



Multilingual Corpus of  
Annotated Spoken Texts

# English

— annotation notes —

June 2016

Nils Norman Schiborr  
University of Bamberg

cite this document as:

Schiborr, Nils N. 2016.  
Multi-CAST English annotation notes.  
In Haig, Geoffrey & Schnell, Stefan (eds.),  
*Multi-CAST (Multilingual Corpus of Annotated Spoken Texts)*.  
(handle) (date accessed.)

file v1.0



Language Archive  
Cologne

## Annotation notes

### 'Gapped' predicates and objects in VP-echo structures

Spoken English makes extensive use of various 'dummy' auxiliaries when identical VPs are repeated (VP-echoes), as in *they paid the full fare and we always have done too*. This strategy is particularly common in replies to questions: *are you studying linguistics? Yes, I am*. In many cases, the substituted phrases include direct objects, at least conceptually. In dealing with constructions of this nature, zero glosses of the kind <0:pred> and <0:p> have been inserted, as in the following examples:

- (1) a. *He said, You can't leave me!*  
 b. # *I* *said*, # *Well*, *I* =*'ll* *have*  
 # pro.1:s\_ds v:pred #ds other\_dp pro.1:a =waux aux  
*to* 0 0  
 lv\_to 0:pred 0:p

*english\_devon01\_036*

- (2) a. *My word was mi bond*,  
 b. # *and* 0 *always* *has* *been* 0  
 # other 0:s other aux cop 0:pred

*english\_kent01\_317–318*

- (3) a. *I said, You, can't you build me one, I said, With three rows?*  
 b. # *Yeah*, 0 *can* 0 0  
 #ds other\_dp 0.1:a aux 0:pred 0:p

*english\_kent02b\_131–132*

### Primary and secondary roles: recipients and others

With some three-participant verbs (including speech, transfer of possession, and a number of others) in English, the recipient, beneficiary, or addressee is treated formally as a direct object, in that it occurs immediately post-verbally and without a preposition. These arguments have been glossed <p>, with the accompanying theme arguments receiving the <p2> gloss, as in (4). Note that recipients, beneficiaries, and addressees that do take a preposition have annotated as <g>.

- (4) # *No-one* *gave* *us* *a* *saw*  
 # indef\_pro.h:a v:pred pro.1:p ln\_deti np:p2

*english\_london01f\_149*

- (5) # *My* *sisters* *gave* *it* *to* *me*  
 # ln\_pro.1:poss np.h:a v:pred pro:p adp pro.1:g

*english\_kent02b\_356*

## Direct speech

As direct speech is syntactically independent and may span several clauses, it has not been annotated as the object of the clauses that introduce or conclude it (e.g. *she said, ...*). In contrast to verbs of speech with NP objects, which may be transitive (e.g. *she said nothing*) or ditransitive (e.g. *she told us a story*), verbs of speech bookending direct speech have hence been annotated as intransitive, as in (6), or transitive if an addressee is present, as in (7). In order to preserve their special status, their subjects have been glossed with the additional subcategorising symbol <\_ds>, yielding <:s\_ds> and <:a\_ds> for intransitive and transitive clauses respectively.

- (6) # *He*            *says,* # *Walk*            *him*    *down*    *the*    *field,*  
 # **pro.h:s\_ds** v:pred #ds vother:pred pro:p adp ln\_det np:g  
 # *he*            *says,* # *And*    *gallop*        *back*  
 # **pro.h:s\_ds** v:pred #ds other vother:pred rv

*english\_kent01\_251*

- (7) # *So*    *she*            *told*    *her*            *groom,*  
 # other **pro.h:a\_ds** v:pred ln\_pro.h:poss np.h:p  
 # *Put*            *the*    *horse*    *in*    *the*    *cart*  
 #ds vother:pred ln\_det np:p adp ln\_det np:g

*english\_kent02a\_047*

## Light verb constructions

Light verb constructions in English combine a semantically bleached light verb (e.g. *do, have, take, etc.*) with a non-verbal element, creating a new predicate with often idiomatic meaning. While light verbs themselves have been glossed simply <v:pred>, their complements have been annotated <xnp:lvc>, indicating that they neither referential and nor the object of the predicate.

- (8) # *But*    *men*    *took*    *pride*    *in*    *the*    *fact*  
 # other np.h:s **v:pred xnp:lvc** adp ln\_det np:other  
 # *that*    *they*        =*'d*        *worked*    *hard*  
 # other pro.h:s =waux v:pred other

*english\_london01e\_125*

## Possessive determiners

English possessive determiners (*my, your*) may occur without an explicitly mentioned possessum, in which case they take on an altered form (*mine, yours*). Both the former and the latter have been annotated as constituents of the possessum-NP (e.g. <ln\_pro:poss>), and for the latter, the omitted possessum has been restored as zero <0>. Compare (9) and (10):

(9) # *She wrote to my father*  
 # pro.h:s v:pred adp ln\_pro.1:poss np.h:g

*english\_kent02a\_045*

(10) # *You know, they want theirs* 0  
 # other other pro:a v:pred ln\_pro:poss 0:p

*english\_kent02b\_181*

### Null subjects and auxiliaries in coordinated clauses

In coordinated clauses of the kind in (11), echoed auxiliaries are omitted in subsequent clauses if the subject not also expressed repeatedly. In such cases, a zero element <0> has been introduced for the subject, but not for unexpressed auxiliaries.

(11) # *I used to go up there*  
 # pro.1:s aux lv\_to v:pred adp other:g  
 # *and* 0 *load it*  
 # other 0.1:a v:pred pro:p

*english\_kent02a\_129*

### Impersonal pronouns

Non-referential constructions using the pronouns *you*, *they*, or *one* have been annotated as <imp\_pro>, as in (12). In order to preserve the formal similarity to their personal counterparts, impersonal *you* and *they* have received additional second and third person glosses, resulting in <imp\_pro.2> and <imp\_pro.h>.

(12) # *If you was a tradesman,*  
 #ac other imp\_pro.2:s cop ln\_deti np.h:pred  
 # *you could get ... a pound a week.*  
 # imp\_pro.2:a aux v:pred ln\_deti np:p adp other

*english\_devon01\_033*

### Tag questions and discourse markers

The verbal element of tag questions has been glossed <other\_tag>. All other kinds of discourse markers have been glossed <other\_dp>.

(13) # *Well, there wadn't the population,*  
 #neg other\_dp other other:predex ln\_det np:s  
*was there?*  
 other\_tag other

*english\_kent01\_232*

**List of corpus-specific GRAID symbols**

<:s_ds>	subject of a verb of speech introducing direct speech, intransitive
<:a_ds>	subject of a verb of speech introducing direct speech, transitive
<:p2>	non-theme NP argument of a ditransitive clause, i.e. recipients, beneficiaries, addressees
<:pred_>	combines with function gloss, chiefly <:obl>, <:g>, <:l>, <:other>, to indicate double incidence of predicate and said function
<:lvc>	non-verbal complement of a light verb
<dem_pro>	demonstrative pronoun
<indef_pro>	indefinite pronoun
<imp_pro>	impersonal pronoun
<refl_pro>	reflexive pronoun
<recip_pro>	reciprocal pronoun
<inter_pro>	interrogative pronoun
<rel_pro>	relative pronoun
<pn_np>	proper name
<_deti>	indefinite determiner, constituent of a NP
<_det>	definite determiner, constituent of a NP
<_dem>	demonstrative determiner, constituent of a NP
<_to>	particle <i>to</i> , constituent of a VP
<_of>	possessive <i>of</i> , constituent of a NP
<aux_neg>	oblique contraction of auxiliary and negator, e.g. <i>don't</i>
<cop_neg>	oblique contraction of copula and negator, e.g. <i>ain't</i>
<other_tag>	verbal element of a tag question
<other_dp>	discourse particle
[INT]	precedes interviewer speech

**References**

Haig, Geoffrey & Schnell, Stefan. 2014. *Annotations using GRAID (Grammatical Relations and Animacy in Discourse): Introduction and guidelines for annotators (version 7.0)*. (<https://lac.uni-koeln.de/en/multicast/>) (accessed 2015-12-30.)