

Mismatches at the syntax-semantics interface

Markus Egg, Universität des Saarlandes
egg@coli.uni-sb.de

Abstract: Recent analyses of mismatches at the syntax-semantics interface investigate e.g. modification of agentive nouns (Larson 1998), quantifying pronouns (Abney 1987), or recursive modification (Kasper to appear). I propose a unified analysis for them that preserves Kasper’s intuitions but is more general because the mismatches are handled locally in the CONT feature. Its pivot is an elaborate syntax-semantics interface that is based on a surface-oriented syntactic analysis. It generalises easily to further syntax-semantics mismatches and mismatches at the morphology-semantics interface described in Müller (2003).

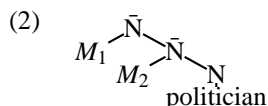
1 The phenomenon

Semantic scope of constituents often depends on their syntactic constellation.¹ In this respect, the syntax-semantics interface (SSI) is *iconic*: Configurational asymmetries of syntactic tree structures are mapped onto semantic asymmetries. The crucial notion here is (unilateral) *c-command*: If a constituent C_1 c-commands a constituent C_2 (but not vice versa), C_1 has wide scope over C_2 .²

Evidence for this iconicity can be found e.g. in cases of multiple modification by scope-bearing modifiers. Here the syntactic order of the modifiers determines their scope. Consequently, switching the order of modifiers around in such cases of multiple modification has an impact on the meaning of these cases. Consider e.g. (1a) and (1b), which differ in the order of the modifiers:

- (1) (a) a former apparent politician (b) an apparent former politician

Their meanings are different, ‘person who used to resemble a politician’ for (1a) and ‘person who resembles someone who used to be a politician’, respectively. This is due to the fact that the preceding modifier M_1 c-commands the following modifier M_2 , but not vice versa. The relevant part of the syntactic structure in (1a) and (1b) can be rendered schematically by (2):



However, in many modification structures there is no such iconicity, because the syntactic asymmetry does not directly map onto a semantic one. In these cases, the modifier has scope (optionally or obligatorily) only over a *part* of the expression it modifies.

As a first example, consider (3a). In its preferred reading, ‘person characterised by beautiful dancing’, the adjective pertains to the *verb stem* only and is in the scope of the affix *-er*. This emerges directly from applying the affix meaning ‘person characterised by X ’ (where X refers to the scope domain of the affix) to the meaning of the stem (after modification by the adjective).

- (3) (a) beautiful dancer (b) everyone in this room

Examples like (3b), where a quantifying pronoun is modified, are equally anti-iconic, because the modifier *in this room* pertains to the *restriction* of the quantification as introduced in the pronoun semantics. I.e., while the semantics of *everyone* is ‘set of properties such that every person has them’, the meaning of (3b) is ‘set of properties such that every person *in this room* has them’. However, (3a) and (3b) differ in that (3b) has only the anti-iconic reading while (3a) is ambiguous because it also has the iconic reading ‘beautiful person characterised by dancing’.

Some conclude from such syntax-semantics mismatches that semantic structure reflects (and is iconic to) a not directly visible layer of syntactic structure like *Logical Form* (Larson 1998; Abney 1987; see also Sag 1997 and Kathol 1999 for further discussion of Abney’s analysis of (3b)). However, the analysis proposed in this paper assumes only a surface-oriented syntactic structure.

Kasper’s (to appear) analysis of modification of modifiers regards them as yet another instance of the syntax-semantics anti-iconicity exhibited by (3). The scope of the modifier of a modifier M may only extend over M but not over the expression modified by M . E.g., the intensionalisation expressed in *seemingly* in (4) relates only to the adjective but not to the noun modified by the full AP. Thus, (4b) refers to people whose being a politician is undisputed, but whose honesty is not:

¹Scope relations of nominal quantifiers among themselves are a well-known exception here.

²C-command relates nodes in a syntax tree. A node A c-commands a node B iff (a) A and B are dominated by the same branching nodes in the tree, (b) A does not dominate B or vice versa, and (c) $A \neq B$.

- (4) (a) seemingly honest (b) seemingly honest politician

Following Kasper (to appear), the modifier of the adjective does not pertain to the whole ('attributive') semantics of the adjective (5a), but only to its 'predicative' part (the property **honest'**). This returns the desired semantic representation (5b) for (4a), a functor intersecting properties P with the property of seeming honest. Note that in this representation the λ -abstracted property P is outside the scope of **seem'**. The meaning of *seemingly* is (5c); here **seem'**(p) is true in a world w iff p is true in all possible worlds where things are as they seem in w . (5c) maps properties P on the property of seemingly having P .

- (5) (a) $\lambda P \lambda x. \mathbf{honest}'(x) \wedge P(x)$
 (b) $\lambda P \lambda x. \mathbf{seem}'(\wedge \mathbf{honest}'(x)) \wedge P(x)$
 (c) $\lambda P \lambda x. \mathbf{seem}'(\wedge P(x))$

I will discuss Kasper's analysis in section 2 before presenting my own analysis in section 3.

2 Kasper's analysis

Kasper divides the attributive meaning of a modifier into its predicative meaning ('inherent content', IC) and the rest ('combinatorial semantics', CS). Modifiers lexically determine the semantics S of the head-adjunct phrase in which they are the head of the adjunct: Their CS specifies the way in which S is composed from the semantic contributions of head and adjunct (e.g., for *honest*, as intersection). However, their own semantic contribution (their IC) cannot fully determine the semantics of the adjunct, since the adjunct as a whole might be a head-adjunct phrase itself, as in (4b).

A modifier M of a modifier M should now affect only the IC of M . This happens in the usual fashion in that the semantics of M' is also the semantics of this local head-adjunct structure. E.g., for *seemingly honest*, the semantics is the one of *seemingly*. In contrast, the CS of M must percolate to the phrase headed by M . Thus, the CS of *seemingly honest* is the one of *honest*.

To this end, Kasper assumes two MOD features ECONT and ICONT for the combinatorial semantics of a modifier M and the semantics of the phrase headed by M , respectively. The ECONT value is specified lexically, but the ICONT value is not. In particular, it is not equated with M 's semantic contribution as specified in its own CONT value. Being head features, ECONT and ICONT percolate from M to the phrase headed by M . This percolation is not affected by modification of M itself, which may only replace the CONT value of M .

The semantics principle then states that the CONT value of a head-adjunct phrase is identical to the MOD|ECONT value of the adjunct. In addition, the MOD|ICONT value of the adjunct is equated with its CONT value. This means that once a modifier has been projected to a full phrase (a precondition for its function as an adjunct), its current CONT value is identical to the semantics of the whole phrase (the ICONT value), because the phrase cannot be extended any further.

However, this analysis predicts that if a modifier may pertain semantically to only part of the expression it modifies syntactically, it must do so. But this is in conflict with cases like (3a), i.e., the analysis cannot be generalised to capture the common ground between (3a) and (4). In section 4, I will propose an analysis of the mismatch that is more flexible than Kasper's yet preserves his insights. The mismatch is handled locally within the CONT feature of linguistic signs.

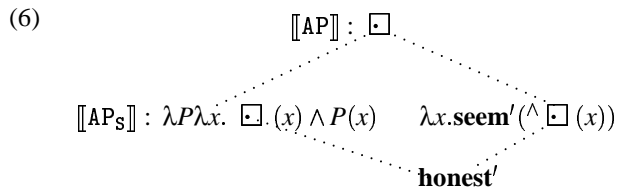
3 The proposed analysis

The pivot of my analysis is the syntax-semantics interface. It models anti-iconic structures as *potential scope ambiguities*. The basic assumption is that the semantic contribution of a constituent C breaks down into a *secondary part* (which ends up in the scope of all constituents that unilaterally c-command C) and a *main part*, whose scope is determined differently. Thus, when C is modified, the modifier outscopes C 's secondary part semantically, but the scope between the modifier and C 's main part is deliberately left open. E.g., for *honest* its inherent content constitutes the secondary, and its combinatorial semantics, the primary part of its semantic contribution. Consequently, in (4a) the adverbial outscopes the IC of *honest*, but the scope of its CS and the adverbial is open. Wide scope of the former is possible, which yields the desired interpretation (5b) for (4a).

The semantic description of these phenomena calls for a suitable *underspecification formalism*, e.g., UDRT (Reyle 1993), MRS (Copestake et al. 1999), or Constraint Language for Lambda Structures (Egg

et al. 2001) (used in an abbreviated form here). Expressions of such a formalism are *constraints* that describe a set of semantic representations (here, λ -terms), one for each reading of a structurally ambiguous expression. Representations described by (or compatible with) a constraint are its *solutions*. (Here we only need *constructive* solutions consisting of the material explicitly mentioned in the constraint.) Constraints are underspecified in that they deliberately abstract away from the differences between their solutions (in particular, w.r.t. scope relations between the fragments). These formalisms allow an adequate representation of structural ambiguity and, what is more, they provide the necessary flexibility in the SSI.

I will now outline the proposed solution with the semantic representation and construction for (4a). The constraint for its meaning is (6). ‘[[C]]’ indicates the main fragment of a constituent C, ‘[[CS]]’, the secondary fragment of C. ‘[[C]]:F’ means ‘main fragment of C is defined as fragment F’:

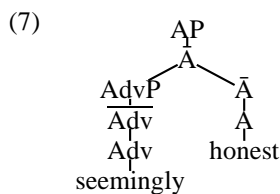


(6) comprises three ingredients, fragments of λ -terms, not yet known parts of these fragments, indicated by ‘holes’ (\square), and *dominance relations* (depicted by dotted lines) that relate fragments to holes. When a fragment is dominated by a hole it is an (im-)proper part of whatever the hole stands for. Dominance relations model scope. Structures like (6) are called *dominance diamonds*.

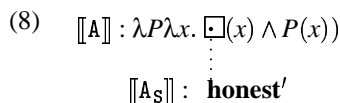
To paraphrase (6), we do not know what the structure as a whole stands for (thus, there is only a hole on top) but both the semantic contribution of modifier (right fragment) and of adjective (combinatorial semantics; left fragment) are its immediate parts. In addition, the adjective’s inherent content (bottom fragment) has narrowest scope, as it is dominated by the other two fragments. These structures are typical for quantifier scope ambiguities, too. This underspecified representation is a simplified form of CLLS.

Resolving the ambiguity in constraints is modelled as adding information monotonically, in particular, by strengthening dominance relations between holes and fragments to *identity*. For (6), there are in principle two choices: Identifying the CS fragment with the top hole, the modifier fragment, with the hole in the CS fragment, and the IC fragment, with the hole in the CS fragment yields (5b). The other choice (starting this procedure with the modifier fragment) is blocked due to the types of the involved fragments: The hole in the modifier fragment cannot be identified with the CS fragment. I.e., there is no danger of unwanted overgeneration ((3b) is handled analogously), while for ambiguous cases like (3a) both choices would return a solution of the constraint.

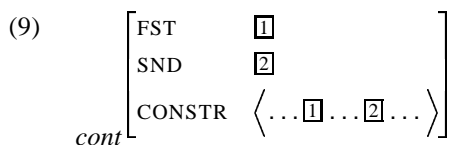
The interface derives the constraint (6) from the syntactic structure for (4a), which is (7):



Deriving constraints like (6) uses lexical entries as the one of *honest*. Here the inherent content of the adjective, which modifiers might pertain to exclusively, is set off in a fragment $\llbracket \text{A}_S \rrbracket$ of its own. The combinatorial semantics of the adjective constitutes the $\llbracket \text{A} \rrbracket$ fragment:



This kind of semantic information is encoded in the *CONT* feature of linguistic signs. Its value, a feature structure of type *cont*, has a list-valued feature *CONSTR* for the constraint itself. Two auxiliary features *FST* and *SND* identify primary and secondary fragment of a constituent among the fragments appearing in *CONSTR* (fragments can be modelled by feature structures, too):



Interface rules specify for a constituent C how the constraints Con_1 and Con_2 of its immediate constituents C_1 and C_2 (these constraints are inherited by C) are combined into a new constraint Con for C . Rules are implemented as phrases that themselves contribute to Con . They combine Con_1 and Con_2 via the FST and SND values of C_1 and C_2 and determine these features for C . This technique is familiar e.g. from semantic construction in MRS (Copestake et al. 1999, p.17).

The modification interface rule is (10): The emerging constituent \bar{x}_1 inherits its main fragment $[[\bar{x}_1]]$ from the modified expression. Its secondary fragment $[[\bar{x}_{1s}]]$ is defined as the modifier fragment $[[Mod]]$ applied to a hole that dominates the secondary fragment $[[\bar{x}_{2s}]]$ of the modified expression. This makes $[[Mod]]$ and $[[\bar{x}_1]]$ scopally ambiguous and yields the bottom half of a dominance diamond. (Recall that $[[\bar{x}_2]]$ dominates $[[\bar{x}_{2s}]]$ [they are fragments of the same constituent] and is equal to $[[\bar{x}_1]]$.) Equating the modifier fragments ($[[Mod]]: [[Mod_s]]$) is not necessary, but facilitates reading.

$$(10) \quad [\bar{x}_1 \text{ Mod } \bar{x}_2] \xRightarrow{\text{(SSS)}} \begin{array}{c} [[\bar{x}_{1s}]] : [[Mod]](\square) \quad [[Mod]] : [[Mod_s]] \quad [[\bar{x}_1]] : [[\bar{x}_2]] \\ \vdots \\ [[\bar{x}_{2s}]] \end{array}$$

The rule that constructs the upper half of the dominance diamond corresponds to the syntax rule $XP \rightarrow \bar{x}$. The main fragment of XP is a hole that dominates both fragments of the \bar{x} constituent:

$$(11) \quad [XP \bar{x}] \xRightarrow{\text{(SSS)}} \begin{array}{c} [[XP]] : \square \\ \vdots \\ [[XP_s]] : [[\bar{x}]] \quad [[\bar{x}_s]] \end{array}$$

Finally, nonbranching \bar{x} constituents inherit their fragments from their heads:

$$(12) \quad [\bar{x} X] \xRightarrow{\text{(SSS)}} [[\bar{x}]] : [X] \quad [[\bar{x}_s]] : [X_s]$$

Semantic construction for *seemingly honest* now uses the lexical entries for *seemingly* (8) and *honest* (13) and rules (10)-(12) to derive the diamond in (6) on the basis of (7).

In the lexical entry for *seemingly*, both fragments are identical; according to (12), this carries over to *seemingly* as \bar{Adv} constituent. Following (11), the constraint for the \bar{AdvP} *seemingly* is (14):

$$(13) \quad [[Adv_s]], [[Adv]] : \lambda P \lambda x. \mathbf{seem}'(\wedge P(x))$$

$$(14) \quad \begin{array}{c} [[AdvP]] : \square \\ \vdots \\ [[AdvP_s]] : \lambda P \lambda x. \mathbf{seem}'(\wedge P(x)) \end{array}$$

Next, (10) combines (8) and (14) into (15), the bottom half of a diamond for the meaning of the \bar{A} constituent *seemingly honest*, before (11) transforms (15) into the full diamond (6).

$$(15) \quad \begin{array}{c} [[\bar{A}]] : \lambda P \lambda x. \square(x) \wedge P(x) \quad [[\bar{A}_s]] : \lambda x. \mathbf{seem}'(\wedge \square(x)) \\ \vdots \\ \mathbf{honest}' \end{array}$$

4 Conclusion and outlook

Modification structures like (3)-(4), where the modifier may have scope only over part of the expression that it modifies syntactically were described as *potential scope ambiguities*.

Many problematic cases of semantic construction can be put down to the same phenomenon, e.g., the repetitive/restitutive ambiguity of *again*-sentences as (16). Its readings are ‘Max manipulates the window, and, as a result, the window is open, and he has done this before’ (repetitive) and ‘Max manipulates the window, and, as a result, the window is open, and it has been open before’ (restitutive). The restitutive reading follows immediately if the modifier *again* pertains only to that part of the meaning of *open* that introduces the result state (the window being open):

$$(16) \quad \text{Max opens the window again}$$

Further cases that suggest themselves for an application of the proposed analysis are the modification of pronouns and proper names like *old Bill* and *poor me*.

This analysis can then be extended to capture the morphosemantic mismatches that Müller (2003) discusses under the heading of ‘bracketing paradoxes’. Consider e.g. separable prefix verbs like *losrennen* ‘start running’. In nominalisations of these verbs by the *Ge...e* circumfix only the verb stem shows up within the circumfix (thus, for *losrennen* the nominalisation is *Losgerenne*). This suggests a morphological structure in which the verb stem combines with the circumfix *before* the prefix is attached. But semantically, the prefix is in the scope of the circumfix. Müller’s solution (which analyses prefixes like *los-* as subcategorised modifiers) can eventually be put down to the same intuitions that motivated Kasper’s and my analysis.

To extend my analysis to such mismatches, I assume a close counterpart to the SSI rule (10) in the morphology-semantics interface: Roughly, prefixation builds semantic structures like (8) in that the main fragment of the affix (also the main fragment of the resulting word) dominates the secondary fragment of its base, while the primary fragment of the base becomes the resulting word’s secondary fragment. In addition, both fragments are identical for roots.

For *Gerenne*, the main fragment comprises a nominalisation operator NOM of the circumfix, the secondary one, the verb stem meaning. Affixation of *Gerenne* by *los-* introduces an additional fragment (with the change-of-state operator BECOME of *los-*) that dominates the verb stem semantics but not the NOM operator. I.e., narrow scope of BECOME w.r.t. NOM is possible, and, indeed, enforced by the types of the fragments involved in the constraint for *Losgerenne*.

Anti-iconicity at the syntax-semantics interface is also very common for affixed words in agglutinating languages like Turkish, e.g. in (17):

- (17) *genç at -li*
young horse equipped with
‘young person with a horse *or* person with a young horse’

In its second reading, the modifier *genç* pertains only to the nominal stem *at*. This case can be handled if the morphosemantics of Turkish affixation returns a semantic structure like (8) for *atl*.

References

- Abney, S. (1987). *The English noun phrase in its sentential aspect*. Ph. D. thesis, MIT.
- Copetake, A., D. Flickinger, and I. Sag (1999). Minimal Recursion Semantics. An introduction. Available from <http://www.cl.cam.ac.uk/~aac10/papers/newmrs.pdf>.
- Egg, M., A. Koller, and J. Niehren (2001). The constraint language for lambda-structures. *Journal of Logic, Language, and Information* 10, 457–485.
- Kasper, R. (to appear). The semantics of recursive modification. *Journal of Linguistics*.
- Kathol, A. (1999). Restrictive modification and polycategorical NPs in English. In G. Bouma et al. (eds), *Constraints and resources in natural language syntax and semantics*, 83–100. Stanford: CSLI Publications.
- Larson, R. (1998). Events and modification in nominals. In D. Strolovitch and A. Lawson (eds), *Proceedings from SALT VIII*, Ithaca. Cornell University.
- Müller, S. (2003). The morphology of German particle verbs: solving the bracketing paradox. *Journal of Linguistics* 39, 275–325.
- Reyle, U. (1993). Dealing with ambiguities by underspecification: construction, representation, and deduction. *Journal of Semantics* 10, 123–179.
- Sag, I. (1997). English relative clause constructions. *Journal of Linguistics* 33, 431–484.