Clausal Coordinative Ellipsis in Hungarian in comparison to Dutch, Estonian and German

Karin Harbusch & István Bátori

Universität Koblenz-Landau, Computer Science Dept.
{harbusch|batori}@uni-koblenz.de

In the following, we prove that the psycholinguistically motivated theoretical framework of clausal coordinate ellipsis generation by Kempen (2009) and its implementation (called ELLEIPO which was originally programmed in JAVA for Dutch and German (Harbusch & Kempen, 2006) and which has been extended for Estonian (Harbusch et al., 2009)) can be easily tailored to Hungarian. Given the fact that the investigated languages belong to two rather different families of languages (Estonian and Hungarian are Finno-Ugric languages whereas Dutch and German belong to the Indo-European languages), not much overlap for the individual elision rules for the two language families would be expected. However, the similarities for Dutch, Estonian and German CCE generation are remarkable. As linguists might object that Estonian is heavily influenced by German, verifying the rules for Hungarian as well underpins the claim that the rules can be generalized for Finno-Ugric languages.

Table 1 delineates the four CLAUSAL COORDINATE ELLIPSIS (CCE) types (GAPPING (including LONG-DISTANCE GAPPING (LDG), SUBGAPPPING and STRIPPING), FORWARD and BACKWARD CONJUNCTION REDUCTION (FCR and BCR) and SUBJECT GAP WITH FINITE/FRONTED VERB (SGF); cf. first column) and illustrates them by Hungarian examples. Table 2 outlines the rule set licensing the individual constructions in Dutch/Estonian/German.

All forms of GAPPING (cf. examples (1) to (4)) are characterized by elision of the posterior member of a pair of LEMMA-IDENTICAL Verbs (N.B. that word order is not essential as in sentence (1); the position of the elided Verb need not be peripheral but is often medial, as in (2) through (4)). Every non-elicided constituent (“REMNANT”) in the posterior conjunct should pair up with a constituent in the anterior conjunct that has the same grammatical function but is not coreferential. Stated differently, the members of such a pair are CONTRASTIVE. In LDG, the remnants originate from different clauses (more precisely: from different clauses that belong to the same SUPERCLAUSE; a superclause is a hierarchy of finite or nonfinite clauses that—with the possible exception of the topmost clause—do not include a Subordinating Conjunction). In FCR, elision affects the posterior token of a pair of LEFT-PERIPHERAL strings consisting of one or more wordform-identical major constituents. BCR is almost the mirror image of FCR as it deletes the anterior member of a pair of RIGHT-PERIPHERAL lemma-identical word strings (‘to-steal’ in (7)); however, BCR may elide PART of a major constituent. SGF elides the Subject of the posterior conjunct in a main clause, when in the anterior conjunct the wordform-identical Subject follows the Finite Verb (Subject-Verb inversion).

In our presentation, we show that Hungarian obeys these rules with few deviations. For instance, example (8) is unacceptable in Dutch; in Hungarian Gapping, superclauses are not relevant (cf. sentence 9—which is ruled out in Dutch, Estonian and German).

(9) A rendőrség [rendőrség] remélzi hogy a tuntetők [tuntetők] hazamennék és [hazamennék]

The police hopes that the demonstrators go home and the police

‘The police hopes that the demonstrators go home and vice versa’

Table 1. CCE examples in Hungarian. Struck-out text represents elisions.

---

1 In the talk, we give a live demo of the ELLEIPO system (see www.uni-koblenz.de/~harbusch/Welcome-to-Elleipo.html).
2 For LEMMA IDENTITY, only the lexical entries (citation form) of the constituents have to be identical. In contrast, WORDFORM IDENTITY requires, in addition, identity of their morphological features. COREFERENTIAL CONSTITUENTS refer to the same discourse entity or entities, irrespective of whether or not they include the same lemma(s).
3 We use the term MAJOR CONSTITUENT of a clause in a broad sense that includes Head Verb (Main, Copula or Auxiliary), Arguments (e.g. Subject, Direct and Indirect Object, and Non-finite Complement Clause), Adjuncts (Adverbial Modifier, including Adverbial Clause), and Subordinating Conjunctions (i.e. the Complementizer in Complement Clauses—that, whether—or the Subordinator in Adverbial Clauses—while, although, when, etc.).
4 N.B. the example works the same when the second Subject is replaced by ‘the organizers’. Thus, the elision does not result from strong semantic expectations overruling detailed syntactic checking.
Table 2. Basic elision conditions for the four clausal coordinate ellipsis (CCE) types in Dutch, Estonian, German and Hungarian.

<table>
<thead>
<tr>
<th>CCE type</th>
<th>Elision conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPPING</td>
<td>Lemma identity of Verb &amp; contrastiveness of remnants</td>
</tr>
<tr>
<td>LONG-DISTANCE GAPPING (LDG)</td>
<td>Gapping conditions in superclause</td>
</tr>
<tr>
<td>SUBGAPPING</td>
<td>Gapping conditions &amp; VP remnant in second conjunct</td>
</tr>
<tr>
<td>STRIPPING</td>
<td>Gapping conditions &amp; Only one non-Verb remnant</td>
</tr>
<tr>
<td>FORWARD CONJUNCTION REDUCTION (FCR)</td>
<td>Form identity &amp; left-peripherality (within clause boundaries) of major clausal constituents</td>
</tr>
<tr>
<td>BACKWARD CONJUNCTION REDUCTION (BCR)</td>
<td>Lemma identity &amp; right-peripherality, possibly disregarding major constituent boundaries</td>
</tr>
<tr>
<td>SUBJECT GAP IN CLAUSES WITH FINITE/FRONTED VERBS (SGF)</td>
<td>Form-identical Subject &amp; first conjunct starting with Verb/Modifier/Adjunct &amp; FCR applied if licensed</td>
</tr>
</tbody>
</table>

References