On Morphological Case and Word-Order Freedom
Author(s): Thomas McFadden

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via http://linguistics.berkeley.edu/bls/.

The Annual Proceedings of the Berkeley Linguistics Society is published online via eLanguage, the Linguistic Society of America's digital publishing platform.
On morphological case and word-order freedom

THOMAS MCFADDEn
University of Pennsylvania

1. The correlation and two ways to explain it
A commonly made cross-linguistic generalization is that languages with extensive case-marking tend also to have greater freedom of word order than languages without. Thus case-marking languages like Latin allow the order of nominal arguments in a sentence to be permuted without changing the basic meaning of the sentence, (discourse-pragmatic factors aside) as in (1) and (2), while swapping the order of arguments in an essentially caseless language like English is not possible without swapping their grammatical roles, as in (3) and (4):

(1) Puella puerum videt.
girl:NOM boy:ACC sees
'The girl sees the boy.'

(2) Puerum puella videt.
boy:ACC girl:NOM sees
'The girl sees the boy.'

(3) The girl sees the boy.

(4) The boy sees the girl.

That a correlation of this sort exists is not really a matter of debate. What is controversial is what form it takes and exactly how we are to explain it.

I will start my discussion by laying out some basic assumptions. First of all, I assume that there is a core of formal linguistic knowledge or grammar which is distinct from the knowledge relating to language use. I further assume that a complete linguistic theory should incorporate tools to deal with both sides of this distinction and provide a principled way to determine which tools should be used to explain a given phenomenon. Thus, in place of the loaded terms competence

---

1 Thanks to Tony Kroch, Dave Embick, Beatrice Santorini and the audience at BLS 29 for their comments and criticisms, and to the following native-speaker informants for their judgments and discussion: Beatrice Santorini, Augustin Speyer, Hans Van de Velde, Hedde Zeijlstra, Sophia Malamud, Anna Nagayeva and Kimiko Nakanishi.
and performance, I will distinguish (narrow) synchronic grammar from the more differentiated (principles of) use, acquisition and change. While a full theory of this type is still a good ways off, the following are reasonable as guidelines. Principles of the synchronic grammar deal with derivation and representation and are stated in formal/structural terms, while principles of use, acquisition and change make reference to the speaker/hearer, to constraints on ease of production, comprehension and acquisition, to social factors, and to the effects of context.

It is clear that the case-marking/word-order freedom correlation (henceforth CWC) could be explained in terms of use, acquisition and change. Namely, word-order variation is easier to interpret, learn, and maintain over time in a language which marks grammatical relations with m-case. However, we can also imagine that this constraint is encoded directly within the narrow synchronic grammar, that the principles of grammar which determine word order make reference to m-case, such that rich case-marking licenses word orders which are impossible in its absence. A number of explanations of this latter type have been proposed over the years, of which I will mention three recent examples here. Roberts (1997), assuming a Minimalist framework, argues that the strength of Case features which drive DP movement is directly related to overt case-marking. Kiparsky (1997) develops a system in which the association of a DP with its θ-role must be guided either by overt case-marking or by canonical ordering. A DP can thus only leave its canonical position if case-marked. Finally, Neeleman and Weerman (1999) propose that Case is a syntactic head K which, when null, as in English, is forced by the ECP to remain in a position properly governed by its licensing head. An overt K is not subject to ECP, thus is free to move.

Now, it's clear that the pressures placed on the speaker-hearer as described above play at least some role in the restrictions on word-order freedom. Thus the burden of proof is arguably on those like Roberts, Kiparsky and Neeleman and Weerman who propose an additional synchronic-grammatical component to the correlation. They must show that such a hybrid explanation can account for facts that a simple use-based explanation misses, and I will argue in this paper that this burden cannot be overcome. Given the breadth of this topic, what I present here will be a preliminary, if rather suggestive, investigation.

2 Theoretical Issues
In this section I discuss three characteristics of the CWC and investigate their implications for a synchronic-grammatical explanation. In the process a series of theoretical problems will be encountered which, I will argue, cannot be overcome.

2.1 The CWC crosses the syntax-morphology interface
A relationship between case and word order must be considered in terms of the syntax-morphology interface. If m-case affects word order within the synchronic grammar, then the syntax must depend (in part) on the morphology. This is not an uncommon assumption, implicit e.g. in the Lexicalist Hypothesis, according to which the morphological properties of a word, including its case-marking, are
already determined at the point when it enters into the syntactic structure. Syntactic processes like those which determine word order could thus be sensitive to morphological factors. However, recent work, e.g. in Distributed Morphology, favors just the opposite view of the interface. There it is hypothesized that the syntax works with underspecified feature bundles, and morpho-phonological material is inserted into those bundles only after Spell-out. This is known as Late Insertion (see e.g. Halle and Marantz 1993; Marantz 1995). But if this is correct, then it is impossible for the presence or absence of case-marking to affect word order, because the markers themselves are not inserted into the structure until after word order has already been determined.

The obvious response to this objection is that what is relevant to word order is abstract Case, which is a syntactic feature and thus can have syntactic consequences, m-case just being the later spelling out of these features. But this too runs into a series of problems. The first is that the simplest hypothesis, that languages with m-case have syntactic Case, while those without it do not, is clearly untenable. Syntactic Case plays a role in DP positioning and licensing that is independent of word-order freedom and important in every language. Thus one would have to assume that abstract Case is universal, but fundamentally different in languages with m-case than in those without. But this would force Case theory to distinguish Latin Case from English Case while claiming that both are Case.

The second problem for a syntactic Case-based account is that there is mounting evidence against a direct connection between m-case and syntactic Case qua positional licensing. The argument boils down to three points which I lay out here in extremely abbreviated form. For the full details I refer the reader to Yip et al. (1987), Marantz (2000) and Schütze (1997). a) The relationship between structural cases and structural positions is not one-to-one, but many-to-many. In most languages, the primary structural cases are assigned not to specific positions, but according to a sequence. Thus in a nominative-accusative language, the highest argument which does not receive a lexical case gets nominative, the next highest accusative. (See Yip et al. 1987 and Marantz 2000 for discussion and analyses of this phenomenon.) This accounts for why underlying objects are marked nominative in passives, unaccusatives and oblique subject constructions, whether they raise to subject position or not. b) It is possible for a position to be assigned m-case without being assigned syntactic Case, as in (5), where einum shows that the subject position gets quirky dative from batna in spite of being empty (Sigurðsson 1991):

(5) að PRO batna veikin einum er erfitt.
   to PRO:DAT recover the-disease alone:DAT.MASC is difficult
   'To recover from the disease alone is difficult.'
c) It is possible for an overt DP to be licensed in a position where m-case is not properly assigned. This is where default case shows up, like the nominative on the left-dislocated DP in (6) (see Schütze (2001) for discussion):

(6) Der/*Dem Hans, mit dem spreche ich nicht mehr.
    the-NOM/*DAT Hans with him-DAT speak I not more
    `Hans, I don't speak with him anymore.'

If case can be assigned by a default rule, then it can't license DPs, because then in any instance where normal case-assignment failed, the default rule would be able to apply, and no DP would ever go unlicensed, rendering the Case filter vacuous.

The third problem with a syntactic Case-based account is that it would not actually constitute a synchronic-grammatical explanation of the CWC (See Bobaljik to appear for similar discussion of theories of verb raising). It claims that word-order freedom depends directly on a special Case feature we can call K'. But of course K' is syntactic, so it cannot in turn depend synchronically on m-case (because in the view adopted here the syntax precedes the insertion of morphophonological material). We would have to say that m-case cues the acquisition of K', i.e. when the learner is presented with rich case-marking, she can conclude that the language must have K' and thus the relevant word-order freedoms. This itself is quite plausible, but it explains the CWC through principles of acquisition, not the synchronic grammar. An account of the CWC which is explicitly based on use and acquisition from the start does not run into any of these problems, because it does not rely on a synchronic causal connection from m-case to syntax.

2.2 The CWC involves optionality

The optionality that word-order freedom represents is notoriously problematic for certain approaches to the CWC. E.g. Roberts' (1997) theory boils down to the simple (and therefore attractive) idea that case-marking drives movement. That is, DPs move to Agr positions in order to unify in some sense with the case-marking which resides there, much as verbs have been argued to raise to positions in I to combine with the tense, mood and agreement affixes that reside there. However, this just predicts that languages with case-marking will have a different rigid word order than those without (e.g. IO-DO instead of DO-IO), while what we want to explain is that such languages allow multiple orders (e.g. IO-DO and DO-IO). Kiparsky (1997) and Neeleman and Weerman (1999), on the other hand, design their theories specifically to derive the optionality effect. For them, m-case does not cause DPs to move, but satisfies obligations which would otherwise have to be met by rigid ordering. Still, the theoretical steps they take to derive the optionality are themselves problematic. E.g. Neeleman and Weerman use the ECP to keep caseless DPs local to their licensing heads, formulating the ECP as a PF filter which rules out offending structures rather than as a principle which steers steps the derivation. However, both filters and the Government relation that the ECP depends on have been abandoned in recent work because of the complexity
they introduce into the grammar. PF filters in particular are dubious because they are inconsistent with the default behavior characteristic of morphology. Similar remarks apply to the theory of Kiparsky (1997), although there matters are less clear because the framework he adopts is novel and less explicit on certain relevant details. Of course, optionality and how speakers deal with it is the proper domain of a theory of language use, so it should not be problematic for a use-based explanation of the CWC.

2.3 The CWC depends on 'richness and 'freedom'
Another familiar question for accounts of the CWC is precisely how ‘rich’ case-marking must be to allow word-order freedom, and how the grammar is to measure this ‘richness’. How we answer this depends on where we encode the CWC. A synchronic-grammatical explanation must define it in purely formal terms that apply universally and cannot use speaker/hearer-based notions like ‘sufficiently distinct’, ‘unambiguous’ or ‘salient’, because such notions are not defined within the synchronic grammar. But a definition that can satisfy these restrictions and account for the cross-linguistic variation on this point turns out to be impossible to formulate.2 First of all, richness cannot be reduced to the simple presence of a marker on a given form. It is typical in case-marking languages that one case in a given paradigm will have a zero marker, usually the nominative or absolutive, yet such forms have the same word-order freedom as ones with clear overt markers. What is actually relevant is distinctiveness between case forms, but even this cannot be determined by simply comparing the forms of a given noun. Latin genu ‘knee’ is the same in the nominative, dative, accusative and ablative, yet is no more restricted in its positional behavior than princeps ‘chief’, which has four distinct forms. So ‘richness’ of morphology must depend not on particular forms, but on distinctions between them, and is determined language by language, not word by word. But this is problematic if we want the synchronic grammar to do the determination, because it implies that it cannot be based strictly on the morphological information present in a given utterance. Rather, it must depend on paradigms having some reality in the synchronic grammar, which is commonly regarded as doubtful (Spencer1991, Bobaljik 2001).

Assuming for the sake of argument that this can be dealt with, we still need a universal formulation for the richness metric. We could propose something like the following: If a language distinguishes subject from object case in at least one number of at least one productive noun inflection class, then it has K'. This is intentionally analogous to a proposal made by Rohrbacher (1994): "A language has V to I raising if and only if in at least one number of one tense of the regular verbs, the person features [1st] and [2nd] are both distinctively marked (p.108)."

---

2 The following owes much to the discussion in Bobaljik (to appear) of the correlation between rich agreement morphology and verb raising. It is telling that both Bobaljik and Alexiadou and Fanselow (2000) come to the conclusion that the former cannot be explained within the synchronic grammar.
However, exceptions to Rohrbacher's principle are easy to find, including Tromsø Norwegian and Kronoby Swedish, which show no person distinctions, yet still have overt verb raising. Indeed, Bobaljik (to appear) has shown that no such definition could handle the variation attested within the Germanic languages, let alone universally. The task is no easier with the CWC, and in Section 3 we will see data that would contradict any definition of the CWC as a strong implication.

But even if we managed to construct a satisfactory definition, we would again have moved the CWC into the acquisition component, as was argued in Section 2.1. Consider what is at issue here. An essentially deductive step must be made from intra-paradigmatic morphological distinctions to the positing of a language-wide syntactic feature K'. Again, the synchronic grammar cannot make this leap, because it involves a syntactic feature depending on morphological information, contrary to the ordering of levels in DM. Rather, this is precisely the sort of thing the language acquisition device does, which constructs a grammar on the basis of a set of input data and the constraints of Universal Grammar. An explicitly used acquisition-based explanation of the CWC again avoids these problems, as it does not require a formal definition for 'rich' morphology. Rather, it depends on the ease with which speaker-hearers use and understand an utterance, thus it can use notions like 'sufficiently distinct' and 'unambiguous', which makes our task much easier. 3

3 Empirical issues
There are thus serious theoretical issues for any explanation of the CWC within the narrow synchronic grammar. Still, the success of any theory hinges on its ability to account for the empirical facts, thus in this section I consider the data that have been used to argue for a synchronic-grammatical explanation.

3.1 How strong is the correlation?
It is clear that certain word-order freedoms do pattern with the presence of case-marking, but is the correlation a bi-conditional, a one-way implication, or just a tendency? There is no shortage of, e.g., languages like Icelandic and Grisons Swiss German which have extensive overt case-marking, yet lack scrambling. Kiparsky (1997) thus argues for a one-way implication:

3 A related issue, which I cannot discuss here due to space constraints, is how a synchronic-grammatical account of the CWC can handle the clear effects of other factors, like agreement and intonation, on word-order freedoms. A use-based account correctly predicts without complication that all such factors which can aid interpretation will support marked word-orders, even in the absence of case-marking, while a synchronic-grammatical account must either formalize the effects of such factors in addition to those of case, or justify a more complicated hybrid account where case-marking constrains word-order freedom in the synchronic grammar but intonation does so in the use component. Similarly, a synchronic-grammatical account must give a principled explanation of why some reordering processes, like scrambling, seem to correlate with rich case-marking, while others, like topicalization, are more freely available.
On morphological case and word-order freedom

The most important point about this relationship is that it is...an exceptionless implication, which however holds in one direction only: lack of inflectional morphology implies fixed order of direct nominal arguments (p. 461, emphasis in original).

This is crucial, because as Kiparsky himself implies, only an actual implication can be properly encoded by a principle of the synchronic grammar. But even the one-way implication cannot be maintained. Dutch has no more case-marking than English, yet, as I will demonstrate in the next subsection, it allows scrambling of one object across another. Thus the CWC is not an implication, but a (strong) tendency. The synchronic-grammatical accounts cannot handle this. On the other hand, a tendential relationship is precisely what a use- and acquisition-based account predicts.

3.2 Dutch scrambling, Russian OVS, and pragmatic markedness

It is frequently reported that Dutch objects can scramble across negation and adverbs but not across other DPs. Kiparsky (1997), e.g., reports the following:

(7) dat Jan zijn vader het boek geeft
    that Jan his father the book gives
    'that Jan gives his father the book'

(8) *dat Jan het boek zijn vader geeft

However, things are not that simple. Zwart (1997) gives the data in (9)-(11) and the explanation below them:

(9) dat Jan Marie (gisteren) het boek gegeven heeft
    that John Mary yesterday the book given has
    'that John gave Mary the book yesterday.'

(10) ??dat Jan het boek Marie gegeven heeft
    that John the book Mary given has
    'that John gave Mary the book.'

(11) dat Jan het boek Marie terug gegeven heeft
    that John the book Mary back given has
    'that John gave the book back to Mary.'

[10] is unacceptable in a neutral stress pattern, i.e. with Marie slightly focused. Almost any marked stress pattern makes [10] acceptable though. Thus, in [11] the particle terug is in focus, and the order of the objects appears to be free (p. 32).

In short, Dutch does allow scrambling of one object across another. A special intonation and an appropriate discourse context are required to license such orders, but this applies to German scrambling as well (see e.g. Lenerz 1977 and Haider 1993 for discussion). There is independent evidence that, where German has DO-IO order without special intonation, a different underlying syntactic structure is involved, which is analogous to the construction where the IO is
marked in Dutch with *aan*, or in English with *to*. Kiparsky (1997) actually comes to this same conclusion and analyzes such examples differently from actual scrambling, following a view that is now fairly standard in analyses of the German double object (see e.g. Wegener 1991, Haider 1993, McFadden 2003). In other words, the 'scrambling' without discourse and intonational motivation that is supposed to occur in German but not in Dutch is not scrambling at all, but alternation in base structures. The difference between German and Dutch is that German morphology spells out the post-accusative argument in the dative case while Dutch spells it out with a preposition. Of course, actual scrambling of DO over IO is much more common in German than in Dutch. This is for the very simple use-based reason that scrambling runs a greater risk of being misunderstood in the absence of case-marking. The lower frequency in Dutch implies higher pragmatic markedness, which we can plausibly interpret as there being a smaller number of contexts in which the scrambled order is felicitous. Dutch speakers presumably have difficulty when presented with this order in grammaticality tests because the correct context is difficult to supply out of the blue, and the marked intonation, unlike case-marking, is not indicated in writing.

Similar remarks are in order for claims that marked word orders are impossible even in languages with case-marking when the case-marking on the specific nouns involved is ambiguous. Jakobson (1936) e.g. claims that OVS order is possible in Russian only in sentences where the case-marking is unambiguous. Thus he gives (12) as OK, because *syna* is marked accusative, but says that (13) is impossible because neither *mat* nor *doč* shows a nominative/accusative distinction:

(12) Syna rodila mat' prošlym letom.
    son:ACC bore mother:N/A last summer
    'The mother bore the son last summer.'

(13) *Doč* rodila mat'.
    daughter:N/A bore mother:N/A
    intended: 'The mother bore the daughter.'

However, according to the native speakers I have consulted, this is actually incorrect for Russian, as are similar claims that have been made for German and other languages of the relevant type. What is true is that this order is marked and requires contextual motivation and/or a marked intonational pattern. Russian speakers thus presumably have the same difficulty with (13) encountered out of the blue that Dutch speakers have with (8).

---

4 The idea is that certain verbs, including *verzeihen* 'forgive (s.o. s.t.)' have a basic structure in which the dative precedes the accusative. Others, including *aussetzen* 'expose (s.o. to s.t.)' have the accusative before the dative (see the cited works for syntactic and semantic tests which distinguish the two classes). Still others, including *geben* 'give' can appear in either frame. This is analogous to English, where verbs like *forgive* can only appear in the double object construction, others like *donate* can only appear in the *to*-dative, and still others like *give* can appear in either.
3.3 Misanalyses
To conclude this survey of empirical data, I will discuss an interesting class of examples which initially seem to support a synchronic-grammatical account of the CWC, but in the end turn out to have alternative analyses.

3.3.1 Ditransitives from Old to Middle English
OE lacked the to-dative construction (e.g. John gave the books to Mary), but allowed either both IO-DO and DO-IO orders without a preposition with the double object. In early ME, as case distinctions were being lost, the to-dative arose, and DO-IO order without a preposition disappeared. Thus it would seem that, when case-marking was lost, so was a certain type of word-order freedom, and the construction with to was created to fill the gap. However, I have argued elsewhere (McFadden, 2002), on the basis of evidence from a corpus study, that there were two distinct double object constructions in late OE and early ME. Sentences with IO-DO order had a structure like the modern double object construction, while at least some sentences with DO-IO order had a structure like the modern to-dative, with dative marking in place of the preposition to, just as is claimed above for German. In other words, DO-IO order was never lost. It simply became the to-dative construction via a morphological - not a syntactic - change.

3.3.2 ‘Semantic’ case DPs
DPs bearing ‘semantic’ or ‘adverbial’ case show a remarkable degree of freedom in their positioning cross-linguistically, frequently beyond that of other case-marked DPs in a given language. In German, e.g., argument DPs are banned from the position following the participle, as in (15), but semantic case-marked DPs, like the accusative of time in (16) are OK here, if somewhat marked:

(14) Er hat letzten Sonntag seinen Bruder besucht.
    he has last Sunday:ACC his brother:ACC visited
    ‘He visited his brother last Sunday.’
(15) *Er hat letzten Sonntag besucht seinen Bruder.
(16) ?Er hat seinen Bruder besucht letzten Sonntag.

It is tempting to think that it is the special case-marking that frees these nominals from normal constraints on DP ordering. But Emonds (1987) and Nikanne (1993) have argued that these are underlyingly PPs realized morphologically as case-marked NPs. This accounts for how they are assigned 0-roles, why they have the same semantic range as overt PPs, a number of restrictions on their behavior with respect to binding and secondary predication and the word-order freedom just noted. As (17) shows, the ability to extrapose is a general property of German PPs:

(17) ?Er hat seinen Bruder besucht im März.
    he has his brother visited in March
‘He visited his brother in March.’

The word order freedom is thus caused by the fact that these constituents are PPs. The case-marking is a symptom of the PP structure, not the cause of the freedom.

3.3.3 Case drop
The clearest possible evidence for a synchronic CWC would be if some word order were only possible in a language in the presence of a case-marker. Examples of this kind have been reported by Lamontagne Travis and (1987) for Japanese and Turkish (their behavior is the same, so I will restrict attention to the former):

(18) John-ga dare-(o) nagutta no?
    John-NOM who-(ACC) hit Q
    ‘Who did John hit?’
(19) Dare-*(o) John-ga nagutta no?

In the unmarked order in (18), the accusative marker -o on the direct object can be dropped, but if the object is fronted, as in (19), it cannot. Lamontagne and Travis thus argue that case is a syntactic head, subject to the ECP when it is phonologically null. The null K is properly governed in (18), but not in (19), because it has moved away from the licensing verb.

However, an alternative analysis is available which is fully consistent with my hypothesis and avoids the theoretical problems of Section 2. Recall that there is nothing in my theory to keep word order from affecting m-case. Like all morphology, it is inserted on the basis of the output of the syntactic derivation. Thus all we need say about (18) and (19) is that in Japanese (and Turkish), a case-marker must be inserted on objects that have moved out of VP (or whatever the precise generalization turns out to be). When the object remains VP-internal, there are two morphological options, -o and -Ø, with the choice between the two being conditioned by other factors, apparently related to focus. In fact, there are reasons to prefer the latter account. In particular, the former would predict object topicalization to be impossible in languages which lack m-case, because there would be no overt version of K available to obviate the effects of the ECP. But of course object topicalization is very much possible in in many languages without m-case, like English and Dutch. Of course, if case-drop depends on word-order, then topicalization in caseless languages presents no problem.

4 Conclusion
So we have seen theoretical and empirical arguments that the CWC is a tendency that results from principles of language use, acquisition and change, not a principle of the narrow synchronic grammar. We can imagine that the correct explanation will be roughly as follows. As a language like Old English with a high degree of word-order freedom lost its case-marking, the more marked word
orders would have become increasingly difficult to interpret correctly. As a result, it would have become increasingly difficult for children to acquire the processes that derived such marked word orders. This would have been helped along by increasing avoidance of the use of marked orders by those who had acquired them, to ensure being understood. The marked orders would have become even more marked, requiring stronger and stronger pragmatic and intonational motivation to ensure their unambiguous interpretation. Eventually, the evidence for the marked orders in the primary linguistic data would have been insufficient to allow children to acquire the word-order freedom, yielding essentially the Modern English situation. Of course, other factors like agreement and intonation can play a significant role in the interpretation of word-order variation, allowing in certain cases the preservation of word-order freedom even in the absence of morphological case-marking, as in Dutch.

References


McFadden, Thomas. 2003. The underlying structures of German inherent Case. Presented at the joint meeting of the FGLS and the SGL.


Thomas McFadden
619 Williams Hall
Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104-6305

tmcfaddel@ling.upenn.edu