

ANNOTATION MANUAL DATIVE ALTERNATION

author: Daphne Theijssen
affiliation: Department of Linguistics, Radboud University Nijmegen
date: April 2009
used on: ICE-GB Corpus
topic: dative alternation

reference:

Daphne Theijssen (2010). Variable selection in Logistic Regression: The British English dative alternation. Thomas Icard and Reinhard Muskens (eds.), *Interfaces: Explorations in Logic, Language and Computation*. Series: Lecture Notes in Computer Science (subseries: Lecture Notes in Artificial Intelligence), volume 6211, Springer, pp. 87-101.

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)

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*).

Construction (nominal)

Construction

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

AnimacyRec (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).

ConcretenessTheme (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 know yet, for example the students that will start next year (without knowing how many and who will subscribe), it is considered inconcrete.

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

GivennessRec/Theme (nominal)

Givenness of recipient and theme

- **g** = given: having a givenness score ≥ 6
- **non** = nongiven: having a givenness score < 6

To establish the givenness score (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 = $0.75*3 + 0.25*$ 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 = $0.75*5 + 0.25*$ 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 or 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

LengthDifference (interval)

Length of recipient and theme in words

$\ln(\text{nr of words in theme}) = \ln(\text{nr of words in recipient})$

Words are all elements separated by white space. We used a Perl script to establish this feature. 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)

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.

PersonRec (nominal)

Person of recipient and theme

- l = local: *I, me, mine, myself, we, us, ours, ourselves, yourself, yourselves, you, yours*
- non = non-local: rest

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.”

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

* e.g. “I promise you ten pounds.”

Or 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 verb 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)

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.”

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*

StructureParallellism (nominal)

Structure parallellism

With the help of a Perl script, we checked whether there is a preceding dative construction used in the same text, maximally 20 clauses before. When this construction is the prepositional dative (P), it is labelled y (yes), when it the double object construction (N), it is labelled n (no).

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