Greville G. Corbett, Surrey Morphology Group
Canonical morphosyntactic features

The canonical approach has been applied to phenomena in syntax, notably agreement (Evans 2003, Corbett 2003, 2006, Comrie 2003, Polinsky 2003, Seifart 2005, Suthar 2006) and to phenomena in morphology, including suppletion (Corbett 2007) and doublets (Thornton 2008). This research has also addressed morphosyntactic features and values. For these, two overarching principles (covering ten converging criteria) have been proposed; namely, a canonical morphosyntactic feature:

1. has robust formal marking;
2. is constrained by simple rules of syntax.

Various non-canonical behaviours have been identified. For example, lack of robust formal marking underlies non-autonomous case values (Zaliznjak 1973). Rather than being restricted to one feature, like case, parallel behaviour recurs in different features (sometimes with different labels). This line of research proved fruitful, but there was a flaw: fully canonical morphosyntactic features are idealizations which appear indistinguishable one from another, since the structural properties that distinguish the features would all be non-canonical. We now tackle that issue, as part of the problem of definitions specified in the call for papers.

To differentiate canonical morphosyntactic features, we examine their interaction with parts of speech. In the canonical situation, the interaction follows these criteria:

C1: **exclusiveness**:
a lexical item belongs to just one part of speech
a value belongs to just one feature.

C2: **exhaustiveness**:
every lexical item has available all values of all features.

C3: **open and closed classes:**
all classes are closed, except the class of lexical items.

C4: **compositionality:**
given the lexical semantics of a lexical item and a specification of its feature values, the meaning of the whole is fully predictable.

The weakenings of these criteria define a space in which we can locate many problematic phenomena of morphosyntax. Here we concentrate on the key issue of differentiating features according to this typology.

The morphosyntactic feature closest to our idealization is **number**. We find instances where different parts of speech have all the number values, as do all of their members (lexical items). Deviations from canonicity, according to the criteria given, allow us to locate the other features. Take a feature which deviates from canonicity in terms of C2 'Exhaustiveness': the lexical entries in one of the relevant parts of speech each select for just one of the values. This gives a clear asymmetry between the parts of speech, and is a representation of **gender**. We may add a second deviation, namely that the number of lexical items is severely restricted (contrary to C3). If we think of

the limited part of speech as being pronouns, we have a representation for **person**. Finally, the last major morphosyntactic features, **case** is a mirror-image of gender. Gender is a classification of nouns, which determines the form of targets, while case requires a classification of governors, which determines the form of their governees.

In summary:

- we make progress on the canonical approach to morphosyntactic features
- we suggest that parallel non-canonical behaviour is found with different features
- we solve the issue of differentiating the canonical morphosyntactic features.