Anti-(Subject-)Control and the Selectional Preferences of Clause-embedding Verbs

Background: This paper is concerned with the correlation between syntactic selectional preferences of clause-embedding verbs and their control behavior in a cross-linguistic perspective. It is well-known that, while (non-)overt pronominal subjects of finite complement clauses can easily refer to referents outside of the superordinate matrix clause, covert pronominal subjects of infinitival complements are usually dependent in their interpretation on arguments in the superordinate clause.

In German, finite complements are usually CPs whereas infinitival clauses can be structurally smaller (depending on their coherence/restructuring properties of the embedding predicate). Based on these premises, Wöllstein (2015) and Brandt et al. (2016) present corpus-based and experimental evidence showing that a certain class of clause-embedding verbs (e.g., *absegnen* 'to approve', *gutheißen* 'to endorse', or *anordnen* 'order' in (1)) systematically resist co-reference of a covert infinitival subject argument with the matrix subject (Anti-(Subject)-Control, henceforth AC).

(1) Otto$_i$ ordnete an, [$_{CP}$ PRO$_{i*/j}$ die Katze zu streicheln]. (Brandt et al. 2016, 81)
Otto ordered the cat to stroke
'Otto ordered to stroke the cat.'

The crucial insight provided in these works for German is that the above mentioned AC predicates although they can embed infinitives, i) show a quantifiable preference for finite over infinitival complements and ii) if they select for an infinitival complement, the structure is necessarily incoherent and the embedded infinitival clause is analyzed as a CP. This correlation between AC and CP selection is not specific to German: We provide data that shows that the selectional preferences for finite complements and the CP status of these complements also holds for other languages (e.g. Polish, Swedish, Dutch, Italian). As discussed in Brandt et al. (2016, 83 ff.), Polish is interesting in that it has a complementizer (*żeby*) that introduces complement clauses to explicitly encode AC, see (2). As shown in (2-a), when *żeby* is present, the embedded subject must not be co-referent with the matrix subject. Moreover, if subject control is excluded by the semantics of the matrix verb (i.e., in the AC case), *żeby* is obligatory (2-b).

(2) a. Piotr$_i$ obiecał **(*żeby)** PRO$_i$ pogłaskać kota.
Piotr promise.PST.PFV ŻEBY stroke.INFV cat
'Piotr promised to stroke the cat.'

b. Piotr$_i$ zarządził, ***(żeby)** PRO$_{*i}$ pogłaskać kota.
Piotr order.PST.PFV ŻEBY stroke.INFV cat
'Piotr ordered to stroke the cat.'

Some Polish verbs can select non-finite complement clauses both with and without *żeby*. This class of verbs comprises volitional verbs, such as *chcieć* 'want' or *woleć* 'prefer', as well as a number of other verbs, such as *pragnąć* 'desire', *zdecydować* 'decide', *lubić* 'like' etc. (cf. Bondaruk 2004, Witkoś 2008, Citko 2012). The facultative use of *żeby* is illustrated in (3), where again a correlation between the presence of *żeby* and the (lack of) control can be observed.

(3) a. Piotr$_i$ chce PRO$_{i/*j}$ wcześnie wstać.
Piotr want.PRS.IPFV early get-up.INFV
'Piotr wants to get up early.'

b. Piotr$_i$ chce, **żeby** PRO$_{*i/j}$ wcześnie wstać.
Piotr want.PRS.IPFV ŻEBY early get-up.INFV
'Piotr wants for others to get up early.'

In addition to existing accounts that analyze *żeby* as a syntactic complementizer (cf. Bondaruk 2004, among many others), we provide grammaticality judgment data showing that infiniti-

val complements introduced by *żeby* always project a full CP. In particular, we show that *żeby* prevents the following operations, all of which are independently known to be blocked by an intervening clause boundary (we omit the data here for reasons of space): **1.** Clitic climbing, **2.** Reflexivization with the reflexive passive morpheme *się*, **3.** Long scrambling, **4.** Licensing of NPIs and the *genitive of negation* by matrix negation, **5.** Anaphoric binding into the embedded constituent. In sum, Polish encodes semantic independence of embedded subjects in a designated element in the syntactic C-projection, thus providing independent evidence for the clausal status of infinitival complements involving AC.

Swedish also has a complementizer (*att*) that introduces both infinitival (4-a) and finite complements (4-b), albeit without the additional AC requirement we observe for *żeby*. In addition, verbs can select for bare infinitives (as in (4-b), marked with italics).

(4) a. Och glöm inte **att** kompostera kaffesumpen.
And forget.IMP not ATT.INF compost.INFV coffee.grounds
'And don't forget to compost the coffee grounds.'

b. Och det gör inget **att** du glömde *skicka julkort*.
And that do.PRS nothing ATT.FIN you forget.PST send.INFV christmas card
'And it doesn't matter that you forgot to send a Christmas card.'

Adopting from the literature (e.g., Platzack 1986, Kalm 2016) the assumption that *att* occupies the syntactic C-head and thus by its presence identifies finite and infinitival complements as CPs, we conducted a study in the 'Språkbanken' corpora to investigate the selectional preferences of embedding verbs with respect to finite complements with *att*, infinitival complements with *att*, and bare infinitives. As a central result, Swedish equivalents of German AC verbs strike out in that i) they show a remarkably low frequency of selected propositional complements when compared to other verb classes, and ii) their selected propositional complements are predominantly finite. Moreover, transitive verbs that semantically require subject control (e.g., *försöka* ('try'), *våga* ('dare')) overwhelmingly select for bare infinitives. Hence, under the assumption that bare infinitives do not project a CP, the Swedish corpus data support the idea of associating AC with CP selection and semantic dependence with 'small' structures.

We can also identify a group of AC predicates in Dutch based on our corpus study in the LASSY Groot corpus (see van Noord et al 2006, 2013). In parallel to the study on Swedish, we searched for clausal complements of AC verbs as well as other verb classes defined in the literature (specifically Broekhuis & Corver 2015) differentiating bare infinitives, *te*-infinitives, *(om)-te*-infinitives and *dat*-clauses. The group of AC verbs i) very rarely occur with clausal complements, and ii) they show a clear preference for finite complements over infinitival ones, which hardly occur (even though they do allow them in principle). In contrast to these AC verbs the subject control verbs that inherently require control (e.g. *proberen* 'to try', *verzuimen* 'to fail') preferably occur with infinitival complements (both with and without the optional complementizer *om*). Thus, the Dutch data also support the generalization that AC verbs can be identified in Dutch based on their selectional preferences.

Conclusion: Corpus evidence shows that in sharp contrast to regular control patterns, the complements of AC verbs are overwhelmingly realized as finite structures. This is in line with the prediction in Rapp et al. (in press) (see also Wöllstein 2015) that infinitival forms are preferably used when the controller is uniquely determined in the matrix clause. This unique identification is present with verbs that semantically require subject control, but not with AC verbs. Additionally, we observe that AC occurs with infinitivals with CP status, which can either be detected by its incoherent/non-restructuring behaviour (German, Polish) or the presence of a complementizer in these infinitives (Polish, Swedish).

References:

Bondaruk, A. (2004): *PRO and Control in English, Irish, and Polish. A Minimalist Analysis*. Lublin: Wydawnictwo KUL.

Brandt, P., B. Trawiński, and A. Wöllstein (2016): (Anti-)Control in German: Evidence from comparative, corpus- and psycholinguistic studies. In I. Reich and A. Speyer (Eds.), *Co-and Subordination in German and other Languages*, *Linguistische Berichte Sonderheft* 21, pp. 77–98.

Broekhuis, H. and N. Corver (2015): Syntax of Dutch: Verbs and Verb Phrases: Volume 2 Comprehensive Grammar Resources. Amsterdam: Amsterdam University Press.

Citko, B. (2012): *Obviation and Control: A view from Polish*. Handout of the talk at *SinFonija* 5, University of Vienna, September 17, 2012.

Kalm, M. (2016): Satsekvivalenta infinitivfraser i svenskan: En synkron och diakron undersökning (Doctoral dissertation, Institutionen för nordiska språk, Uppsala universitet).

van Noord, G., I. Schuurman & V. Vandeghinste (2006): Syntactic annotation of large corpora in STEVIN. In: Proceedings of LREC'06, Genoa, Italy, 1811– 1814.

van Noord, G., G. Bouma, F. Van Eynde, D. de Kok, J. van der Linde, I. Schuurman, E. Tjong Kim Sang, & V. Vandeghinste (2013) Large Scale Syntactic Annotation of Written Dutch: Lassy. In: P. Spyns & J. Odijk (eds.) *Essential Speech and Language Technology for Dutch: resources, tools and applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 147–163.

Platzack, C. (1986): COMP, INFL, and Germanic word order. In Topics in Scandinavian syntax, Studies in Natural Language and Linguistic Theory, pp. 185–234. Dordrecht: Reidel.

Rapp, I., K. Laptieva, A. Koplenig & S. Engelberg (in press): Lexikalisch-semantische Passung und argumentstrukturelle Trägheit - eine korpusbasierte Analyse zur Alternation zwischen dass-Sätzen und zu-Infinitiven in Objektfunktion. Deutsche Sprache 17.3

Witkoś, J. (2008): CP-infinitives in Polish and two theories of control: A reply to the analysis by Anna Bondaruk. *Research in Language* 6, pp. 217–259.

Wöllstein, A. (2015): Grammatik-explorativ. In Eichinger, L. (ed.), Sprachwissenschaft im Fokus, Jahrbuch des Instituts für Deutsche Sprache, 93–120. Berlin and New York: Mouton de Gruyter.