

ANNOTATION MANUAL DATIVE ALTERNATION

author: Daphne Theijssen
affiliation: Department of Linguistics, Radboud University Nijmegen
date: August 2008
used on: ICE-GB Corpus
topic: dative alternation

AnimacyRec/Theme (nominal)

Animacy of recipient and theme

- a = animate: human + animal
- in = inanimate: not human or animal, includes companies and organizations, though it depends on the context: e.g. The police are giving out fines (animate => policemen) and Send your complaint to the police (animate => people working at police station) vs. More tax money should be spent on the police. (inanimate => institution).

CaseID (nominal)

Case identification codes

For instance, case names are S1B-009_12:2:A_1, meaning:

- S: spoken (vs. W written)
- 1: type 1 = dialogues (see <http://www.ucl.ac.uk/english-usage/ice/textcats.htm>)
- B: type B = public (see <http://www.ucl.ac.uk/english-usage/ice/textcats.htm>)
- 009: text number 009
- 12: sentence number 12
- A: speaker A (only in spoken data)
- 1: clause 1 (following our extraction script)

ClauseFunction (nominal)

Clause function

Only the main clause is labelled m = main, the rest is embedded. Clauses preceded by subordinators are always ‘embedded’. An embedded clause has either a phrasal = eph or a clausal = ecl function. Examples:

- phrasal:
“when you have a vague feeling [you 've lent a book to somebody]” (clause is part of the NP “feeling”)
- clausal:

“Well we’ll ... ask one of the consultatives there [to send you an appointment]”
(clause is the complement of the verb ‘to ask’)

ClauseMode (nominal)

Clause mode

The clause is either int = interrogative, imp = imperative or d = declarative.

ClauseWO (nominal)

Clause word order

The word order of the clause is either fth = fronting of theme, freq = fronting of recipient, unm = unmarked or rev = reversed. Examples:

- fronting of theme:

“But uh [hardbacks]_{THEME} I wouldn't lend to [anyone]_{RECIPIENT}”

- fronting of recipient:

“The Family Practitioner Committee (FPC) to [whom]_{RECIPIENT} [your form]_{THEME} will be sent may however check your claim at a later date.”

- unmarked:

“Do you send [electronic mail]_{THEME} to [people that you know]_{RECIPIENT} though” (thus including subject-verb inversion, as in questions or marked expressions)

- reversed:

“Tonight I want to explain to[you]_{RECIPIENT} [why they are there and what this conflict is about <, >]_{THEME}” (unmarked but theme and recipient have been reversed)

ClEmbClauseImportance (nominal)

Clausal embedded clause importance

A clausal embedded clause is either a = an adjunct (can be left out without making the sentence ungrammatical) or c = a complement (obligatory in the syntactic structure).

For main clauses and phrasal embedded clauses, this feature is na = non-applicable.

CollostrStrength (continuous)

Collostructional Strength

Gries and Stefanowitsch (2004) introduce the collostructional strength: the bias of a verb towards one of the two constructions. It can be calculated by counting, for each construction (in this case NP NP and NP PP): the number of occurrences of a verb with the construction and the number of other verbs with the construction. A table with the found numbers for the

verb and other verbs in the constructions is fed to the Fisher exact test (we used R). The p-value indicates the collostructional strength for that verb.

Next, we calculated the expected number of verbs with the construction (proportion of construction in the data times the number of occurrences of the verb). Next, the bias can be established (being the construction for which the number of verbs found is higher than the expected number). We redefine the collostructional strength in such a way that it indicates both the strength and the direction of the bias of the verb: NP PP: $1 - p$; NP NP: $-1 * (1 - p)$. A value near 0 indicates that there is no or only a small bias. A value near -1 indicates a strong bias towards NP NP, a value near 1 towards NP PP.

ConcretenessRec/Theme (nominal)

Concreteness of recipient and theme

Objects that have a known physical size, are labelled c = concrete, others are in = inconcrete. Organisations are typically inconcrete when they are meant in the animate sense (actually meaning the people in the organisation). The personal pronouns 'you' and 'us' are only 'concrete' when they indicated a known group of people, for example the students in the class room. When the group is not known yet, for example the students that will start next year (without knowing how many and who will subscribe), it is considered inconcrete.

Construction (nominal)

Construction

- N = NP NP (double object construction)
- P = NP PP (prepositional *to*-dative)

DefinitenessRec/Theme(nominal)

Definiteness of recipient and theme

- d = definite:
 - * head is preceded by definite article ("the")
 - * head is (preceded by) demonstrative pronoun ("these"/"this"/"those"/"that")
 - * head is (preceded by) interrogative or (nominal) relative pronoun (e.g. "who", "whose")
 - * head is (preceded by) possessive pronoun (e.g. "my", "mine")
 - * head is a reciprocal pronoun (e.g. "each other")
 - * head is a reflexive pronoun (e.g. "himself")
 - * head is a personal pronoun (e.g. "her")
 - * head is proper noun (names of companies, people, etc.)
- in = indefinite:
 - * head is preceded by indefinite article ("a", "an")
 - * head is (preceded by) (non)assertive pronoun (e.g. "any", "somebody")

- * head is (preceded by) universal pronoun (e.g. "both", "every")
- * head is (preceded by) negative pronoun (e.g. "no")
- * head is (preceded by) quantifying pronoun (e.g. "a bit")
- * head is (preceded by) pronoun "one"
- * head is not preceded by a determiner, and is not a proper noun

EmbClauseType (nominal)

Embedded clause type

An embedded clause is either sub = subordinate, rel = relative or irel = independent relative. Difficult are the to-infinitival clauses, such as those below:

- 1) “*And when you 're budding of course you want to put three four or five buds onto the top of that stalk to make sure that wh if they do grow even if one or two die you 've still got plenty left [to give you a good head]*”
- 2) “*Now this within your genes is the information to uhm <.,> [to tell cells in your body what to do]*”
- 3) “*And I lose no opportunity [to point this out to my colleagues within the European Community]*”

The (simplified) rule of thumb we applied is to check whether the to-infinitival clause can be transformed into a relative clause:

- 1) relative (“*plenty left that give you a good head*” = okay, so empty subject = preceding np)
- 2) relative (“*information that tells cells in your body what to do*” = okay, so empty subject = preceding np)
- 3) subordinate (“*opportunity that points this out to my colleagues within the European Community*” = nonsense, so subordinate)

For main clauses, this feature is na (non-applicable).

GivennessRec/Theme (continuous)

Givenness of recipient and theme

For givenness (discourse accessibility), we followed the taxonomy introduced in Grondelaers, Stefan, and Kris Heylen 2003. *Towards a "new" given-new taxonomy*. Paper presented at the 8th International Cognitive Linguistics Association Conference, Logroño (La Rioja).

Our interpretation is as follows:

- 1) to be created and unconstrained (indefinite NPs, including abstract notions such as *tribute* as in “*Michael Heseltine pays tribute to Mrs <<w>> Thatcher 's remarkable premiership <.,>*”)
- 2) to be created but constrained (indefinite NPs preceded by an adjective or followed by a modifier without an anchor element)
- 3) to be created on the basis of an anchor element (indefinite NPs followed by a modifier with an anchor element), score = $0.75 * 3 + 0.25 * \text{the givenness of the anchor element}$
- 4) available in the encycopaedic context or mentioned >20 clauses before (long definite NPs, proper nouns)

- 5) inferrable from an anchor element (long definite NP of NP with possessive pronoun, e.g. “*my mother*”), score = $0.75 * 5 + 0.25 * \text{the givenness of the anchor element}$
- 6) available in the context <20 and >1 clauses before (long definite NPs, often with demonstrative)
- 7) inferrable from entity <20 clauses before on the basis of stereotypes (short definite NPs and the use of ‘you’ and ‘us’ in the meaning of ‘one’)
- 8) available in the same of preceding clause (often 3rd person pronouns)
- 9) available in the speech context (often 1st and 2nd person pronouns, only used in dialogues, hardly ever in written text, unless the writer refers to the book itself for example)

In case of real doubt:

- if the object is definite: 4
- if the object is indefinite: 2

LengthAdverbChars/Words (discrete)

Length of intervening adverbials in characters and words

Any adverbial between the **theme** and the **recipient**.

NB: This does not mean that all intervening words are automatically included:

1. in clauses with a mark word order, the subject and/or verbs may appear between the recipient and the theme
2. sometimes elements should be ignored, e.g. consider the following clause:

While she 's While she 's doing that could you tell me what the <, > you were speaking a bit of nucleated form what the E in transcription what the C stands for <, >

The theme, recipient and verb have been marked below:

While she 's While she 's doing that could you [tell]_{VERB} [me]_{RECIPIENT} what the <, > you were speaking a bit of nucleated form [what the E in transcription what the C stands for]_{THEME} <, >

The following text can be found between the recipient and the theme and is therefore a potential ‘intervening adverbial’:

While she 's While she 's doing that could you tell me [what the <, > you were speaking a bit of nucleated form] what the E in transcription what the C stands for <, >

In ICE, the part *what the <, >* is marked as an UNTAG, with feature ‘ignore’:

While she 's While she 's doing that could you tell me [what the <, >]_{IGNORE} you were speaking a bit of nucleated form what the E in transcription what the C stands for <, >

This makes sense since it is the start of the theme before the speaker realises he/she wants to insert an adverbial. Therefore, we consider the following part of the clause the ‘intervening adverbial(s)’:

While she 's While she 's doing that could you tell me what the <, > [you were speaking a bit of nucleated form]ADVERBIAL what the E in transcription what the C stands for <,, >

The text we ignored is that:

- between [] and < >
- all text parsed with 'ignore'
- punctuation marks (they are marked with ascii codes, so we checked which actually represent letters and which punctuation marks, and only ignored the last category)

The remaining text is used to determine the length in words and characters:

Words are all elements separated by white space.

Characters are all elements except the white spaces.

LengthRec/ThemeChars/Words (discrete)

Length of recipient and theme in characters and words

The text we ignored is that:

- between [] and < >
- all text parsed with 'ignore'
- punctuation marks (they are marked with ascii codes, so we checked which actually represent letters and which punctuation marks, and only ignored the last category)

The remaining text is used to determine the length in words and characters:

Words are all elements separated by white space.

Characters are all elements except the white spaces.

Medium (nominal)

Medium

- s = spoken (as coded in ICE-GB)
- w = written (as coded in ICE-GB)

NumberRec/Theme (nominal)

Number of recipient and theme

- s = singular: singular nouns and mass nouns
- p = plural: plural nouns

For words such as 'that', 'what' and other words of which the number is not directly clear, the antecedent was checked.

PersistenceRec/Theme (discrete)

Persistence or recipient and theme

Persistence is described in: Stefan Grondelaers and Dirk Speelman. Constituent ordering in presentative sentences. *Corpus Linguistics and Linguistic Theory*. 2008.

It is defined as the number of clauses that the object (or a synonym of the object) is used after its mentioning. Interventions of single clauses are allowed, but are not counted.

In cases such as:

“*[Well yours in Cambridge have always told me the same story [that mine have]]*”

the object of the clause (“*the same story that mine have*”) includes a clause itself (“*that mine have*”). When this clause does not include the object, we allowed it to be skipped without losing the right of the intervening clause. When it does include the object (as in this case with “*mine*” which can be rewritten as “of me”), it is counted. This may seem arbitrary, but it made most sense when looking at the data. On the one hand, we do not want to ‘punish’ an object when the following object consists of (a) clause(s), and on the other hand, we do not want to exclude mentionings in these clausal objects.

PersonRec/Theme (discrete)

Person of recipient and theme

- 1: "I", "me", "mine", "myself", "we", "us", "ours", "ourselves"
- 2: "yourself", "you", "yours"
- 3: rest

Polarity (nominal)

Polarity

All clauses are p = positive, except those in which the verb is modified by ‘not’, ‘n’t’, ‘never’ or adverbs with similar meanings, such as ‘hardly’. These are labelled n = negative.

PronominalityRec/Theme (nominal)

Pronominality of recipient and theme

- p = pronominal: headed by a pronoun (head = pronoun)
- non = nonpronominal: not headed by a pronoun (head = not a pronoun)

SemVerbClass (nominal)

Semantic verb class

t = transfer_of_possession:

The theme is literally moved from the actor to the recipient

* e.g. “I gave you a letter.”, “I’ll send you a prescription.”

f = future_transfer_of_possession:

The theme will be literally moved from the actor to the recipient in the future

* e.g. “I promise you ten pounds.”

p = prevention_of_possession:

The theme is prevented from literally moving from the actor to the recipient

* e.g. “They denied me my salary.”

c = communication:

The actor communicates something to the recipient in one of the following ways:

- 1) The verb is a communication verb¹ and there is literal communication (verbally or physically)
* e.g. “I told him a story”, “I showed him the picture”
- 2) The verb is not a communication verb but the theme is a form of communication
* e.g. “I gave him a message” (compared to “I gave him a letter”, which is transfer_of_possession!)
- 3) The verb is a communication verb but there is no literal communication, but the actor is a means of communication
* e.g. “The letter told him so.”, but also “The laryngograph shows us that...”
because the goal of the laryngographs is to communicate to us what the sound is like (compared to “The rocks show that water used to run there” in which the goal of the rocks is not to communicate this to us)

¹ Communication verbs: *advise, assure, command, communicate, confirm, demonstrate, dictate, disclose, explain, express, fax, formulate, guarantee, illustrate, mouth, order, persuade, play (of sounds), present, promise (when interchangeable with ‘tell’), quote, read, report, show, teach, tell, wish, write*

a = abstract:

The verb belongs to one of the above classes but there is a figurative meaning

* e.g. “I gave it some thought.”, “The rocks show that water used to run here.”

StructureParallellism (nominal)

Structure parallellism

For this purpose we included all cases, also those with unexpected word order (for instance PP NP), with particle verbs, with clausal objects and with verb senses only present in one of the two constructions (in our data set).

For each case, we checked whether there is a preceding construction used in the same text, maximally 20 clauses before. When this construction is the same as that under consideration, we label it s, when it is different d and when no construction is present in the maximally 20 preceding clauses, it is nf (not found).

A text is defined as those with the same text and subtext identification code in ICE-GB (see Case identification code).

Verb (nominal)

Verb in original form

The verb in its original form (e.g. *given*).

VerbLemma (nominal)

Verb lemma

The lemma of the verb (e.g. *give* for *given*).

VerbSense (nominal)

Verb sense

This consists of the verb lemma plus its semantic class (e.g. *give* with class 'transfer of possession' is *give_t*).

Voice (nominal)

Voice

The voice of the clause is either p = passive or a = active.