Computationalism as a Philosophy of Science in Cognitive Science

Cem Bozşahin

Computer Engineering & Cognitive Science
Middle East Technical University (METU), Ankara
bozsahin@metu.edu.tr

Philosophy and Cognition Day, March 8, 2008

Cognitive Problems

General domains, general solutions (?)

Unique domains, unique solutions (Cognitivism)

General domains, unique solutions (Computationalism)

Cognitivism in Psychology:

Stagewise development of cognitive structures in the child (Piaget, 1952)

Computationalism in Psychology:

Different stages in development are attained by increasing access of the child to **the same structure**, by adoption of new **information processing** strategies (Janet Fodor 1975).

Some Piagetian stages

Period of Sensorimotor activity

Stage of reflexes

Stage of primary circular reactions

Stage of coordination of secondary circular reactions

Period of Operational thought

Period of Formal operations

Cognitivism in Linguistics:

Word learning and grammar learning are different problems, because words and phrases are different.

Nouns first?

Verbs first?

Computationalism in Linguistics:

Acquiring a grammar starts with simple and short things, which **tend** to be words.

Short strings first?

Unambiguous strings first?

Conceptual primitives in Linguistics:

Grammatical subject is a special category, because it is available universally.

Computationalism in Linguistics:

Grammatical subjects are part of the full interpretability problem, therefore not very special in the beginning.

Cognitivism in Philosophy of Mind:

Brain is not just any computer. Formal operations alone cannot bring semantics and intelligence (Searle, 1980).

Computationalism in Philosophy of Mind:

Understanding and intelligence builds on top of syntax-semantics correspondence problem, modulo the neural subtrate.

Computationalism in the narrow sense is **not** instrumentalist.

Not reductionist: uniqueness of solutions.

Difference in Philosophy of Science:

Computationalism introduces weak bias. (Lappin and Shieber, 2007)

(uncomplicated, uniform, task-general methods; specific solutions)

Cognitivism introduces strong bias.

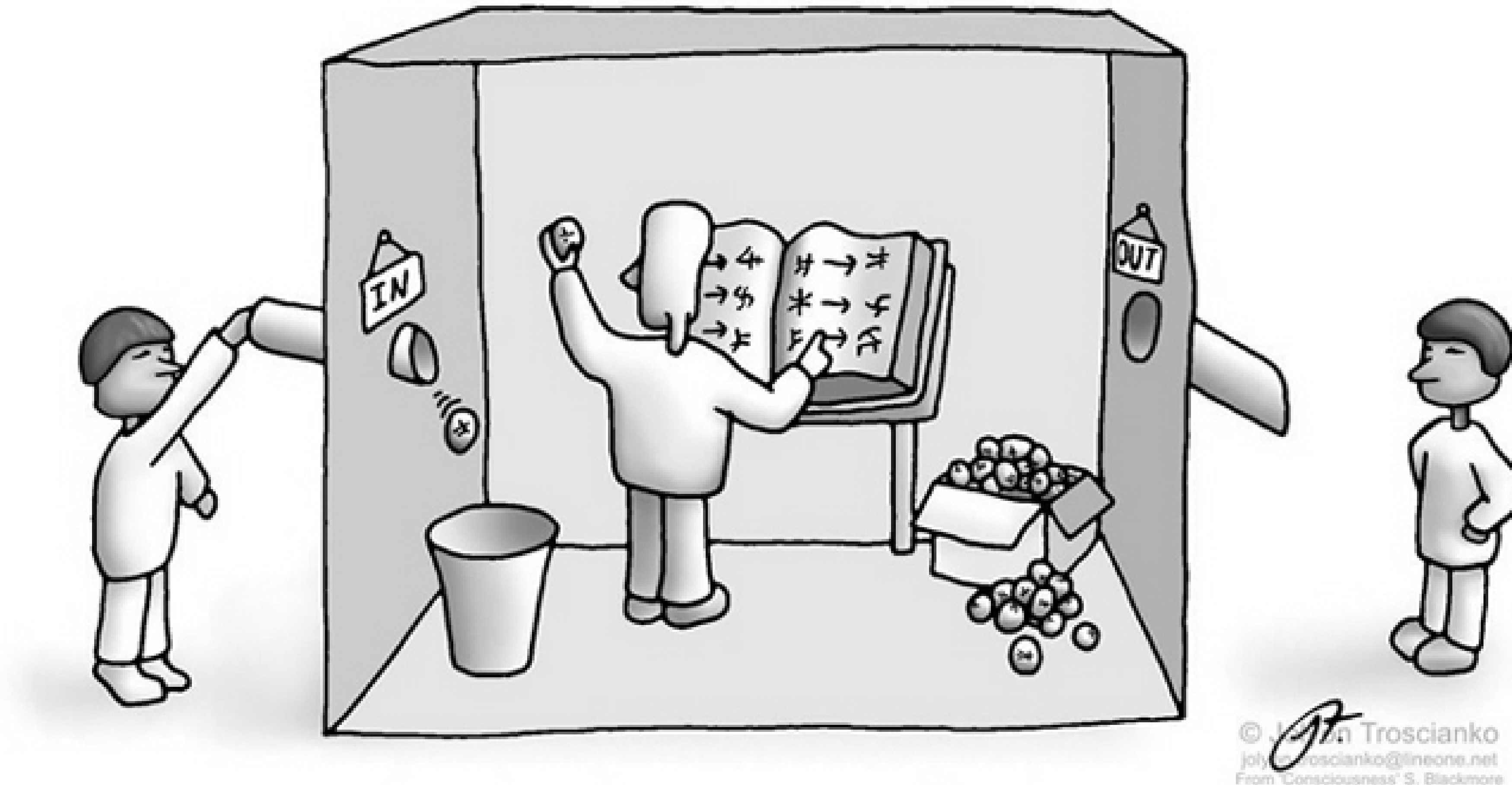
Some examples

Searle and Understanding (Bozsahin, 2006)

Word and grammar learning (Coltekin and Bozsahin, 2007)

Grammatical subjects and full interpretation

(Bozsahin and Steedman, in prep.)


IN
OUT
© J…n Troscianko
From 'Consciousness' S. Blackmore

Strong AI according to Searle (1980):

“the computer is not merely a tool in the study of mind; rather the appropriately programmed computer really is a mind in the sense that computers given the right programs can be literally said to understand and have cognitive states.”

Searle claims that such a purely formalist account of mind is not possible.

Searle assumes AI relates mind, language and verbal behaviour à la Turing.

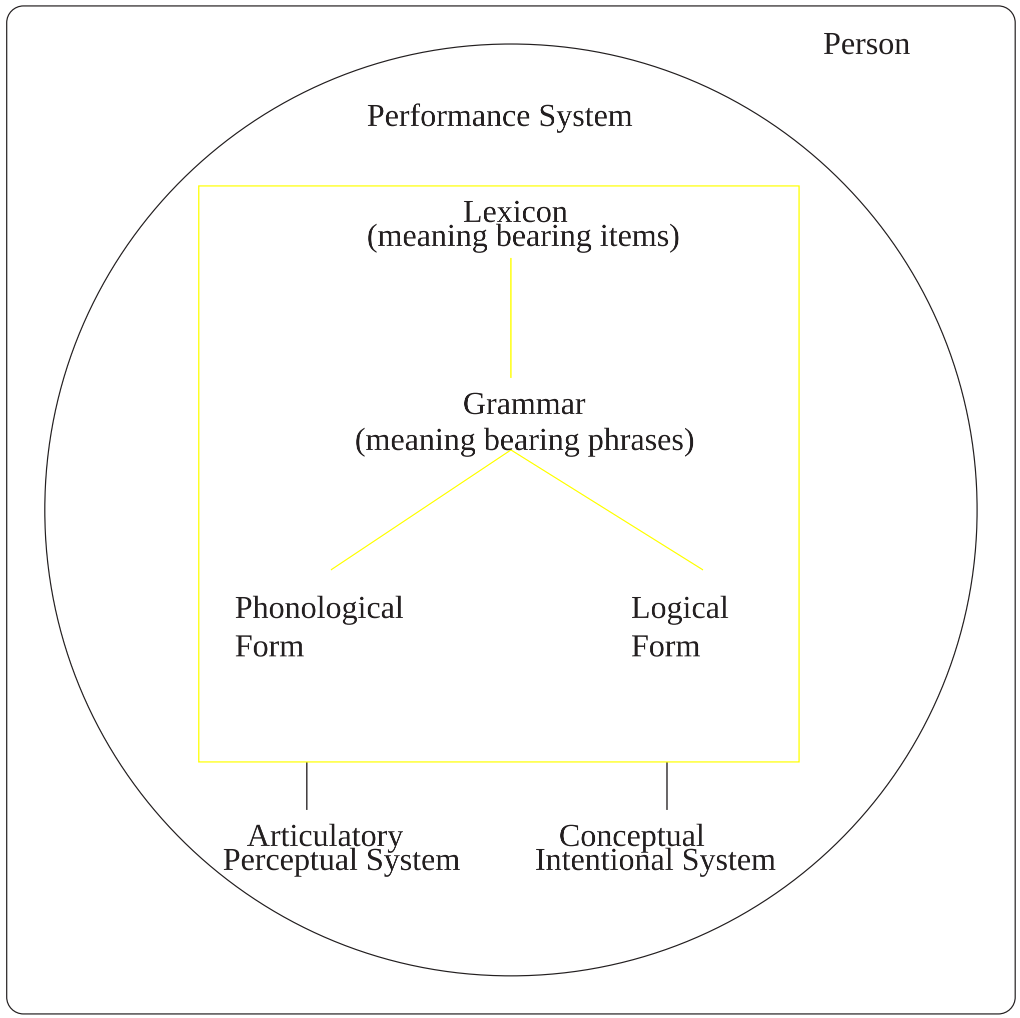
It does, and that’s probably wrong.

Unlike Turing, who only used verbal behaviour as a convenient conduit, Searle refers to **knowledge of language** of **native** speakers.

A linguistic argument would also question his Chinese Room setup.

The issue would not be whether Chinese Room's internal states has semantics (say by a causal connection to external states),

but whether what's inside the room can possibly (even) approximate a native speaker's linguistic competence.


Person
Performance System
Lexicon
(meaning bearing items)
Grammar
(meaning bearing phrases)
Phonological
Form
Logical
Form
Articulatory
Perceptual System
Conceptual
Intentional System

What is Chinese in the Chinese room is the

(1) **database**

(2) fragments of the **program** that contains Chinese symbols and their abstractions.

The program cannot be of **infinite** size.

The correspondences in the program cannot be **phrase-to-phrase matchings**.

Therefore the program must contain finitely characterizable symbols and their **program-internal** abstractions.

In other words, a **generative grammar** of Chinese.

In the thought experiment we must assume that the program contains a generative grammar because we can

"suppose also that the programmers get so good at writing the programs that from the external point of view—that is, from the point of view of somebody outside the room in which I am locked—my answers to the questions are indistinguishable from those of **native** Chinese speakers." (Searle, 1980).

The database would minimally contain Chinese vocabulary, and perhaps a large inventory of expressions based on symbols in the program.

This too must be finite to fit into the room.

We thus have a generative linguistic system of grammar and a lexicon housed in the room.

The experimental setup of Chinese Room is inconsistent:

1) We are **forced** to assume that there is a generative grammar inside the room.

2) The room cannot use it for semantic interpretation (it's assumed to have access to symbols—their shapes, but not to their meaning.)

3) All generative grammars are semantically interpretable, because their product, a **structural description**, is there solely to provide a full array of phonetic and semantic interpretation.

Linguists conceive language understanding as an **interface problem** of connecting internal (linguistic) meanings with external (anchored) meanings.

Just as strong AIers conceive understanding as situating the program+substrate in a conceptual system.

Then, a computational system can **in principle** be made to face the same conditions as the child for understanding the connections between sounds and meanings.

Zettlemoyer and Collins (2005) report an experiment in statistical learning of generative grammars by machine.

The training data (for the machine) is sound-meaning pairs, in which syntax is a **hidden variable**,

There is no external access to the internal states of a program such as Searle's.

Therefore, the input to the room **must** be sound-meaning pairs in order for computation to take place inside the room.

Conclusions-1

Severed semantics in the room is not fair.

Experimental setup is linguistically inadequate.

A way to argue **against** strong AI: a bona fide computational system can model a native speaker **only by chance**.

If understanding is an interface problem, the interfaces must be given a fair chance computationally.

Human understanding is probably unique.

So is bee navigation.

These are domain-specific solutions, but the problem is the same.

Word Learning as a Search Problem

How can the meaning and category of words arise in the mind of a child?

Identifying which meanings go with which substrings in speech.

Lexicon & Grammar acquisition: continuous problem space?

Can we start **without** the morpheme hypothesis?

(it assumes words are qualitatively different)

CDS in CHILDES (5 occurences; range 2;0–4;0):

oyuncak	**-lar**	**-ın**	**-ı**
toy	-PLU	-POSS.2s	-ACC
‘your toys (ACC)’			

syllables: **o·yun·cak·la·rı·nı**

Cues for syllables, and word boundaries (e.g. Jusczyk, 1999; Thiessen and Saffran, 2003)

No cues for morpheme boundaries

56% of nouns in CHILDES database are inflected

CDS contains words as complex as

el-in-de-ki-ler-le

(‘with the ones in your hand’).

Child speech in CHILDES (5 occurences; range 3;0–4;0):

oyuncak	**-lar**	**-ım**	**-la**
toy	-PLU	-POSS.1s	-INST

‘with my toys’

Transcription is to conventional form or target morphology.

Three models

1) **Syllable-based model SBM**: An LF is associated with a contiguous sequence of syllables*

2) **Morpheme-based model MBM**: An LF is associated with a cont. sequence of morphemes

3) **Random model RM**: Randomly-selected contiguous substrings are associated with an LF.

These models are inspired by Zettlemoyer and Collins (2005).

*Jack, Reed and Waller 2006 is another syllable-meaning model

Only 23% of the syllables in nouns are also morphemes in CHILDES.

If we only match boundaries, the overlap is 57%.

araba-lar (car-PLU), with the syllables

a·ra·ba·lar

2 matches out of 4 syllables

Learning a grammar

Input: (PF, LF) pairs* **(arabalara,** $to'plu'car'$**)**

Output: Lexical hypotheses (a lexicalised grammar of nouns)

Lexical hypothesis: pairing of contiguous substrings with a syntactic type and a semantic type (i.e. category)

The syllable model shows 77% overlap of lexical hypotheses with the morpheme model.

*see Tenenbaum and Xu 2000 for Bayesian learning of concepts

How many possibilities per input word?

pisilere := $to'(plu'\,cat')$ requires

4 morpheme-LF pairings to be considered in the morpheme model,

8 syllable-LF pairings in the syllable model.

Average number of pairings are respectively 3.24 and 5.63 in CHILDES.

An adult word such as **kitabındakilerdeki** would require 49 and 343 pairings.

Left unconstrained, the number of pairings is unknown, and most likely to be computationally unmanageable.

The bias: Universal Rules and Principles (CCG; Steedman 2000)

Why weak bias? The solution is domain-specific, but the problem is not (Steedman, 2002)

Rules as constraints

(1)a. Forward Application:

$$\textbf{X}/\textbf{Y}{:}\, f \quad \textbf{Y}{:}\, a \quad \Rightarrow \quad \textbf{X}{:}\, fa \qquad (>)$$

b. Backward Application:

$$\textbf{Y}{:}\, a \quad \textbf{X}\backslash\textbf{Y}{:}\, f \quad \Rightarrow \quad \textbf{X}{:}\, fa \qquad (<)$$

A CCG Derivation from REF model

(2)

$$
\begin{array}{lll}
\textit{pisi} & \textit{ler} & \textit{e} \\
\text{kitty} & \text{-PLU} & \text{-DAT} \\
\hline
\mathbf{N}: cat' & \mathbf{N}\backslash\mathbf{N}: \lambda x.plu'\,x & \mathbf{N}_{\mathrm{dat}}\backslash\mathbf{N}: \lambda y.to'\,y \\
\hline
\multicolumn{2}{c}{\mathbf{N}: plu'cat'} \; < & \\
\hline
\multicolumn{3}{c}{\mathbf{N}_{\mathrm{dat}}: to'(plu'cat')} \; < \\
\multicolumn{3}{c}{\text{'to the kitties'}}
\end{array}
$$

Syllables: pi·si·le·re

Principles

The Principle of Categorial Type Transparency (PCTT): (Steedman, 2000)

"For a given language, the semantic type of the interpretation together with a number of language-specific directional parameter settings uniquely determines the syntactic type of a category."

The Principle of Consistency (PC):

“All syntactic combinatory rules must be consistent with the directionality of the principal functor.”

Principles at work for *pisiler*

(3a) violates PC

Others violate PCTT.

(3)a. {*pisiler* := **N**, *pisi* := **N**$\backslash$**N**, *ler* := **N**} (*)

b. *ler* := **N**$\backslash$**N**: $plu'_{(t)}$ (*)

c. *tut* (catch) := **S**$\backslash$**NP**$\backslash$**NP**: $\lambda x.catch'_{(e,t)}x$ (*)

Learning the hypotheses

Start with an initial lexicon. Take a new (PF,LF) pair as input word.

Generate all hypotheses (closure of rules and principles over the lexicon).

Update Bayesian probabilities of lexical items, or add new ones.

Repeat with updated lexicon and new input.

Test measures over the grammars.

Lexicon	#of items	precision	recall	f-score
L_r	1041	100.00	100.00	100.00
L_m	1040	99.61	99.51	99.55
L_s	909	81.73	71.37	76.19
L_{rm}	1697	51.73	83.57	63.90

Results of the $10\times$ recognition tests.

Lexicon	precision		recall		f-score	
	μ	σ	μ	σ	μ	σ
L_r	87.00	1.63	92.90	1.66	89.83	0.79
L_m	86.70	1.42	92.90	1.66	89.67	0.66
L_s	84.20	4.32	72.80	2.04	78.02	2.06
L_{rm}	57.10	3.35	82.00	2.98	67.21	1.71

With varying degrees of success, all three models learn a **syntactic type** for the (LF, PF) pair, i.e. a fragment of word **grammar**.

MBM approximates an adult reference word grammar as expected.

The syllable model SBM is not too far behind (77%),

at least in the circumstances where word-level ambiguity is kept to a minimum.

This is promising for grounding early development of language in perception.

Shorter strings engender less number of hypotheses.

Unambiguous strings engender less number of hypotheses (Steedman and Hockenmaier, 2007).

This seems orthogonal to verb-first or noun-first categorical bias.

We would expect word learning to interact with cross-situational learning (Siskind 1995, 1996) and the scenario above.

Conclusions-2

Re-look at stages of development from this perspective.

Some short forms **might be** child-verbs (**cookie** means 'give'?)

Nouns-first theory (Gentner, 1982) can show

short verbs (of simple concepts) are not learned early,

the impediment to learning long nouns is not categorical.

20-22 month-old Mandarin children seem to show no N/V bias (Tardif, 1996).

We are tempted to believe there is no categorical bias,

maybe a computational one (besides the world, of course).

Keren's first words (Dromi, 1987) Hebrew (Israel)

Age m(d)	Child's word	conven. form					
10(12)	haw	(?)	a dog's bark	12(16)	hita	(?)	going out for a walk
11(16)	?aba	(aba)	Father	12(18)	tiktak	(?)	sound of clock
11(17)	?imaima	(?)		12(19)	cifcif	(?)	bird's tweet
11(18)	ham	(?)	said while eating	12(20)	hupa	(?)	accom. making sudden contact w/ground
12(3)	mu	(?)	a cow's moo	12(23)	dio	(dio)	giddi up
12(3)	?ia	(?)	a donkey's bray	12(25)	hine	(hine)	here
12(8)	pil	(pil)	an elephant	12(25)	?ein	(?ein)	all gone
12(11)	buba	(buba)	a doll	12(25)	na?al	(na?al)	a shoe
12(13)	pipi	(pipi)	urine	12(25)	myau	(?)	a cat's meow

Tad's first words (Gentner, 1982) AmE

Age (m.)					
11	dog	16	eye	19	down
12	duck	18	cow		boo
13	daddy		bath		bottle
	yuk		hot		up
	mama		cup		hi
	teh (teddy bear)		truck		spoon
	car	19	kitty		bye
14	dipe (diaper)		pee pee		bowl
	toot toot (horn)		happy		uh oh
	owl		oops		towel
15	keys		juice		apple
	cheese		TV		teeth

Are subjects special?

There is a notion of grammatical subject in every language.

Not all subjects are agents. *The ship sank.*

Universal tendencies of subjects. (Keenan and Comrie, 1977)

Only subjects resist true reflexivisation. **Heself saw John.*

Relational Grammar and LFG consider them special.

Word order typologies

Subject-Object-Verb (SOV) is most common.	$\approx$ 40-45%
SVO is second.	$\approx$ 38-40%
VSO is distant third.	15%

Intransitives: SV, SV and VS resp.

Cross-situational learning of subjects

Eat cookies.

eat := **S**:$eat'\,cookie'$

cookies := **S**: $eat'\,cookie'$

eat := **S**:$cookie'\,eat'$

cookies := **S**: $cookie'\,eat'$

eat cookies := **S**: $cookie'\,eat'$

eat cookies := **S**: $eat'\,cookie'$

eat:= **N**: $cookie'$ cookies := **S**$\backslash$**N**: $\lambda x.eat'\,x$

eat := **S**$/$**N**: $\lambda x.cookie'\,x$ cookies := **N**: eat'

eat := **S**$/$**N**: $\lambda x.eat'\,x$ cookies := **N**: $cookie'$

The following hypotheses are not possible for formal reasons.

eat := $\mathbf{S}\backslash\mathbf{N}$: $\lambda x.cookie'\,x$ cookie := $\mathbf{N}$: eat'

eat := $\mathbf{N}$: $cookie'$ cookie := $\mathbf{S}/\mathbf{N}$: $\lambda x.eat'\,x$

eat := $\mathbf{N}$: eat' cookie := $\mathbf{S}/\mathbf{N}$: $\lambda x.cookie'\,x$

eat := $\mathbf{S}\backslash\mathbf{N}$: $\lambda x.eat'\,x$ cookie := $\mathbf{N}$: $cookie'$

These hypotheses are not possible for substantive reasons.

eat := $\mathbf{N}$: $\lambda x.eat'\,x$ cookies := $\mathbf{S}\backslash\mathbf{N}$: $\lambda y.cookie'\,y$

eat := $\mathbf{N}$: $\lambda x.cookie'\,x$ cookies := $\mathbf{S}\backslash\mathbf{N}$: $\lambda y.eat'\,y$

eat := $\mathbf{S}/\mathbf{N}$: $\lambda x.eat'\,x$ cookies := $\mathbf{N}$: $\lambda y.cookie'\,y$

eat := $\mathbf{S}/\mathbf{N}$: $\lambda x.cookie'\,x$ cookies := $\mathbf{N}$: $\lambda y.eat'\,y$

eat := $\mathbf{S}/\mathbf{N}$: eat'_0 cookies := $\mathbf{N}$: $cookie'$

eat := $\mathbf{N}$: eat' cookies := $\mathbf{S}\backslash\mathbf{N}$: $cookie'{}_0$

In both cases, what limits hypotheses is full interpretability.

Eat veggies.

More cookies.

Eating without cookies sieves wrong assumptions about cookies.

Cookies without eating sieve wrong assumptions about eating.

Wrong assumptions can still survive.

eat := $\mathbf{S}/\mathbf{N}$: $\lambda x.eat'\, cookie'\, x$ veggies := $\mathbf{N}$: $veggie'$

eat veggies := $\mathbf{S}$:$eat'\, cookie'\, veggie'$

eat := $\mathbf{N}$:$eat'\, cookie'$ veggies := $\mathbf{S}\backslash\mathbf{N}$:$\lambda x.veggie'\, x$

a. *Liz eat cookies.*

b. *Liz eat herself.*

Is *Liz* like *cookie* in previous experiences?

eat := $\mathbf{(S/N)/N}$? (a) is not interpretable

eat := $\mathbf{(S\backslash N)\backslash N}$ (a–b) & prev. expr. not int.

eat := $\mathbf{N}$ Not consistent with all experiences

eat := $\mathbf{(S\backslash N)/N}$ consistent with all experiences

$\mathbf{S\backslash N}$ and $\mathbf{S/N}$ mean different things in English.

Interaction among experiences provides enough cues to sieve out wrong syntactic hypotheses.

Consistent exposure to SVO and SV can lead to $\backslash$**N** as subject,

assuming full (right or wrong) interpretability of experiences.

Grammatical relations emerge due to hypothesis revision (under feedback) for full interpretability.

Steedman (2000) shows how these emergent properties satisfy subject-related biases.

Conclusions

The key to honing cognitive traits seems to be **interaction**

Computationalism offers transparent ways to conceive form-meaning correspondence as the **conduit** of interaction

to make an attempt to understand **continuity** in problem spaces.

Problems look similar computationally, solutions are domain-specific.

Language (humans, bird songs)		Understanding
Vision	Planning	Path finding
Music		Cross-situational learning
Cybernetic serendipity		Self-organising cyber communities

Gordon Pask   1928–1996

*References

Bozsahin, Cem. 2006. “Grammars, programs, and the Chinese room.” In *Proc. of the Int. European Conf. on Computing and Philosophy*. Trondheim, Norway.

Coltekin, Çağrı, and Cem Bozsahin. 2007. “Syllable-based and Morpheme-based Models of Bayesian Word Grammar Learning from CHILDES Database.” In *Proc. of the 29th Annual Meeting of Cognitive Science Society*. Nashville, TN.

Dromi, Esther. 1987. *Early lexical development*. Cambridge University Press.

Gentner, Dedre. 1982. “Why nouns are learned before verbs: Linguistic Relativity versus Natural Partitioning.” In Stan A. Kuczaj II, ed., *Language Development, vol.2: Language, Thought and Culture*, 301–334. Hillsdale, New Jersey: Lawrence Erlbaum.

Jack, Kris, Chris Reed, and Annalu Waller. 2006. “From Syllables to Syntax: Investigating Staged Linguistic Development through Computational Modelling.” In *Proc. of the 28th Annual Conference of the Cognitive Science Society*. Vancouver.

Jusczyk, P. W. 1999. “How Infants Begin to Extract Words from Speech.” *Trends in Cognitive Science*, 3, 323–328.

Keenan, Edward, and Bernard Comrie. 1977. “Noun Phrase Accessibility and Universal Grammar.” *Linguistic Inquiry*, 8, 63–99.

Lappin, Shalom, and Stuart Shieber. 2007. “Machine learning theory and practice as a source of insight into universal grammar.” *J. Linguistics*, 47, 1–34.

Piaget, Jean. 1952. *The Origins of Intelligence in Children*. New York: Norton.

Searle, John R. 1980. “Minds, Brains and Programs.” *The Behavioral and Brain Sciences*, 3, 417–424.

Siskind, Jeffrey. 1995. “Grounding Language in Perception.” *Artificial Intelligence Review*, 8, 371–391.

Siskind, Jeffrey. 1996. “A Computational Study of Cross-Situational Techniques for Learning Word-to-Meaning Mappings.” *Cognition*, 61, 39–91.

Steedman, Mark. 2000. *The Syntactic Process*. Cambridge, MA: MIT Press.

Steedman, Mark. 2002. “Plans, Affordances, and Combinatory Grammar.” *Linguistics and Philosophy*, 25, 723–753.

Steedman, Mark, and Julia Hockenmaier. 2007. “The Computational Problem of Natural Language Acquisition.” Ms., University of Edinburgh.

Tardif, Twila. 1996. “Nouns are not always learned before verbs: Evidence from Mandarin speakers’ early vocabularies.” *Developmental Psychology*, 32, 3, 497–504.

Tenenbaum, Joshua B., and Fei Xu. 2000. “Word learning as Bayesian inference.” In *Proc. of the 22nd Annual Conf. of the Cognitive Science Society*. Philadelphia.

Thiessen, Erik D., and Jenny R. Saffran. 2003. “When Cues Collide: Use of Stress and Statistical Cues to Word Boundaries by 7- to 9-Month-Old Infants.” *Developmental Psychology*, 39, 4, 706–716.

Zettlemoyer, Luke S., and Michael Collins. 2005. “Learning to Map Sentences to Logical Form: Structured Classification with Probabilistic Categorial Grammars.” In *Proc. of the 21st Conf. on Uncertainty in Artificial Intelligence*. Edinburgh.