

# COSC451: Artificial Intelligence

## Lecture 1: Language and Sensorimotor Cognition

Alistair Knott

Dept. of Computer Science, University of Otago

# This year's 451 course

The theme of the course this year: 'the neuroscience of language'.

The big question we'll consider:

- How is language implemented in the brain?

We'll investigate this by looking in detail at a smaller question:

- How does language interface with the sensorimotor system?

First: why is this an important question?

## Answer 1: an 'Artificial Intelligence' perspective

Q: How should we represent the meaning of the following piece of text?

*The dog, whose name was Anax, a distinguished and unusual collie with blue eyes, and long silver-grey fur blotched with black and white, came bounding back to the two women.*

## Answer 1: an 'Artificial Intelligence' perspective

Q: How should we represent the meaning of the following piece of text?

*The dog, whose name was Anax, a distinguished and unusual collie with blue eyes, and long silver-grey fur blotched with black and white, came bounding back to the two women.*

Logic:

dog(x1)      fur(x2)      has(x1, x2)      silver-grey(x2)

## Answer 1: an 'Artificial Intelligence' perspective

Q: How should we represent the meaning of the following piece of text?

*The dog, whose name was Anax, a distinguished and unusual collie with blue eyes, and long silver-grey fur blotched with black and white, came bounding back to the two women.*

Logic:

dog (x1)      fur (x2)      has (x1, x2)      silver-grey (x2)

Is this really good enough?

# Embodied AI

One proposal is that a representation of a ‘proposition’ must be grounded in our perceptual or motor systems.

- We don’t know the meaning of the word *dog* unless we can recognise an actual dog in the actual world.
- Ditto *fur*, *silver-grey*...
- Ditto *bounded*...

The proposal in **embodied AI** is that we need to look seriously at how information is delivered by the sensorimotor system.

## Answer 2: a ‘cognitive science’ perspective

We know quite a lot about how **low-level** sensorimotor processes work in the brain.

- Perception, motor control

But we don't know so much about **high-level** representations.

- E.g. of whole events or propositions.

Maybe studying how SM representations are converted into language can help us learn about high-level neural representations.

# A starting point

Take a simple example event: **a man grabs a cup.**

This event can be studied from two points of view:

- Linguists study the syntax and semantics of sentences which describe the event (e.g. *The man grabbed a cup*).
- Psychologists study how the event can be recognised, and how the described action can be performed.

Do these areas of study overlap?

# The language/sensorimotor interface

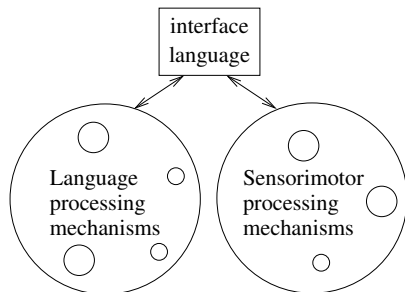
Clearly we can *convert* sensorimotor representations into linguistic ones (and vice versa).

- Because we can talk about what we see;
- Because we can execute verbal instructions.

Question: how much work is involved in this conversion?

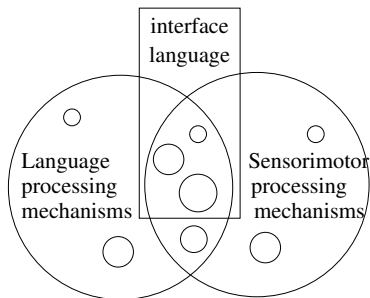
## Two suggestions

Language and SM processing are **modules**



Semantic representations **abstract away** from details of SM processing

Language and SM processing **share mechanisms**



Semantic representations **retain** details of SM processing

# Motivations for the 'shared mechanisms' hypothesis

A priori arguments: good, but vague

- Evolutionary arguments
- Parsimony
- 'Embodied' accounts of meaning

Empirical experiments: specific, but only about single words

- Pulvermüller *et al.* (1999) ERP study: concrete nouns activate visual cortices; action verbs activate motor/premotor cortices
- Pulvermüller *et al.* (2005): TMS over hand and leg areas differentially affects processing of hand and leg verbs
- Pulvermüller *et al.* (2005b) MEG study: leg and mouth verbs activate corresponding areas of motor cortex

# Compositionality in language and SM cognition

It's not surprising that individual words access SM representations.

A more interesting question: what's the relationship between **combinatorial mechanisms** in language and in the SM domain?

- There must be mechanisms for combining primitive **SM** concepts representing actions and objects to create event representations.
- Q: Do these bear any resemblance to syntactic mechanisms in language?

# Methodology 1: starting with concrete sentences

My approach: start by looking at sentences which describe events we can apprehend (or bring about) using ‘well-understood’ SM mechanisms.

We know a lot about the SM mechanisms involved in:

- attention to / categorisation of concrete objects
- control/perception of reach-to-grasp actions

So *The man grabbed a cup* is a good place to start.

## Methodology 2: comparing syntactic and SM models

My approach:

- Choose a syntactic model of the cup-grabbing sentence (motivated on purely linguistic grounds)
- Develop a SM model of the processes involved in grabbing a cup, or in observing a cup-grabbing event (motivated from SM psychology / neuroscience)
- Look for **formal similarities** between these models.

If there are nontrivial similarities, this is support for the 'shared mechanisms' hypothesis.

## An objection

Surely syntactic reps can't be closely related to SM mechanisms, because syntax is orthogonal to concreteness.

*The man grabbed a cup*

*The company acquired a subsidiary. . .*

Linking syntax to SM mechanisms forces us to adopt a particular account of the semantics of abstract sentences, in which abstract concepts are grounded in concrete SM ones.

- There is enough support for this account to make it worth looking *in detail* at concrete sentences.

## Another objection

Surely syntactic reps can't be closely related to SM mechanisms, because SM mechanisms are *universal*, but syntactic structures (on the surface) are not.

To maintain a strong version of the shared mechanisms hypothesis, we must adopt a syntactic theory that posits an 'underlying' level of syntactic structure:

- which is reasonably invariant over translation;
- which represents the meaning of sentences.

This idea of an 'underlying' level of syntactic structure is a cornerstone of the **generative grammar** theories developed by Noam Chomsky.

# Summary of the course

Part 1: I'll introduce a model of the SM processes involved in experiencing a reach-to-grasp action.

Part 2: I'll introduce a version of a recent syntactic theory of Chomsky's, called **Minimalism** (Chomsky, 1995).

Part 3: I'll argue that the underlying syntactic structure of *The man grabbed a cup* (according to Minimalism) can be understood as a description of the SM processes involved in experiencing the described event.

- Then I'll present a neural network model of language learning and language production.

# Today's lecture: an overview of the argument

- 1 A Minimalist account of *The man grabbed a cup*
- 2 A sensorimotor model of reach-to-grasp actions
- 3 A SM characterisation of LF (Version 1)
- 4 A working memory model of reach-to-grasp actions
- 5 A SM characterisation of LF (Version 2)

# Outline of today's lecture

- 1 A Minimalist account of *The man grabbed a cup*
- 2 A sensorimotor model of reach-to-grasp actions
- 3 A SM characterisation of LF (Version 1)
- 4 A working memory model of reach-to-grasp actions
- 5 A SM characterisation of LF (Version 2)

# Phonetic form and logical form

In Minimalism, a sentence needs to be described at two different syntactic levels.

- **Phonetic form (PF)** represents the surface form of the sentence.
- **Logical form (LF)** is ‘the level of syntactic representation which interfaces with the semantic system’.

I’ll assume that all translations of *The man grabbed a cup* have the same LF.

# Components of the Minimalist model

Both LF and PF are composed of **phrases**.

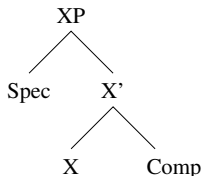
The Minimalist model provides:

- A definition of what phrases are
- A generative mechanism for forming (and altering) phrases.

The generative mechanism (supplemented by language-specific parameters) generates the sentences in a given language.

# Definition of phrases

The basic unit of phrase structure is the **X-bar schema**.



Each word in a sentence contributes an XP structure.

- The **head** of the structure (X) is the word itself.
- The structure also has slots for a **specifier** (Spec) and a **complement** (Comp).
- These slots can (recursively) be occupied by other XPs.

# The generative mechanism

The generative mechanism builds a **phrase structure** (out of XPs).

There are two sorts of operation:

- In **phrase formation operations**, XPs are created and joined together.
- In **movement operations**, elements move from one position in the phrase structure to another.

# Generating PF and LF structures

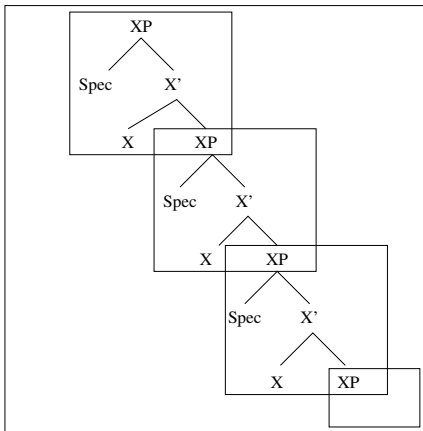
The sequence of movement operations is divided into two subsequences.

- First there's a sequence of **overt movement operations**
- Then there's a sequence of **covert movement operations**.

The PF of a sentence is **read off** the phrase structure after the overt movement operations. The LF is created by executing the remaining covert movement operations.

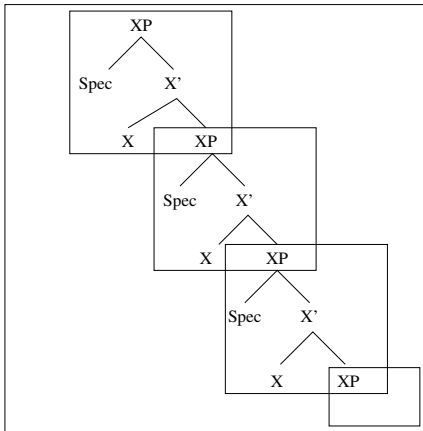
# An example sentence: *The man grabbed a cup*

Phrase formation operations create the following structure:



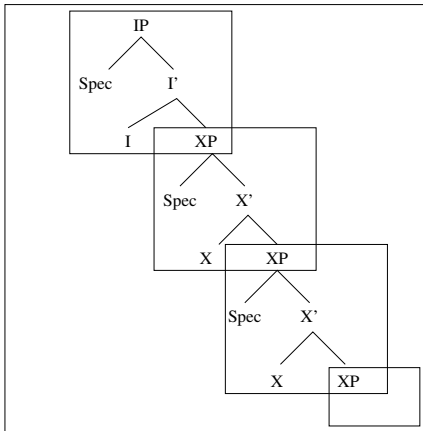
# An example sentence: *The man grabbed a cup*

Each XP is of a different type.



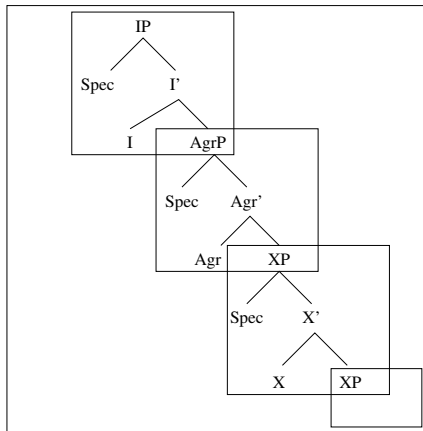
An example sentence: *The man grabbed a cup*

IP is associated with the subject. . .



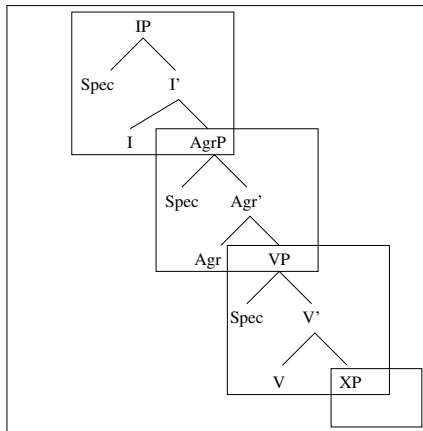
# An example sentence: *The man grabbed a cup*

AgrP is associated with the object. . .



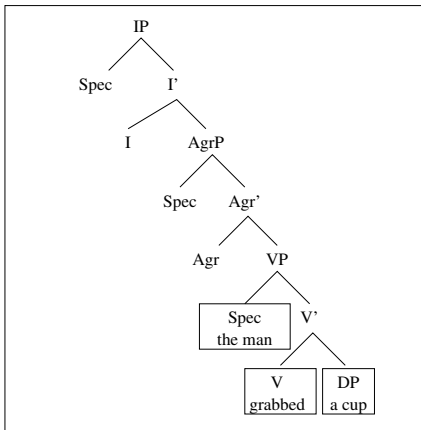
# An example sentence: *The man grabbed a cup*

VP is associated with the verb.



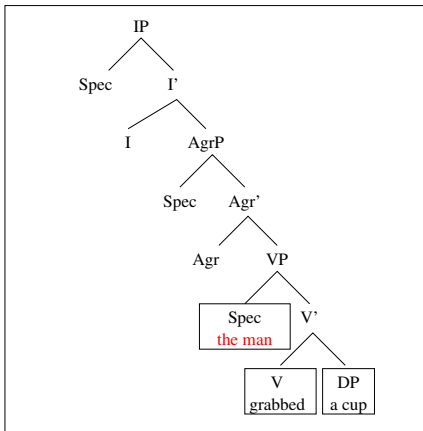
# An example sentence: *The man grabbed a cup*

VP is headed by a verb, and introduces two **DPs**:



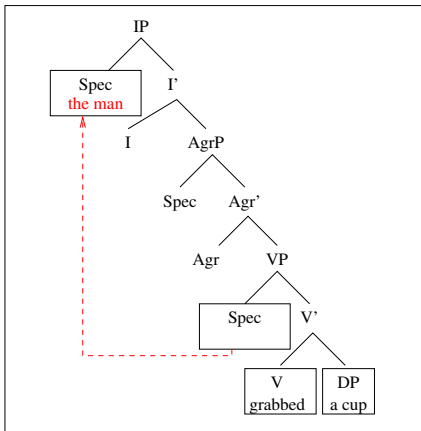
# An example sentence: *The man grabbed a cup*

Movement operations are of two types. One is **DP movement**:



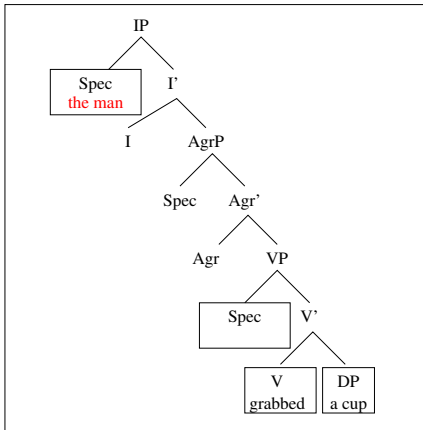
# An example sentence: *The man grabbed a cup*

The subject DP raises to [Spec,IP] to get **Case** (assigned by I)



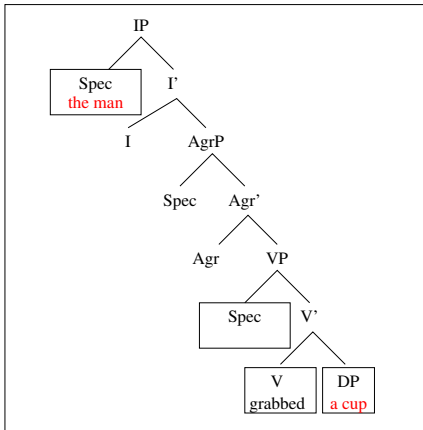
# An example sentence: *The man grabbed a cup*

The subject DP raises to [Spec,IP] to get **Case** (assigned by I)



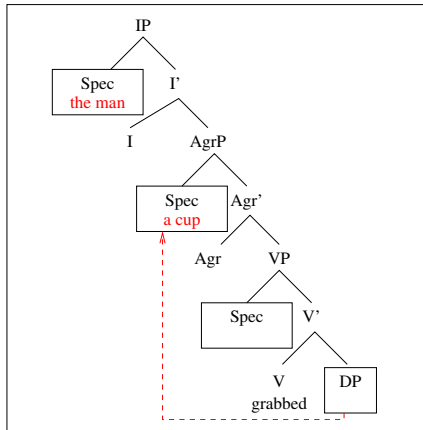
# An example sentence: *The man grabbed a cup*

The object DP raises to [Spec,AgrP] to get **Case** (assigned by Agr)



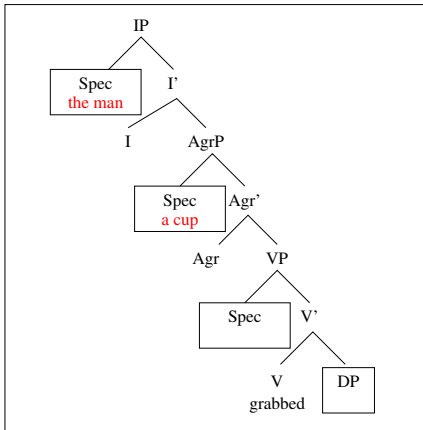
# An example sentence: *The man grabbed a cup*

The object DP raises to [Spec,AgrP] to get **Case** (assigned by Agr)



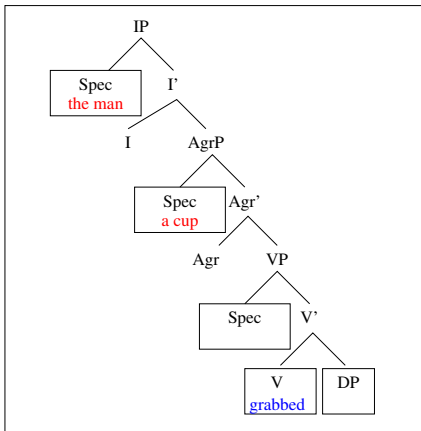
# An example sentence: *The man grabbed a cup*

The object DP raises to [Spec,AgrP] to get **Case** (assigned by Agr)



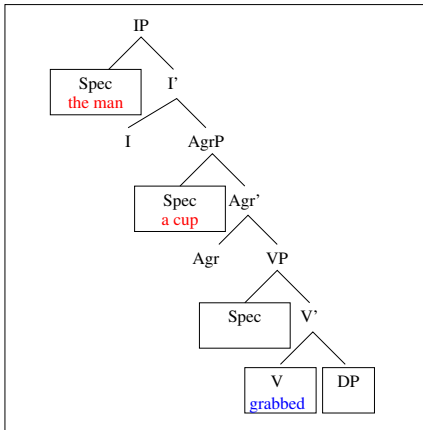
# An example sentence: *The man grabbed a cup*

The other type of movement is **head movement**.



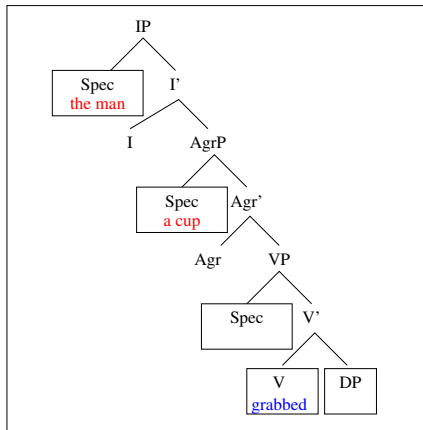
# An example sentence: *The man grabbed a cup*

The verb has **morphology** agreeing with subject and object.



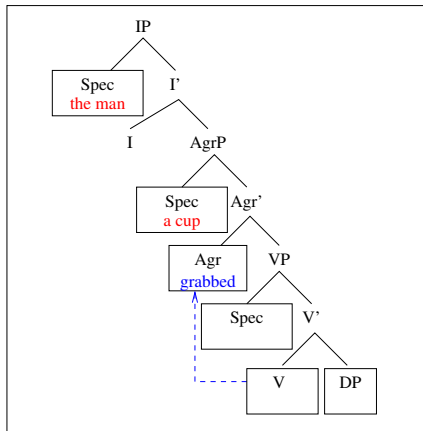
# An example sentence: *The man grabbed a cup*

It raises to the heads of AgrP and IP to 'check' this morphology.



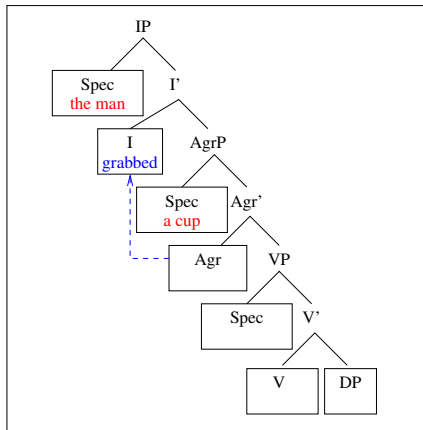
# An example sentence: *The man grabbed a cup*

It raises to the heads of AgrP and IP to 'check' this morphology.



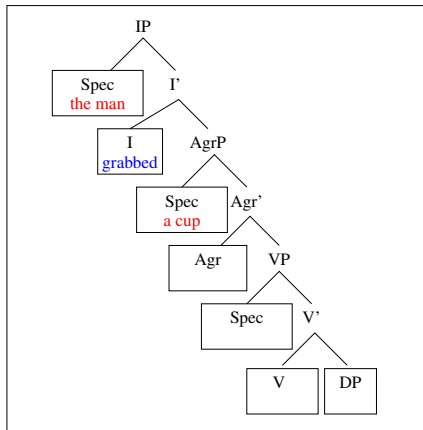
# An example sentence: *The man grabbed a cup*

It raises to the heads of AgrP and IP to 'check' this morphology.



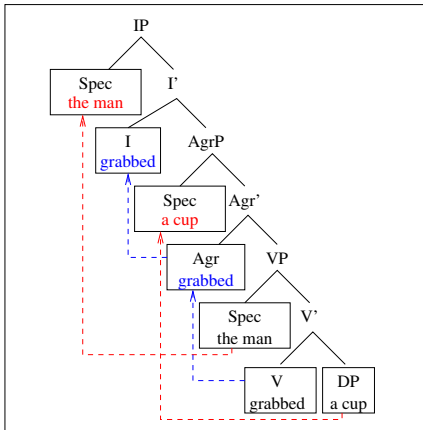
# An example sentence: *The man grabbed a cup*

It raises to the heads of AgrP and IP to 'check' this morphology.



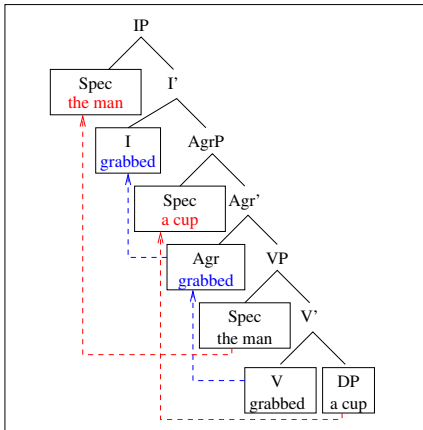
# An example sentence: *The man grabbed a cup*

Moved elements remain linked to their previous position(s):



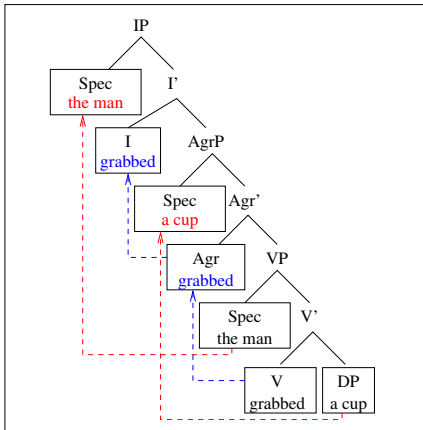
An example sentence: *The man grabbed a cup*

Some movement operations are **overt**, and some are **covert**:



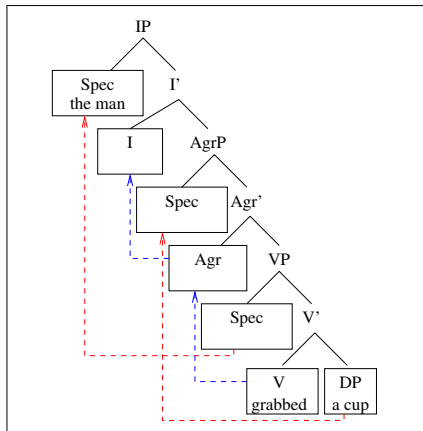
# An example sentence: *The man grabbed a cup*

Languages have different patterns of overt/covert movement:



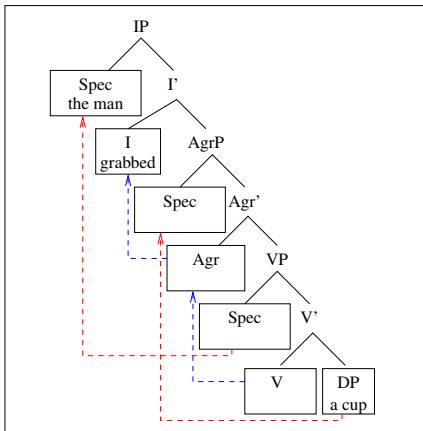
# An example sentence: *The man grabbed a cup*

English PF looks like this:



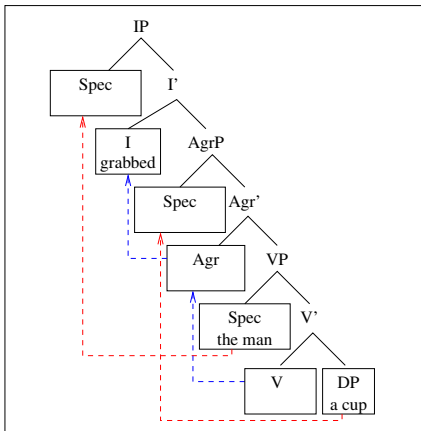
An example sentence: *The man grabbed a cup*

French/Italian PF looks like this:



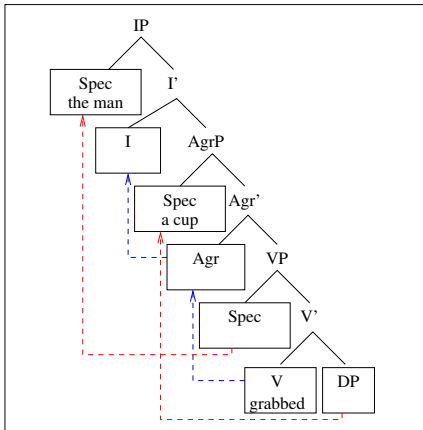
An example sentence: *The man grabbed a cup*

Māori PF looks like this:



# An example sentence: *The man grabbed a cup*

(Perhaps) Japanese PF looks like this:



# Outline of today's lecture

- 1 A Minimalist account of *The man grabbed a cup*
- 2 A sensorimotor model of reach-to-grasp actions**
- 3 A SM characterisation of LF (Version 1)
- 4 A working memory model of reach-to-grasp actions
- 5 A SM characterisation of LF (Version 2)

# Preliminaries for a SM model of reach-to-grasp events

We need a model which covers both perception and execution of reach-to-grasp actions.

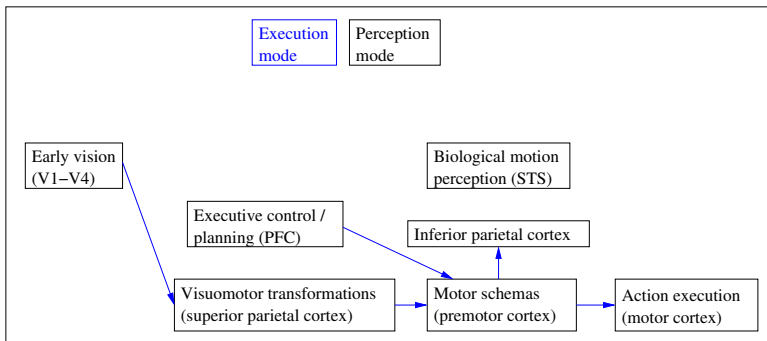
Note:

- Syntax does not signal whether a sentence reports an action from the perspective of the agent or of an observer.
- If syntax is a description of SM processing, this means the models of action perception and action execution must be very similar.
- This is very much in line with current SM models.

*The **mirror system hypothesis**: the neural apparatus used to perceive a reach-to-grasp action overlaps extensively with the apparatus used to execute the action.*

# Neural circuitry for grasp action execution/perception

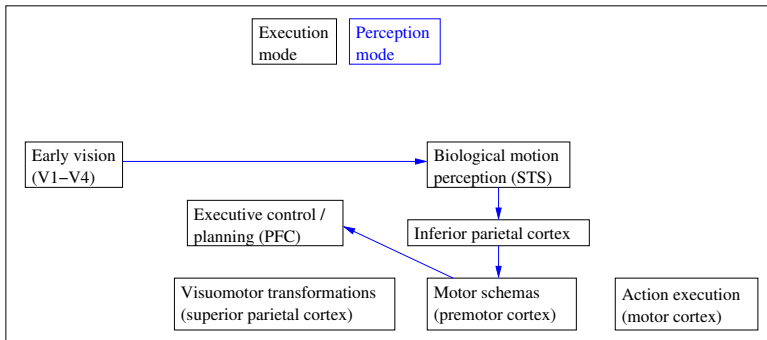
While perceiving and executing a reach/grasp action make use of the same *neural representations*, they differ in the way these representations are *connected*.



FARS (1998), Iacoboni (2001), Oztop & Arbib (2002), Rizzolatti & Matelli (2003), Keysers & Perrett (2004)

# Neural circuitry for grasp action execution/perception

While perceiving and executing a reach/grasp action make use of the same *neural representations*, they differ in the way these representations are *connected*.



FARS (1998), Iacoboni (2001), Oztop & Arbib (2002), Rizzolatti & Matelli (2003), Keysers & Perrett (2004)

## Action perception / action execution modes

How does an 'observer' know whether a motor programme evoked in motor cortex / PFC is being observed or executed?

Proposal: before an observer can use the mirror system, he must configure it for action observation or action execution.

- I.e. he must decide whether to *do* an action or *observe* an action.

## Action perception / action execution modes

How does an ‘observer’ know whether a motor programme evoked in motor cortex / PFC is being observed or executed?

Proposal: before an observer can use the mirror system, he must configure it for action observation or action execution.

- I.e. he must decide whether to *do* an action or *observe* an action.

Proposal: deciding to observe and deciding to act are both forms of ‘attention to the agent’.

- ‘Deciding to observe’ happens when we attend to an external agent.
- ‘Deciding to act’ is a form of ‘attention to oneself’.

# Operation 1: a 'mode-setting' attention to the agent

Proposal: observing or executing a reach-to-grasp action must have sequential structure.

The first operation is an action of attention to the agent.

The meaning of subsequent activity in the mirror system is 'deictically referred' to the prior mode-setting operation. (C.f. Ballard *et al*, 1997.)

## Operation 2: attention to the cup

An agent must attend to a target object before it can be reached for.

## Operation 2: attention to the cup

An agent must attend to a target object before it can be reached for.

An observer watching a reach-to-grasp action probably attends to the target before 'classifying the action'.

- Empirical evidence: observers watching an agent reach for a target saccade to the target well before the agent's hand reaches it (Flanagan & Johansson, 2003)
- Computational models of action recognition tend to assume knowledge of target object location/affordances (see e.g. Oztop & Arbib)



## Operation 3: execution/monitoring of a motor programme

Operations 1 and 2 are 'actions of attention', to agent and patient.

Operation 3 is a motor action. (Or the simulation of one.)

- It is extended in time.
- It has the character of a dynamical system.
- It probably involves forward and inverse internal models.

## Endpoint of the reach-to-grasp action

At the end of a (successful) action, the agent has a stable grasp on the cup.

Note that this is a new *attentional* state, as well as a new world state.

- A grasp action is an **attentional action** as well as a ‘substantive’ action.
- A stable grasp provides an opportunity to *integrate* haptic and visual modalities.

Visual perception of contact also triggers learned haptic associations (see e.g. Keyers, 2004).

# Actions and their reafferent feedback

A SM action signal should be distinguished from the perceptual signals which result from its execution.

For attentional actions:

- An action signal might involve the posting of a ‘search goal’, and a direction of attention.
- Reafferent feedback is a (bottom-up) sensory representation of an object.

For motor actions:

- An action signal might involve the activation of a motor programme in (pre)motor cortex or PFC;
- Reafferent feedback might include a representation of the agent *as an agent*.

# Summary

SM 'experience' of a reach-to-grasp action has a characteristic serial structure, consisting of a sequence of attentional/motor operations, evoking a sequence of reafferent attentional states.

The structure is similar whether the action is perceived or executed.

I'll summarise the sequence in each case.

# The sequence for perception of a reach-to-grasp


# The sequence for perception of a reach-to-grasp

Operation 1     $O$  attends to an external agent, configuring his mirror system circuit for action perception.

# The sequence for perception of a reach-to-grasp

Operation 1    *O* attends to an external agent, configuring his mirror system circuit for action perception.

State 1        *O* receives reafferent feedback from this operation; the percept 'man'.

# The sequence for perception of a reach-to-grasp

Operation 1	<i>O</i> attends to an external agent, configuring his mirror system circuit for action perception.
State 1	<i>O</i> receives reafferent feedback from this operation; the percept 'man'.
Operation 2	<i>O</i> establishes joint attention with the agent, and attends to another object (the cup).

# The sequence for perception of a reach-to-grasp

Operation 1	<i>O</i> attends to an external agent, configuring his mirror system circuit for action perception.
State 1	<i>O</i> receives reafferent feedback from this operation; the percept 'man'.
Operation 2	<i>O</i> establishes joint attention with the agent, and attends to another object (the cup).
State 2	<i>O</i> receives feedback from this operation: the percept 'cup'.

# The sequence for perception of a reach-to-grasp

Operation 1	<i>O</i> attends to an external agent, configuring his mirror system circuit for action perception.
State 1	<i>O</i> receives reafferent feedback from this operation; the percept 'man'.
Operation 2	<i>O</i> establishes joint attention with the agent, and attends to another object (the cup).
State 2	<i>O</i> receives feedback from this operation: the percept 'cup'.
Operation 3	<i>O</i> initiates a process of biological motion classification, which results in the action 'grab' being activated in <i>O</i> 's premotor cortex.

# The sequence for perception of a reach-to-grasp

Operation 1	$O$ attends to an external agent, configuring his mirror system circuit for action perception.
State 1	$O$ receives reafferent feedback from this operation; the percept 'man'.
Operation 2	$O$ establishes joint attention with the agent, and attends to another object (the cup).
State 2	$O$ receives feedback from this operation: the percept 'cup'.
Operation 3	$O$ initiates a process of biological motion classification, which results in the action 'grab' being activated in $O$ 's premotor cortex.
State 3	As a corollary of this process, $O$ re-attends to the agent <i>as an agent</i> .

# The sequence for perception of a reach-to-grasp

Operation 1	<i>O</i> attends to an external agent, configuring his mirror system circuit for action perception.
State 1	<i>O</i> receives reafferent feedback from this operation; the percept 'man'.
Operation 2	<i>O</i> establishes joint attention with the agent, and attends to another object (the cup).
State 2	<i>O</i> receives feedback from this operation: the percept 'cup'.
Operation 3	<i>O</i> initiates a process of biological motion classification, which results in the action 'grab' being activated in <i>O</i> 's premotor cortex.
State 3	As a corollary of this process, <i>O</i> re-attends to the agent <i>as an agent</i> .
State 4	<i>O</i> re-attends to the cup, in the course of perceiving the agent establishing a stable grasp.

# The sequence for execution of a reach-to-grasp


# The sequence for execution of a reach-to-grasp

Operation 1    *A* attends to himself, configuring his mirror system circuit for action execution.

# The sequence for execution of a reach-to-grasp

Operation 1	<i>A</i> attends to himself, configuring his mirror system circuit for action execution.
State 1	<i>A</i> receives reafferent feedback that this operation succeeded.

# The sequence for execution of a reach-to-grasp

Operation 1	<i>A</i> attends to himself, configuring his mirror system circuit for action execution.
State 1	<i>A</i> receives reafferent feedback that this operation succeeded.
Operation 2	<i>A</i> selects an object to reach for (the cup), and hence executes an action of attention to the cup.

# The sequence for execution of a reach-to-grasp

Operation 1	<i>A</i> attends to himself, configuring his mirror system circuit for action execution.
State 1	<i>A</i> receives reafferent feedback that this operation succeeded.
Operation 2	<i>A</i> selects an object to reach for (the cup), and hence executes an action of attention to the cup.
State 2	<i>A</i> receives feedback from this operation: the percept 'cup'.

# The sequence for execution of a reach-to-grasp

Operation 1	A attends to himself, configuring his mirror system circuit for action execution.
State 1	A receives reafferent feedback that this operation succeeded.
Operation 2	A selects an object to reach for (the cup), and hence executes an action of attention to the cup.
State 2	A receives feedback from this operation: the percept 'cup'.
Operation 3	A selects an action category ('grab') and begins to execute the grab action.

# The sequence for execution of a reach-to-grasp

Operation 1	A attends to himself, configuring his mirror system circuit for action execution.
State 1	A receives reafferent feedback that this operation succeeded.
Operation 2	A selects an object to reach for (the cup), and hence executes an action of attention to the cup.
State 2	A receives feedback from this operation: the percept 'cup'.
Operation 3	A selects an action category ('grab') and begins to execute the grab action.
State 3	A receives feedback from the process of action execution, which includes a percept of himself 'as an agent'.

# The sequence for execution of a reach-to-grasp

Operation 1	A attends to himself, configuring his mirror system circuit for action execution.
State 1	A receives reafferent feedback that this operation succeeded.
Operation 2	A selects an object to reach for (the cup), and hence executes an action of attention to the cup.
State 2	A receives feedback from this operation: the percept 'cup'.
Operation 3	A selects an action category ('grab') and begins to execute the grab action.
State 3	A receives feedback from the process of action execution, which includes a percept of himself 'as an agent'.
State 4	A re-establishes the cup in the haptic modality.

# SM signals during experience of a reach-to-grasp

	Transient signals	
	Action signals	Reafferent signals
	<i>attend_agent</i>	<i>attending_to_agent</i>
	<i>attend_cup</i>	<i>attending_to_cup</i>
	<i>grasp</i>	<i>attending_to_agent</i>
		<i>attending_to_cup</i>

# Outline of today's lecture

- 1 A Minimalist account of *The man grabbed a cup*
- 2 A sensorimotor model of reach-to-grasp actions
- 3 A SM characterisation of LF (Version 1)**
- 4 A working memory model of reach-to-grasp actions
- 5 A SM characterisation of LF (Version 2)

# A SM characterisation of LF (Version 1)

My proposal (roughly):

*The LF of 'The man grabbed a cup' describes the sequence of SM operations/states evoked during experience of a cup-grabbing event.*

- Each XP denotes a SM operation, and its reafferent consequence. (These might be thought of jointly as a 'context-update' operation.)
- A right-branching structure of XPs denotes a sequence of these operations.

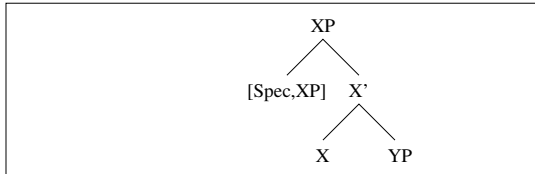
# SM interpretation of a single XP schema

- Each XP denotes a SM operation, and its reafferent consequence. (These might be thought of jointly as a ‘context-update’ operation.)

# SM interpretation of a single XP schema

- Each XP denotes a SM operation, and its refferent consequence. (These might be thought of jointly as a ‘context-update’ operation.)

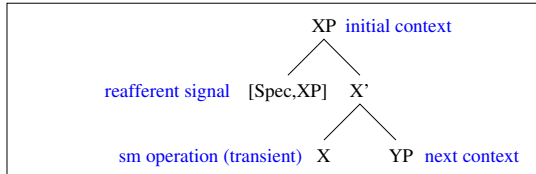
Each item in the XP schema has a SM interpretation.



# SM interpretation of a single XP schema

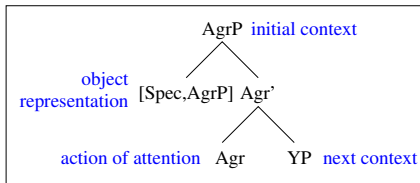
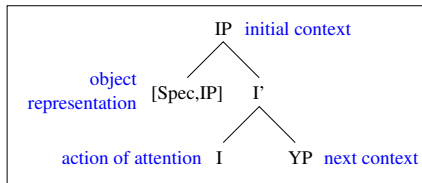
- Each XP denotes a SM operation, and its reafferent consequence. (These might be thought of jointly as a 'context-update' operation.)

Each item in the XP schema has a SM interpretation.



# XPs denoting attentional operations

IP and AgrP denote attentional operations, and their reafferent consequences.

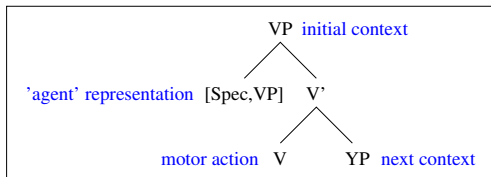


In these XPs:

- The head (which 'checks' a piece of verb morphology) denotes an **action of attention**.
- The specifier (the position occupied by a DP) denotes the sensory representation of an object.

# XPs denoting motor operations

VP denotes a motor operation, and its refferent consequences.



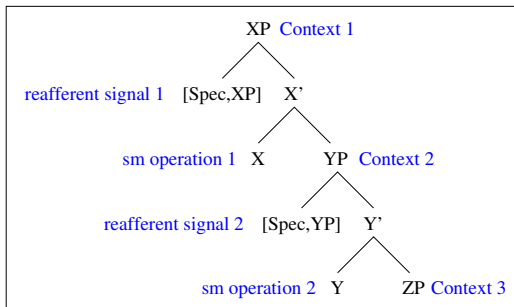
In VP:

- The head (which contributes a verb stem) denotes a **motor programme**.
- The specifier (occupied by another noun phrase) denotes the sensory representation of an agent.

## A right-branching structure of X-bar schemas

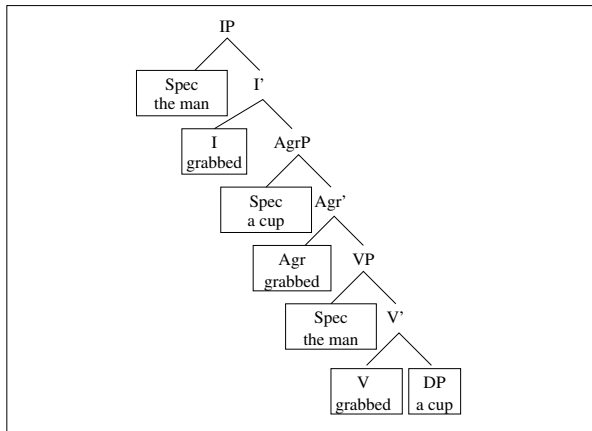
In a right-branching X-bar structure, the next context of one XP is the initial context of its complement XP.

So a right-branching structure of XPs describes *successive iterations* in the sequence.



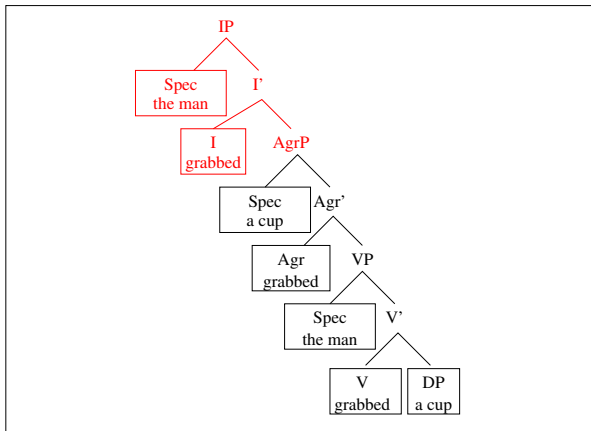
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



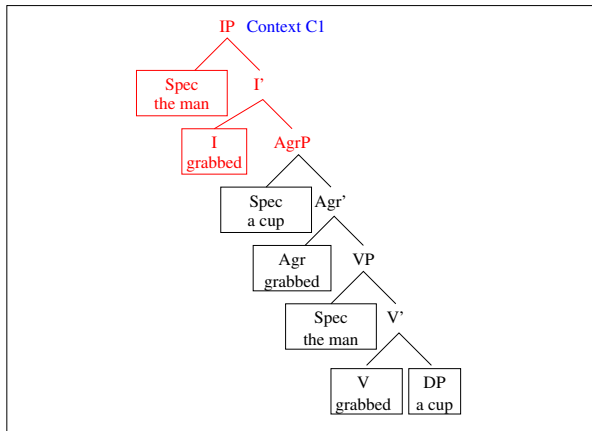
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



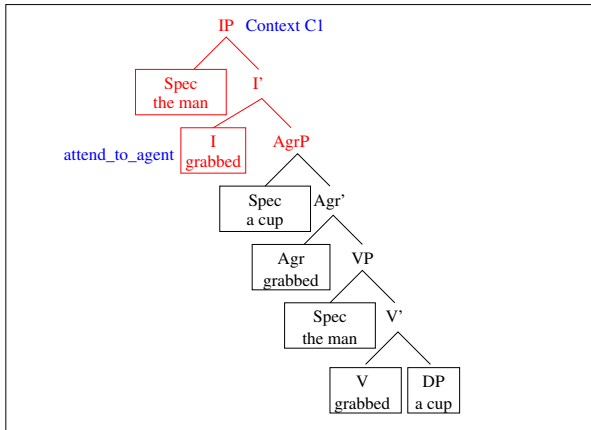
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



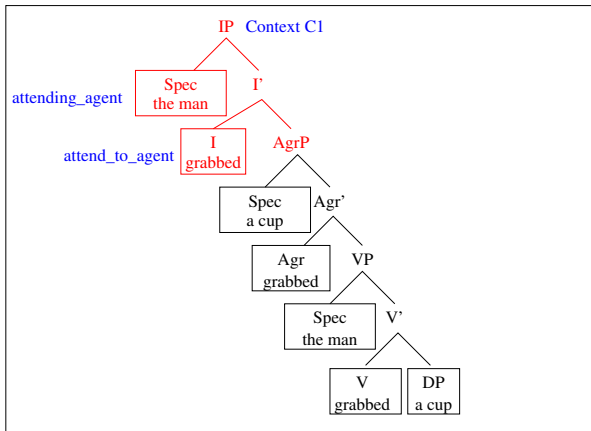
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



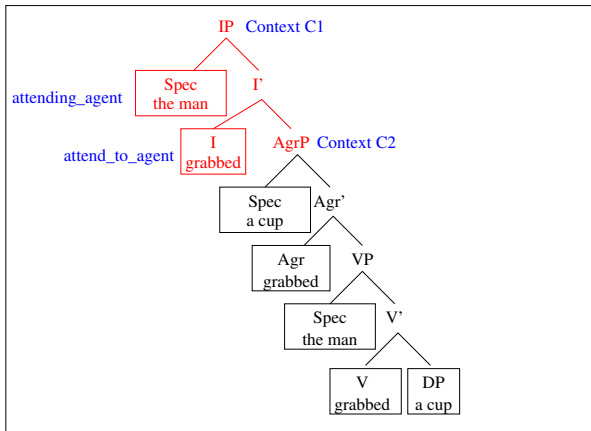
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



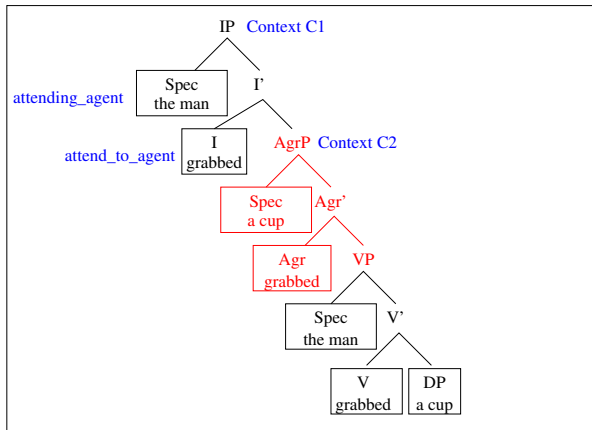
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



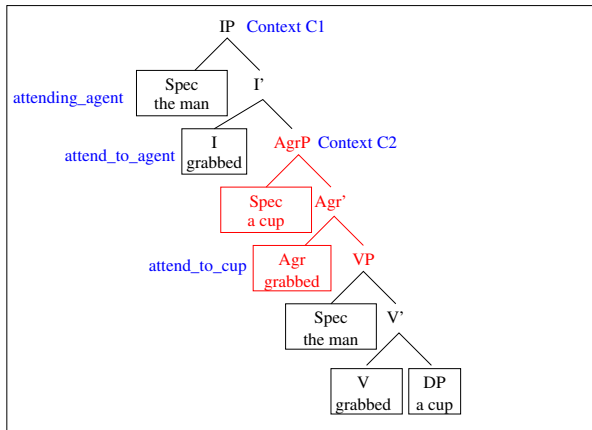
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



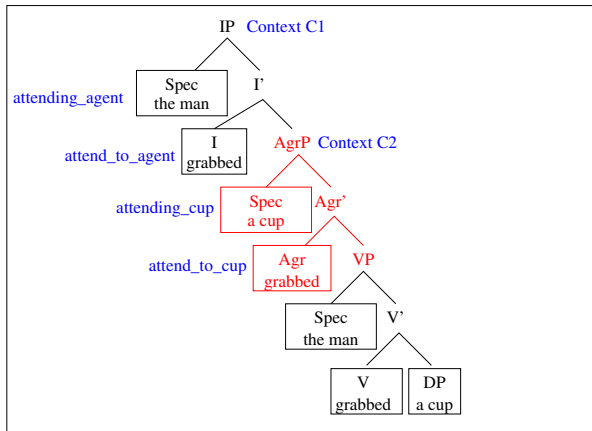
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



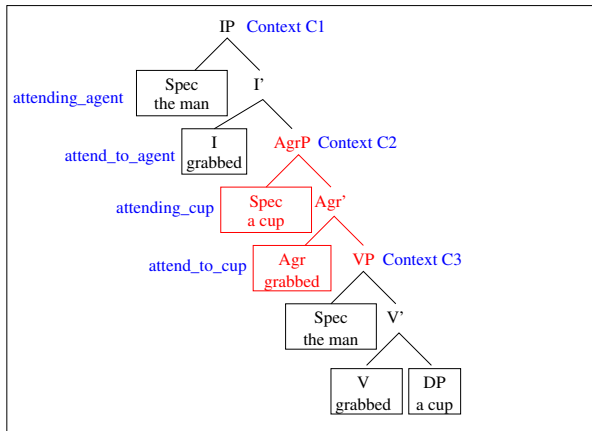
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



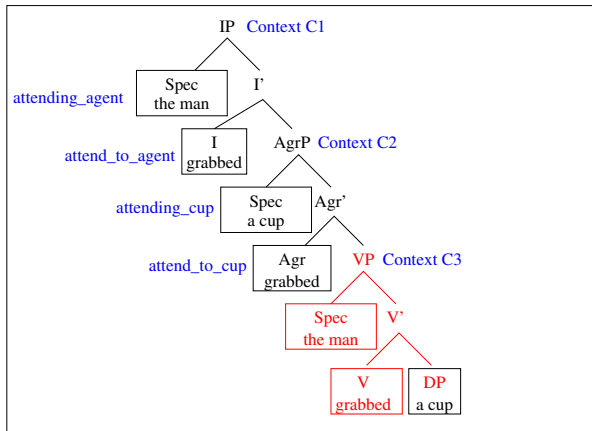
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



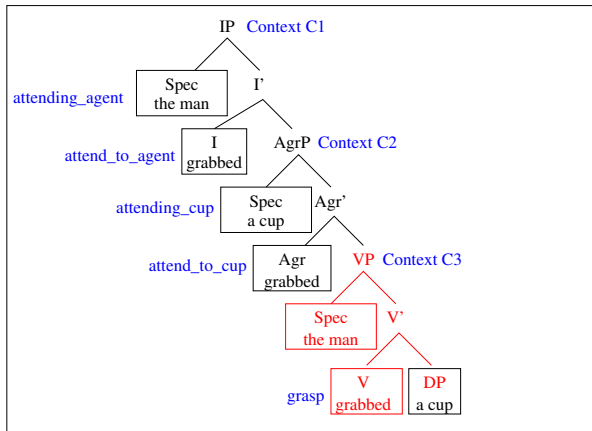
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



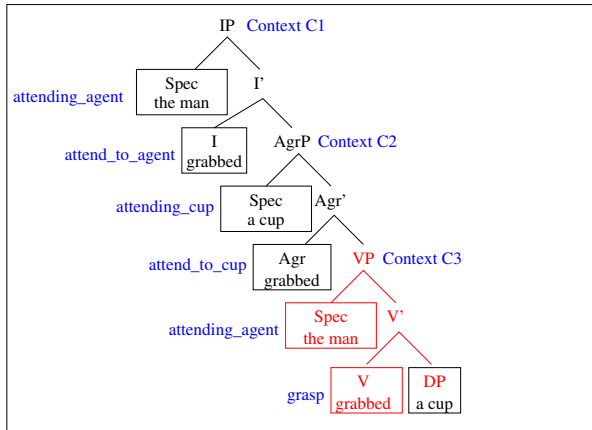
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



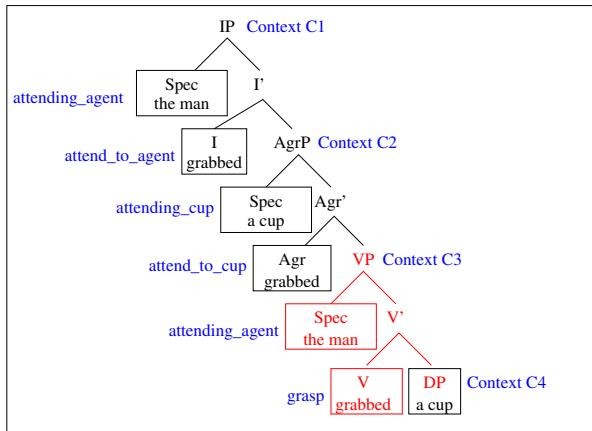
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



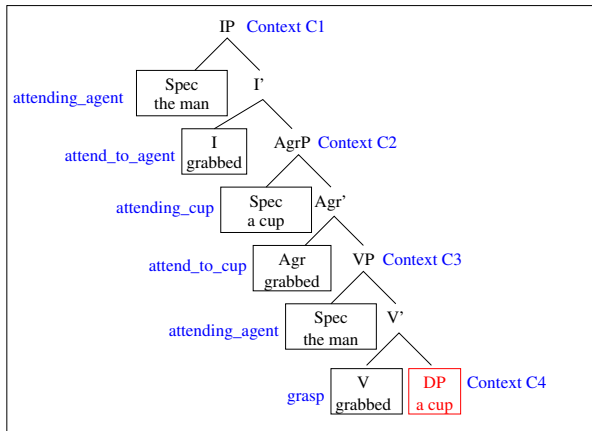
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



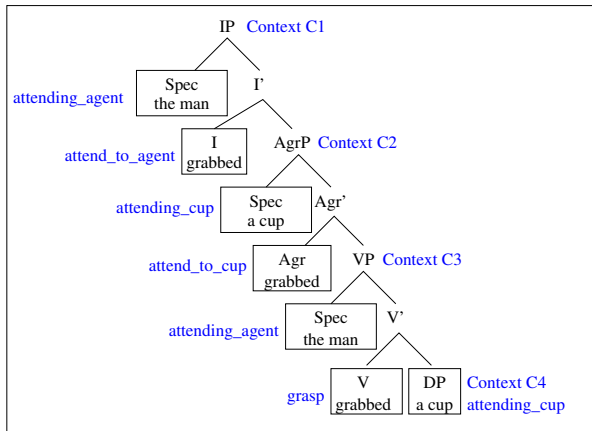
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



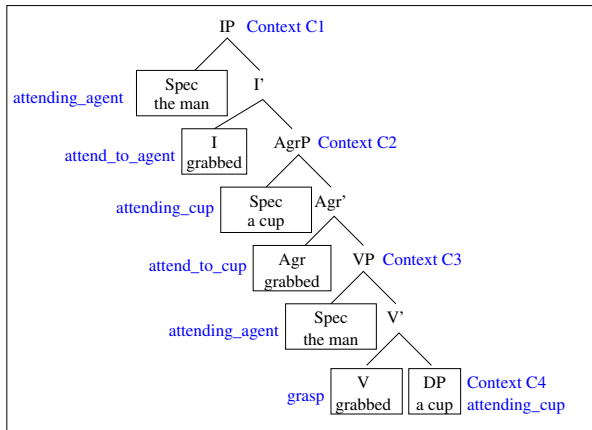
# SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



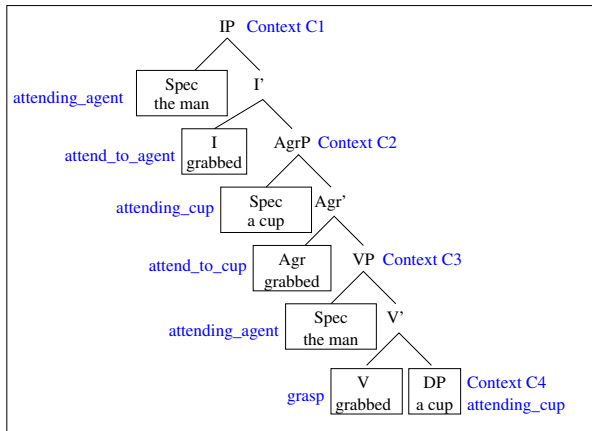
# SM interpretation of the LF of *The man grabbed a cup*

So: what's the SM characterisation of **movement**?



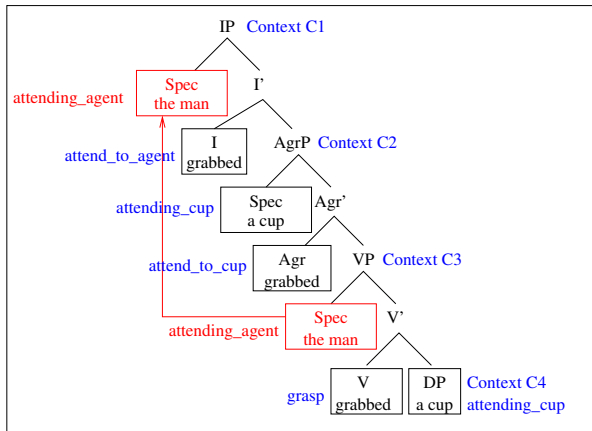
# SM interpretation of the LF of *The man grabbed a cup*

DP raising reflects operations of *re-attention to agent and patient*.



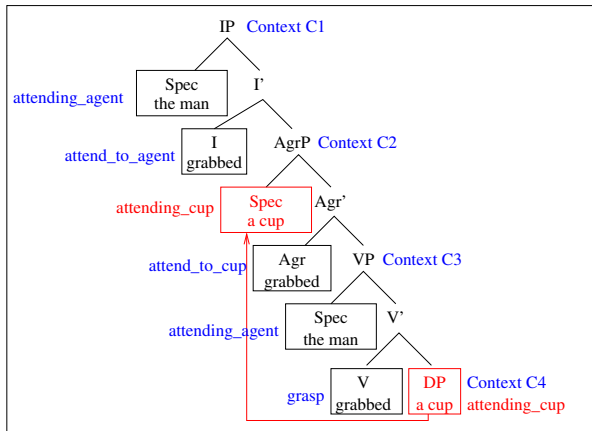
# SM interpretation of the LF of *The man grabbed a cup*

DP raising reflects operations of *re-attention to agent and patient*.



# SM interpretation of the LF of *The man grabbed a cup*

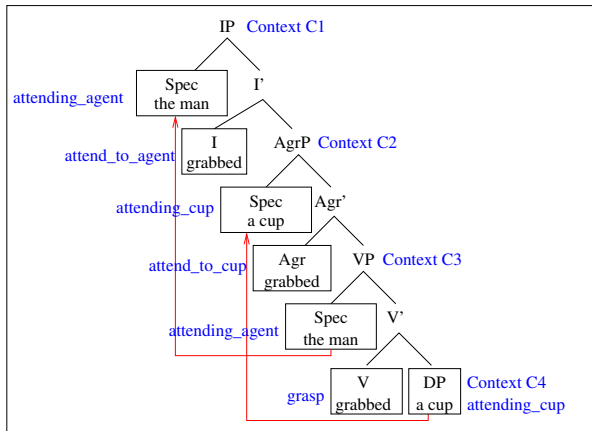
DP raising reflects operations of *re-attention to agent and patient*.



# SM interpretation of the LF of *The man grabbed a cup*

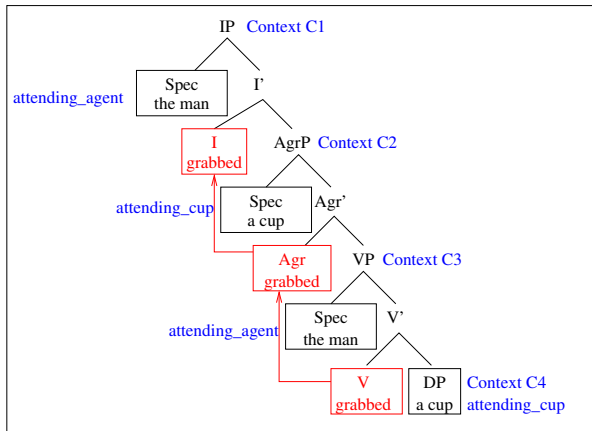
SM interpretation of 'DPs must raise to check Case':

Objects must be attended to before being involved in action monitoring.



# SM interpretation of the LF of *The man grabbed a cup*

To discuss V-Agr-I raising, I must extend the model a little.



# Outline of today's lecture

- 1 A Minimalist account of *The man grabbed a cup*
- 2 A sensorimotor model of reach-to-grasp actions
- 3 A SM characterisation of LF (Version 1)
- 4 A working memory model of reach-to-grasp actions**
- 5 A SM characterisation of LF (Version 2)

# A working memory model of reach-to-grasp actions

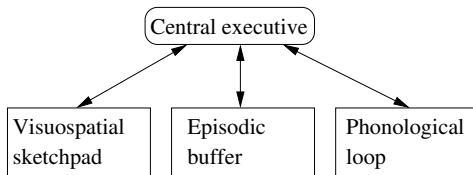
It's far too simple to relate syntactic structure *directly* to SM structure.

- Utterances are not a reflex of SM experience.
- 'Preverbal messages' are held in **working memory (WM)**.

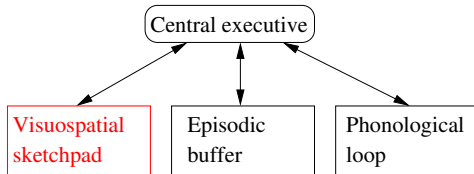
The next few slides outline a model of working memory for experienced events.

- Main idea: a cup-grabbing episode is held in working memory as a *planned SM sequence*, in which all three SM actions are represented in parallel.

# Baddeley's model of WM

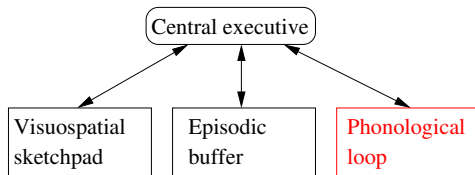


# Baddeley's model of WM



The visuospatial sketchpad: a working memory for **visual patterns**.

# Baddeley's model of WM

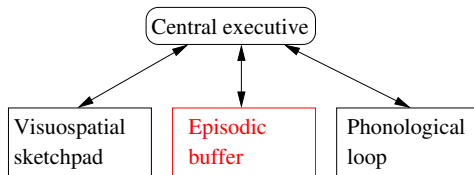


The phonological loop: holds a short sequence of **words or phonemes**.

This is where 'surface utterances' are buffered.

- There are well-established models of how sequences are prepared/rehearsed in the phonological loop.

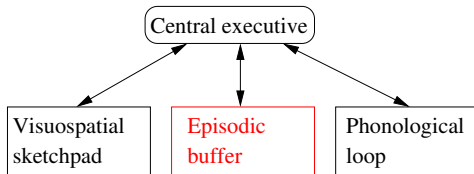
# Baddeley's model of WM



The episodic buffer: a special form of storage for 'episodes'.  
(Baddeley, 2000)

This is where 'preverbal messages' are held.

# Baddeley's model of WM



The episodic buffer encodes episodes as *prepared SM sequences* which can be *internally replayed*.

- The purpose of the internal replay operation is to allow an episode stored in WM to be transmitted to more permanent hippocampal memory (see e.g. Lisman & Otmakhova, 2001; Burgess & Hitch, 2005).
- The hippocampus is widely believed to store sequences, and to support their replay (see e.g. Lee and Wilson, 2002).

## Prepared action sequences in prefrontal cortex

Prepared action sequences are maintained in prefrontal cortex (PFC) (see e.g. Barone & Joseph, 1989; Petrides, 1991).

There are several good neural models of sequential action planning in PFC.

- One common feature of these models is that all the actions in a prepared sequence are active *simultaneously* in PFC, both prior, during and after action execution. (See e.g. Averbeck *et al.*, 2002.)

During internal replay of a prepared sequence in PFC, we will then have a mixture of tonically active and transient SM signals.

## SM signals during replay of the 'cup-grab' sequence

Sustained PFC signal	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>1</sub>	<i>attend_agent</i>	<i>attending_to_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>2</sub>	<i>attend_cup</i>	<i>attending_to_cup</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>3</sub>	<i>grasp</i>	<i>attending_to_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>4</sub>		<i>attending_to_cup</i>

# Outline of today's lecture

- 1 A Minimalist account of *The man grabbed a cup*
- 2 A sensorimotor model of reach-to-grasp actions
- 3 A SM characterisation of LF (Version 1)
- 4 A working memory model of reach-to-grasp actions
- 5 A SM characterisation of LF (Version 2)**

# A SM characterisation of LF (Version 2)

My proposal (more precisely):

*The LF of 'The man grabbed a cup' describes the sequence of SM signals evoked during experience of a cup-grabbing event.*

# A SM characterisation of LF (Version 2)

My proposal (more precisely):

*The LF of 'The man grabbed a cup' describes the sequence of SM signals evoked during experience of a cup-grabbing event **as internally replayed from working memory.***

## A SM characterisation of LF (Version 2)

My proposal (more precisely):

*The LF of 'The man grabbed a cup' describes the sequence of SM signals evoked during experience of a cup-grabbing event **as internally replayed from working memory.***

The idea of a *replayed* sequence permits a SM account of V-Agr-I movement.

# A SM interpretation of V-Agr-I movement

Recall:

- Verb morphology denotes actions of attention to the agent or patient: it 'belongs' at I (or Agr).
- The verb stem denotes a motor action: it 'belongs' at V.

So how come:

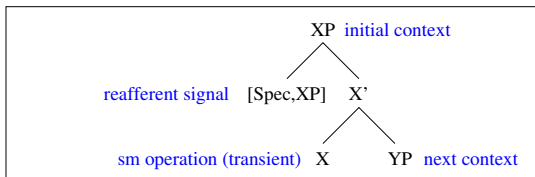
- the verb morphology can be *attached to* the verb?
- the inflected verb can appear at I *or* at V?

Attentional/motor actions seem able to appear *out of sequence*.

# A SM interpretation of V-Agr-I movement

Proposal:

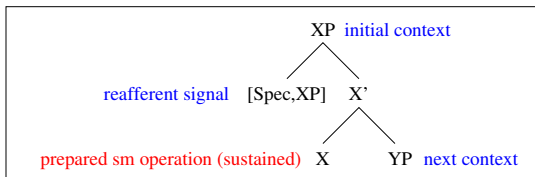
- SM operations (attentional and motor) feature in LF as *prepared operations*, rather than transient signals.



# A SM interpretation of V-Agr-I movement

Proposal:

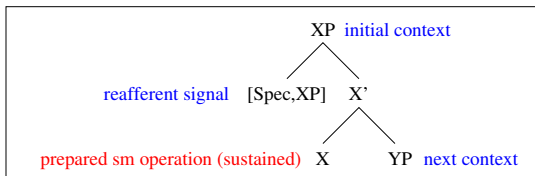
- SM operations (attentional and motor) feature in LF as *prepared operations*, rather than transient signals.



# A SM interpretation of V-Agr-I movement

Proposal:

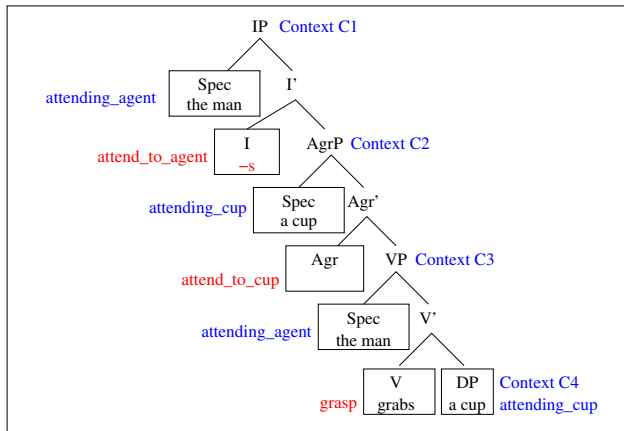
- SM operations (attentional and motor) feature in LF as *prepared operations*, rather than transient signals.



- Note: each prepared operation is active at each stage of the replayed sequence.

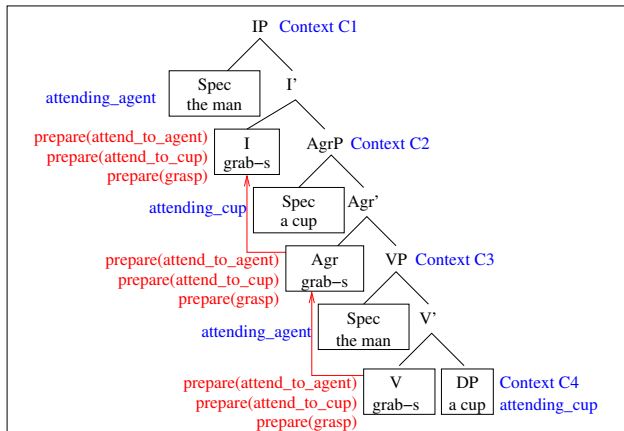
# A SM interpretation of V-Agr-I movement

Original definition:



# A SM interpretation of V-Agr-I movement

New definition, supporting head movement:



# Deciding when to read out phonological items

Here's the sequence of SM signals again:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_1$	<i>attend_agent</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_3$	<i>grasp</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_4$		<i>attending_cup</i>

# Deciding when to read out phonological items

Only some of these SM areas have interfaces with phonology:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_1$	<i>attend_agent</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_3$	<i>grasp</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_4$		<i>attending_cup</i>

# Deciding when to read out phonological items

A selective read-out for English SVO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_1$	<i>attend_agent</i>	<b>man</b>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_3$	<i>grasp</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_4$		<i>attending_cup</i>

# Deciding when to read out phonological items

A selective read-out for English SVO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_1$	<i>attend_agent</i>	<b>man</b>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<b>grabs</b> ↓ ↓	$C_3$	<i>grasp</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_4$		<i>attending_cup</i>

# Deciding when to read out phonological items

A selective read-out for English SVO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_1$	<i>attend_agent</i>	<b>man</b>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<b>grabs</b> ↓ ↓	$C_3$	<i>grasp</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_4$		<b>cup</b>

# Deciding when to read out phonological items

A selective read-out for Māori VSO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_1$	<i>attend_agent</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_3$	<i>grasp</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	$C_4$		<i>attending_cup</i>

# Deciding when to read out phonological items

A selective read-out for Māori VSO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>1</sub>	<i>attend_agent</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>2</sub>	<i>attend_cup</i>	<i>attending_cup</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>3</sub>	<i>grasp</i>	<i>attending_agent</i>
<i>plan</i> <sub>attend_agent/attend_cup/grasp</sub> ↓ ↓	<i>C</i> <sub>4</sub>		<i>attending_cup</i>

# Deciding when to read out phonological items

A selective read-out for Māori VSO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<p><b>grabs</b></p> <p>↓</p> <p>↓</p>	$C_1$	<i>attend_agent</i>	<i>attending_agent</i>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_3$	<i>grasp</i>	<i>attending_agent</i>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_4$		<i>attending_cup</i>

# Deciding when to read out phonological items

A selective read-out for Māori VSO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<p><b>grabs</b></p> <p>↓</p> <p>↓</p>	$C_1$	<i>attend_agent</i>	<i>attending_agent</i>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_3$	<i>grasp</i>	<b>man</b>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_4$		<i>attending_cup</i>

# Deciding when to read out phonological items

A selective read-out for Māori VSO structure:

Sustained signals	Transient signals		
	Context signals	Action signals	Reafferent signals
<p><b>grabs</b></p> <p>↓</p> <p>↓</p>	$C_1$	<i>attend_agent</i>	<i>attending_agent</i>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_2$	<i>attend_cup</i>	<i>attending_cup</i>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_3$	<i>grasp</i>	<b>man</b>
<p><i>plan</i><sub>attend_agent/attend_cup/grasp</sub></p> <p>↓</p> <p>↓</p>	$C_4$		<b>cup</b>

# A neural network model of language learning

An infant needs to learn how to read out a rehearsed SM sequence as a sequence of phonological words.

There are two steps to this:

- Learning a mapping from individual SM signals to individual words;
- Learning which SM signals to read out, and which to ignore.

We'll look at the areas of the brain where these processes happen, and develop a neural network model of them.

# Reading

The textbook for the course is one I've just written: *Sensorimotor cognition and natural language syntax*. It's up on the course website.

- For this lecture: Chapter 1.
- For next lecture: Chapter 2 Sections 1-4. (Read in advance!)