Non-configurational Case Assignment in HPSG

Adam Przepiórkowski

Seminar für Sprachwissenschaft
Universität Tübingen, Germany

Email: adamp@sfs.nphil.uni-tuebingen.de

20th May 1996

Despite the initial claim that "there is no separate theory of case (or Case)" and that case should be "simply treated as part of subcategorization" (Pollard and Sag, 1994, ch. 1, p. 30) it soon became clear that a separate Case Principle is needed not only for Icelandic (Pollard and Sag, 1994, p. 30, fn. 25), but also for German (Heinz and Matiasek, 1994)[1], Polish (Przepiórkowski, 1996) and, indeed, English (Grover, 1995, ch. 2). In the first part of this paper we reiterate these arguments for a separate case module within HPSG.

The second, crucial part of this paper concerns ways of formalizing such a Case Principle. Pollard (1994) notes that this principle would have to either be configurational (i.e., refer to the DTRS attribute) or apply defaults, neither of which seems to be desirable. Indeed, the analyses of German, Polish and English cited above are configurational, while the one presented in Sag *et al.* (1992) conceptually employs defaults. Another problem faced by the analyses of Heinz and Matiasek (1994) and Przepiórkowski (1996) is making them compatible with the traceless dialect of HPSG, cf. e.g. Sag (1995) and Sag and Fodor (1994), without introducing considerable redundancy in the formulation of the Case Principle.[2] In this paper we present a non-configurational, traceless and non-redundant formalization of HPSG case theory, specifically, we propose that the locus of the Case Principle be ARG-S (i.e., SUBCAT). Our analysis is based on the observation that the Case Principle, which assigns morphological case to structural NP arguments, can be stated as a constraint on ARG-S as long as for each such NP it is known whether it has been 'cancelled' (syntactically realized) from the corresponding VALENCE feature. Crucially, the case marking does not depend on the way a given NP has been realized, be it via the Subcat Principle, or via an extraction lexical rule (cf. (5) below).

We formalize these observations by introducing a type *arg*, appropriate for features ARG with values of type *synsem*, and REALIZED with values of type BIN (i.e., + or −). Moreover, we assume that the VALENCE features take values of type listof(*arg*), and that ARG-S is canonically the concatenation of the VALENCE values.

(1) *arg*: $\begin{bmatrix} \text{ARG } \textit{synsem} \\ \text{REALIZED } \textit{bin} \end{bmatrix}$

(2) lexical *cat*: $\begin{bmatrix} \text{VALENCE} & \begin{bmatrix} \text{SUBJ } \boxed{1}\text{listof}(\textit{arg}) \\ \text{SPR } \boxed{2}\text{listof}(\textit{arg}) \\ \text{COMPS } \boxed{3}\text{listof}(\textit{arg}) \end{bmatrix} \\ \text{ARG-S} & \boxed{1} \oplus \boxed{2} \oplus \boxed{3} \end{bmatrix}$

We assume that the basic descriptions of lexical entries do not specify the REALIZED value of their complements, although the CELR and SELR specify the arguments they remove from

[1] Also Gerdemann (1994) and Pollard (1994) give arguments for a Case Principle for German.
[2] Such a partly redundant extension is provided for English in (Grover, 1995, ch. 2).

VALENCE as REALIZED+. Also the Subcat Principle specifies as REALIZED+ these VALENCE arguments, which are realized through the schemata. On the other hand, any lexical item subcategorizing for an unsaturated phrase marks the unrealized arguments of this phrase (i.e., those present in its VALENCE) as REALIZED−. In all three cases these specifications "percolate" to ARG-S via structure sharing between ARG-S and VALENCE, cf. (2).

In the remainder of the paper we show how the case phenomena from Polish, German and English described in the initial part of this paper can be accounted for by the mechanism presented above. For example, the clause of the Case Principle for Polish responsible for structural complements of verbs (cf. Przepiórkowski (1996)) can be formalized[3] as in (3–4) below.[4]

(3)
$$\left[\begin{array}{l}\textsc{neg}\ -\\ \textsc{arg-s}\ \boxed{1}_{ne\_list} \oplus \langle \left[\begin{array}{l}\textsc{arg}\ \mathrm{NP}[str]\\ \textsc{realized}\ +\end{array}\right] \rangle \oplus \boxed{2}_{list}\end{array}\right] \rightarrow \left[\ \textsc{arg-s}\ \boxed{1} \oplus \langle \left[\ \textsc{arg}\ \mathrm{NP}[sacc]\right] \rangle \oplus \boxed{2}\right]$$

(4)
$$\left[\begin{array}{l}\textsc{neg}\ +\\ \textsc{arg-s}\ \boxed{1}_{ne\_list} \oplus \langle \left[\begin{array}{l}\textsc{arg}\ \mathrm{NP}[str]\\ \textsc{realized}\ +\end{array}\right] \rangle \oplus \boxed{2}_{list}\end{array}\right] \rightarrow \left[\ \textsc{arg-s}\ \boxed{1} \oplus \langle \left[\ \textsc{arg}\ \mathrm{NP}[sgen]\right] \rangle \oplus \boxed{2}\right]$$

With these constraints, the only difference between the case assignment of (5b) and (5c) below is a different way of checking REALIZED+ feature on the NP[*str*] complement: in (5b) it is through the Subcat Principle, while in (5c) it is via the CELR.

(5) a. Janek lubi Marię.
John$_{nom}$ likes Mary$_{acc}$.

b. Janek nie lubi Marii.
John$_{nom}$ NEG likes Mary$_{gen}$.

c. Marii Janek nie lubi.
Mary$_{gen}$ John$_{nom}$ NEG likes.

The REALIZED feature is used similarly to account for the case of subject-to-object raised complements in English.

(6) Mary expects him to be promoted.

In examples such as (6) the postulates found in the HPSG literature that the case of the subject of *verbum infinitum* is underspecified or accusative fall out from the theory presented above. The matrix verb which takes an unsaturated complement marks the unrealized VALENCE (and, through structure sharing, ARG-S) elements of this complement as REALIZED− making them immune to the Case Principle and in this sense underspecified; this is so with the subject of the copula in (6). On the other hand, the matrix verb also takes the NP[*str*] complement, which is marked as REALIZED+ by the Subcat Principle. Now the Case Principle comes into play and assigns accusative case to this NP, and—via structure sharing of *synsems*—also to the subject of the copula; in this sense this subject is accusative. Note that on this account the case values of complements of verbs never depend on verbal form; e.g., subjects of English verbs are always marked in the lexicon as structural. It is only the Case Principle that is responsible for the accusative case of the raised subject in (6) above and the nominative in (7) below.[5]

(7) It was decided that he be promoted.

In summary, we present a formalization of the Case Principle as used, e.g., by Heinz and Matiasek (1994), Przepiórkowski (1996) and Grover (1995), which avoids the drawbacks of those analyses, i.e., which is non-configurational, traceless and non-redundant.

[3] The type *str* (structural) is a supertype of *snom*, *sacc* and *sgen* in Polish; see Przepiórkowski (1996) for details.
[4] Note that (4) takes care of the genitive of negation here.
[5] See (Grover, 1995, ch. 2).

References

Gerdemann, D. (1994). Complement inheritance as subcategorization inheritance. In Nerbonne *et al.* (1994), pages 341–363.

Grover, C. (1995). *Rethinking Some Empty Categories: Missing Objects and Parasitic Gaps in HPSG.* Phd dissertation, University of Essex. Available from ftp://ftp.cogsci.ed.ac.uk/pub/grover/thesis.ps.gz.

Heinz, W. and Matiasek, J. (1994). Argument structure and case assignment in German. In Nerbonne *et al.* (1994), pages 199–236.

Nerbonne, J., Netter, K., and Pollard, C., editors (1994). *German in Head-Driven Phrase Structure Grammar*, Stanford, CA. Center for the Study of Language and Information.

Pollard, C. (1994). Toward a unified account of passive in German. In Nerbonne *et al.* (1994), pages 273–296.

Pollard, C. and Sag, I. A. (1994). *Head-driven Phrase Structure Grammar.* Chicago University Press.

Przepiórkowski, A. (1996). Case assignment in Polish: Towards an HPSG analysis. In C. Grover and E. Vallduví, editors, *Edinburgh Working Papers in Cognitive Science: Studies in HPSG*, pages 195–233. Centre for Cognitive Science, University of Edinburgh.

Sag, I. A. (1995). Constraint-based extraction (without a trace). Unpublished manuscript.

Sag, I. A. and Fodor, J. D. (1994). Extraction without traces. In R. Aranovich, W. Byrne, S. Preuss, and M. Senturia, editors, *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, volume 13, Stanford University. CSLI Publications/SLA.

Sag, I. A., Karttunen, L., and Goldberg, J. (1992). A lexical analysis of Icelandic case. In I. A. Sag and A. Szabolcsi, editors, *Lexical Matters*, number 24 in Lecture Notes, pages 301–318. Center for the Study of Language and Information, Stanford.