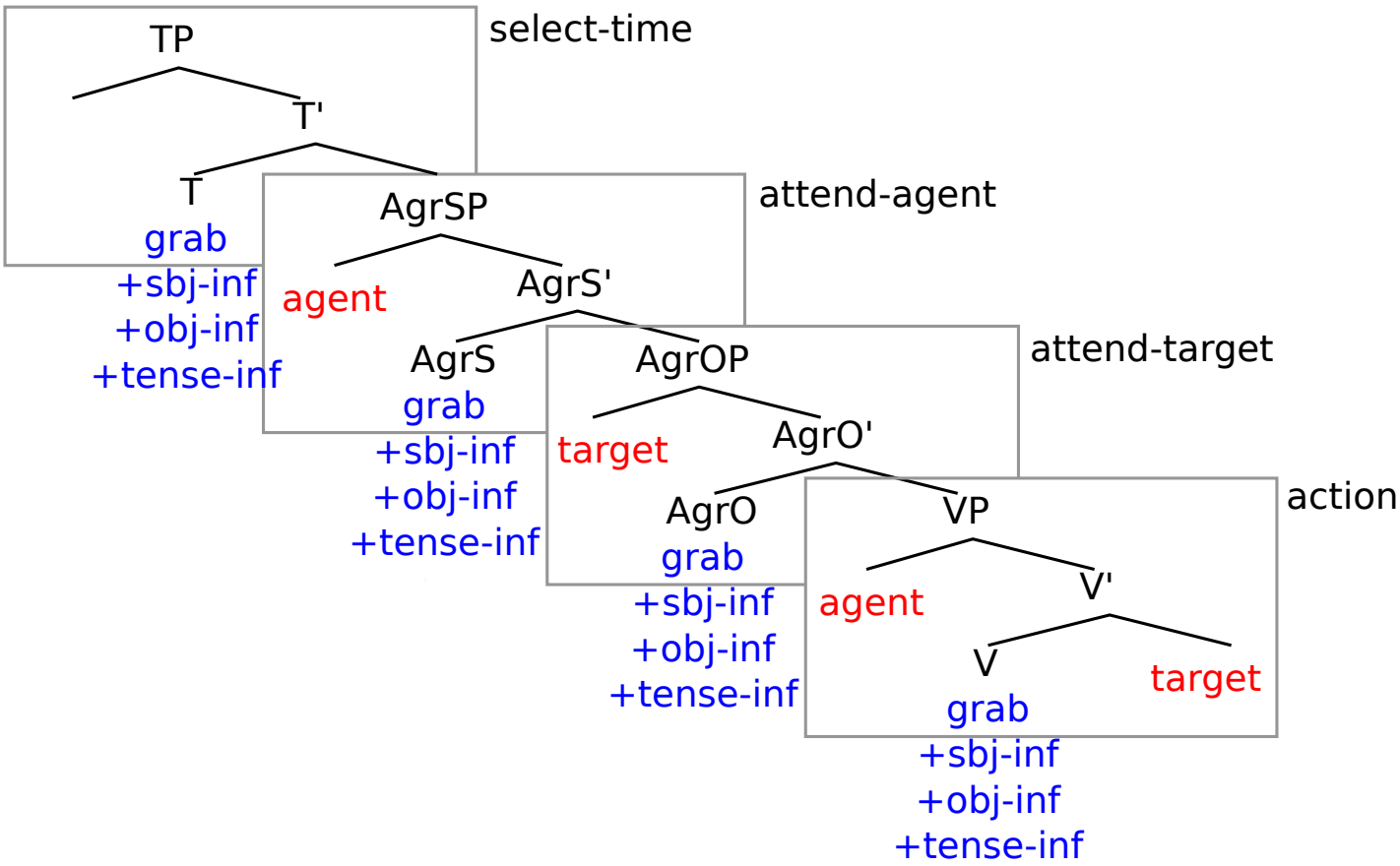


In syntactic terms

If heads are read from the planning medium, we can read out information about *each* SM operation at *each* head position.



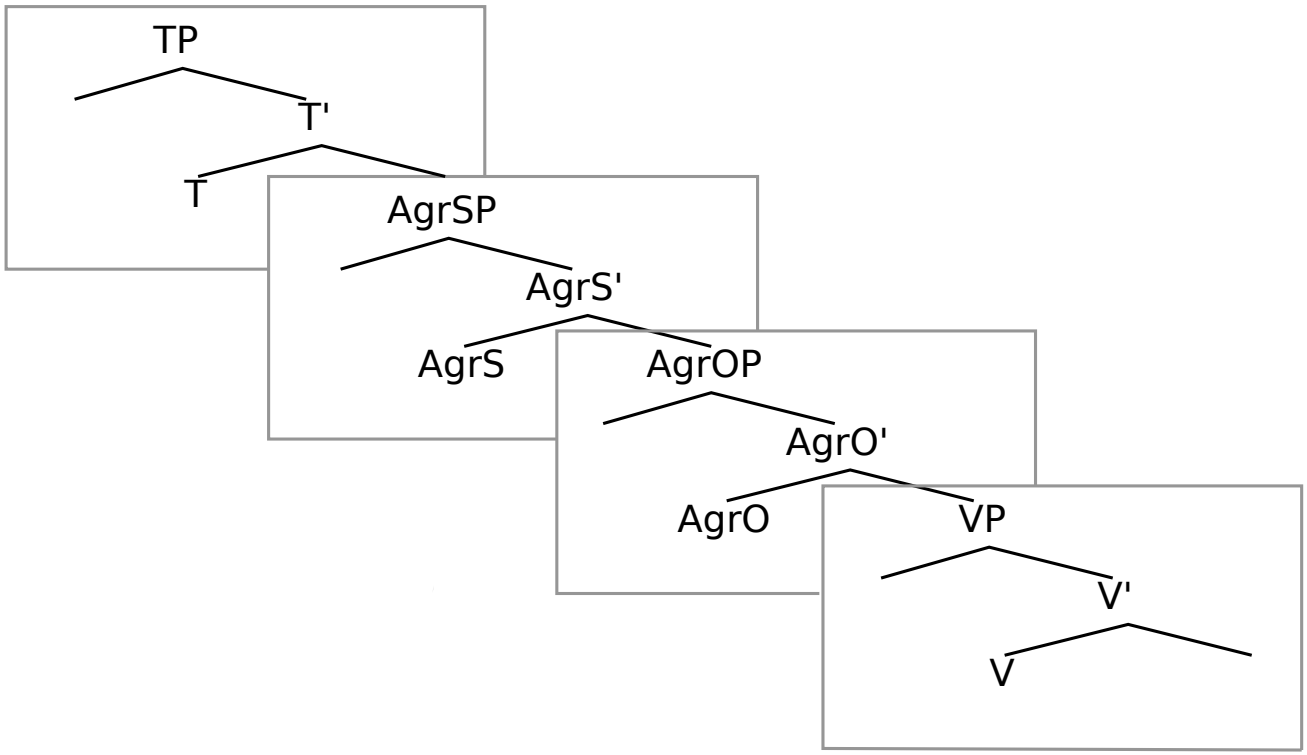
In syntactic terms



In SM terms

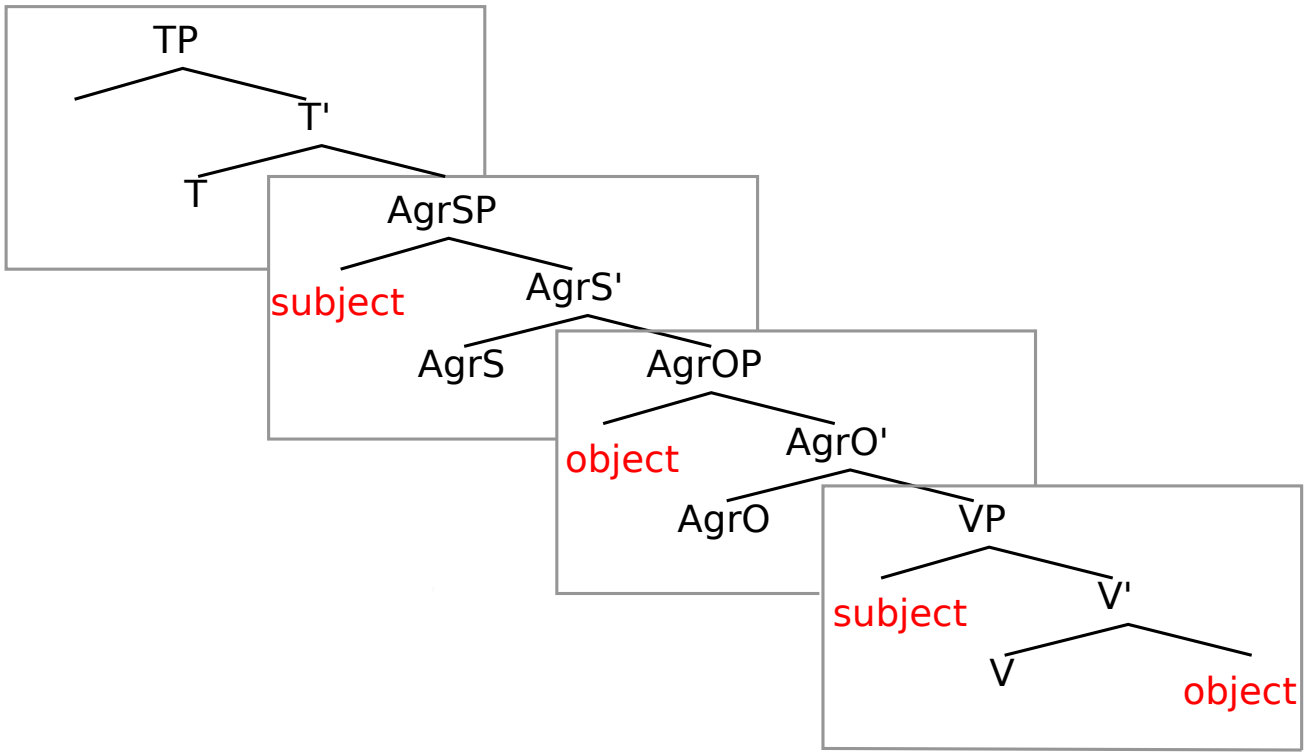
Sustained signals	Transient signals		
		SM operation	Reafferent signal
select_time / *attend_agent* / *attend_target* / *action*			
select_time / *attend_agent* / *attend_target* / *action*			*agent*
select_time / *attend_agent* / *attend_target* / *action*			*target*
select_time / *attend_agent* / *attend_target* / *action*			*agent*
			target

Syntactic phenomena in the scope of our model



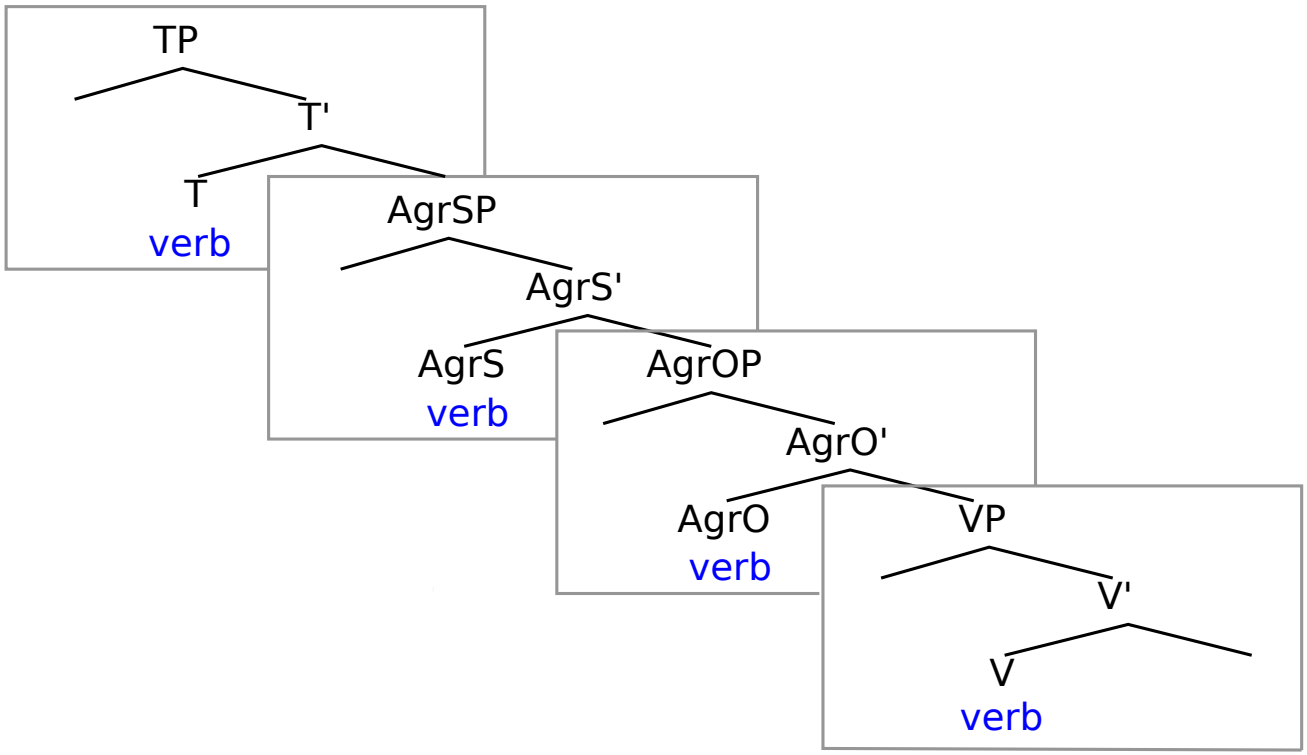
Syntactic phenomena in the scope of our model

Different subject/object positions



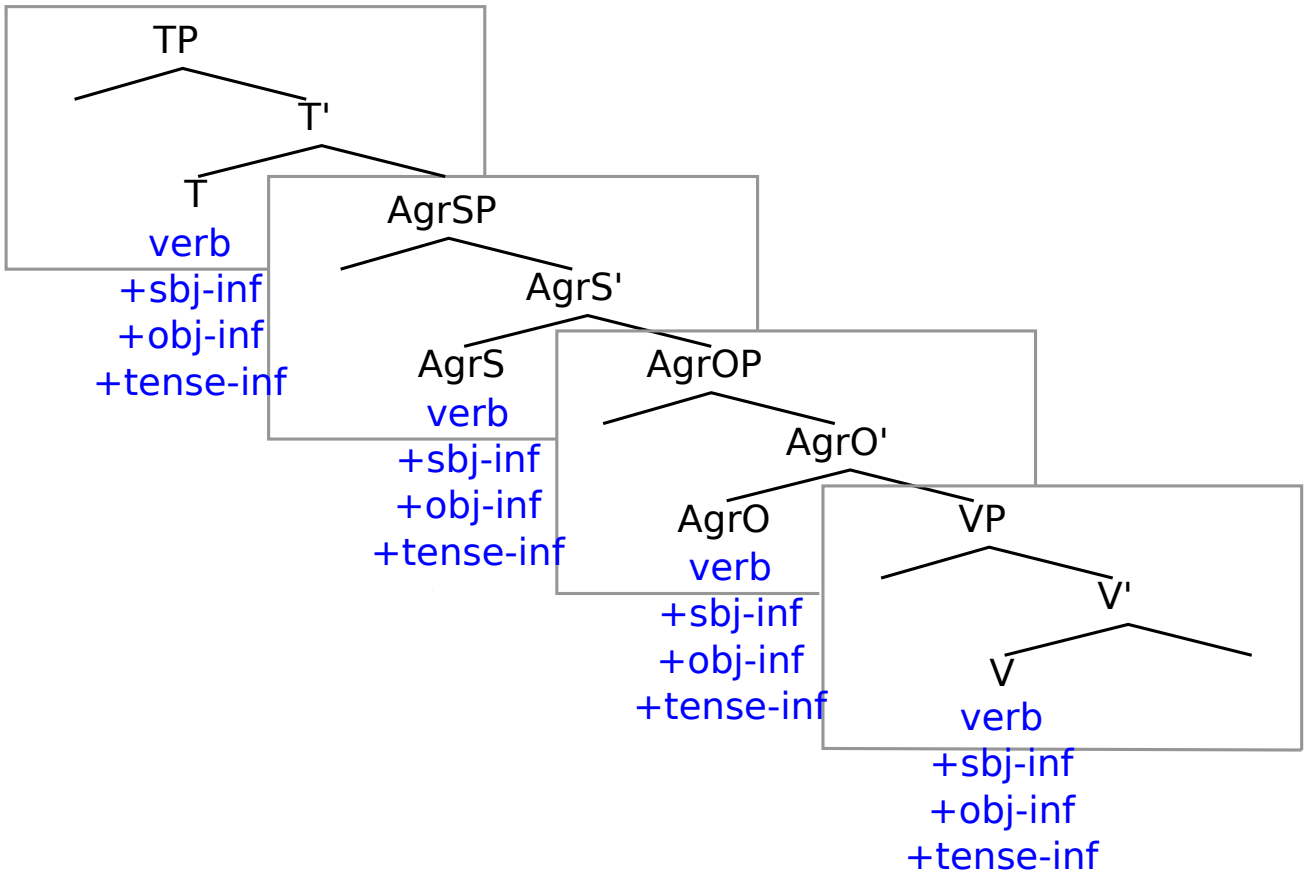
Syntactic phenomena in the scope of our model

Different verb positions



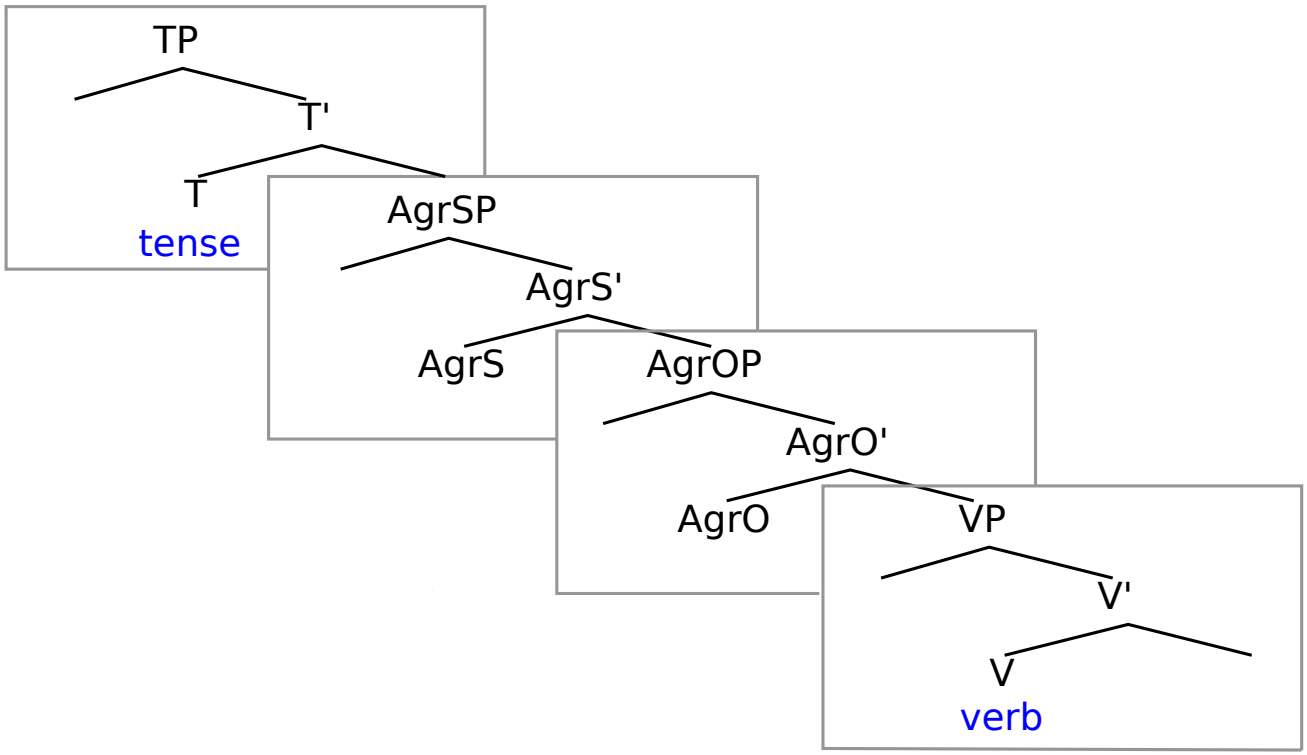
Syntactic phenomena in the scope of our model

Different patterns of verb inflection



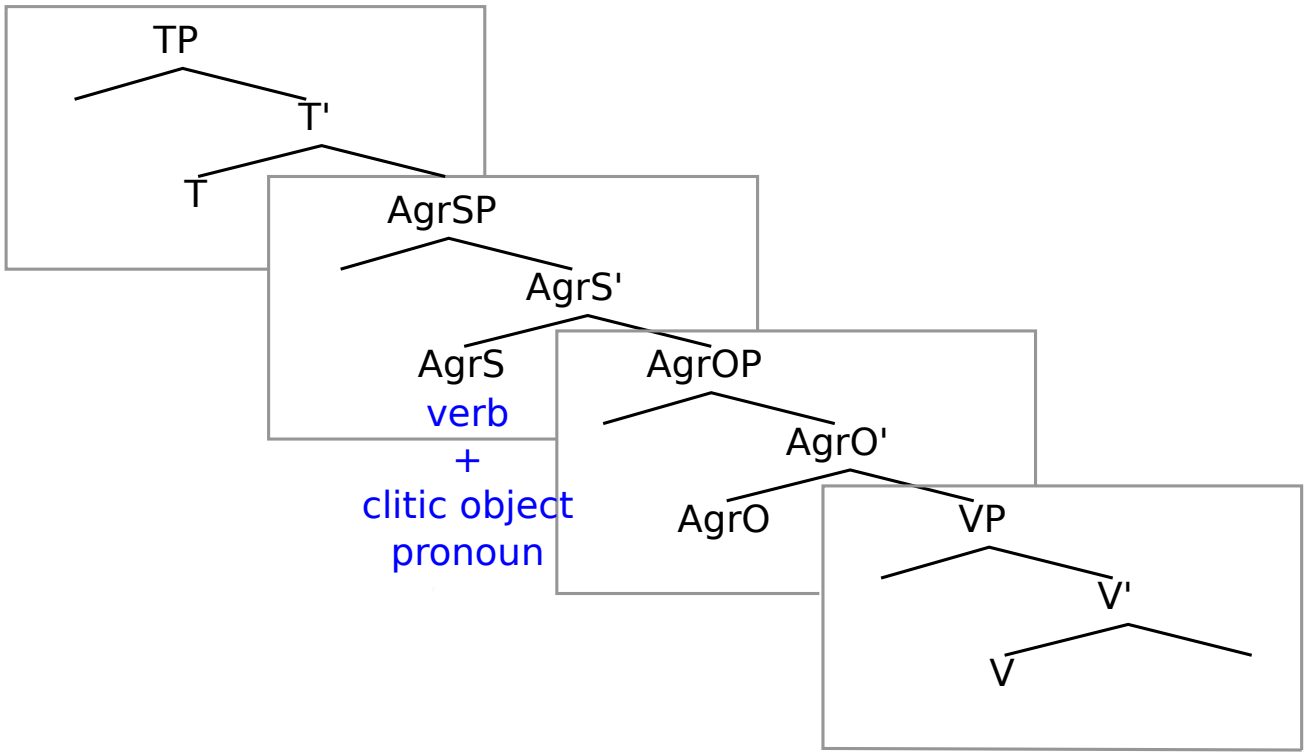
Syntactic phenomena in the scope of our model

Head features signalled separately from verbs (e.g. tense markers)



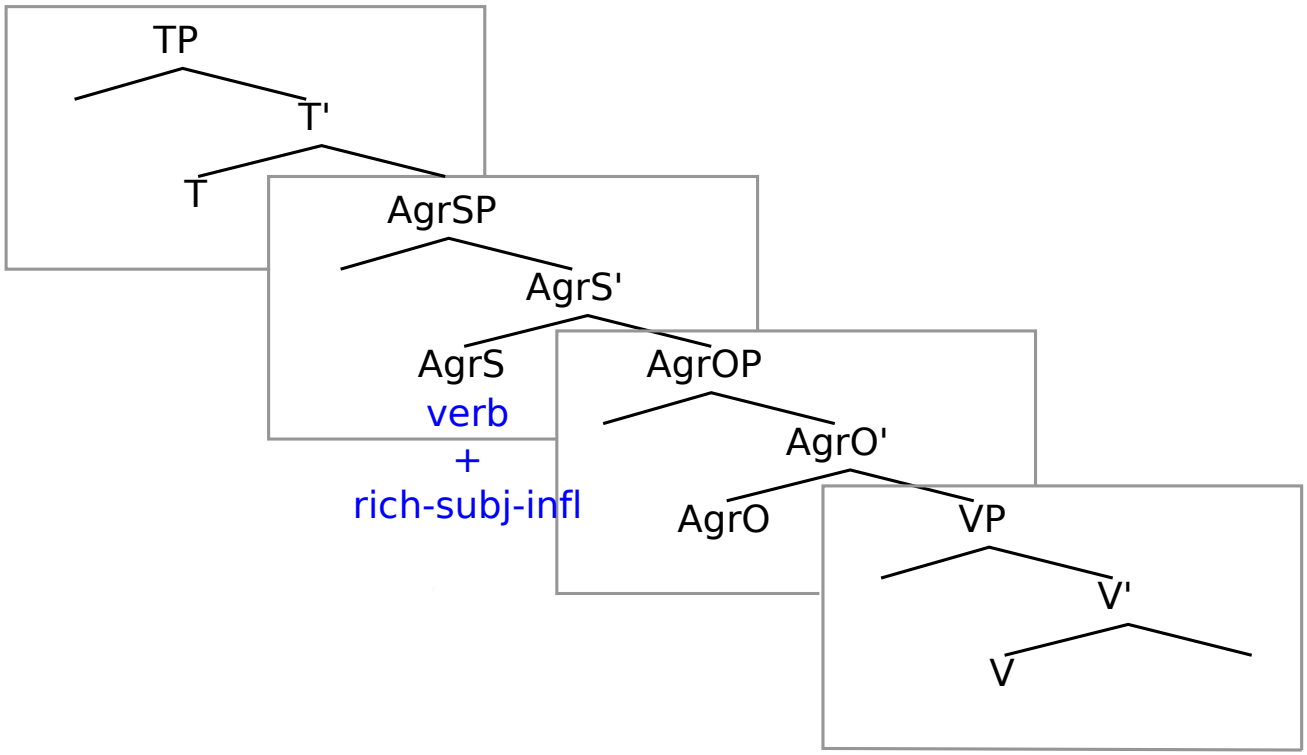
Syntactic phenomena in the scope of our model

Clitic pronouns



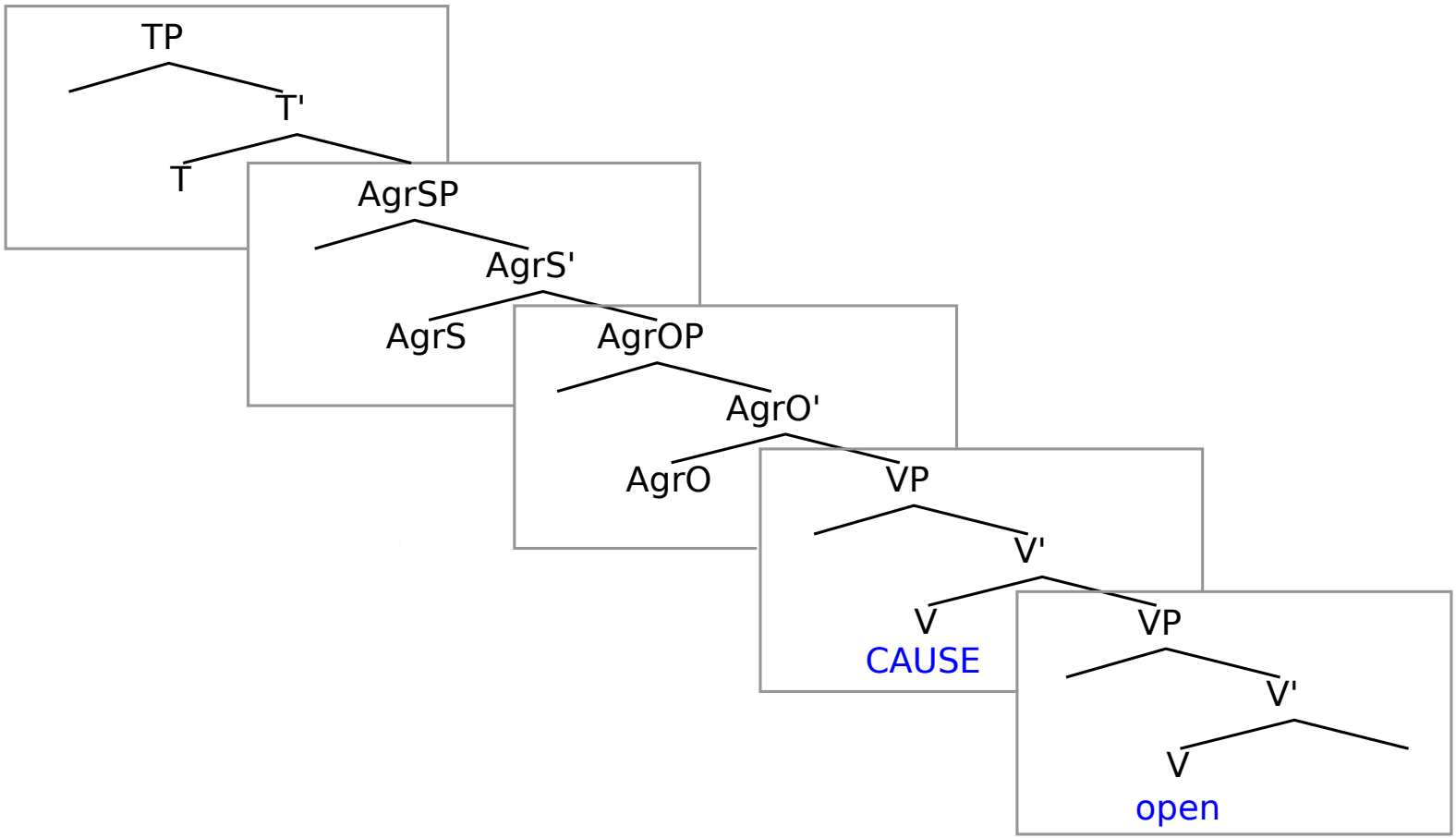
Syntactic phenomena in the scope of our model

Dropped subjects

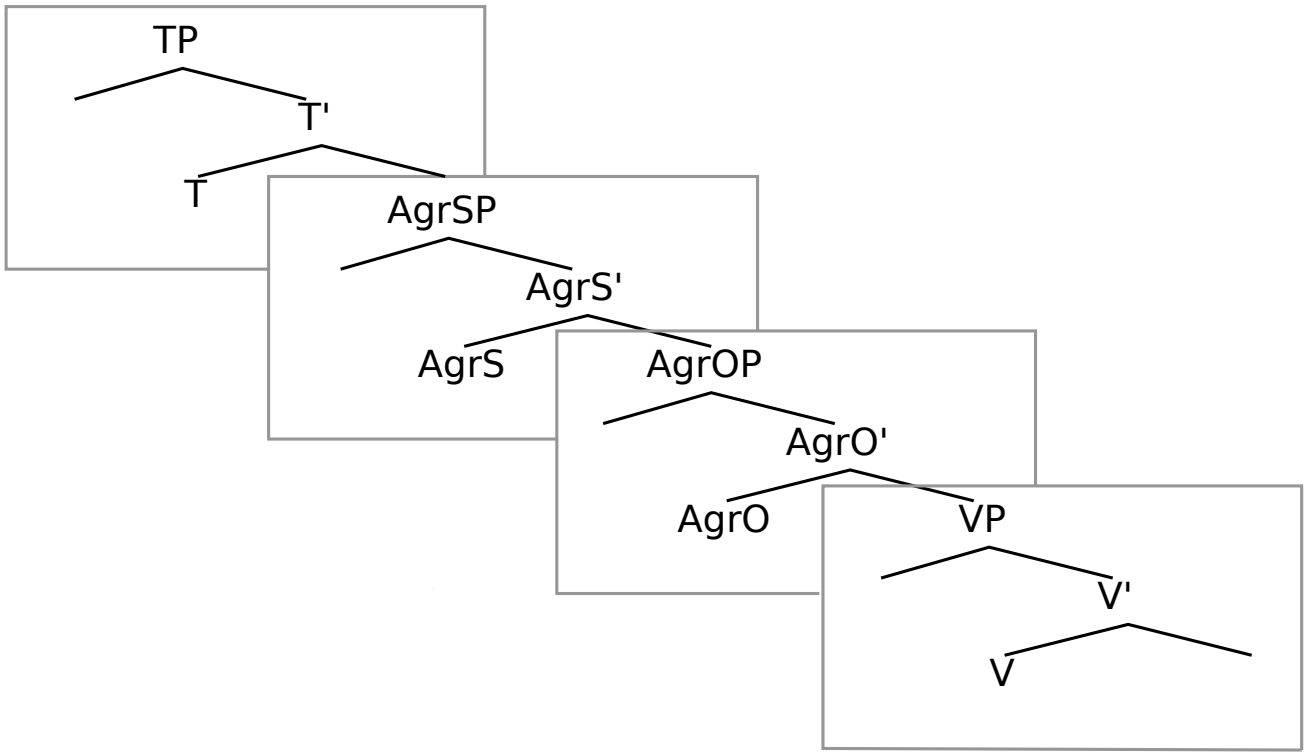


Syntactic phenomena in the scope of our model

Causative constructions

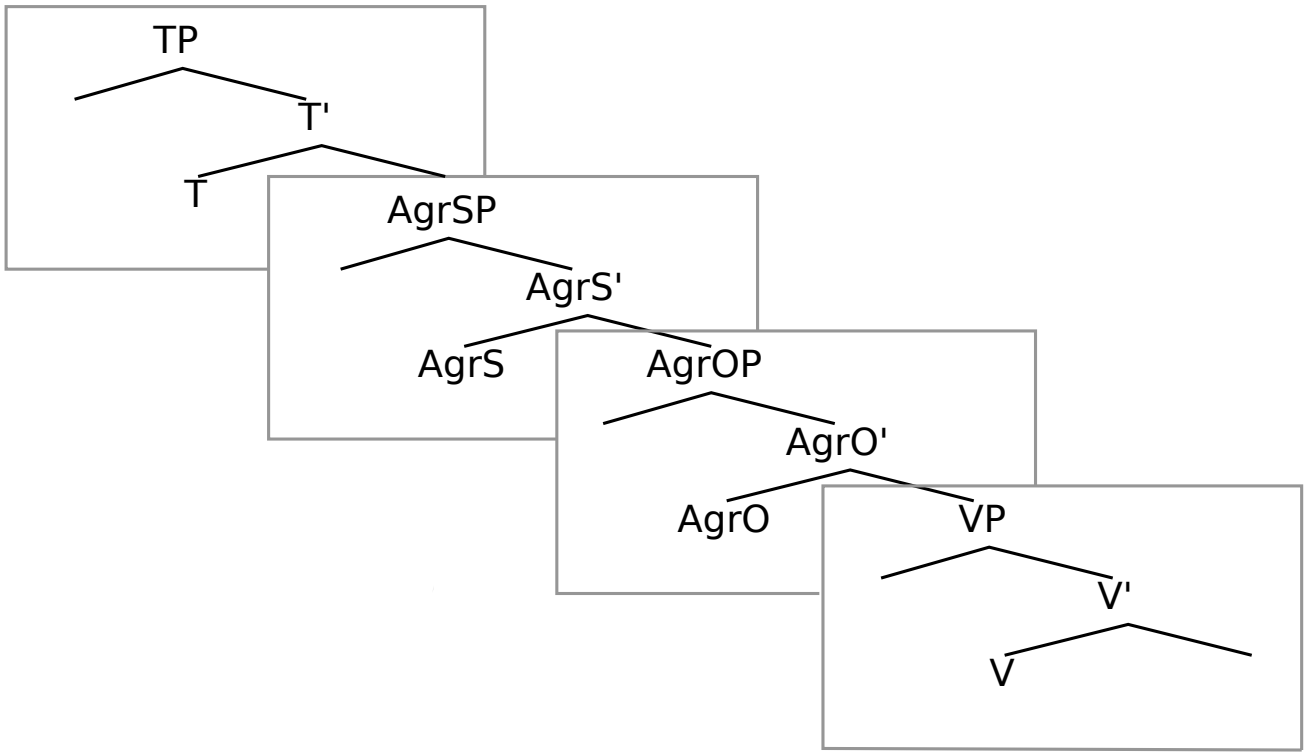


Hard phenomena for a Chomskyan model



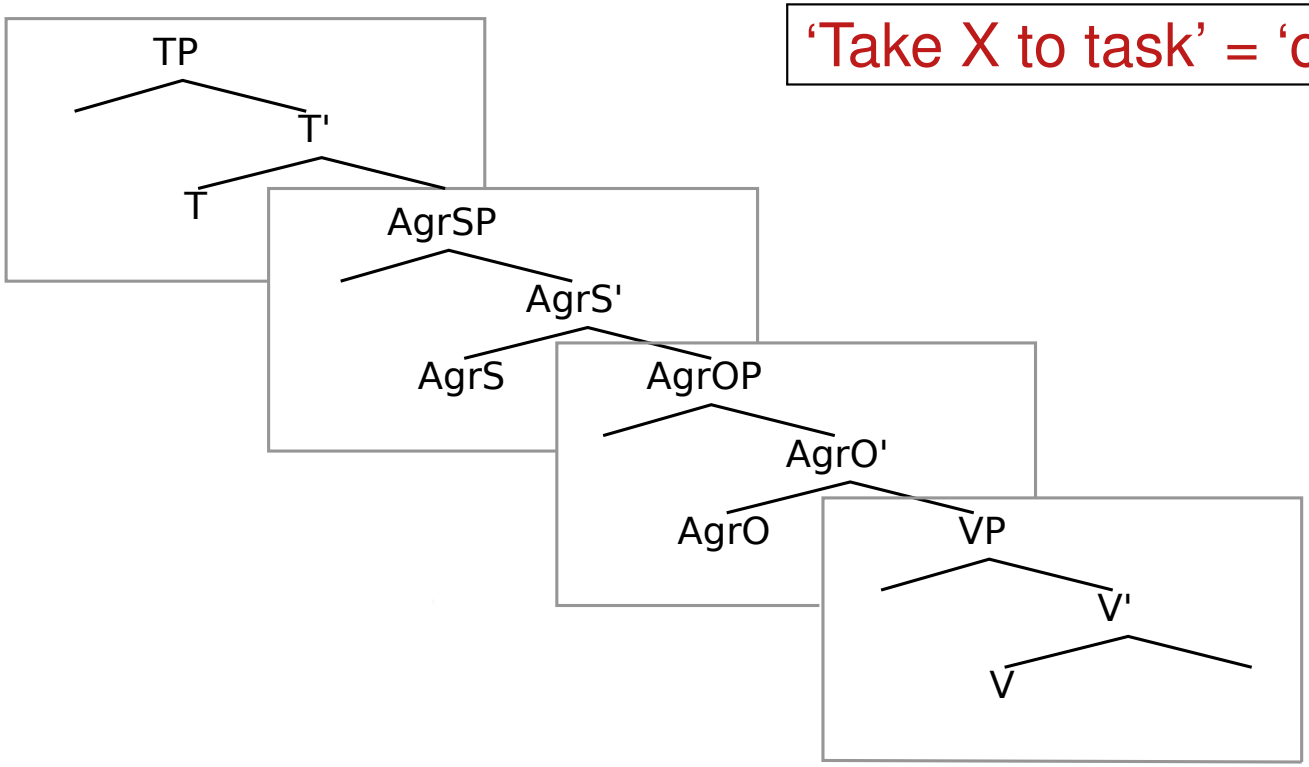
Hard phenomena for a Chomskyan model

Surface constructions in language—'Idioms'.



Hard phenomena for a Chomskyan model

Surface constructions in language—‘Idioms’.



‘Take X to task’ = ‘criticise X’

A neural network for sentence generation

We built a network that learns to map rehearsed SM routines onto surface sentences.

- The network simulates how a baby learns its mother tongue.
- It's trained on pairs of ⟨SM-sequence, word-sequence⟩.
- Three training languages were tried: English, Māori, Slovak.

A neural network for sentence generation

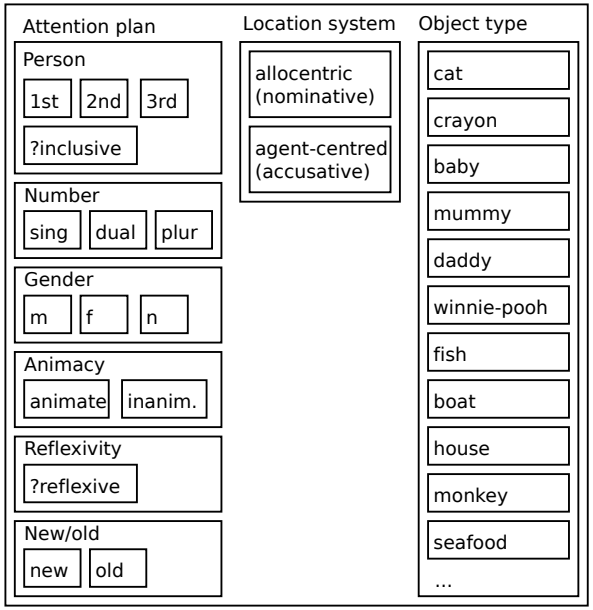
There are three parts to the model.

- A word production network, that learns to map SM signals onto surface words (stems/inflections).
- A control network, that learns *when* words should be pronounced.
- A word sequencing network, that learns surface patterns in sentences, and completes patterns when they are predictable.

SM inputs to the network

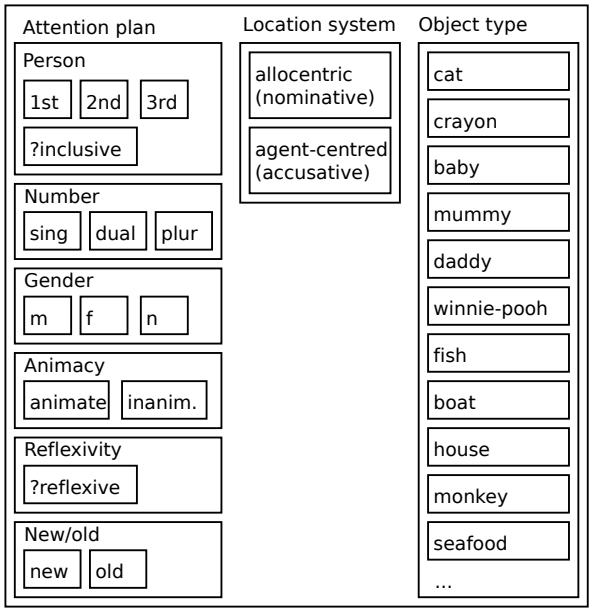
SM inputs to the network

An area holding a transient object representation. . .

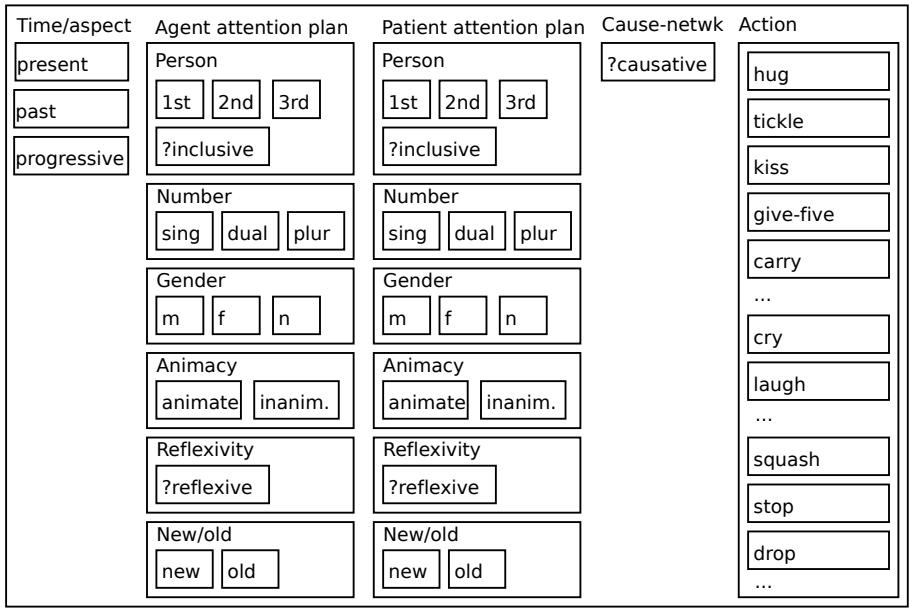


SM inputs to the network

An area holding a sustained WM episode representation. . .



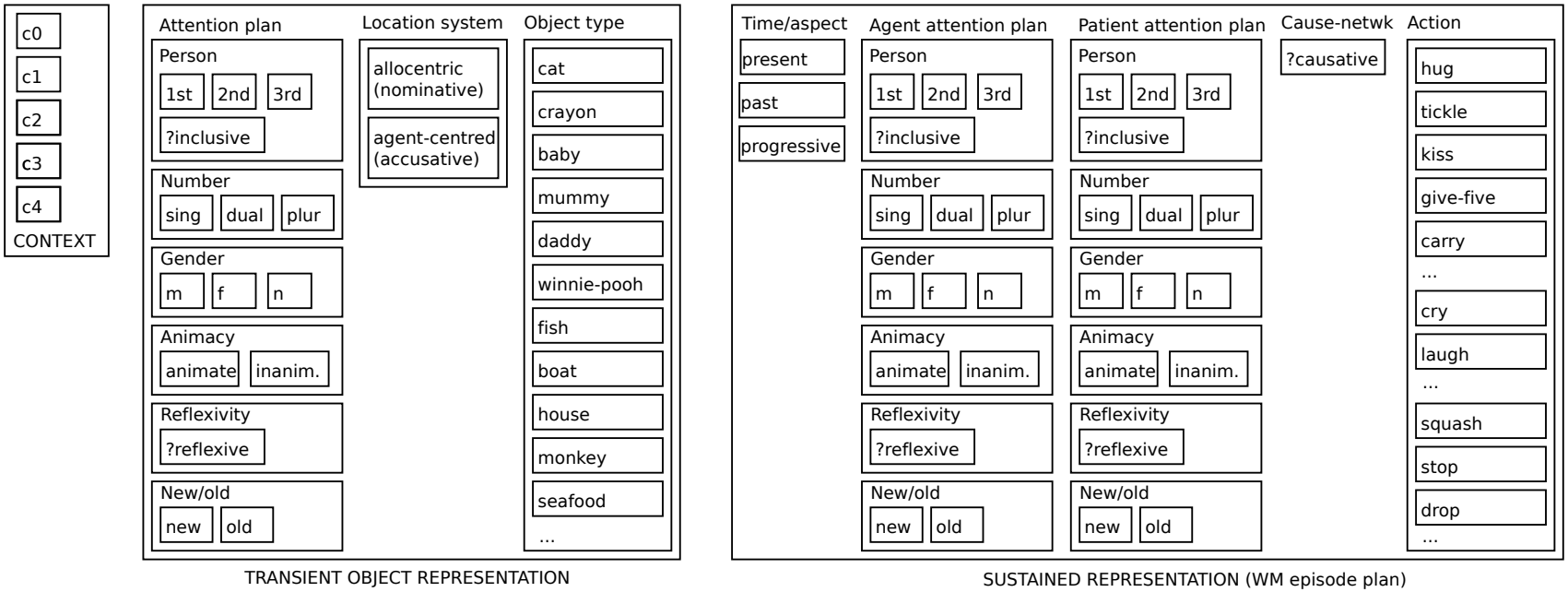
TRANSIENT OBJECT REPRESENTATION



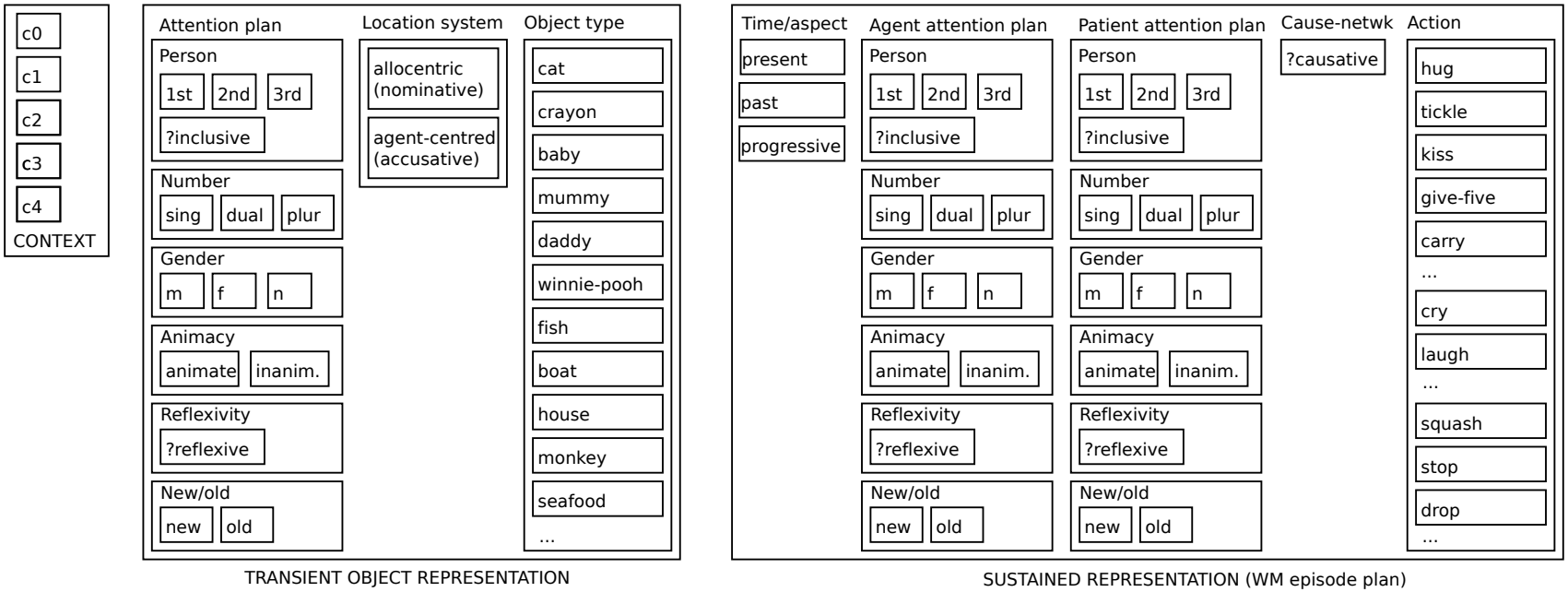
SUSTAINED REPRESENTATION (WM episode plan)

SM inputs to the network

An area holding the current SM context.

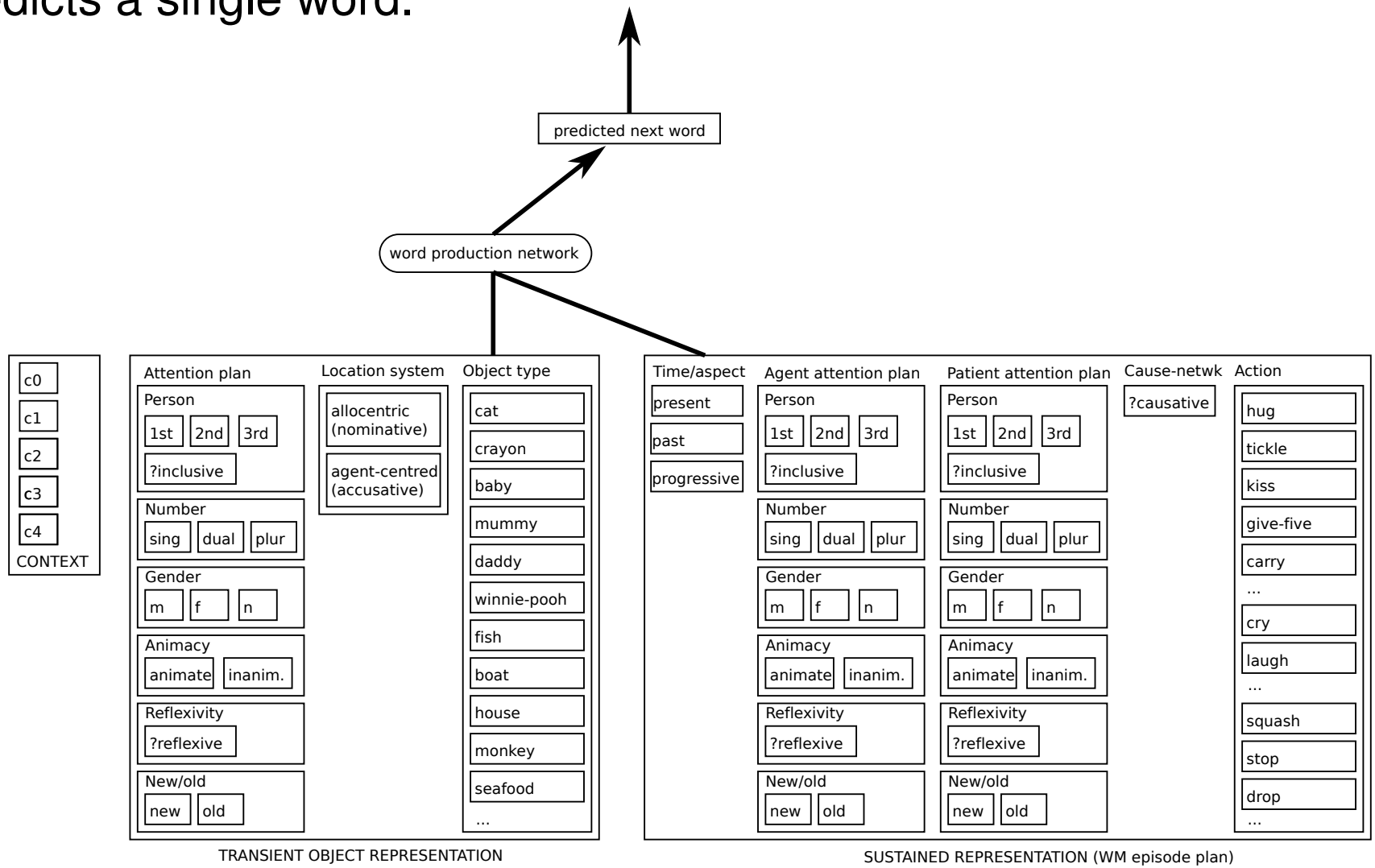


Network architecture



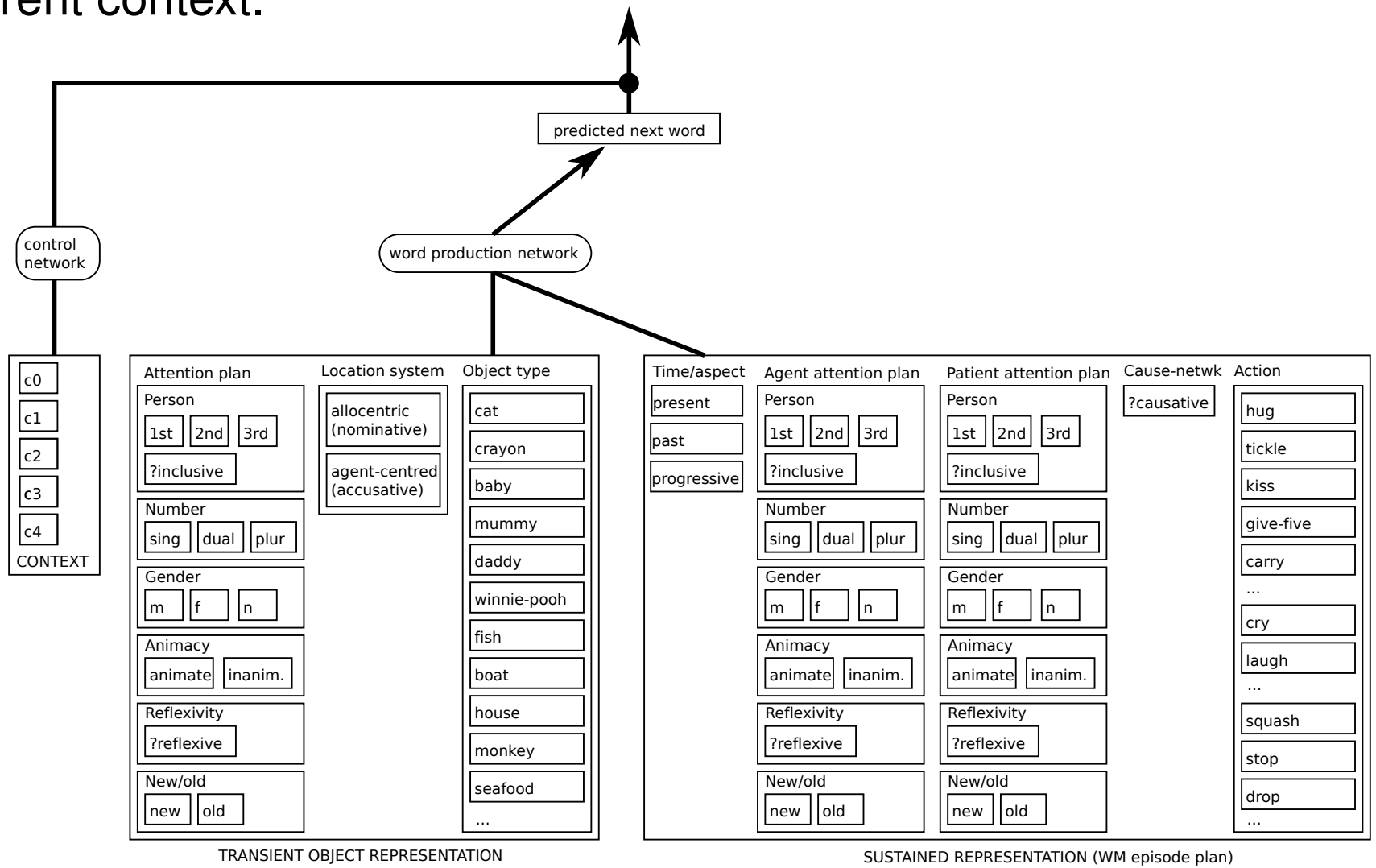
Network architecture

The word production network takes input from each area in turn, and predicts a single word.



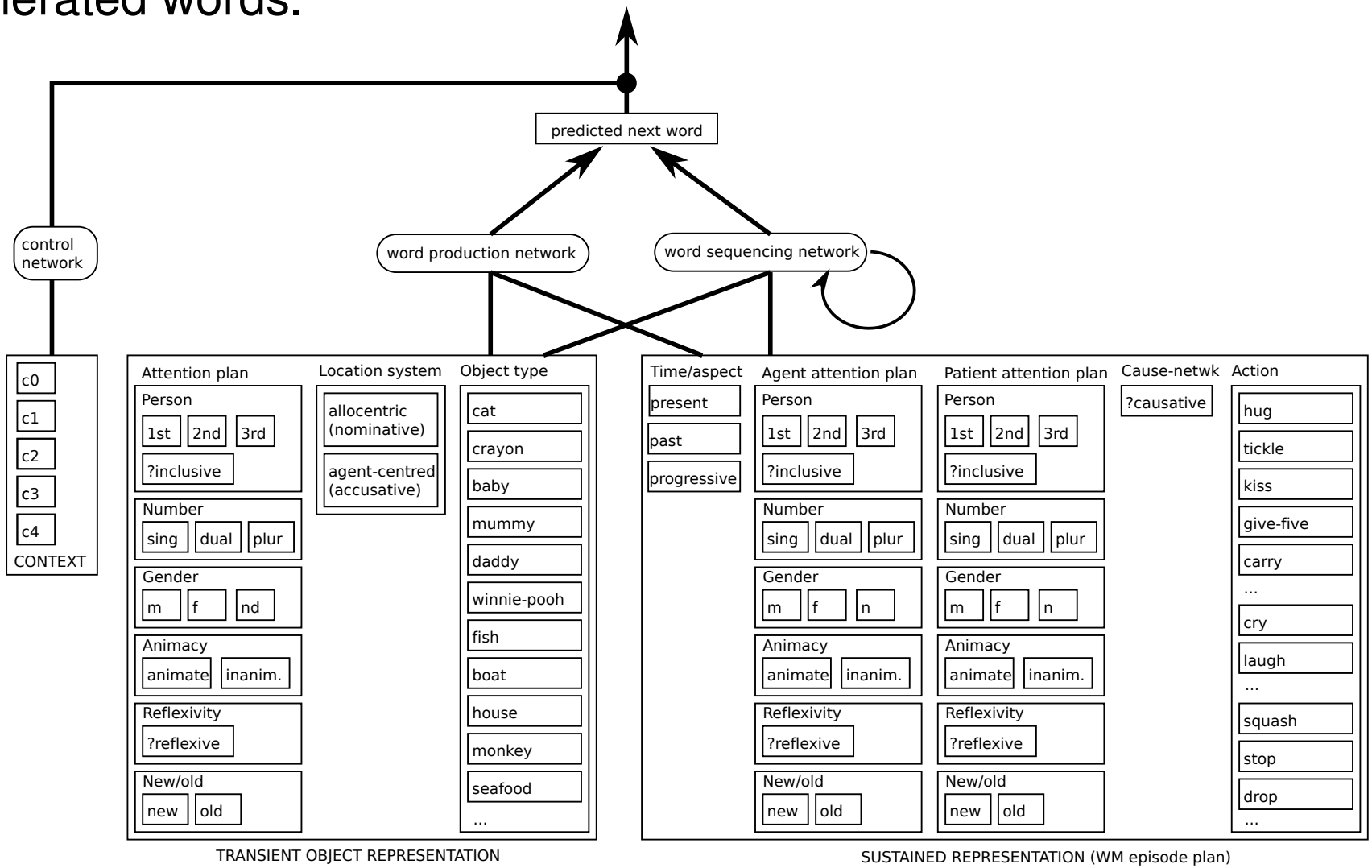
Network architecture

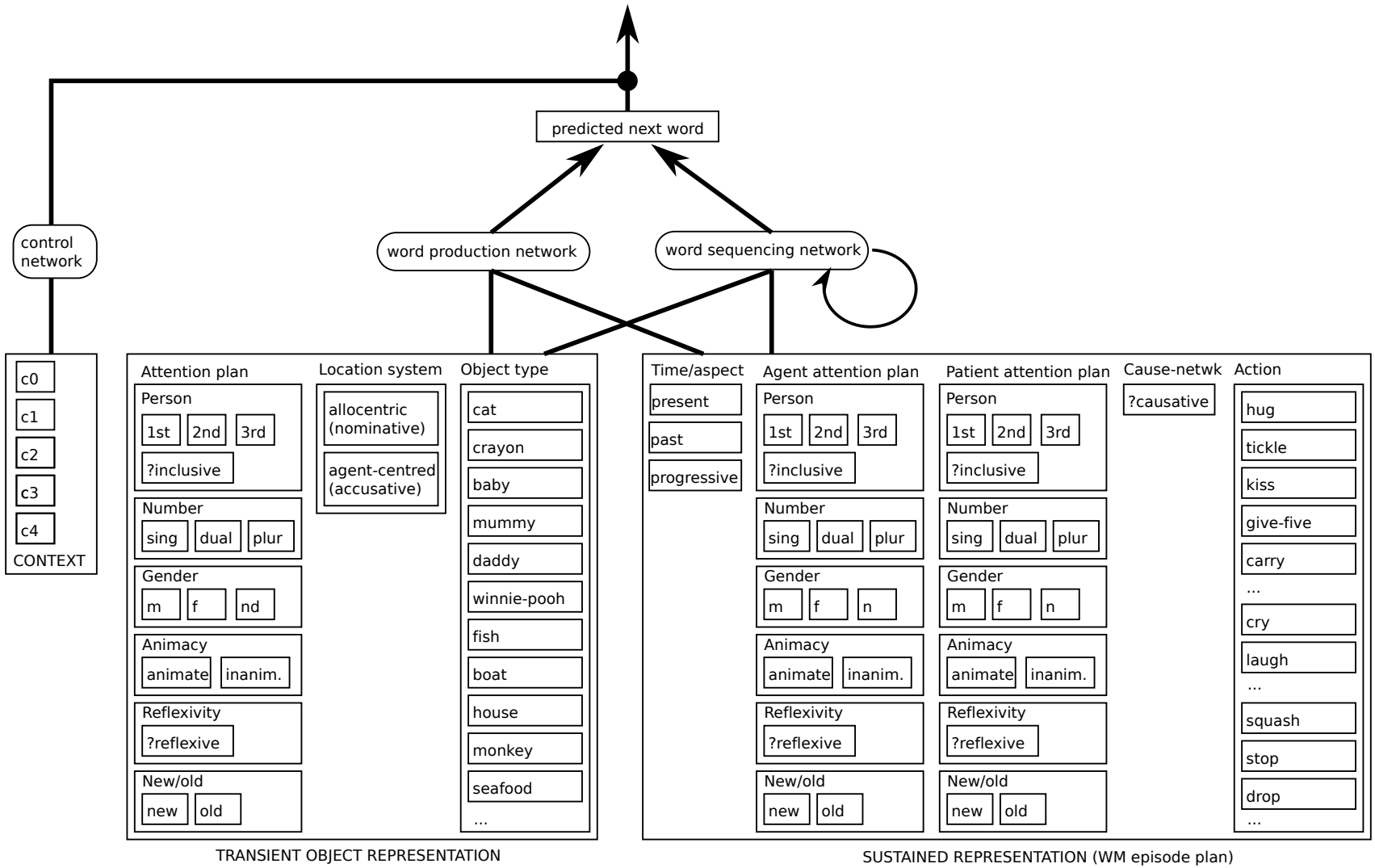
The control network decides whether to *pronounce* the word in the current context.



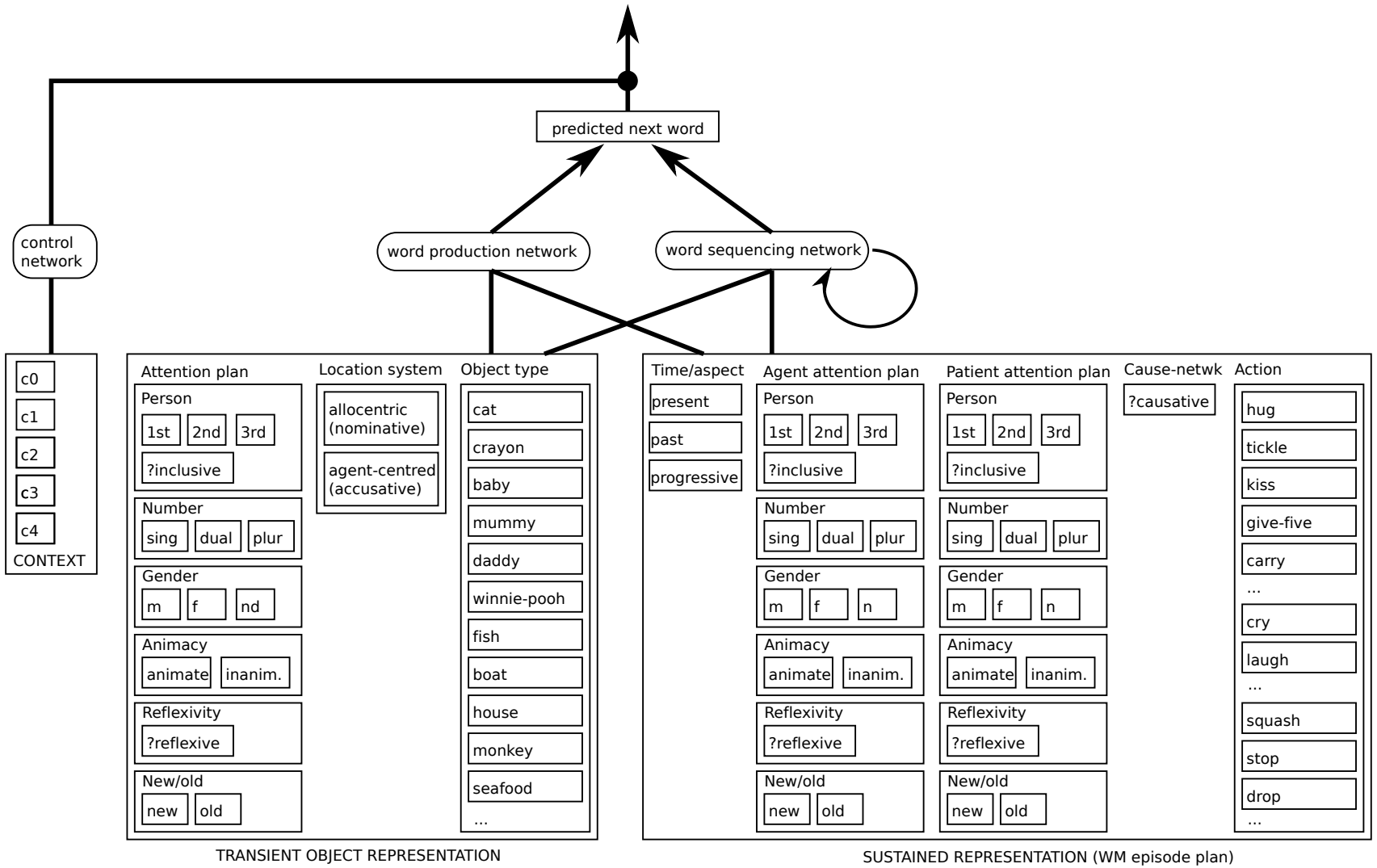
Network architecture

The word sequencing network produces idiomatic continuations of generated words.



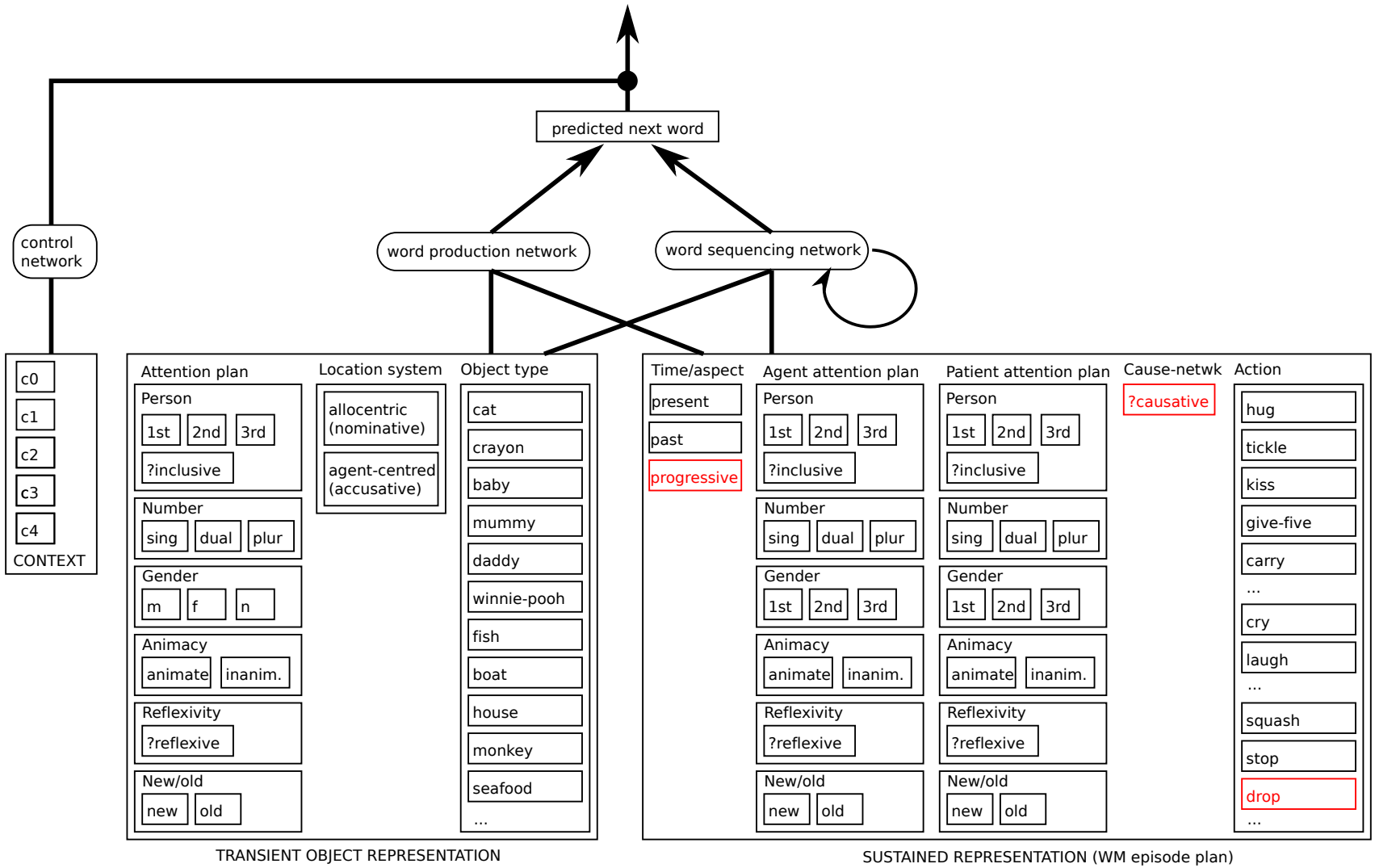

predicted next word
control network
word production network
word sequencing network
c0
c1
c2
c3
c4
CONTEXT
Attention plan
Person
1st
2nd
3rd
?inclusive
Number
sing
dual
plur
Gender
m
f
nd
Animacy
animate
inanim.
Reflexivity
?reflexive
New/old
new
old
Location system
allocentric (nominative)
agent-centred (accusative)
Object type
cat
crayon
baby
mummy
daddy
winnie-pooh
fish
boat
house
monkey
seafood
...
TRANSIENT OBJECT REPRESENTATION
Time/aspect
present
past
progressive
Agent attention plan
Person
1st
2nd
3rd
?inclusive
Number
sing
dual
plur
Gender
m
f
n
Animacy
animate
inanim.
Reflexivity
?reflexive
New/old
new
old
Patient attention plan
Person
1st
2nd
3rd
?inclusive
Number
sing
dual
plur
Gender
m
f
n
Animacy
animate
inanim.
Reflexivity
?reflexive
New/old
new
old
Cause-netwk
?causative
Action
hug
tickle
kiss
give-five
carry
...
cry
laugh
...
squash
stop
drop
...
SUSTAINED REPRESENTATION (WM episode plan)

E whaka makere ana te ngeru i ngā ika



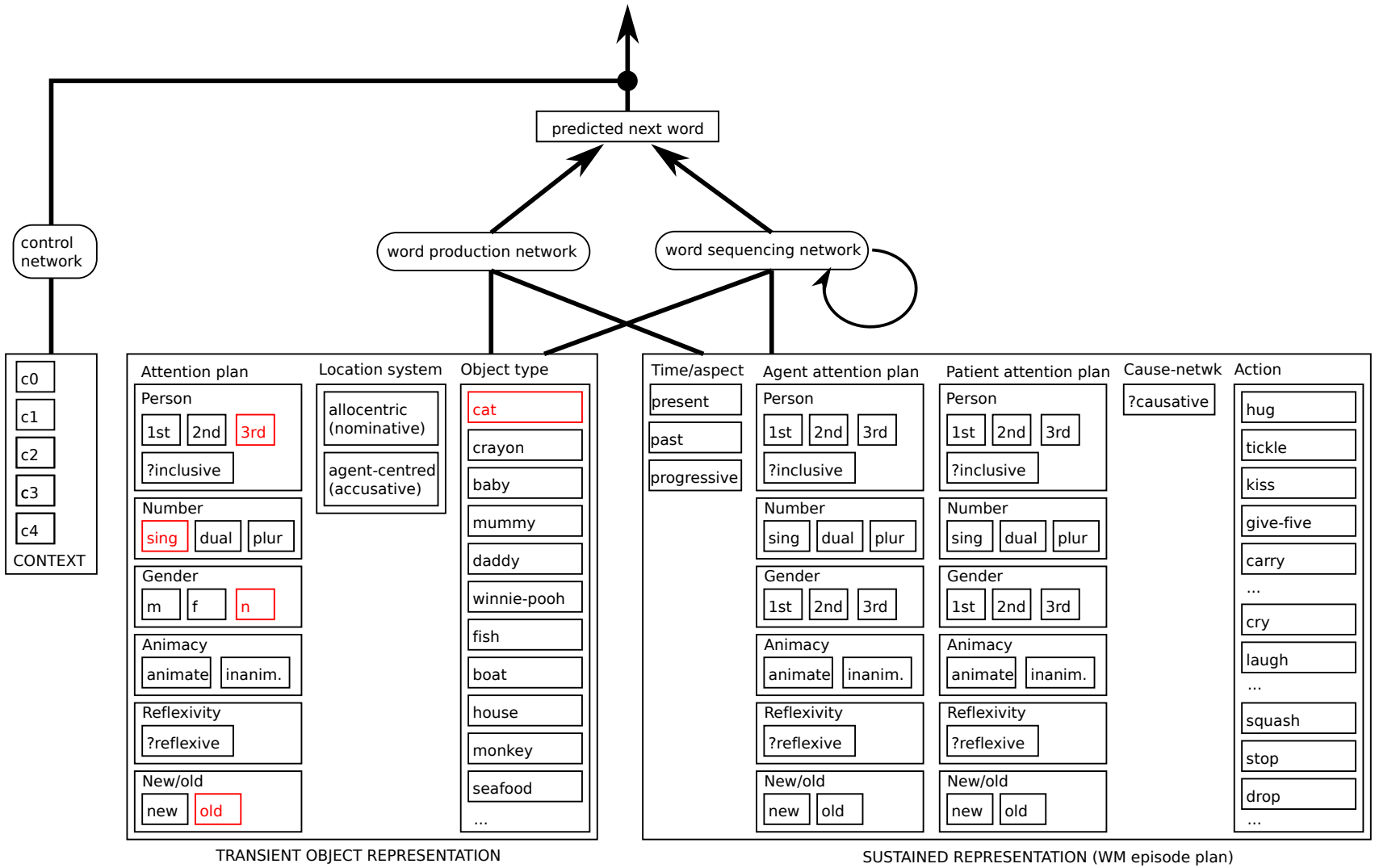
E whaka makere ana te ngeru i ngā ika

E, whaka, makere, ana. . .

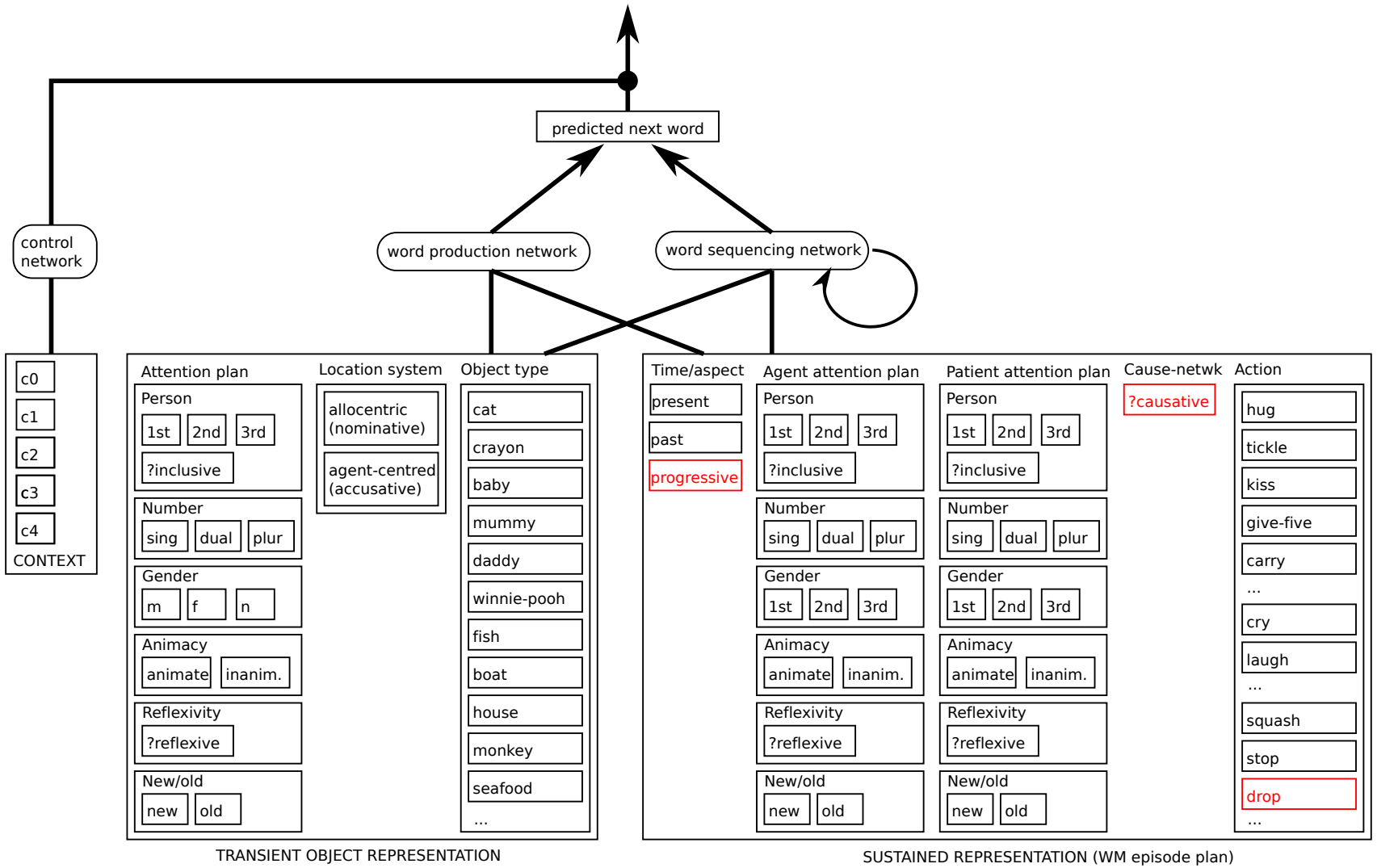


E whaka makere ana te ngeru i ngā ika

te, ngeru. . .

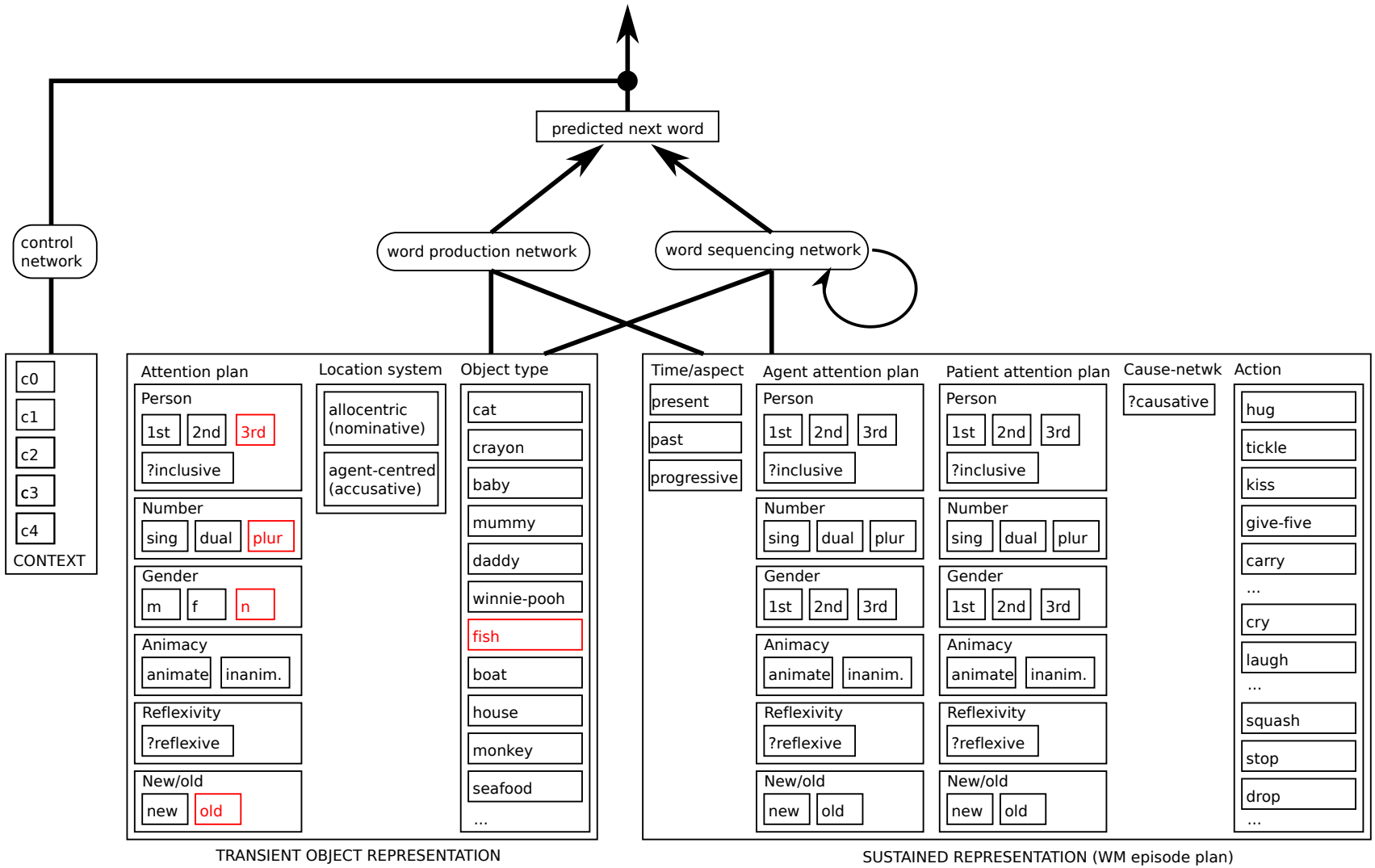


E whaka makere ana te ngeru i ngā ika

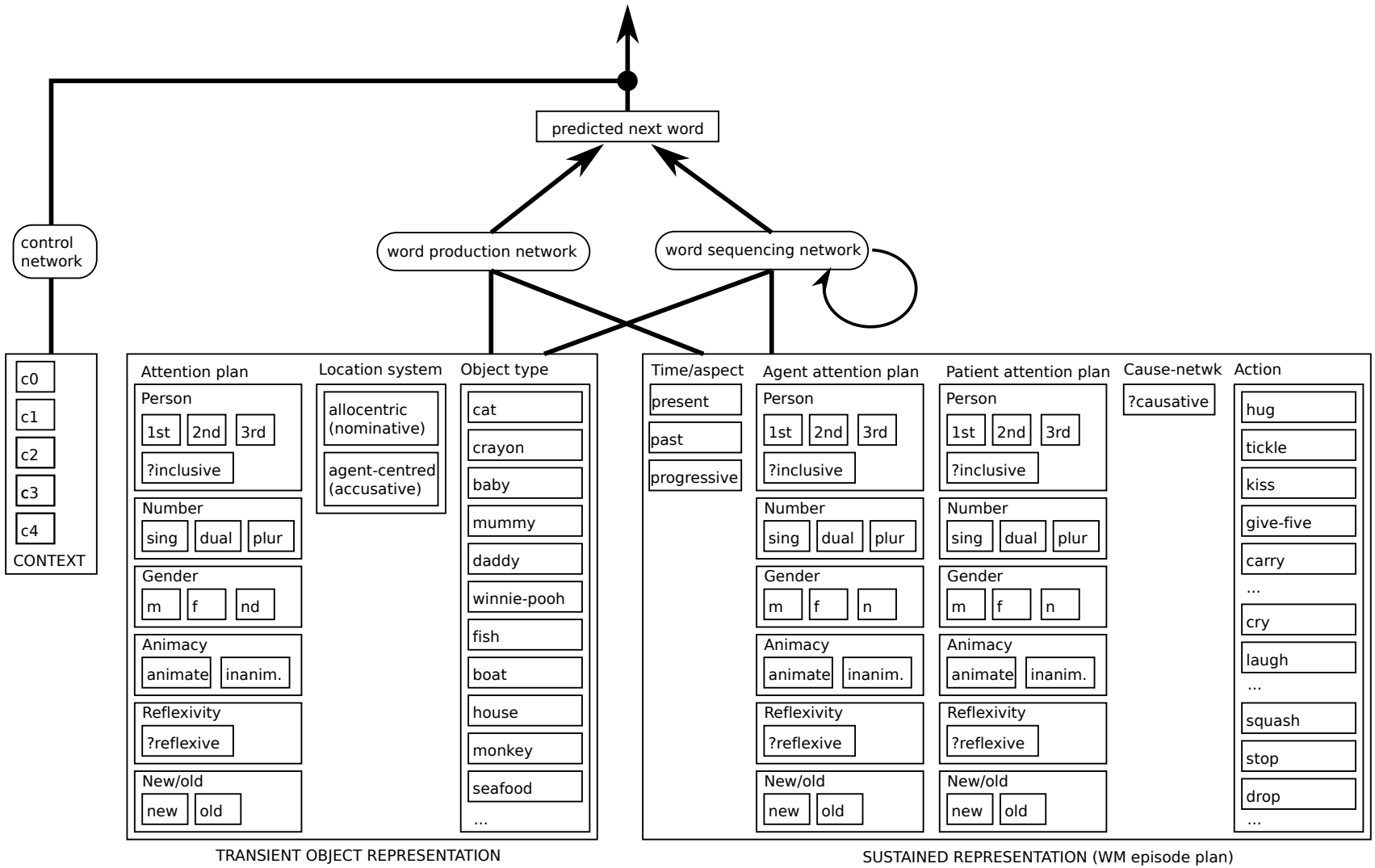


E whaka makere ana te ngeru i ngā ika

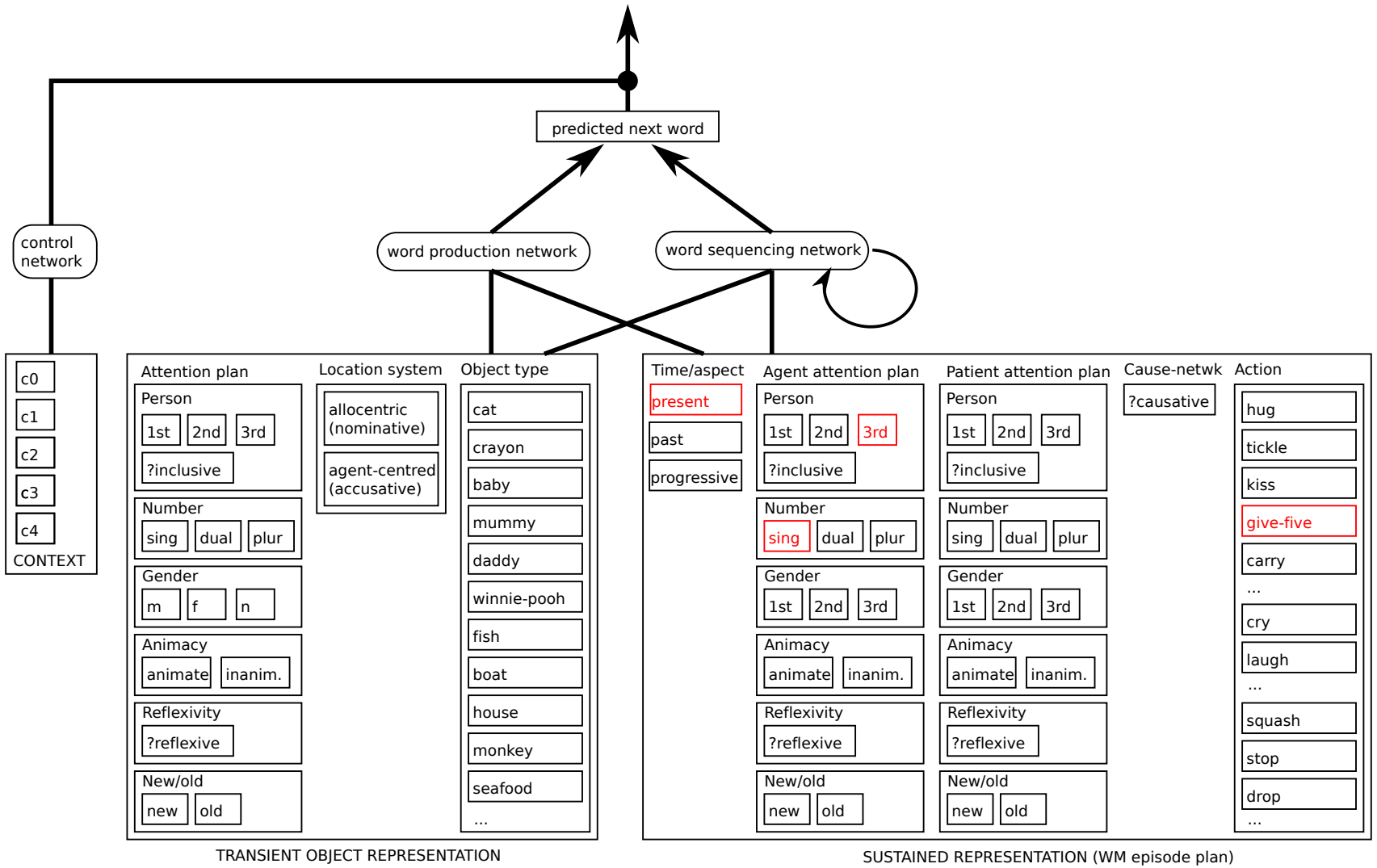
i, ngā, ika.



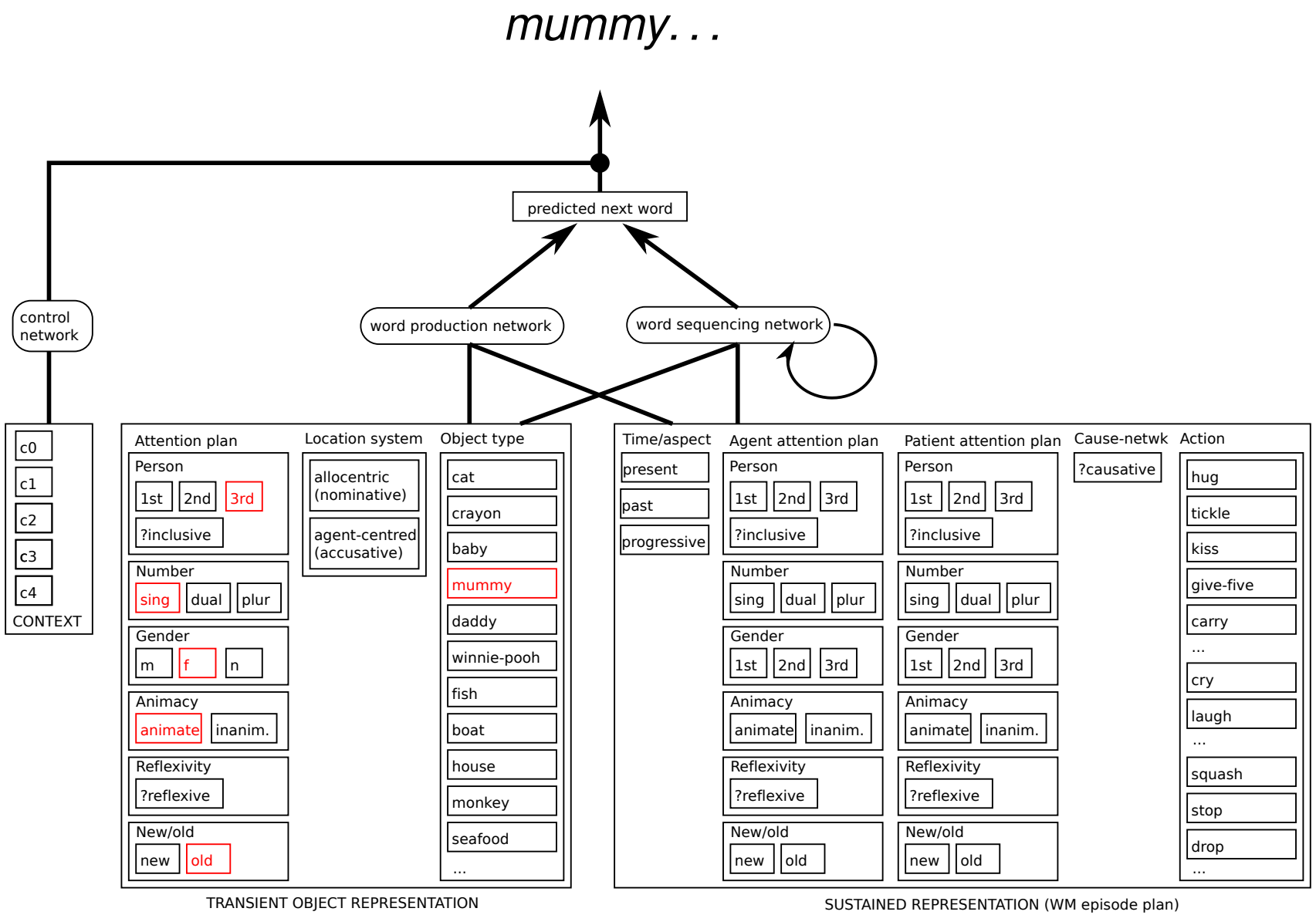
Mummy gives the baby five



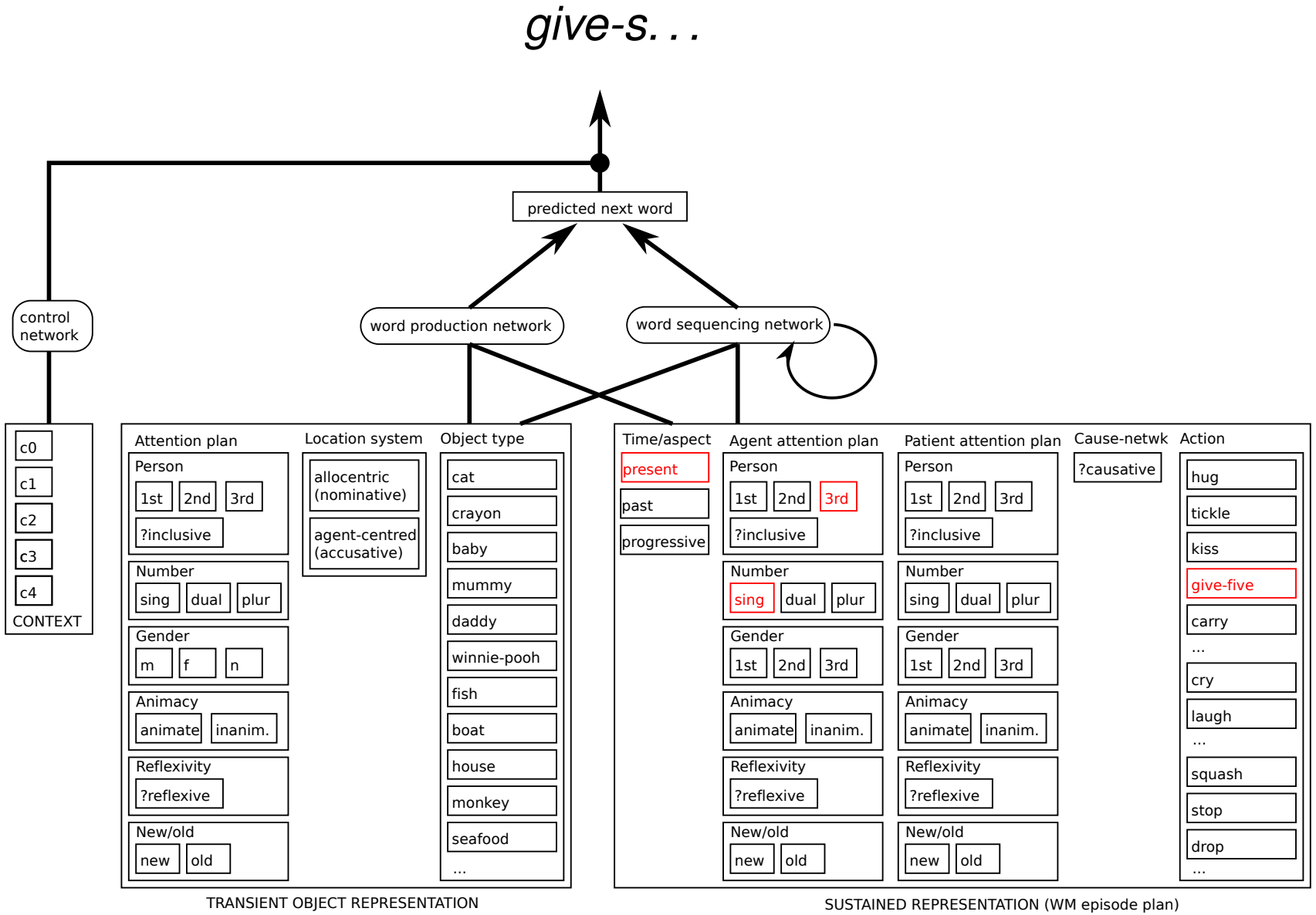
Mummy gives the baby five



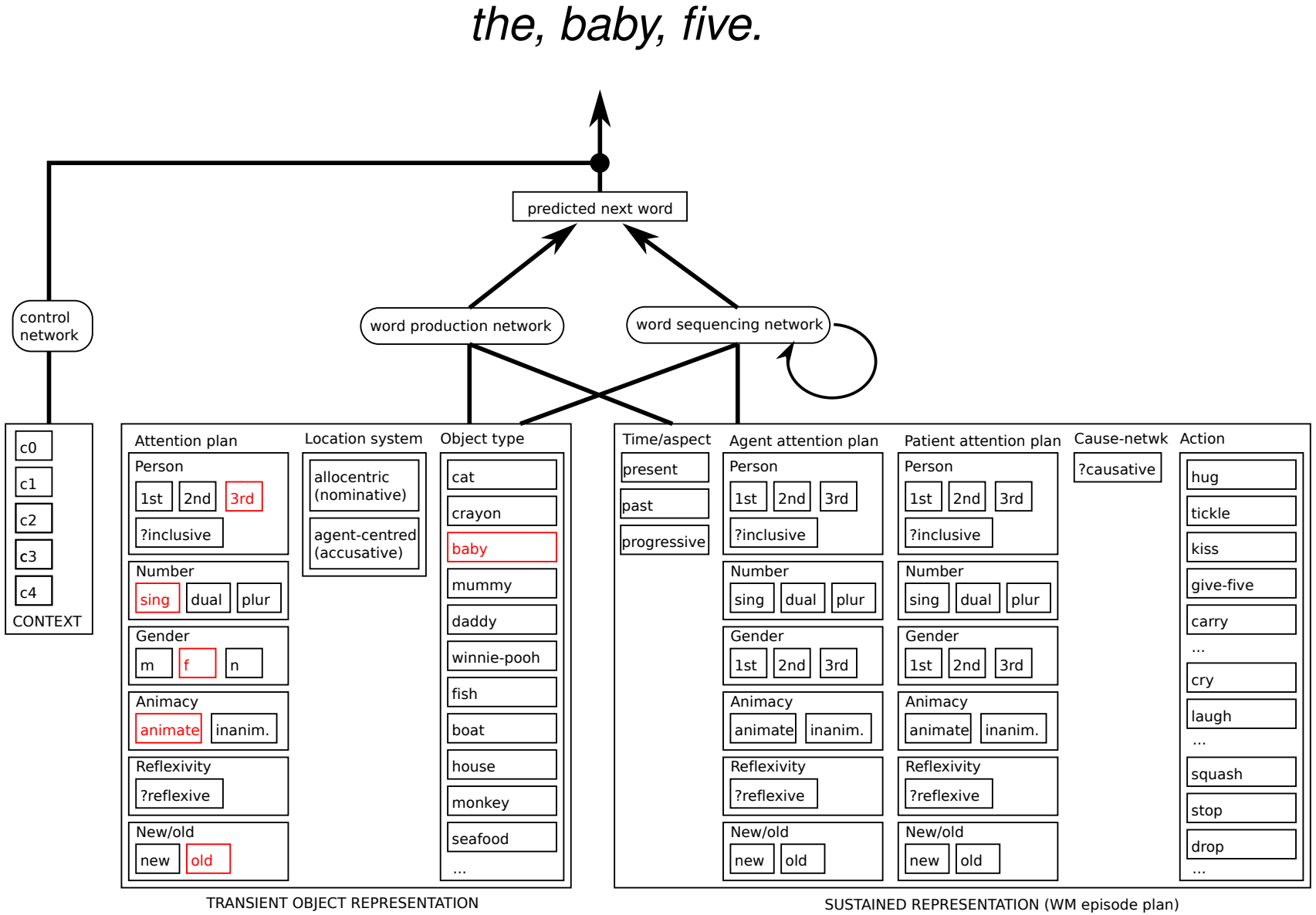
Mummy gives the baby five



Mummy gives the baby five



Mummy gives the baby five



Training and evaluation of the network

In each language:

- We trained the network on 10,000 sentences, in 16 epochs.
- Then we tested on 1000 ‘seen’ and 1000 ‘unseen’ sentences.

Results:

	English	Māori	Slovak
Performance for seen sentences	99.5%	99.2%	99.9%
Performance for new sentences	99.1%	98.5%	99.0%

Training and evaluation of the network

In each language:

- We trained the network on 10,000 sentences, in 16 epochs.
- Then we tested on 1000 ‘seen’ and 1000 ‘unseen’ sentences.

Results:

	English	Māori	Slovak
Performance for seen sentences	99.5%	99.2%	99.9%
Performance for new sentences	99.1%	98.5%	99.0%

In each language, the system begins by learning single words, then idiomatic phrases, and finally, full syntax.