

# ESL-Soar: An Architectural Account of Errors in Foreign Language Learning

Julie Van Dyke  
Department of Computer Science  
Carnegie Mellon University  
vandyke@cs.cmu.edu

## Articles are hard!

### Polish:

Nasza rodzina ma psa i kanarka w domu  
*[Our family has a dog and a canary.]*

Pies jest dobrym towarzyszem.  
*[The dog is a good companion.]*

### Responses:

Dog is good companion.

Dog is a good companion.

A dog is the good companion.

A dog is a good companion.

The dog is the good companion.

The dog is a good companion.

## But rule governed.

[+/- SR, +/-HK]

[+ Specific Referent] [+ Hearer Knowledge] --> the

- a. Unique referents
- b. Referent physically present or already mentioned.
- c. Specific referents assumed known to hearer.

[+ Specific Referent] [- Hearer Knowledge] --> a or 0

- a. First mention in a discourse

[- Specific Referent] [- Hearer Knowledge] --> the, a, 0

- a. Equative NPs (He's a nice man.)

[- Specific Referent] [+Hearer Knowledge] --> the, a, or 0

- The/A/0 Lion(s) is/are a beautiful animal.

## But if you're Polish,

1. W pokoju siedziala dziewczyna.  
<in room sit-past+3pfem girl-nom-sg>

- a. Wszedl chłopiec.  
<come-past+3pmasc boy-nom-sg>
- b. \*Chłopiec wszedl.

2. W pokoju siedziala chłopiec.  
<in room sit-past+3pmasc boy-nom-sg>

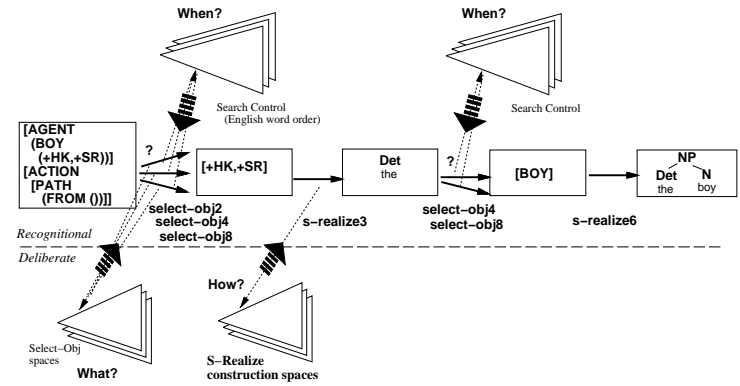
- a. Wszedl chłopiec.  
<come-past+3pmasc boy-nom-
- b. Chłopiec wszedl.

Rule 1: Referring (Given) items are first, non-referring (New) items are last.

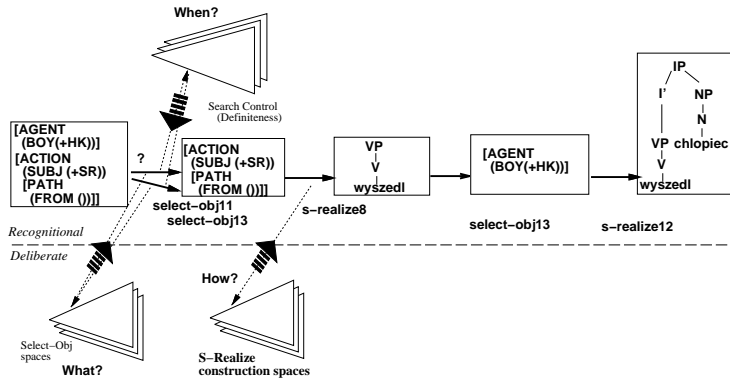
# The rules are harder to learn.

- Language architecture = “topology of chunked knowledge”
- Architectural Mismatch

# English Generation



# Polish Generation



# Predicted Errors

1. Articles omitted  
Transfer of L1 architecture
2. Overgeneralization  
[HK] shifted first (more salient)
3. Non-standard or inconsistent rules  
Knowledge continually shifting towards target

## Future

- Evaluate predictions
  - Laotian (Huebner, 1983)
  - Polish data analysis in progress
  - Chinese, Japanese
- Identify which architectural mismatches matter
- Cognitive evidence for embedded comprehension in NL-Soar generation

## Bibliography

Huebner, T (1983). *A longitudinal analysis of the acquisition of English*. Karoma Publishers, Ann Arbor.

Duskova, L. (1969). On sources of errors in foreign language learning. *International Review of Applied Linguistics in Language Teaching*, 7:11-36.

Szwedek, A. (1976). *Word order, sentence stress, and reference in English and Polish*. Linguistic Research, Edmonton, Alberta.

VanDyke, J. and Lehman, J. (1997) An Architectural Account of Errors in Foreign Language Learning. Proceedings of the 19th Annual Conference of the Cognitive Science Society. Stanford, August 7-12.

Young, R. (1995). Form-function relations in articles in English Interlanguage. In Bayley, R and Preston, D., editors. *Second Language Acquisition and Linguistic Variation*, pages 135-175. John Benjamins, Amsterdam.