

# CUE-BASED ACQUISITION AND INFORMATION STRUCTURE DRIFT

Marit Westergaard  
*University of Tromsø – CASTL, NORMS*  
[marit.westergaard@hum.uit.no](mailto:marit.westergaard@hum.uit.no)  
<http://www.hum.uit.no/a/westergaard/>

DGFS, UNIVERSITÄT SIEGEN, FEBRUARY 27-MARCH 2, 2007  
WORKSHOP ON *THE ROLE OF INFORMATION STRUCTURE IN LANGUAGE CHANGE*

## 1. INTRODUCTION

Within a model of language acquisition and change which recognizes the importance of cues and ‘micro-cues’ in the input (e.g. Lightfoot, 2006, Westergaard 2006, forthcoming/a), this paper argues that patterns of information structure can be a factor causing word order change. Examples are taken from mixed grammars which allow two subject or object positions. In these systems children are exposed to the cue for a particular word order, e.g. OV or verb-second (V2), in only some of the relevant contexts in the primary linguistic data (PLD). The choice of word order in such systems is often (partly) dependent on information structure. This may cause a natural increase of one word order in language use, i.e. the E-language which constitutes the input to children in the acquisition process, and a corresponding drop in frequency of the other word order, which may then be vulnerable to change.

## 2. WORD ORDER VARIATION AND INFORMATION STRUCTURE

✚ Common within generative syntactic theory to assume that grammars may only consist of major parameters to be learnable.

✚ However: Considerable word order variation found in many languages.

(1) Peter **will never** drink this wine. /Peter **is never** late.  
Peter **never drinks** this wine. /\*Peter **drinks never** this wine.

(2) When **will Peter** drink this wine?  
On weekends **Peter will** drink this wine. /\*On weekends **will Peter** drink this wine.

(3) She looked **up the word**. /She looked **the word/it up**.

(4) I gave **Peter a glass of wine**. /I gave **a glass of wine to Peter**.

✚ In mixed grammars, patterns of information structure tend to govern the choice of the two word orders (Bresnan & Nikitina 2003, Bresnan et al 2007), e.g.

- Double object constructions, DP-DP preferred when IO is short and informationally given (often a pronoun) and DO correspondingly long and informationally new; DP-PP when informational status of the two objects is the other way around.
- 94% of all datives in a corpus predicted by a variety of factors: discourse accessibility, length, definiteness, animacy, etc.

(5) The movie gave <sub>DP</sub>[**me**] <sub>DP</sub>[**the creeps**]. /\*The movie gave the creeps to me.

(6) Stories like these must give <sub>DP</sub> [**the creeps**] <sub>PP</sub>[**to people whose idea of heaven is a world without religion...**]

✚ Word order variation in Norwegian:

❖ Two subject positions visible in questions/non-subject-initial declaratives w/negation:

- (7) Dette kan **Peter ikke** gjøre./Dette kan **ikke Peter** gjøre. **Neg-S/S-Neg**  
*this can Peter not do*  
 ‘This Peter can’t do.’

- Westergaard & Vangsnes (2005): informationally given subjects precede, informationally new/focused subjects follow negation/sentence adverbs.

- (8) Dette kan **du ikke** gjøre. **S<sub>PRO</sub>-Neg**  
*this can you not do*

- (9) \*?Dette kan **ikke du** gjøre./Dette kan **ikke DU** gjøre.  
*this can not you do*

❖ Object shift: Weak pronouns precede negation, full DPs+strong pronouns follow.

- (10) Peter så **ikke boka**. **Neg-O**  
*Peter saw not book.DEF*

- (11) Peter så **den ikke**./Peter så **ikke DEN**. **O<sub>PRO</sub>-Neg**  
*Peter saw it not /Peter saw not that*

❖ Two subject positions in embedded clauses, before and after sentence adverbs/negation. Weak pronouns virtually impossible in low position. Garbacz (2004): Neg-S extremely rare in written and spoken corpora (0.9%-7% of *that*-clauses).

- (12) Vi vet at **studentene ikke** ville drikke denne vinen. **S-Neg**  
*we know that student.DEF/PL not would drink this wine*

- (13) Vi vet at **ikke studentene** ville drikke denne vinen. **Neg-S**  
*we know that not student.DEF/PL would drink this wine*  
 ‘We know that the students wouldn’t drink this wine.’

❖ Tromsø dialect: Both V2 and non-V2 in questions with monosyllabic *wh*-constituents (Vangsnes 2005, Westergaard 2003, 2005b) – historical development V2 ⇒ non-V2.

- (15) kor er mitt fly? (INV, file Ole.17) **V-S**  
*where is my plane*  
 ‘Where is my plane?’

- (16) kor vi lande henne? (INV, file Ole.17) **S-V**  
*where we land LOC*  
 ‘Where do we land?’

- Westergaard (2003): V2 with informationally new/focused subjects (often a full DP), non-V2 with informationally given subjects (often a pronoun)

✚ Russian OV vs. VO dependent on discourse factors, pronominal objects virtually always in OV constructions, Diakonova (2003):

- (17) Ivan potseloval tsarevnu. **VO**  
*Ivan kissed princess.ACC*  
 ‘Ivan kissed the princess.’

- (18) Ivan tsarevnu potseloval. **OV**  
*Ivan kissed princess.ACC*

‘It was the princess that Ivan kissed.’ (from Diakonova, 2003: 11)

### 3. DIACHRONIC WORD ORDER VARIATION

- ✚ OV vs. VO in Old Icelandic (Hróarsdóttir 2004): OV chosen if object was given information (often a pronoun); VO if object was heavy and/or conveyed new information (often a full DP).

- (19) að hann hafi **hana** drepð OV (Old Icelandic)  
*that he had her killed*  
 ‘that he had killed her’ (from Hróarsdóttir 2004: 141)
- (20) hvört hann vilji ei kaupa **þræla** VO  
*whether he wanted not buy slaves*  
 ‘whether he didn’t want to buy slaves’ (from Hróarsdóttir 2004: 145)

- ✚ Mixed VO/OV also in the history of English (e.g. Pintzuk 1991, 2005, Roberts 1997, etc.). Various factors argued to play a role, e.g. OV appears if object is negated/quantificational. Nevertheless, certain preferences presumably subject to information structure: Pronouns virtually always in OV, heavy DPs strongly preferred in VO structures.

- (21) We ne magan **eow** neadian OV (OE)  
*we NEG can you constrain*  
 ‘We cannot constrain you. . .’ (from Pintzuk, 2005: 129)
- (22) Se wolde gelytlian **þone lyfigendan hælend** VO  
*he would diminish the living saviour*  
 ‘He would diminish the living saviour . . .’ (from Pintzuk 2005: 117)

- ✚ Similar situation in Middle Norwegian, Sundquist (2002).

- ✚ Choice of V2 vs. non-V2 in declaratives in Old and Middle English (OE/ME) dependent on information structure: non-V2 (XSV) preferred if subject is informationally given (often a pronoun); V2 (XVS) if the subject was focused or new information (often a full DP), see Bech (2001) and Westergaard (2005a).

- (23) Him **geaf** þa **se cync** twa hund gildenra pænenga. V2 (OE)  
*him gave then the king two hundred golden pennies*  
 ‘Then, the king gave him two hundred pence in gold.’
- (24) Hiora untrymnesse **he sceal** ðrowian on his heortan. Non-V2  
*their weakness he shall atone in his heart*  
 ‘He shall atone in his heart for their weakness.’ (from Haeberli 2002: 88-90)

- ✚ Two subject positions in embedded clauses in OE (above and below certain adverbs, *þa/þonne*), van Kemenade & Los (2006): Higher subject position preferred when subject is informationally given (often a pronoun); lower subject position when subject is new or focused (normally a full DP).

- (25) He ne mihte swapeah æfre libban, þeah ðe **he hine þa** ut alysd. (OE)  
*he not-could nevertheless ever live, though that they him then released*  
 ‘Nevertheless, he could not live forever, though they then released him’
- (26) Gif him **þonne God** ryhtlice & stræclice deman wile.  
*if him then God justly and strictly judge will*  
 ‘if God will then justly and strictly judge him’ (from van Kemenade & Los 2006: )

- ✚ Old Swedish: Verb movement (V-to-I) in embedded clauses if subject is a full DP, lack of verb movement with pronominal subjects.

- (27) æn min guḅ **brytar eigh** niḅar þin guḅ  
*if my god breaks not down your god*  
 ‘if my god does not break down your god.’ (from Sundquist 2002: 211)
- (28) æn han **eigh bannape** ihesus namne  
*if he not cursed Jesus name*  
 ‘if he did not curse Jesus’ name.’ (from Sundquist 2002: 227)

✚ Historical development:

- Icelandic/English/Mainland Scandinavian: OV ⇒ VO.
- English/Norwegian dialects: V2 ⇒ non-V2.
- Embedded subject positions in OE: Mainly high position found in ME.
- Embedded subject positions in Old Swedish: Loss of V-to-I, the word order linked to pronominal subjects survives.
  - High subject positions & low object positions survive, i.e. subject position preferred for informationally given subjects (pronouns) and object position linked to heavy or new/focused elements.
- Due to external effects? – e.g. Hroarsdottir (2004), Sundquist (2002).
- Unrelated developments – due to preferences in the acquisition process?

#### 4. CHILD LANGUAGE DATA

**Table 1: Overview of Norwegian corpus of child language, Tromsø dialect.<sup>1</sup>**

| Name of Child | Age            | Files     | Child Utterances |
|---------------|----------------|-----------|------------------|
| Ina           | 1;8.20-3;3.18  | Ina.01-27 | 20,071           |
| Ann           | 1;8.20-3;0.1   | Ann.01-21 | 13,129           |
| Ole           | 1;9.10-2;11.23 | Ole.01-22 | 13,485           |
| Total         |                |           | 46,685           |

✚ Children distinguish between informationally given and new subjects in the two types of *wh*-questions from early on, with the same patterns for subject/verb types as in adult data.

- (29) kor **er Ann sin dukke** hen? (Ann.04, age 1;11.0) **V2**  
*where is Ann POSS doll LOC*  
 ‘Where is Ann’s doll?’
- (30) ka **du skal** finne? (Ina.05, age 2;0.5) **Non-V2**  
*what you shall find*  
 ‘What do you want to find?’

✚ Similar patterns found in certain non-target-consistent (non-V2) non-subject-initial declaratives, where there is no input for this distinction.

- (31) her **er sekken**. (Ann.03, age 1;10.2) **V2**  
*here be.PRES backpack.DEF*  
 ‘Here is the backpack.’
- (32) der **Ann har** et. (Ann.08, age 2;1.28) **Non-V2**  
*there Ann have.PRES one*  
 ‘There Ann has one.’  
 Target form: Der har Ann et.

<sup>1</sup> Apart from 10 files collected and transcribed by the author, the corpus has been collected by Merete Anderssen.

✚ Delay attested in ‘subject shift’ constructions (Westergaard forthcoming/b):

- (33) har **ikkje han** fota her? (Ina.13, age 2;5.25) **Neg-S<sub>PRO</sub>**  
*have.PRES not he feet here*  
 ‘Doesn’t he have feet here?’  
 Target form: Har han ikkje fota her?

✚ Early preference for low subject position in embedded clauses, cf. examples (12)-(13):

- (34) ...ho si at **ikkje det** er min kjæreste. (Ina.27, age 3;3.18) **Neg-S**  
*...she say.INF/PRES that not it be.PRES my sweetheart*  
 ‘She says that it isn’t my sweetheart.’  
 Target form: Ho sir at det ikkje er min kjæreste.

✚ Delay attested in object shift constructions:

- (35) eg finn <**ikkje han**> [>]. (Ina.13, age 2;5.25) **Neg-O<sub>PRO</sub>**  
*I find.PRES not him*  
 ‘I can’t find him.’  
 Target form: Eg finn han ikkje.

- Similar delay attested for object shift in Swedish (Josefsson 1996), and for object scrambling in Dutch and German (Schaeffer 2000, Barbier 2000).

✚ Russian VO/OV (Diakonova 2003): Both word orders produced from early on, occasional examples of VO with pronominal objects at an early stage.

- Children have an early sensitivity to patterns of information structure.
- Children have an early preference for the lower position of two available ones (due to economy of movement). This could be a factor for OV ⇒ VO – but would have nothing to say about subjects, where typically the high position survives.
- Generally, target word order falls into place around age 2;6-3 (subject shift) - for certain constructions somewhat later (object shift, object scrambling).

✚ Perhaps an answer may be sought in the E-language input (PLD)?

## 5. CUE-BASED ACQUISITION AND FREQUENCIES IN CHILD-DIRECTED SPEECH

✚ Cue-based acquisition and change, Lightfoot (2006): Children scan the PLD for designated cues, which are abstract pieces of structure, e.g.  $_{VP}[DP V]$  or  $_{CP}[XP V]$ .

- Structures with a low input frequency may be ignored by children in the acquisition process and disappear from the I-language grammar of the next generation.
- Language change is not gradual, but ‘catastrophic’, reflecting I-language differences between the grammars of two generations.

✚ ‘Micro-cues’ in acquisition, Westergaard (2006, forthcoming/a, submitted):

- Different word orders often distinguished by information structure and various syntactic ‘micro-cues’, e.g. different initial elements in V2 constructions (*pa/ponne* in OE/ME, different *wh*-elements in Norwegian dialects), or different types of object (negated/quantificational) in OV/VO constructions.
- Children easily learn exceptions to general rules, if variation is based on natural classes of linguistic categories (e.g. different types of verbs, different *wh*-elements),

or if exceptions to the general rules are relatively frequent in the input (e.g. *kanskje* ‘maybe’ as initial element allowing non-V2 in Scandinavian declaratives).

- Gradualism in historical data presumably the result of ‘many small catastrophes’.

✚ Patterns of information structure may gradually cause a statistical shift in the frequency of a particular cue in the input to children.

- Subjects tend to be given information, while objects more often express new/focused information.

**Table 1: Overview of the realization of subjects and objects in a sample of Norwegian conversational speech (speakers 119 and 120 in the Nota corpus).**

|          | <b>Pronouns<br/>(pers/refl)</b> | <i>det</i> (‘it/that’) | <b>DPs/clauses</b> | <b>Misc/Other</b> | <b>Total</b> |
|----------|---------------------------------|------------------------|--------------------|-------------------|--------------|
| Subjects | 340/0 (57.8%)                   | 204 (34.7%)            | 35/0 (6.0%)        | 9 (1.5%)          | 588 (100%)   |
| Objects  | 11/22 (15.4%)                   | 37 (17.3%)             | 104/38 (66.4%)     | 2 (0.9%)          | 214 (100%)   |

- Disregarding *det* ‘it|that’ (tends to be expletive/referential as subjects, demonstratives as objects), subjects realized by pronouns 10 times more often than by full DPs (340 vs. 35).
- Objects realized by DPs/clauses 4-5 times more often than as pronouns (142 vs. 33).

**Table 2: Overview of the realization of subjects and objects in a sample of Norwegian child-directed speech (INV in Ole.15).**

|          | <b>Pronouns<br/>(pers/refl)</b> | <i>det</i> (‘it/that’) | <b>DPs/clauses</b> | <b>Misc/Other</b> | <b>Total</b> |
|----------|---------------------------------|------------------------|--------------------|-------------------|--------------|
| Subjects | 272/0 (67.5%)                   | 99 (24.6%)             | 29/0 (7.2%)        | 3 (0.7%)          | 403 (100%)   |
| Objects  | 26/3 (16.4%)                    | 40 (22.6%)             | 74/32 (59.9%)      | 2 (1.1%)          | 177 (100%)   |

- Again, disregarding *det* ‘it|that’, subjects are realized by pronouns about 10 times more often than by full DPs (272 vs. 29).
- Objects are realized by DPs/clauses 3-4 times more often than as pronouns (106 vs. 29).

**Table 3: Overview of the realization of subjects and objects in a sample of English child-directed speech (MOT in Adam.10).**

|          | <b>Pronouns<br/>(pers/refl)</b> | <i>it/there</i> | <i>that</i> | <b>DPs/<br/>clauses</b> | <b>Misc/<br/>Other</b> | <b>Total</b> |
|----------|---------------------------------|-----------------|-------------|-------------------------|------------------------|--------------|
| Subjects | 193/0 (65.4%)                   | 16/2 (6.1%)     | 46 (15.6%)  | 35 (11.9%)              | 3 (1.0%)               | 295 (100%)   |
| Objects  | 7/1 (6.2%)                      | 14/0 (10.8%)    | 10 (7.7%)   | 96 (73.8%)              | 2 (1.5%)               | 130 (100%)   |

- Disregarding *that*, subjects are realized by pronouns or *it/there* 6 times more often than as full DPs (211 vs. 35).
- Objects are realized by DPs/clauses 4 times more often than as pronouns (96 vs. 22).
- Word orders linked to given (pronominal) subjects will be naturally frequent in the E-language – order linked to new/focused (full DP) subjects correspondingly infrequent.
- Word orders linked to given (pronominal) objects will be naturally infrequent – word orders linked to new/focused (full DP) objects correspondingly frequent.

## 6. INFORMATION STRUCTURE DRIFT

- Children easily learn different word orders that are dependent on information structure, e.g. V2 vs. non-V2 in Norwegian, VO vs. OV in Russian, etc.
- E-language pattern: Subjects predominantly given info, objects new/focused info.
- Structures with an extremely low input frequency vulnerable to change.
- Input varies! Children typically get input from many different *adult* and *child* speakers. Some children exposed to extremely few instances of the word order used with informationally new subjects/informationally given objects – may develop a grammar with only one subject or object position.

## 7. SUMMARY/CONCLUSION

- + When syntax allows two word orders, information structure patterns will typically distinguish between them.
- + Mixed word order systems easily learnable by children ⇒ early sensitivity to information structure.
- + Many word order changes in historical data involve survival of subject positions that are reserved for informationally given subjects and object positions that are typically used for informationally new objects.
- + In the E-language that children are exposed to, subjects are typically given info and objects new ⇒ causes a higher frequency of certain word orders in the input (Information Structure Drift).
- + External factors may reverse such developments.

## REFERENCES

- Barbier, I. 2000. "An Experimental Study of Scrambling and Object Shift in the Acquisition of Dutch." In S. M. Powers and C. Hamann (eds.), *The Acquisition of Scrambling and Cliticization*, 41-69. Dordrecht: Kluwer Academic Publishers.
- Bech, Kristin. 2001. *Word Order Patterns in Old and Middle English: A Syntactic and Pragmatic Study*. Doctoral dissertation, University of Bergen.
- Bresnan, Joan and Tatiana Nikitina. 2003. 'On the gradience of the dative alternation.' Ms., Stanford University.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina and Harald Baayen. 2007. 'Predicting the Dative Alternation.' In G. Boume, I. Kraemer and J. Zwarts (eds.) *Cognitive Foundations of Interpretation*. Amsterdam: Royal Netherlands Academy of Science, pp. 69--94.
- Diakonova, Marina. 2003. *The Acquisition of Word Order in English and Russian*. MA thesis, University of Tromsø.
- Garbacz, Piotr. 2004. 'Bisatsstrukturer som man kan inte bortse från (om bisatsordföljd i de fastlandsskandinaviska språken) [Embedded structures that cannot be disregarded [on embedded clause word order in Mainland Scandinavian].' Ms., University of Lund.
- Haeberli, Eric. 2002. 'Inflectional Morphology and the Loss of Verb-Second in English.' In D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*, pp. 88-106, Oxford/New York: Oxford University Press.
- Hróarsdóttir, Thorbjörg. 2004. Cues and expressions. *Nordlyd* 32.1: *Tromsø Working Papers in Language Acquisition*, 135-155.
- Josefsson, G. 1996. "The Acquisition of Object Shift in Swedish Child Language." In C. E. Johnson and J. H. V. Gilbert (eds.), *Children's Language* 9, 153-165. Mahwah: Lawrence Erlbaum.

- Kemenade, Ans van and Bettelou Los. 2006. 'Discourse adverbs and clausal syntax in Old and Middle English.' In Ans van Kemenade and Bettelou Los (eds.), *The Handbook of the History of English*, 224-248. Malden, MA and Oxford: Blackwell.
- Lightfoot, David. 2006. *How New Languages Emerge*. Cambridge University Press.
- NoTa – Norsk Talespråkskorpus, Oslodelen [Corpus of Spoken Norwegian, the Oslo part], Tekstlaboratoriet, ILN, University of Oslo - <http://www.tekstlab.uio.no/nota/oslo/index.html>.
- Pintzuk, Susan. 1991. *Phrase Structures in Competition: Variation and Change in Old English Word Order*. Ph. D. dissertation, University of Pennsylvania.
- Pintzuk, Susan. 2005. 'Arguments against a universal base: evidence from Old English.' *English Language and Linguistics* 9.1, 115–138.
- Roberts, Ian G. 1997. 'Directionality and word order change in the history of English.' In Ans van Kemenade and Nigel Vincent (eds.), *Parameters of Morphosyntactic Change*, 396-426. Cambridge: Cambridge University Press.
- Schaeffer, J. 2000. *The Acquisition of Direct Object Scrambling and Clitic Placement: Syntax and Pragmatics*. Amsterdam/Philadelphia: John Benjamins.
- Sundquist, John D. 2002. *Morphosyntactic Change in the History of the Mainland Scandinavian Languages*. PhD dissertation, Indiana University.
- Vangsnes, Øystein A. 2005. 'Microparameters for Norwegian *wh*-grammars.' *Linguistic Variation Yearbook* 5, pp. 187-226. Amsterdam and Philadelphia: John Benjamins.
- Westergaard, Marit R. 2003. 'Word Order in *Wh*-Questions in a North Norwegian Dialect: Some Evidence from an Acquisition Study.' *Nordic Journal of Linguistics* 26.1, 81-109.
- Westergaard, Marit R. 2005a. 'Norwegian child language and the history of English: The interaction of syntax and information structure in the development of word order.' In Kevin McCafferty, Tove Bull and Kristin Killie (eds.), *Contexts - Historical, Social, Linguistic. Studies in Celebration of Toril Swan*, 293-410. Bern: Peter Lang.
- Westergaard, Marit R. 2005b. 'Optional word order in *wh*-questions in two Norwegian dialects: A diachronic analysis of synchronic variation.' *Nordic Journal of Linguistics* 28.2, 269-296.
- Westergaard, Marit R. 2006. 'Triggering V2: The amount of input needed for parameter setting in a Split-CP model of word order.' In Adriana Belletti, Elisa Bennati, Cristiano Chesi, Elisa DiDomenico and Ida Ferrari (eds.), *Language Acquisition and Development: Proceedings of GALA 2005*, 658-671. Cambridge: Cambridge Scholars Press.
- Westergaard, Marit R. Forthcoming/a. 'Learning and Unlearning V2: On the Robustness of the Triggering Experience in a Historical Perspective.' In *Proceedings of Incontro di Grammatica Generativa* 32. Florence: Edizioni dell'Orso.
- Westergaard, Marit R. Forthcoming/b. 'Verb Movement and Subject Placement in the Acquisition of Word Order: Pragmatics or Structural Economy?' In Pedro Guijarro-Fuentes, Pilar Larranaga and John Clibbens (eds.), *First Language Acquisition of Morphology and Syntax: Perspectives across languages and learners*. [Language Acquisition and Language Disorders], John Benjamins.
- Westergaard, Marit R. Submitted. 'Many Small Catastrophes: Gradualism in a Microparametric Perspective.' Submitted to volume on diachronic generative syntax (Oxford University Press).
- Westergaard, Marit R. & Øystein A. Vangsnes. 2005. '*Wh*-questions, V2, and the left periphery of three Norwegian dialects.' *Journal of Comparative Germanic Linguistics* 8, 117-158.