

Towards a treatment of word order errors

When to use deep processing – and when not to

Vanessa Metcalf and Detmar Meurers

The Ohio State University

NLP in CALL Workshop
CALICO, May 16, 2006

Word order errors &
Processing regimes

Vanessa Metcalf and
Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs
Adverb placement

Summary

References



1 / 18

Background

The topic

- ▶ ICALL research has focused on diagnosing word-based learner errors (e.g. morpho-syntax, missing/extra word).
- ▶ 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?

Where we are coming from

- ▶ We're interested in linking CL research to genuine foreign language teaching (FLT) needs.
- ▶ We focus on developing web-based systems in support of traditional teaching, essentially an intelligent workbook approach.

Word order errors &
Processing regimes

Vanessa Metcalf and
Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs
Adverb placement

Summary

References



2 / 18

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.
 - ▶ Target:
 - (1) *He made his son clean the room.*
 - ▶ Learner production:
 - (2) *He get to cleaned his son.*

Example from Hiroshima English Learners' Corpus (HELIC 1998)

Word order errors &
Processing regimes

Vanessa Metcalf and
Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs
Adverb placement

Summary

References



3 / 18

Approaches to diagnosis 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., translation exercises (Heift 2001)

Word order errors &
Processing regimes

Vanessa Metcalf and
Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs
Adverb placement

Summary

References



4 / 18

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 (i.e., *mal*-rules)
 - manipulation of chart edges, the hypotheses introduced by phrase structure rules in a chart parser (Reuer 2003)

Word order errors & Processing regimes

Vanessa Metcalf and
Doreen Neuhues

Background

The topic
Word order and FLT

Approaches

List and match

Deep processing

Our perspective and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



5 / 18

Approaches to diagnosing word order errors

Deep processing: Downside

- ▶ 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, the modular unit licensed by a phrase structure rule.
- ▶ Licensing more word orders causes explosion of search space since word order is tied to 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

Vanessa Metcalf and
Doreen Neuhues

Background

The topic
Word order and FLT

Approaches

List and match

Deep processing

Our perspective and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



6 / 18

Our perspective and approach

- ▶ Word order error 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

Vanessa Metcalf and
Doreen Neuhues

Background

The topic
Word order and FLT

Approaches

List and match

Deep processing

Our perspective and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



7 / 18

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

Vanessa Metcalf and
Doreen Neuhues

Background

The topic
Word order and FLT

Approaches

List and match

Deep processing

Our perspective and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



8 / 18

Phrasal verbs

Linguistic description

- ▶ 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**

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



9 / 18

Phrasal verbs

Pedagogical relevance of particle placement

- ▶ English learners make errors in particle placement:
 - (7) a. * *so they give up it*
 - b. * *we will not catch up it.*
 - c. * *food which will build up him*
 - d. * *rather than speed up it.*

Examples from the Chinese Learner English Corpus (CLEC 2004)
- ▶ Learners may 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

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



10 / 18

Phrasal verbs

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.

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



11 / 18

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.
 - ▶ Desired error diagnosis (and feedback) is one-to-one with those patterns.
- ⇒ Particle placement is an example of a word order type for which instance-based regular expression matching is effective.

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



12 / 18

Adverb placement in English

- ▶ Pedagogical relevance
 - ▶ The rules governing adverb placement are difficult to articulate and master.
 - ▶ English has many different adverbs, and the word order possibilities depend on adverb subclass distinctions.
 - ▶ Many adverb placements are not right or wrong, but more or less natural.
- ▶ Students misplace adverbs in English
 - they **cannot already** live without the dope.*
 - There **have been already** several campaigns held by 'Outdoor'.*
 - while any covert action **brings rarely** such negative connotations.*
 - 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

Vanessa Mitchell and
Domenico Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



13 / 18

Adverb placement

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

Vanessa Mitchell and
Domenico Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



14 / 18

Adverb placement

Analysis

- ▶ Adverb placement can be described in terms of linear order with respect to constituents.
 - Sid*₁ *might*₂ *be*₃ *taking*₄ *his mother*₅ *to the store*₇.
 - clause-initial
 - preceding a finite auxiliary
 - preceding a nonfinite auxiliary
 - preceding a main verb
 - preceding an NP complement
 - preceding a PP complement
 - following the VP
- ▶ For each adverb subclass, we rate the positions in terms of acceptability (good, bad, awkward).
- ▶ In contrast to phrasal verbs, a characterization in terms of individual words is not possible.

Word order errors & Processing regimes

Vanessa Mitchell and
Domenico Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



15 / 18

Adverb placement

Processing the example exercises

- ▶ Instance-based shallow processing 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.
 - ▶ No reference to syntactic structure, which is needed for
 - ▶ identification of possible placements,
 - ▶ error diagnosis, and
 - ▶ content of feedback.
- ▶ Deep processing
 - ▶ Parsing identifies the necessary sentence structure.
 - ▶ The lexicon of the grammar allows us to model adverb classes.

Word order errors & Processing regimes

Vanessa Mitchell and
Domenico Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective
and approach

Two types of
word order errors

Phrasal verbs

Adverb placement

Summary

References



16 / 18

Adverb placement

Deep processing in our prototype

- ▶ In the implemented prototype, we parse sentences with all envisaged adverb placements, using an HPSG (Pollard and Sag 1994) grammar implemented in the TRALE system (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:** specifies what category the adverb combines with
 - ▶ **POSTHEAD:** specifies whether the adverb occurs before or 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.
- ▶ Given the lexical nature of the encoding of the adverb position, it is not dependent on local trees.

Word order errors & Processing regimes

Vanessa Metcalf and Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective and approach

Two types of word order errors

Phrasal verbs

Adverb placement

Summary

References



17 / 18

References

CLEC (2004). Chinese Learner English Corpus. Web interface to Corpus. .

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>.

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.

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>.

Odlin, Terence (1989). *Language Transfer: Cross-linguistic influence in language learning*. New York: Cambridge University Press.

Word order errors & Processing regimes

Vanessa Metcalf and Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective and approach

Two types of word order errors

Phrasal verbs

Adverb placement

Summary

References



18 / 18

Summary

When to use deep processing

- ▶ 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 which can only be obtained through deep analysis

When to use shallow processing instead

- ▶ lexical material and erroneous placements are predictable and listable
- ▶ there is no grammatical variation
- ▶ error patterns correspond directly to intended feedback

Word order errors & Processing regimes

Vanessa Metcalf and Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective and approach

Two types of word order errors

Phrasal verbs

Adverb placement

Summary

References



18 / 18

Odlin, Terence (2003). Cross-linguistic Influence. In Catherine Doughty and Michael Long (eds.), *Handbook on Second Language Acquisition*, Oxford: Blackwell, pp. 436–486.

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.

Pollard, Carl and Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press.

Reuer, Veit (2003). Error recognition and feedback with Lexical Functional Grammar. *CALICO Journal* 20(3), 497–512. http://www.cl-kl.uni-osnabrueck.de/~vreuer/publ/calico03_reuer.pdf.

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

Word order errors & Processing regimes

Vanessa Metcalf and Detmar Meurers

Background

The topic
Word order and FLT

Approaches

List and match
Deep processing

Our perspective and approach

Two types of word order errors

Phrasal verbs

Adverb placement

Summary

References



18 / 18