The grammaticalization of K(case)Ps within Minimalism: formalism vs functionalism, synchrony vs diachrony:

Keith Tse
University of Manchester
keith.tse@balliol-oxford.com

Summary:
Roberts and Roussou (2003) argue that the cross-linguistic distribution of grammaticalization is due to ‘simplification’, and the grammaticalization of case-markers displays it since there is a loss of Agree relations. Synchronic ‘simplicity’ explains diachronic trends, and formalism and functionalism are not mutually exclusive.

Introduction:
Grammaticalization occurs cross-linguistically and is a challenge for Lightfoot (1999, pp. 148-149, 166-173; 2006) which predicts that language evolution should not have cross-linguistic trends. Roberts and Roussou (R & R) (2003, pp. 2-7) propose that grammaticalization is a natural kind of change that occurs cross-linguistically. They analyse the grammaticalization of three functional categories: auxiliary verbs (T), complementisers (C) and determiners (D). There is another functional category (K(case)). In this paper, I propose to expand R & R’s hypotheses by testing them on the grammaticalization of K(case), and I shall illustrate the theoretical relationships between synchrony and diachrony and between formalism and functionalism.

Section 1: Generative models of language change:
Section 1.1: Lightfoot (1999, 2006):
Lightfoot argues that grammar is moulded in first language acquisition, which is the locus for language change (Lightfoot 1999, pp. 60-74; 2006, pp. 10-15, 88-89). There are three components:
1) internal grammar (IG)
2) universal principles and parameters of grammar (UG)
3) trigger experience in the form of primary linguistic data (PLD).
IG is formed by children analysing their PLD and setting the parameter values of their UG:
   a) Linguistic triggering experience (genotype → phenotype)
   b) Primary linguistic data (Universal Grammar → internal grammar)¹
As UG is a genetic constant, the source for language change lies in the PLD and in how children (re-)analyse it in language acquisition.²

Section 1.2: ‘Re-analysis’ in grammaticalization:
The classic example of ‘re-analysis’ in grammaticalization is English going to > gonna:³ a) there are examples (purposive directional constructions with non-finite complements) where the old (going to denoting movement and purpose) and new (gonna denoting futurity) interpretations co-exist⁴ b) there is a context (the absence of an overt directional phrase) in which the old interpretation is weakened and the new one strengthened⁵ c) the outcome of ‘re-analysis’ is identified in examples where only the new interpretation is possible/likely.⁶

---

³ Hopper and Traugott (H & T) (1993:2-4).
⁴ ‘the change occurs only in a very local context, that of purposive directional constructions with non-finite complements... I am going to marry Bill (i.e. I am leaving/travelling to marry Bill)’ (H & T (1993:2)).
⁵ b) ‘... there is an inference of futurity from purposives... In the absence of an overt directional phrase, futurity can become salient.’ (H & T (1993:3)).
⁶ ‘the re-analysis is discoverable... only when the verb following be going to is incompatible with a purposive meaning, or at least unlikely in that context... I am going to like Bill, I am going to go to London...’ (H & T (1993:3)).
Section 1.3: ‘Re-analysis’ in generative models of language change:
Both Lightfoot (1999, 2006) and R & R (2003) argue that in language acquisition ‘cues’ express parameter values.\(^7\) Steps a), b) and c) are therefore all ‘cues’, and b) is the exact point of parameter resetting where the previous parameter value (going as a lexical verb construed with to as a preposition) is dropped in favour of the new parameter value (gonna as an auxiliary verb). Lightfoot asserts that language evolution is random because he argues that PLD is language-specific and unpredictable. He makes no comment on how PLD shifts through time.\(^8\) In Lightfoot’s model, the cross-linguistic distribution of ‘cues’ is random. Grammaticalization occurs cross-linguistically and is hence incompatible with Lightfoot’s model. R & R (2003:14-17) introduce a learning device in language acquisition that chooses the ‘simpler’ alternative in ambiguous ‘cues’,\(^9\) and since they argue that grammaticalization always leads to ‘simpler’ structures, it is a natural mechanism in language acquisition that can occur cross-linguistically.\(^10\) They define ‘simplicity’ as the reduction of ‘formal feature syncretisms’, namely ‘the presence of more than one formal feature in a given structural position: H [+F, +G...]’,\(^11\) and they discover three types of grammaticalization:\(^12\)

1) \([xp \ Y=X \ [yp\...\ y...]] > [xp \ Y=X \ [yp\...\ y...]]\)  
2) \([xp \ X=F \ [yp\...\ y...]] > [xp \ X=F \ [yp\...\ y...]]\)  
3) \([xp \ Y=P X\ ... \ [\ ...\ y...]] > [xp \ Y=X \ [\ ...\ y...]]\)  

The first (1) and third types (3) involve the loss of Move and the introduction of Merge to the grammaticalized item in a higher position. The second type (2) involves the loss of Agree and an upward shift of features to the grammaticalized item. R & R represent grammaticalization thus:

\[
\text{XP} \\
\text{Y=X} \quad ... \\
\text{YP} \\
\text{Y} \quad ... \\
\]

In all three types, features in a lower syntactic position (Y) are shifted upwards to a higher position (X). Such is R & R’s characterisation of grammaticalization.

Section 2: K(case):
The earliest postulation of K(case) as a functional category was proposed by Lamontagne and Travis (L & T),\(^14\) who note that when nominal complements are adjacent to their head predicates, their morphological case-endings can be optionally dropped (5a, 6a, 7a), but when they are not adjacent, their morphological case-endings are obligatory (5b, 6b, 7b):

Japanese:

5a) John-ga dare(-wo) nagutta no?  
John-NOM who-ACC hit Q  

---


\(^8\) They permit an appropriately contingent account of why the change took place... the expression of the cues changed in such a way that a threshold was crossed and a new grammar was acquired. That is as far as this model goes, and it has nothing to say about why the distribution of cues should change.’ (my italics) (Lightfoot (1999:166)).

\(^9\) It is not clear whether this learning device is part of UG or not, since Vincent & Borjars (2010:280, 293) consider it as part of UG whereas van Gelderen (2011:9) attributes it to principles that are not specific to UG. Either way this learning device plays a prominent role in Minimalism.


\(^12\) R & R (2003:198-199).


5b) dare*(-wo) John-ga nagutta no?
   who-ACC John-NOM hit Q
   ‘Who did John hit?’

Turkish:

6a) Hasan dünn (bu) pasta(-yi) ye-di
    Hasan yesterday this cake-ACC eat-PAST
6b) Hasan *(bu) pasta*(-yi) dünn ye-di
    Hasan this cake-ACC yesterday eat-PAST
   ‘Hasan ate (this) cake yesterday.’

Welsh:

7a) mae ’r dyn wedi gweld ci
    is the man after seeing dog
    ‘The man has seen a dog.’
7b) gwelodd y dyn gi
    saw the man dog
   ‘The man saw a dog.’

This distribution resembles other functional categories e.g. complementisers, which are also omissible when they are adjacent to their head predicates (8a, 9a) and not elsewhere (8b, 9b-d):

Japanese:

8a) Mary-ga kinoo John-ni Koobe-ni iku (te)
    Mary-NOM yesterday John-DAT Kobe-DIRECTIONAL go COMP
    yuuteta (koto)
    was.saying fact
8b) Mary-ga kinoo Koobe-ni iku * (te) John-ni
    Mary-NOM yesterday Kobe-DIRECTIONAL go COMP John-DAT
    yuuteta (koto)
    was.saying fact
   ‘Mary said to John yesterday that she was going to Kobe.’

English:

9a) John believes (that) Mary will win.
9b) John believes wholeheartedly *(that) Mary will win
9c) That Mary will win, John believes with all his heart.
9d) *Mary will win, John believes with all his heart.

L & T therefore postulate a functional category for morphological case called K(case) on the left-edge of DPs (11), just like complementisers are postulated on the left-edge of TPs (12):

11)

\[
\begin{array}{c}
\text{KP} \\
\text{K} \\
\text{DP} \\
\text{D} \quad \text{NP} \\
\text{N} \quad \text{...}
\end{array}
\]

---

17 L & T (1986:54) analyse Welsh gi as a variant of ci and argue that this voicing (ci > gi) is equivalent to morphological case since it is obligatory when the complement is not adjacent to the head predicate (7b).
18 L & T (1986:54).
Van Kemenade and Vincent argue that functional categories host functional morphology and lexical categories move to them in order to ‘pick up’ their morphology\textsuperscript{22} e.g. Infl (=T), which hosts verbal morphology and causes lexical verbs to move from V to T.\textsuperscript{23} In the case of K, van Kemenade and Vincent argue that K hosts case morphology and NPs/DPs with morphological case move to K in order to have their morphological case licensed.\textsuperscript{24}

**Section 2.2: the grammaticalization of K(case):**

As K represents morphological case, any morpheme that is equivalent to morphological case can be analysed as K i.e. a case-marker.\textsuperscript{25} In this section, I identify some case-markers in Latin/Romance.

**Section 2.2.1: lexical prepositions and functional prepositions:**

Cinque argues that there are two types of prepositions, lexical and functional.\textsuperscript{26} Huddleston and Pullum define functional prepositions in English as those that are obligatorily selected by the head predicate e.g. of in they disposed of the box (*they disposed the box).\textsuperscript{27} Furthermore, of in this construction does not co-vary with other prepositions spatially: *they disposed at / below / on / through / under the box, which is a contrast to lexical prepositions that do: put it under / above / near the table.\textsuperscript{28} This suggests that of in English dispose of is a non-spatial complement of the head verb (dispose). All this conforms with generative approaches that define functional prepositions as being part of subcategorisation and not having argument structure or spatial features\textsuperscript{29} e.g. English rely on his help:

1)\[
\text{VP} \\
\text{V} \quad \text{K(case) P} \\
\text{rely} \quad \text{K(case) on} \\
\text{DP} \quad \text{his help}\]

The complement (his help) of the head verb (rely) is non-spatial (‘instrument/theme’) and on here is analysed as a case-marker (K) for this complement.\textsuperscript{31} This is supported by the fact that many functional prepositions are equivalent to morphological case in other languages:\textsuperscript{32}

2)\[
\text{milit-} \quad \text{Graec-i} \\
\text{soldier-NOM.PL} \quad \text{Greek-NOM.PL} \quad \text{super-at-i} \\
\text{sunt} \quad \text{overcome-PERF.PTCP-NOM.PL} \\
\text{be.PRES.3PL} \quad \text{Roman-is} \\
\text{Roman-ABL.PL} \quad \text{‘The Greek soldiers were beaten by the Romans (=Romanis).’ (my brackets) (Latin)}\textsuperscript{33}
\]

\textsuperscript{22} Van Kemenade and Vincent (1997:6-7).
\textsuperscript{24} Van Kemenade and Vincent (1997:20).
\textsuperscript{25} van Kemenade and Vincent (1997:18ff).
\textsuperscript{26} Cinque (2010:3-11).
\textsuperscript{27} Huddleston and Pullum (2002:647).
\textsuperscript{28} Huddleston and Pullum (2002:647).
\textsuperscript{30} Abraham (2010:272).
\textsuperscript{33} Abraham (2010:272).
Romanis has ablative case (-is) and is rendered into English by the preposition by (by the Romans). English by is therefore equivalent to the Latin ablative case and can be analysed as a functional preposition (K).

2.2.2: Latin/Romance functional prepositions:
In Romance, there are non-spatial functional prepositions that correspond to Latin morphological cases e.g. Romance ad, which marks the third non-spatial argument ('recipient', 'beneficiary') and corresponds to the Latin dative case e.g.

3) les tengo que regalar
    que regal-ar
    them.DAT have-1SG.PRES that give-INF
    a los niños
    to DEF.ART.MASC.PL child-MASC.PL
    la bicicleta nueva
    DEF.ART.FEM.SG bicycle.FEM.SG new-FEM.SG

'I have to give the children a new bike.' (Spanish)

Latin ad is originally a lexical spatial preposition denoting 'direction' and is used with trivalent verbs in Plautus (254-184 BC): 4a)

4a) qui ae ad patr-em v-is
    which-N.PL.ACC to father-ACC.SG want-PRES.2SG
    nuntiari-report-INF-PASS
    '... the things which you want to be reported in the direction of your father.'
    (Plautus, Captivi 360)

ad assigns morphological case (accusative) to its nominal complement (patr-em), and so there is an Agree relation between them. Its complement has morphological case (patr-em) and is therefore a KP which contains a Move relation between K and its DP complement (see section 2.1). The ad-PP (ad patrem) modifies the verb (nuntiari) and should be analysed as its adjunct:

---

37 I have placed the indirect object (les... a los niños) in SpecV, following Larson (1988), who places the English direct object in SpecV and the indirect object as the complement of V (Larson (1988:339, 342ff)). In order to derive the correct word order in Spanish, I analyse the direct object (la bicicleta nueva) as the sister of V and the indirect object as the specifier of V. This reversal has been suggested by Hale and Keyser (2002:160-163).
As *ad* marks the ‘direction’ in which the message is to be conveyed (*quae ad patrem... nuntiari* ‘the things which... to be announced in the direction of your father’), its complement (*patrem* ‘father’) can be re-analysed as the ‘recipient/beneficiary’ of the verb i.e. its indirect object. The *ad*-PP can therefore be re-analysed as a dative KP headed by *ad* in SpecV (see section 2.2.2, ex. 3 and footnotes 40 and 42):

---

39 In the same passage, there is a very similar example where the dative (*patri* ‘to the father’), the default case for marking indirect objects in Latin (Blake (1994:6)), is actually used with the same predicate (*vis... nuntiari* ‘you want... to be announced’):

1) numquid aliu-d v-is patr-i
whether another-N.SG.ACC want-PRES.2SG father-DAT.SG
nunti-ar-i report-INF-PASS
‘whether you want another thing to be reported to your father.’ (Plautus, Captivi 400)

The Latin *ad*-PP in 4a) is therefore functionally very similar to the Latin dative case.
4aii) is ‘simpler’ than 4ai), since the Agree relation ([i-P], [u-K]) between the lexical preposition *ad* and its KP complement (*patre-em*) is lost. The features associated with morphological case ([i-K], [u-D]) are shifted upwards from the complement of *ad* (*patre-em*) in an adjunct position to *ad* itself in a base-generated complement position (SpecV). However, Adams and Pinkster both point out that in this particular example *ad* retains its full directional force, as the recipient (*patrem* ‘father’) is most likely not immediately present in this scene and the message has to be transported to him, which implies spatial ‘direction’. This *ad*-PP is therefore stronger than the morphological dative in footnote 42 and cannot be equated with it yet. This is step a).

Step b) occurs when the spatial force of *ad* is weakened, and this is found in very late Latin (e.g. 4b)

et dix-erunt *ad* Petr-um
and say-PERF.3PL to Peter-ACC.SG
et *ad* reliqu-os apostol-os
and to rest-ACC.PL apostle-ACC.PL
quid faci-emus
what do-FUT.1PL
‘... and they said to Peter and to the rest of the Apostles what we shall do...’

*(The Latin Vulgate Bible, Actus Apostolorum 37, 6th century AD)*

---

Like 4a), the complements of ad (Petrum… reliquos apostolos ‘Peter and the rest of the apostles’) denote the ‘direction’ in which the message was conveyed (4bi) and can hence be re-analysed as the ‘recipients’ of the main verb (dixerunt ‘they said…’) (4bii). Here there is no question that the ‘recipients’ (Petrum… reliquos apostolos ‘Peter and the rest of the apostles’) are immediately present in the scene, and so there is very little (if any) difference between the ad-PPs (4bi) and the dative (4bii) here. 4bii), the ‘simpler’ analysis, comes through.

Step c) occurs in the Merovingian documents (7th-8th century AD). Adams argues that the use of ad-PPs with trivalent verbs is mainly attested with verbs of saying. In the Merovingian documents, Adams (2011:266-267).
there are the first attestations of *ad*-PPs being used with other types of trivalent verbs e.g. verbs of giving/showing:  
4c) \( \text{ad monasthryio condona-verant} \)  
\( \text{AD monastery present-PLUPERF.3PL} \)

‘they had presented to the monastery...’ *(Merovingian documents, 29.11)*

The spread of *ad*-PPs to other trivalent verbs suggests that they have already been grammaticalized as dative KPs. The grammaticalization of Latin/Romance *ad* therefore conforms to R & R’s ‘simplicity’ and ‘upward feature analysis’.

**Section 2.3: cross-linguistic distribution:**

It is pleasing to see that the grammaticalization of Latin/Romance *ad* as K, a new functional category (see introduction), conforms to R & R’s hypotheses. This predicts that it should have cross-linguistic counterparts that also undergo ‘structural simplification’ (see section 1.3), which is indeed borne out e.g. English to:

5) \( \text{his suna twegen mon brohte to } \text{þæm cyninge} \)

*His sons brought one of the two in the direction of that king*. *(Two of the Saxon Chronicles Parallel 86.26, 894 AD)*

to is a lexical spatial preposition marking ‘direction’ and to *þæm cyninge* is hence a PP-adjunct to the main verb (*brohte* ‘brought’) (5a). Its complement (*þæm cyninge* ‘that king’) is therefore re-analysable as the ‘recipient/beneficiary’ of the main verb, in which case to is re-analysed as K(dative), \(^{44}\) a complement within the VP (5b) (see footnote 40):

---

\(^{43}\) Vielliard (1927:201).

\(^{44}\) The same usage is found synchronically with the same head predicate (*broht* ‘bring’) and the same complement (*þæm cyninge* ‘that king’) where the morphological dative is used:

1) \( \text{þa teþ hie brohton sume } \text{þæm cyninge} \)

*They brought some of those teeth to that king.* *(King Alfred, Orosius 18.1, 849-899 AD)*

The to-PP in 5) is therefore functionally equivalent to the English dative case, like Latin/Romance *ad* (see footnote 42).
5b) is ‘simpler’ than 5a), since, like Latin/Romance ad (4a-b), the Agree relation between the lexical preposition (to) and its KP complement (pæm cyninge) and the Move relation between K and its DP complement are lost, and since Old English preposition to subcategorises for complements with morphological case (pæm cyninge), the K features already exist in the original ‘cue’ (5a) and are ‘upwardly shifted’ from the complement of to (pæm cyninge) in adjunct position to to itself in complement position (5b). With the subsequent loss of spatial force, to is re-analysed as K(dative), as in modern English.  

45 Traugott (1972:80-81).
Section 3: Vincent & Borjars (2010):
Vincent & Borjars (V & B) argue that formalism and functionalism should not be seen as mutually exclusive in language change.\(^{47}\) Both Latin/Romance *ad* and English to go through ‘structural simplification’, namely ‘reduction of feature syncretisms’ (see sections 1.3, 2.2.2, 2.3). R & R’s synchronic definition of ‘simplicity’, therefore, accounts for a diachronic trend, namely the cross-linguistic distribution of grammaticalization. Synchrony and diachrony can therefore be combined in the Minimalist account of grammaticalization. Furthermore, while R & R’s ‘simplicity’, a formalist consideration (see footnote 51), holds for these cross-linguistic examples, the ‘cues’ in the PLD, which are functionalist and data-based (see footnote 52), are not cross-linguistically random: both Romance *ad* (4a) and English to (5a) denote ‘direction’ and ‘recipient/beneficiary’ simultaneously, which is a strong cross-linguistic trend,\(^{48}\) and both lose their spatial force in ‘re-analysis’ (4b, 5b). Formalism and functionalism, therefore, account for different yet related aspects of ‘re-analysis’ and are hence not mutually exclusive. The role played by functionalist factors in grammaticalization further attests to the fact that pragmaticalization\(^{49}\) plays a significant role in grammaticalization.

Conclusions:
It is pleasing to see that my analysis of the grammaticalization of a new functional category (K(case)) conforms to the previous literature about Minimalism and grammaticalization (R & R (2003)), which is powerful evidence in support of their hypotheses. Furthermore, this case-study allows me to verify another hypothesis (V & B (2010)), which attests to the interface between several mutually unexclusive methodologies in linguistics (synchrony/diachrony, formalism (theory)/functionalism (data), pragmaticalization/grammaticalization).

Literature:

\(^{47}\) Formalism is defined as ‘a property of a theoretical system’ (V & B (2010:283)) and is said ‘to model this data in terms of the innate asymmetries of Universal Grammar (UG) (‘simplicity’ being preferred in language acquisition (see section 1.3)) (my brackets) (V & B (2010:280)), while functionalism ‘relates internal aspects of language to the external context of language use’ (V & B (2010:283)) and ‘seeks to explain these diachronic patterns (i.e. cross-linguistic distribution) with reference to discourse and interpersonal communication strategies rather than in terms of an innate UG’ (my brackets) (V & B (2010:280)).


\(^{49}\) Pragmaticalization is defined as ‘the fossilization of discourse strategies in syntactic and morphological structure’ (Traugott and Heine (1991:2, 5)) and is closely related to functionalism (see footnote 51).


