

Towards a treatment of word order errors in Computer-Aided Language Learning

When to use deep processing — and when not to

Detmar Meurers and Vanessa Metcalfe

The Ohio State University

Large-scale Grammar Development and Grammar Engineering
Research Workshop of the Israel Science Foundation
University of Haifa, Israel, 25.–28. June, 2006

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalfe

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

A downside of mail-rules

Our perspective and approach

Two types of word order errors

Prasal verbs

Adverb placement

Summary

References



1 / 22

Background

- ▶ NLP technology can be used in Computer-Aided Language Learning tools that
 - provide individual feedback on learner errors,
 - foster learner awareness of language forms & categories.
- ▶ Very few ICALL systems are used in FLT practice today (Nagata 2002; Heift 2001).
- ▶ Problem: lack of interdisciplinary research combining computational, linguistic, and FLT/SLA expertise.
- ▶ Our general approach:
 - Link CL research to genuine FLT needs.
 - Develop task-based systems in support of traditional teaching, essentially an intelligent workbook approach.
 - ▶ TAGARELA System for Portuguese (Amaral and Meurers 2005, 2006) → integration into Portuguese Language Program at OSU in Spring 07
 - ▶ WERTi System for English (Metcalfe and Meurers 2006) → started prototype development

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalfe

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

A downside of mail-rules

Our perspective and approach

Two types of word order errors

Prasal verbs

Adverb placement

Summary

References



2 / 22

Background

From word-based to word-order errors in ICALL

- ▶ ICALL research has largely focused on diagnosing word-based learner errors (i.e., morpho-syntax).
- ▶ Such approaches can rely on parsing algorithms to reign in the recursive potential of natural language.
- ▶ How about word order mistakes, a type of error regularly produced by language learners?

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalfe

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

A downside of mail-rules

Our perspective and approach

Two types of word order errors

Prasal verbs

Adverb placement

Summary

References



3 / 22

Background

Word order and Foreign Language Teaching

- ▶ It is hard to learn word order:
 - Language learners are known to produce a range of word order errors (cf., e.g., Odlin 1989).
 - Word order differs significantly across languages → transfer errors (cf., e.g., Selinker 1972; Odlin 2003)
- ▶ It is important to master word order, especially since word order errors can significantly complicate comprehension.
 - Example from Hiroshima English Learners' Corpus:
 - (1) *He get to cleaned his son.*
→ *He get his son to cleaned.*
 - Exercise target:
 - (2) *He made his son clean the room.*

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalfe

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

A downside of mail-rules

Our perspective and approach

Two types of word order errors

Prasal verbs

Adverb placement

Summary

References



4 / 22

Approaches to diagnosing word order errors

Instance-based list and match

- ▶ Basic idea: Match user input with listed expected forms.
 - ▶ matching all or some words,
 - ▶ with a complete or partial order,
 - ▶ based on surface forms or lemmata.
- ▶ Strength: simple and efficient processing
- ▶ Weakness: lack of generalization over tokens and patterns
 - ▶ All words for which order is to be checked must be known.
 - ▶ All grammatical orders must be preenviaged and listed.

→ works well for heavily constrained activities,

- ▶ e.g., "Build a Sentence" or "Translation" exercises in German Tutor (Heift 2001)

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of *mal*-rules

Our perspective and approach

Two types of word order errors

Preasal verbs
Adverb placement

Summary

References



5 / 22

Approaches to diagnosing word order errors

Deep processing: Basics

- ▶ Use grammars, which are compact representations of the wide range of lexical and word order possibilities.
- ▶ Efficient parsing algorithms are available to license a potentially infinite set of strings based on finite grammars.
- ▶ The additional erroneous word orders can be captured by:
 - ▶ extra phrase structure rules (so-called *mal*-rules, cf. e.g., Heift 1998; Fortmann and Forst 2004)
 - ▶ manipulation of chart edges, the hypotheses introduced by phrase structure rules in a chart parser (Reuer 2003)

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of *mal*-rules

Our perspective and approach

Two types of word order errors

Preasal verbs
Adverb placement

Summary

References



6 / 22

Approaches to diagnosing word order errors

Deep processing: A downside of *mal*-rules

- ▶ Phrase structure grammars express two things at once
 - ▶ generative potential (resource sensitivity, combinatorics)
 - ▶ word order regularitiesand both are determined at the level of a local tree.
- ▶ Licensing more word orders can significantly increase the search space since the word order possibilities are directly tied to the combinatorics.
- ▶ Only local reordering between sisters in a local tree are achievable through *mal*-rules.
Ex. Extending the word order options of $S \rightarrow NP VP$ by adding $S \rightarrow VP NP$ licenses a. and b., but not c.

- (3) a. *Mary [loves cats].*
- b. ** [loves cats] Mary.*
- c. ** loves Mary cats.*

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of *mal*-rules

Our perspective and approach

Two types of word order errors

Preasal verbs
Adverb placement

Summary

References



7 / 22

Our perspective and approach

- ▶ Word order errors are not uniform:
 - ▶ some involve lexical triggers (one of a finite set of words is known to occur) or indicative patterns,
 - ▶ others can only be spotted with deeper analysis.
- ▶ FLT activities are not uniform:
 - ▶ some can be set up to include specific lexical material or patterns,
 - ▶ in others it is hard to control lexical and structural variation.

⇒ Activity-based ICALL systems need a flexible approach to word order error detection and diagnosis.

- ▶ We want to argue for:
 - ▶ choosing processing methods depending on targeted word error type and activity design
 - ▶ in deep processing: moving beyond local trees as the units corresponding to errors

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of *mal*-rules

Our perspective and approach

Two types of word order errors

Preasal verbs
Adverb placement

Summary

References



8 / 22

Two types of word order errors

- ▶ We explore two aspects of English grammar with interesting word order properties:
 - phrasal verbs
 - adverbs
- ▶ For each, we describe
 - linguistic properties,
 - exercises supporting awareness of the relevant word order patterns, and
 - the processing needed for those exercises.

Word order errors & Processing regimes

Detmar Meures and
Veronica Meisat

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mal-rules

Our perspective and approach

Two types of word order errors

Phrasal verbs
Adverb placement

Summary

References



9 / 22

Phrasal verbs

Pedagogical relevance of particle placement

- ▶ English learners make errors in particle placement:
 - (7) a. * *so they give up it*
 - b. * *food which will build up him*
 - c. * *rather than speed up it.*

Examples from the Chinese Learner English Corpus (CLEC 2004)
- ▶ Learners also avoid using phrasal verbs:
 - Liao and Fukuya (2002) show that Chinese learners of English avoid phrasal verbs; similar research for other L1.
 - We also found patterns of avoidance in the CLEC:
 - heavy use of pattern that is always grammatical
 - little use of patterns restricted to certain verb & object types

Word order errors & Processing regimes

Detmar Meures and
Veronica Meisat

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mal-rules

Our perspective and approach

Two types of word order errors

Phrasal verbs
Adverb placement

Summary

References



11 / 22

Phrasal verbs

- ▶ Separable phrasal verbs
 - Particles can precede or follow a full NP object.
 - (4) a. **wrote down the number**
 - b. **wrote the number down**
 - Particles must follow a pronominal NP object.
 - (5) a. * **wrote down it**
 - b. **wrote it down**
- ▶ Inseparable phrasal verbs
 - Particles always precede any NP object.
 - (6) a. **ran into {my neighbor, her}**
 - b. * **ran {my neighbor, her} into**

Word order errors & Processing regimes

Detmar Meures and
Veronica Meisat

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mal-rules

Our perspective and approach

Two types of word order errors

Phrasal verbs
Adverb placement

Summary

References



10 / 22

Phrasal verbs

Example exercise tasks

Part 1 of the exercise targets lexical particle choice:

Complete the following sentence:

Please turn the radio _____ a little. It's too loud.

Part 2 targets particle placement (and pronoun choice).

Now, replace the object with a pronoun:

Please turn down the radio a little. It's too loud.

→ *Please _____ a little. It's too loud.*

Word order errors & Processing regimes

Detmar Meures and
Veronica Meisat

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mal-rules

Our perspective and approach

Two types of word order errors

Phrasal verbs
Adverb placement

Summary

References



12 / 22

Phrasal verbs

Processing the example exercises

- ▶ We target two possible error patterns:
 - separable-phrasal-verb < particle < pronoun
(8) * **wrote down it**
 - inseparable-phrasal-verb < NP < particle
(9) a. * **ran my neighbor into**
b. * **ran her into**
 - ▶ Regular expression matching with those patterns is sufficient to capture the targeted errors.
 - The relevant words (or strings) to be matched are specified in the activity model.
 - ▶ Desired error diagnosis and feedback is one-to-one with those patterns.
- ⇒ Particle placement is an example for a word order phenomenon which can adequately be diagnosed based on a shallow analysis.

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisler

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mail-rules

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



13 / 22

Adverb placement in English

- ▶ English has many different adverbs, and the word order possibilities depend on adverb subclass distinctions.
 - ▶ The rules governing adverb placement are difficult to articulate and master.
 - ▶ Many adverb placements are not right or wrong, but more or less natural.
 - ▶ Students frequently misplace adverbs
- (10) a. *they **cannot already** live without the dope.*
b. *There **have been already** several campaigns held by 'Outdoor'.*
c. *while any covert action **brings rarely** such negative connotations.*
d. *It seems that the Earth **has still** a lot to reveal . . .*
- Examples from Polish part of Int. Corpus of Learner English (PICLE 2004)

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisler

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mail-rules

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



14 / 22

Adverb placement

Example exercise tasks

Task 1:

Find and move any misplaced adverbs:

(11) *She has finished almost her breakfast.*

Task 2:

Add the given adverb to the sentence:

Adverb: *slowly*

(12) *Taking his visitor by the arm, he walked her along the corridor.*

(Example taken from British National Corpus)

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisler

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mail-rules

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



15 / 22

Adverb placement

Processing the example exercises

- ▶ Instance-based matching is inadequate:
 - Many placements throughout a sentence are possible.
 - Targeted errors are predictable, but numerous.
 - Generalizations about the many adverbs of English and the subclasses they form are lost.
 - Reference to syntactic structure is needed for
 - ▶ identification of possible placements,
 - ▶ error diagnosis, and
 - ▶ content of feedback.
- ▶ Deep processing
 - Parsing can identify the necessary sentence structure.
 - The lexicon of a grammar supports modeling adverb classes.

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisler

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of mail-rules

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



16 / 22

Adverb placement

Combining native and interlanguage patterns

- ▶ We need to model a learner grammar which combines
 - native English patterns with
 - anticipated interlanguage patterns.
- ▶ Word orders not licensed by the space between native and interlanguage patterns should be excluded, to support efficient processing.
- ▶ The combination of native and interlanguage patterns should not result in spurious ambiguities (i.e., same word order licensed by different structures).

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of ml-rules

Our perspective
and approach

Two types of
word order errors

Prosal verbs

Adverb placement

Summary

References



17/22

Adverb placement

Targeted word orders

- ▶ Adverb placement can be described in terms of linear order with respect to constituents.

(13) ₁ Sid ₂ might ₃ be ₄ taking ₅ his mother ₆ to the store ₇.

1. clause-initial
2. preceding a finite auxiliary
3. preceding a nonfinite auxiliary
4. preceding a main verb
5. preceding an NP complement
6. preceding a PP complement
7. following the VP

- ▶ This is the basic picture; the situation is more complex in the presence of negative auxiliaries or passive sentences.
- ▶ For each adverb subclass, we rate the positions in terms of acceptability (good, bad, marked).

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of ml-rules

Our perspective
and approach

Two types of
word order errors

Prosal verbs

Adverb placement

Summary

References



18/22

Adverb placement

Deep processing in prototype

- ▶ In the implemented prototype, we parse sentences with all envisaged adverb placements, using an HPSG grammar implemented in the TRALE system (MILCA environment; Meurers, Penn and Richter 2002).
- ▶ We encode the actual adverb position through the value of two features in the lexical entry of the adverb:
 - MOD: what category the adverb combines with
 - POSTHEAD: whether the adverb occurs before/after the head
- ▶ The lexical subclass of the adverb and its position is passed up and encoded as part of the overall structure, where it can inform negative or positive feedback.

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of ml-rules

Our perspective
and approach

Two types of
word order errors

Prosal verbs

Adverb placement

Summary

References



19/22

Adverb placement encoding in the prototype

The lexical principle constraining and recording adverb position

(word, synsem:head: (adv, mod:synsem))

*>

```
synsem:head: (mod:Mod,  
              posthead:Where,  
              output_info:[position:adv_placement(Mod,Where)]).
```

```
fun adv_placement(+,-,-).
```

```
adv_placement(@clause, minus, pre_clause) if true.
```

```
adv_placement(@fin_aux, minus, pre_finite_aux) if true.
```

```
adv_placement(@nfin_aux, minus, pre_nonfinite_aux) if true.
```

```
adv_placement(@main_vp, minus, pre_main_verb) if true.
```

```
adv_placement(@np_comp, plus, pre_np_comp) if true.
```

```
adv_placement(@pp_comp, plus, pre_pp_comp) if true.
```

```
adv_placement(@fin_vp, plus, post_finite_vp) if true.
```

Word order errors & Processing regimes

Detmar Meurers and
Veronica Meisard

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of ml-rules

Our perspective
and approach

Two types of
word order errors

Prosal verbs

Adverb placement

Summary

References



20/22

Adverb placement and beyond

- ▶ Adverb position is constrained and recorded using a lexical principle, i.e., not in terms of a local tree.
 - ▶ Such lexicalization is appropriate for words which are fixed by the activity model.
 - ▶ Phrases (e.g., NPs) not targeted by an activity can be pre-processed by a chunker/supertagger to keep a limited lexicon across a range of contextualized activities.
- ▶ Argument reordering encoded parallel to optional complement selection in MERGE (Meurers et al. 2003).
- ▶ Outlook:
 - ▶ For local tree-based word order phenomena (e.g., SOV → VOS) *mal*-rules can be used.
 - ▶ For other word order phenomena, a formalism that supports word order domains beyond local trees (e.g., GIDL, Daniels and Meurers 2004) can be used.

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalf

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A database of mal-rules

Our perspective and approach

Two types of word order errors

Prasal verbs
Adverb placement

Summary

References



21 / 22

Summary

- ▶ When to use instanced-based matching:
 - ▶ lexical material and erroneous placements are predictable and listable
 - ▶ there is no grammatical variation
 - ▶ error patterns correspond directly to intended feedback
- ▶ When deep processing is preferable:
 - ▶ possible correct answers are predictable but not (conveniently) listable for a given activity
 - ▶ predictable erroneous placements occur throughout a recursively built structure
 - ▶ feedback is desired which requires linguistic information about the learner input that can only be obtained through deep analysis
- ▶ Lexicalization of word order options can be an attractive, modular alternative to *mal*-rule based encodings.

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalf

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A database of mal-rules

Our perspective and approach

Two types of word order errors

Prasal verbs
Adverb placement

Summary

References



22 / 22

References

- Amaral, Luiz and Detmar Meurers (2006). Where does ICALL Fit into Foreign Language Teaching? CALICO Conference. May 19, 2006. University of Hawaii.
- Amaral, Luiz Alexandre and Walt Detmar Meurers (2005). Towards Bridging the Gap between the Needs of Foreign Language Teaching and NLP in ICALL. In António Pedros-Gascon (ed.), *Electronic Proceedings of the 8th Annual Symposium on Hispanic and Luso-Brazilian Literatures, Linguistics, and Cultures*.
- CLEC (2004). Chinese Learner English Corpus. Web interface to Corpus.
- Daniels, Mike and W. Detmar Meurers (2004). A grammar formalism and parser for linearization-based HPSG. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING-04)*. Geneva, pp. 169–175.
- Fortmann, Christian and Martin Forst (2004). An LFG grammar checker for CALL. In *InStIL/ICALL 2004 Symposium on Computer Assisted Learning, NLP and speech technologies in advanced language learning systems*. Venice, Italy: International Speech Communication Association (ISCA).
- Heift, Gertrud D. (1998). *Designed Intelligence: A Language Teacher Model*. Ph.D. thesis, Simon Fraser University.
- Heift, Trude (2001). Intelligent Language Tutoring Systems for Grammar Practice. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 6(2). <http://www.ualberta.ca/~german/ejournal/heift2.htm>.
- HELIC (1998). Hiroshima English Learners' Corpus. Data available on webpage. <http://home.hiroshima-u.ac.jp/d052121/eigo2.html>.

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalf

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A database of mal-rules

Our perspective and approach

Two types of word order errors

Prasal verbs
Adverb placement

Summary

References



23 / 22

- Liao, Yan D. and Yoshinori J. Fukuya (2002). Avoidance of phrasal verbs: The case of Chinese learners of English. *Second Language Studies* 20(2), 71–106.
- Metcalf, Vanessa and Detmar Meurers (2006). When to Use Deep Processing and When Not To – The Example of Word Order Errors. Pre-conference Workshop on NLP in CALL – Computational and Linguistic Challenges. CALICO 2006. May 17, 2006. University of Hawaii.
- Meurers, W. Detmar, Kordula De Kuthy and Vanessa Metcalf (2003). Modularity of grammatical constraints in HPSG-based grammar implementations. In Melanie Siegel, Frederik Fouvry, Dan Flickinger and Emily Bender (eds.), *Proceedings of the ESSLLI '03 workshop 'Ideas and Strategies for Multilingual Grammar Development'*. Vienna, Austria. <http://ling.osu.edu/~dm/papers/meurers-dekuthy-metcalf-03.html>.
- Meurers, W. Detmar, Gerald Penn and Frank Richter (2002). A Web-based Instructional Platform for Constraint-Based Grammar Formalisms and Parsing. In Dragomir Radev and Chris Brew (eds.), *Effective Tools and Methodologies for Teaching NLP and CL*. pp. 18–25. <http://ling.osu.edu/~dm/papers/ac102.html>.
- Nagata, Noriko (2002). BANZAI: An Application of Natural Language Processing to Web based Language Learning. *CALICO Journal* 19(3), 583–599. <http://www.usfca.edu/japanese/CALICO02.pdf>.
- Odlin, Terence (1989). *Language Transfer: Cross-linguistic influence in language learning*. New York: Cambridge University Press.
- Odlin, Terence (2003). Cross-linguistic Influence. In Catherine Doughty and Michael Long (eds.), *Handbook on Second Language Acquisition*, Oxford: Blackwell, pp. 436–486.

Word order errors & Processing regimes

Detmar Meurers and Vanessa Metcalf

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A database of mal-rules

Our perspective and approach

Two types of word order errors

Prasal verbs
Adverb placement

Summary

References



24 / 22

PICLE (2004). Polish portion of the International Corpus of Learner English. Web interface to Corpus.

http://elex.amu.edu.pl/przemka/concord2advr/search_adv_new.html.

Reuer, Veit (2003). Error recognition and feedback with Lexical Functional Grammar. *CALICO Journal* 20(3), 497–512.

http://www.cl-ki.uni-osnabrueck.de/~vreuer/publ/calico03_reuer.pdf.

Selinker, Larry (1972). Interlanguage. *International Review of Applied Linguistics* 10(3), 209–231.

Word order errors & Processing regimes

Dietmar Meurers and
Susanna Winkler

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing
A downside of rule rules

Our perspective and approach

Two types of word order errors

Phrasal verbs
Adverb placement

Summary

References

