



**Australian Government**

**Department of Defence**

Defence Science and  
Technology Organisation

# Lexpresso: a Controlled Natural Language

Adam Saulwick

Defence Science and Technology Organisation

Fourth Workshop on Controlled Natural Language (CNL 2014)  
Galway 20–22 August 2014

**DSTO**



# Outline

Use case

Input / Output & Interaction

System architecture & module functions

Input – CNL Sensor

Deep Lexpresso

Semantic formalism

Output – CNL Effector

Syntactic structures

Semantic structures

Classification examples

PENS applied

Conclusion

# Lexpresso for natural interaction with Consensus



**Australian Government**  
**Department of Defence**  
Defence Science and  
Technology Organisation

- ▶ Bidirectional natural language interface to a prototype, agent-based, high-level information fusion system
- ▶ Lexpresso bridges the natural-language/formal-language gulf
- ▶ Requirement for time-critical updates of information
- ▶ Limitations: domain, expressiveness & coding effort



## Lexpresso's purpose

Human users need to communicate with Consensus in a natural & intuitive way. Lexpresso provides:

- ▶ Bidirectionality – input & generation capabilities
- ▶ Human users able to query current & historical real-world (potentially) far-flung events
- ▶ Answers:
  - formulated as coherent natural English situation reports
  - describe transit or spatiotemporal interaction of observed maritime, land- &/or air-based platforms
  - report social relationships between people inferred from certain text descriptions [ST14]
  - optionally delivered by a Virtual Adviser & coordinated with replayed events on 3D geospatial display



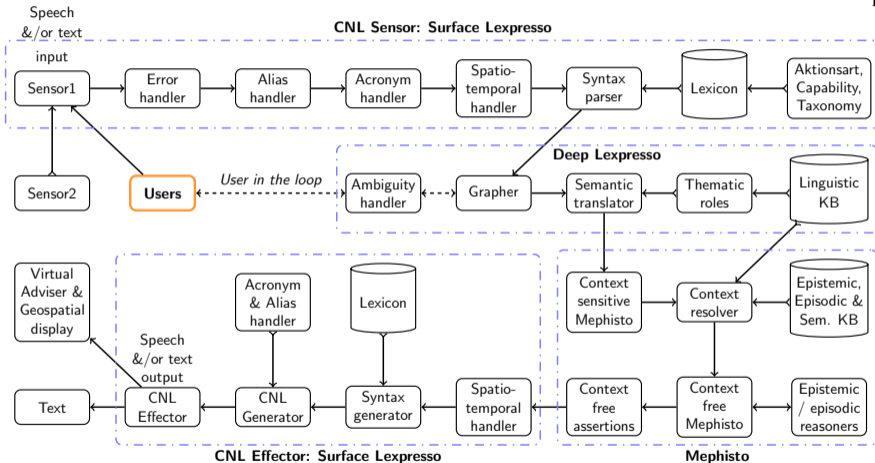
# Input / Output and (distributed) Interaction

- ▶ Humans interact with Virtual Advisers via Lexpresso
- ▶ Virtual Advisers speak in Lexpresso





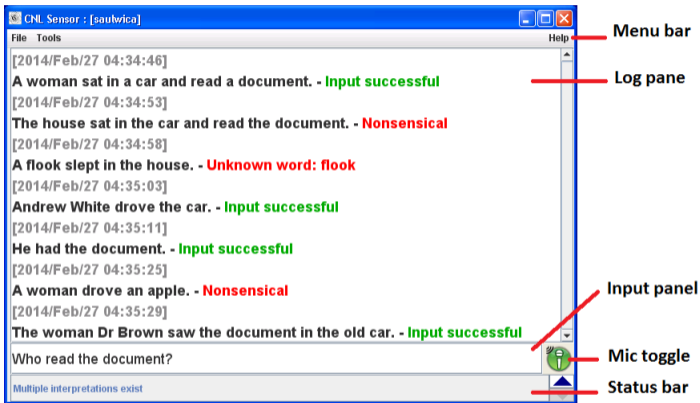
# Modular System Architecture



Lexpresso system architecture



# CNL Sensor UI

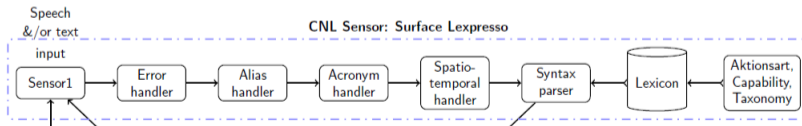


CNL Sensor: showing sample text (with timestamps, proper names, person title & anaphoric resolution), input panel (with possible query), colour-coded feedback messages in log pane, microphone toggle button (on) for speech input & status message



# CNL Sensor Modules

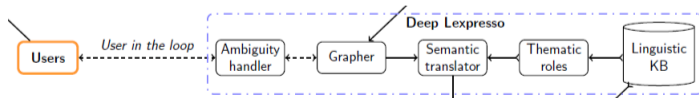
- ▶ Sensor modules process surface language
  - Error handler
  - Alias handler
  - Acronym handler
  - Spatiotemporal handler
  - Syntax parser
  - Lexicon
  - Aktionsarten, Capability, Taxonomy





# Deep Lexpresso

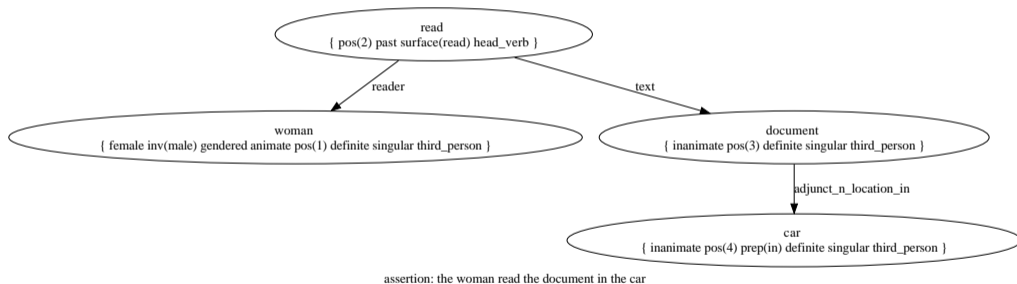
- ▶ Deep modules process ambiguity by drawing on different types of linguistic knowledge
- ▶ Deep syn-sem structure transformations
  - Grapher
  - Ambiguity handler
  - Semantic translator
  - Thematic roles
  - Linguistic knowledgebase



# Graph representation



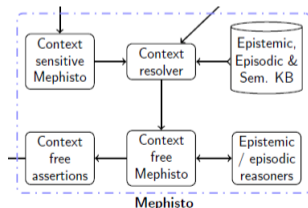
- Structures generated as graphs for disambiguation





# Semantic formalism – inter-language

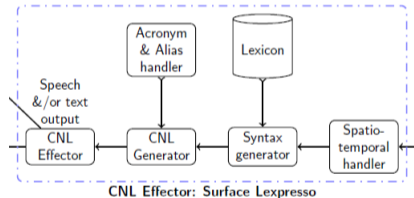
- ▶ Mephisto – formal semantic language independent inter-lingua [LN08, ST14]
- ▶ Five tiered ontology of the domain
- ▶ Propositional logic
- ▶ Modules transform linguistic forms into unambiguous context-free formal semantic correspondences
  - Context sensitive Mephisto
  - Context resolver
  - Epistemic, Episodic & Semantic Knowledgebase & Reasoners
  - Context free Mephisto
  - Context free assertions



# CNL Effector Modules



- ▶ Effector modules generate surface language from Mephisto constructs
- ▶ CNL Output Modules
  - Spatiotemporal handler
  - Syntax Generator
  - CNL Generator
  - CNL Effector



# Syntactic structures

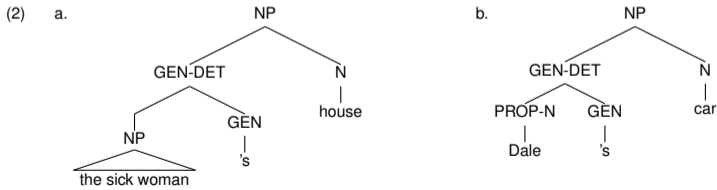
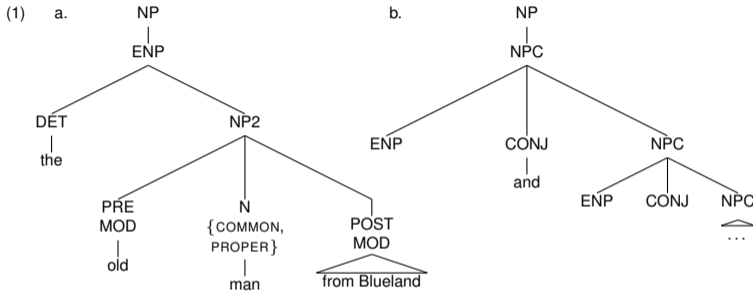


**Australian Government**  
**Department of Defence**  
Defence Science and  
Technology Organisation

- ▶ Possible syntactic structures
  - Declaratives
  - Interrogatives
  - Directives
  - Indirect speech



# Noun Phrases



# Semantic Structures



Australian Government  
Department of Defence  
Defence Science and  
Technology Organisation

- ▶ Possible semantic structures
- ▶ Kuhn's [Kuh13] classification criteria – Precision, Expressiveness, Naturalness & Simplicity (PENS)
- ▶ 1-5 scale, low to high
  - Universal quantification over individuals
  - Binary or higher relations
  - Multiple universal quantification
  - *If-then* conditionals
  - Weak & strong negation
  - Second-order universal quantification
  - Existential quantification, equality, speech acts



# Classification examples 1

## ► Universal quantification over individuals

(3) Women stand.

```
all([skc2], woman(@ (skc2, t_3, s_2), [female, plural, ...])  
=> stands(@ (skc2, t_3, s_2), [general_habitual, ...])).
```

## ► Binary or higher relations

(4) All women always read all documents.

```
all([skc81, skc82, t_81], ((woman(@ (skc81, t_81, s_81), [...]) &  
document(@ (skc82, t_81, s_82), [...]))  
=> reads(@ (skc81, t_81, s_81), @ (skc82, t_81, s_82), [...])))).
```



## Classification examples 2

### ► *If-then* conditionals

(5) If all women did not see the car then all women did not see the driver.

```
all([skc81], ((woman(@(skc81,t_81,s_81), [...]) &
car(@(skc82,t_81,s_82), [...])) =>
~sees(@(skc81,t_81,s_81),@(skc82,t_81,s_82)))) =>
all([skc81], ((woman(@(skc81,t_81,s_81), [...]) &
driver(@(skc84,t_81,s_84), [...])) =>
~sees(@(skc81,t_81,s_81),@(skc84,t_81,s_84), [...])))).
```

# Classification examples 3



Australian Government  
Department of Defence  
Defence Science and  
Technology Organisation

## ► Negation

(6) The woman did not read the document.

```
woman (@ (skc81, t_22, s_81), [definite, ...]),  
document (@ (skc07, t_22, s_07), [definite, ...]),  
~reads (@ (skc81, t_22, s_81), @ (skc07, t_22, s_07), [past, ...]).
```

## ► Second-order universal quantification, see (4)



# Classification examples 4

## Other determinants of expressiveness

### ► Existential quantification

(7) The woman stood in the house.

```
animate(@ (skc2,t_4,s_2) ), female (@ (skc2,t_4,s_2) ),  
before (t_4, inv1 (timestamp (2014,6,2,1,3,48), timestamp (2014,6,2,1,3,48))),  
location_in ([stands (@ (skc2,t_4,s_2) )], @ (skc3,t_4,s_3) ),  
woman (@ (skc2,t_4,s_2), [animate, definite, singular, ...]),  
house (@ (skc3,t_4,s_3), [definite, singular, prep (in)]),  
stands [@ (skc2,t_4,s_2)], [past, ...])).
```

### ► Equality

(8) Andrew White is the Prime Minister.

```
Andrew_White (@ (skc6,t_10,s_6), [...]),  
prime_minister (@ (skc7,t_10,s_7), [...]),  
identical [@ (skc6,t_10,s_6), @ (skc7,t_10,s_7)].
```

# Classification examples 5



**Australian Government**  
**Department of Defence**  
Defence Science and  
Technology Organisation

## Speech acts

### ▶ Directives

(9) Show merchant ship situation report on MR41\_PAN-EAV.

### ▶ Indirect speech

(10) Michael said that the woman read the document.

# PENS applied



**Australian Government**  
**Department of Defence**  
Defence Science and  
Technology Organisation

- ▶ Tentative classification of Lexpresso as:  $P^{3-4} E^4 N^{4-5} S^3$ 
  - Precision—reliably & semi-deterministically interpretable  $P^{3-4}$
  - Fairly high expressiveness  $E^4$
  - Fair degree of naturalness  $N^{4-5}$
  - Simplicity  $S^3$

# Conclusion






**Australian Government**  
**Department of Defence**  
Defence Science and  
Technology Organisation

- ▶ Brief introduction to Lexpresso
  - Use case & purpose
  - I/O capability
  - System architecture
  - Main syntactic and semantic features
- ▶ Assessed against PENS system [Kuh13]
  - Tentative classification  $P^{3-4} E^4 N^{4-5} S^3$
  - Reliably or perhaps deterministically interpretable language
  - High expressiveness
  - Considerable naturalness, and
  - Requires lengthy description of syntax and semantics



# References I

-  T. Kuhn.  
A survey and classification of controlled natural languages.  
*Computational Linguistics*, pages 121–170, June 2013.
-  D.A. Lambert and C. Nowak.  
The mephisto conceptual framework.  
Technical Report DSTO-TR-2162, Defence Science and Technology Organisation, 2008.
-  A. Saulwick and K. Trentelman.  
Towards a formal semantics of social influence.  
*Knowledge-Based Systems*, in press, 2014.