

The grammaticalization of formal nouns and nominalizers in Chinese, Japanese and Korean.

Andrew Simpson (SOAS) & Xiu-Zhi Zoe Wu (USC)

1. Introduction

This paper focuses on the syntactic status of three elements which can be shown to have strikingly similar distributions in Chinese, Japanese and Korean, the morphemes *de*, *no* and *kes*, and attempts to see how both the similarities and the differences in the distribution of these elements can be given a unified account. The suggestion is made that the similarities and the differences in the patterns found can be systematically explained if it is assumed that the three elements are all essentially of the same type in origin but synchronically in different stages of grammaticalization and syntactic change. The paper also introduces and argues for a formal movement-based model of grammaticalization and shows how such an approach is able to capture the apparently gradual nature of syntactic change in a possibly more precise way than traditional feature-based cline descriptions. Applied here to the *de/no/kes* paradigm, the approach allows for a new understanding of the development of formal nouns and functional categories in the NP/DP, and also results in an interesting prediction concerning the existence of double nominalizers which appears to be borne out by various cross-linguistic patterns.

The structure of the paper is as follows. First of all in section 2 we point out the clear similarities in distribution which are found with *de*, *no* and *kes*, as well as noting the important differences which also occur. In section 3 we then consider what one might initially conclude about the categorial status of *de*, *no* and *kes* on the basis of diachronic and synchronic evidence, and the hypothesis is put forward that these elements might lie at different points in a grammaticalization chain relating nouns to nominalizers and determiner-like elements. In section 4 we reflect on how grammaticalization processes may in general be modelled and propose that occurrences of syntactic category change may in fact frequently be the result of movement and re-analysis taking place in the functional structure projected by a lexical element. Returning to the *de/no/kes* patterns in section 5, such an approach is then shown to allow for an account which reconciles the full range of similarities and differences in the distribution of these elements. Finally in section 6 we consider how the occurrence of certain related patterns of change in other languages seems to clearly vindicate the treatment of *de/no/kes* and support the general model of grammaticalization suggested in the paper.

2. Formal nouns and nominalizers – *de*, *no* and *kes*

2.1. The distribution of Chinese *de*

In Mandarin Chinese a morpheme pronounced as /dʅ/ transcribed in Pinyin as *de* occurs in a variety of constructions. The element *de* regularly occurs in headed relative clauses as in (1) as well as headless relative clauses of the type in (2):

- (1) wo mai de shu
I buy DE book
'the book I bought'

- (2) ta zongshi ting [wo shuo de]
 he always listen I speak DE
 'He always obeys me.'

In children's Mandarin Chinese one also finds head-internal relative clauses with the element *de* present, as reported in Chiu (1998)

- (3) ganggang nei-ge daxiang tiao de
 just-now that-Cl elephant jump de
 'the elephant which just jumped'

The element *de* also has a genitive case-like function indicating possession, as seen in example (4):

- (4) wo de shu
 I DE book
 'my book'

Finally *de* occurs sentence-finally often accompanied by the copula *shi* in a cleft-like constructions illustrated in (5) and in pseudo-cleft-like constructions such as (6):

- (5) wo shi zuotian lai de
 I BE yesterday came DE
 'It was yesterday that I arrived.'

- (6) wo mai de shi shu
 I buy DE BE book
 'What I bought was a book.'

This distribution of the functional element *de* in Chinese shows an interesting and strong similarity to patterns found with Japanese *no* as shown below.¹

2.2. The distribution of Japanese *no*

First of all, as shown in (7) below, the Japanese morpheme *no* occurs in a genitive case-function clearly similar to Chinese (4):

- (7) Taro no hon
 Taro NO book
 'Taroo's book'

¹ Note that there are in fact a number of functional elements in Mandarin Chinese all having the same pronunciation /dʅ/, and that in addition to the environments detailed in (1-6) an element pronounced as /dʅ/ occurs in resultative constructions, potential 'V-*de*-V' forms and in the so-called 'descriptive *de*' constructions. However, if one considers other Chinese dialects such as Cantonese, Shanghainese and Taiwanese, it is found that the elements which occur in these latter constructions all have quite different pronunciations, while the morphemes occurring in the environments in (1-6) maintain a uniform pronunciation (e.g. /gʅ/ in Cantonese, /ʅ/ in Taiwanese). Consequently it is commonly accepted that the elements occurring in the environments in (1-6) are (at least diachronically) the same element and different from other morphemes with a similar pronunciation.

The element *no* also occurs with the copula in a sentence-final cleft-like form which closely resembles the Chinese *shi-de* cleft seen in (5) in both meaning and use (see e.g. Kitagawa & Ross 1982):

- (8) (watashi-wa) kinoo kita no desu
 (I-TOP) yesterday came NO BE
 ‘I came *yesterday*/It was yesterday that I came.’

No furthermore occurs in pseudo-clefts such as (9), resembling Chinese (6):

- (9) (watashi-ga) katta no wa hon desu
 (I-NOM) bought NO TOP book BE
 ‘What I bought was a book.’

When one considers the patterns found in relative clauses, the results are particularly interesting and rather different from Chinese, but in a potentially revealing way. In adult headed relative clauses *no* does not appear, unlike Chinese (compare example (1)):

- (10) watashi-ga katta (*no) hon
 I-NOM bought NO book
 ‘the book (which) I bought’

The element *no* does however occur in adult headless relative clauses and also head-internal relative clauses as illustrated in (11) and (12):

- (11) kono-hon-wa watashi-ga katta-no ja-nai
 this-book-TOP I-NOM bought-NO BE-NOT
 ‘This book is not the one which I bought.’

- (12) keikan-wa [doroboo-ga ginkoo-kara dete-kita]-no-o tsukamaeta
 policeman-TOP thief-NOM bank-from exit-came NO-ACC arrested
 ‘The policeman arrested the robber who came out from the bank.’
 (Murasugi 1997)

No furthermore significantly also occurs in children’s headed relative clauses as illustrated in (13)

- (13) ohana motteru no wanwa
 flower holding NO doggie
 ‘the doggie (which is) holding the flower’ (Murasugi 1997)

Finally, *no* is found as a fairly clear nominalizer in examples such as (14):

- (14) [Taroo-ga tsuita]-no-o shiranai
 Taroo-NOM arrived-NO-ACC not-know
 ‘I didn’t know that Taroo had arrived.’

2.3. The distribution of Korean *kes*

Turning now to Korean, here one finds another morpheme with a distribution strikingly similar to Chinese *de* and Japanese *no* - the element *kes*. Considering first of all relative clauses, *kes* seems to show a patterning which is fully parallel that of Japanese *no*. As with Japanese *no*, *kes* does not occur in adult headed relative clauses (15), but it is found in adult headless relative clauses (16) and head-internal relative clauses (17):

(15) [Manho-ka sa-n] (*kes) chayk
Manho-NOM bought KES book
'the book which Manho bought'

(16) ku-chayk-un [nae-ka sa-n] -kes-i anieyo
that-book-TOP I-NOM bought KES-NOM not-BE
'That book is not the one I bought.'

(17) [chayk pily-e ka-n]- kes nayil kac-ko o-kyess-upni-ta
book borrow go-PAST KES tomorrow bring-ing come-FUT-POL-DEC
'I will bring back the book that I borrowed tomorrow.'
(Whitman, Lee & Lust 1991)

Furthermore, as with Japanese *no*, *kes* is in fact also found in Korean children's (headed) relative clauses as shown in (18):

(18) kuriko [appa ssu-nun]-ke ankyeng-un..
then papa wear-ing KES glasses-TOP
'And the glasses that papa wears..' (Whitman, Lee & Lust 1991)²

Like Japanese *no* and Chinese *de*, Korean *kes* is additionally present in pseudo-clefts, as in (19):

(19) [nae-ka san]-kes-un chayk ieyo
I-NOM bought KES-TOP book BE
'What I bought was a book.'

Kes also occurs in a sentence-final pattern with the copula in a way which is syntactically similar to the sentence-final cleft-like copula patterns in Chinese (5) and Japanese (8); this is illustrated in (20):

(20) Yongguk-ulo ttona-ss-ul-ke-eyo
England-to leave-PAST-IRR-KES-BE
'He must have left for England.'

Finally, *kes* functions to allow clauses to be embedded as arguments of verbs and be case-marked, this being similar to the nominalization function of Japanese *no* in examples such as (14):

² Note that *kes* here appears in its reduced form pronounced as /kɨ/. This pronunciation is common if there is no case particle following *kes*.

- (21) [totuk-i ton-ul hunchin]-kes-un yeki lopute ta
 robber-NOM money-ACC stole KES-TOP here from BE
 Lit: ‘The robber’s stealing the money was from here.’

One clear difference between Korean *kes* and Japanese *no* and Chinese *de* is that *kes* unlike *no* and *de* does not occur in a genitive case-like function. This role is instead instantiated by a different element *uy*:

- (22) Manho uy chayk
 Manho GEN book
 ‘Manho’s book’

For convenience, the full range of patterns which have been noted above with *de*, *no* and *kes* are now summarized in (23) below, together with the relevant example numbers:

(23)

	relative clauses				pseudo-cleft	cleft/copula	genitive	clausal ‘nominalizer’ function
	adult headed	child headed	head-internal	headless				
Chinese	✓ (1)	✓ (1)	✓(children) (3)	✓ (2)	✓ (6)	✓ (5)	✓ (4)	
Japanese	Ψ (10)	✓ (13)	✓ (12)	✓ (11)	✓ (9)	✓ (8)	✓ (7)	✓ (14)
Korean	Ψ (15)	✓ (18)	✓ (17)	✓ (16)	✓ (19)	✓ (20)	Ψ (22)	✓ (21)

From (23) it can be seen that alongside a small number of differences, there is indeed a tremendous similarity in the distribution of *de*, *no* and *kes*, which seems to suggest that these elements should *in some way* be analyzed as being largely of the same type. The interesting challenge which is posed by the collective distribution of *de*, *no* and *kes* is how to arrive at a unified analysis of these elements which both captures the important generalizations observed and also accounts for the potentially significant differences in (23). In section 3 we now attempt to develop such an account critically suggesting that *de*, *no* and *kes* are all at somewhat different stages of development along a single path of grammaticalization.

3. Grammaticalization and the *de/no/kes* paradigm

3.1. Further synchronic and diachronic clues concerning the syntactic identity of *de*, *no* and *kes*

If the largely parallel distribution of *de*, *no* and *kes* might seem to indicate some kind of identity relation among the three morphemes, one obviously needs to ask what kind of syntactic element *de*, *no* and *kes* could possibly all instantiate. Initially this is rather unclear and while all three elements would indeed seem to perform a largely

functional role in the constructions in which they appear, it is not obvious what type of syntactic category these functional roles might actually correspond to. Here a reflection on various other synchronic and diachronic aspects of the patterning of each of the three elements is able to provide a certain amount of further useful information relevant to their current syntactic status.

Of the three elements in the paradigm, the categorial status of Korean *kes* would at first sight seem to be the clearest. *Kes* is often considered to be a noun with the simple meaning ‘thing’ and appears to have such a meaning when modified just by a demonstrative as in (24):

- (24) ku-kes-i
 that-KES-NOM
 ‘that thing’

However, *kes* is not a fully regular type of noun and has a distribution which is significantly more restricted than other nouns. Critically *kes* requires some type of modification and unlike other nouns is not free to occur unmodified, as seen in the contrast in (25):

- (25) chayk-un/*kes-un chayk-sang-wi-e issoyo
 book-TOP/KES-TOP table-top-on be
 ‘The book(s)/thing(s) are on the table.’

This modification furthermore has to be of a specific type and it is not possible for *kes* to occur modified just by a following universal quantifier as in (26):

- (26) *kes-i ta yogi issoyo
 KES-NOM all here be
 intended: ‘The things are all here/All the things are here.’

Kes as a noun meaning ‘thing’ is also rather exceptional in that it cannot be modified and licensed just by numerals, as shown in (27):

- (27) *kes-i du-kae chayk-sang-wi-e issoyo
 KES-NOM two-CL table-top-on be
 intended: ‘There are two things on the table.’

Consequently, although it is initially not unnatural to consider *kes* as a type of noun/N⁰ (meaning broadly ‘thing’), *kes* also seems to show important restrictions which distinguish it from more fully regular nouns. *Kes* might therefore seem to have more of the status of what Matsushita (1930), Kuno (1973) and others have (in Japanese) referred to as ‘formal nouns’, nouns which have very little or sometimes no substantive content and which occur predominantly or even exclusively in functional syntactic roles.

Moving on to the other two members of the *de/no/kes* group, if it is possible to assume that Korean *kes* is at least in some degree nominal/an N⁰ on the basis of occurrences such as (24), the categorial identity of Chinese *de* and Japanese *no* turns out to be rather harder to establish. Considering first Chinese, the element *de* is not able to occur modified by demonstratives in the way that Korean *kes* is, as shown in

(28), and *de* does not show any other signs of being a noun either, not being countable nor occurring in isolation as seen in (29) and (30):

- (28) *zhei de
this DE
- (29) *liang-ge de
two-CL DE
- (30) *de (dou) zai zhe-li
DE all be here

It might therefore seem that *de* is unlikely to be an N^0 and its synchronic categorial status consequently remains an open issue. Being a functional element with a rather unclear identity it is reasonable however to assume that *de* has reached this stage through some kind of grammaticalization from an earlier more transparent source, and a consideration of the possible origins of *de* provides more insight into its current functional status.

Although it is not in fact fully clear exactly how *de* came into modern Chinese, there is a widespread belief that it simply replaced the earlier classical Chinese morpheme *zhi*. This element *zhi* occurred in the same range of syntactic positions as *de*, linking two nominals in the genitive-like function and also linking a clause to a noun in what is often referred to as a nominalization (relative clause) function, as illustrated in (31) and (32)

- (31) Wang zhi zhu-chen
king ZHI minister
'Your majesties ministers' (Mengzi 1A/7)
- (32) wu duo ren zhi jun
insult rob people ZHI ruler
'a ruler who insults and robs his people' (Mengzi 4A/17)

Significantly *zhi* also had a wider distribution than modern *de* and it is this wider distribution which provides critical clues to the categorial identity of *de* and *zhi*. In addition to the above patterns which mirror those with *de*, *zhi* is also documented as occurring as a pronominal and as a demonstrative as shown in (33) and (34):

- (33) Gong yue, shi qi sheng ye yu wu tong wu. Ming zhi yue Tong.
duke say be in-birth too and I same matter call him say Same
'The duke said: 'This one in his birth is the same substance as me.' He named him Tong (Same).' (Zuozhuan Huan 6/5)
- (34) zhi er chong you he zhi
these two worm again what know
'And what do these two worms know?' (Zhuangzi 1/10)

Both demonstratives and pronouns have a much clearer linguistic identity, functioning to pick out and refer to some other entity in the discourse and having in common the property of being referentially fully definite. Syntactically because of the definiteness

inherent in demonstratives and pronouns, both are assumed to be elements which normally occur in the D⁰ position in DPs, the formal locus of referentiality.

Demonstratives are also elements which have been frequently noted to undergo certain change over time, often developing into determiners, pronouns (as for example Latin *ille* which provided both the modern French determiner *le* (masculine) ‘the’ and the masculine pronoun *il* ‘he’) and other functional morphemes with less clear identities. It is then quite plausible that the original source of modern Chinese *de* may in fact have been this earlier demonstrative element *zhi* and that *zhi* as a demonstrative gradually gave rise to the wider distribution of *zhi* and later *de* in which the functional role and syntactic type of *zhi/de* became less immediately obvious.

If it is indeed supposed that modern Chinese *de* originated as a demonstrative in the D⁰ head of the DP and gradually lost its former simple demonstrative role, such an assumption will allow for a rather natural explanation of certain other important aspects of the distribution of *de* (and *zhi*). The first of these is the fact that *de* (and *zhi*) are the relevant elements used in some way to form relative clauses in Chinese. Although it is synchronically not fully clear what this relativizing role is with *de*, cross-linguistically it is well-attested that demonstratives or determiner-like elements developed from demonstratives do indeed frequently occur in a very similar kind of function in relative clauses, as illustrated in (35-7):³

- | | | | | | |
|------|---|-----------------------------------|--|-----------------------|------------------------------|
| (35) | tanay
yesterday | awa:
house | awu:w- pu -l
I-saw- the -in | ciyawx
I-will-sing | DIEGUENO |
| | ‘I will sing in the house I saw yesterday.’ | | | | (Keenan 1985) |
| | | | | | |
| (36) | Mary
Mary | owiza
quilt | wa
a | kage
make | ki
the |
| | he | ophewathu
I-buy | that | LHAKOTA | |
| | ‘I bought the quilt that Mary made.’ | | | | (Williamson 1987) |
| | | | | | |
| (37) | te winike
the man | te mac’a
the who | la
past | smah
hit | te Ziake
the Ziak |
| | ‘the man who Ziak hit’ | | | | TZELTAL

(Keenan 1985) |

This common cross-linguistic use of D⁰-type elements to form relative clauses therefore adds support to the hypothesis that *de* originated as a D⁰ demonstrative form *zhi* and remains as a rather different functional instantiation of D⁰.

Secondly *de* has been often typologically referred to as a nominalizing morpheme (see e.g. Paris 1979, Li & Thompson 1981). As such a term describes a particular function rather than a categorial type, the precise categorial status of nominalizers is actually not immediately clear. What the term nominalizer simply indicates is that the output of a linguistic form to which a nominalizer is applied will be somehow ‘nominal’ in its identity. Noting that both NPs and DPs are characterized as being nominal, it is possible that a nominalizer could in theory be either some kind of N⁰ which applies to an input and creates an NP, or alternatively a type of D⁰ which converts a non-nominal input form into a DP. Significantly for present purposes it is important to note that clear D⁰ determiner elements are indeed often used in a nominalization function as seen in example (38) from Spanish:

³ Note that (36) is a head-internal relative clause.

- (38) [el _{CP} que Juan haya ganado el concurso]] garantiza nuestro triunfo
the that Juan has won the competition guarantees our triumph
 ‘That Juan has won the competition guarantees our victory.’

Consequently if *de* is assumed to occur in some way as a clausal nominalizer, the observation that other clausal nominalizers may indeed be D⁰ determiners adds further plausibility to the hypothesis that *de* is a D⁰ type element.

A third important aspect of *de*'s distribution is that it occurs in a genitive-like function licensing a possessor-possessee relation (example (4) and (31) with *zhi*). Concerning genitive case, it is commonly believed that the markers of such case occur located in the D⁰ head position of a DP, this explaining the apparent complementary distribution of genitive case and determiner elements as seen in (39):

- (39) a. John's book
 b. the/that book
 c. *John's the/that book

If Chinese *de* is indeed a genitive marker, it is therefore natural for it to be in the D⁰ head of DP position from a cross-linguistic point of view. Furthermore in modern Chinese it can be noted that an interesting alternative to the use of *de* as a genitive marker with pronominal possessors is to use the *demonstrative* element *zhe* ‘this’:

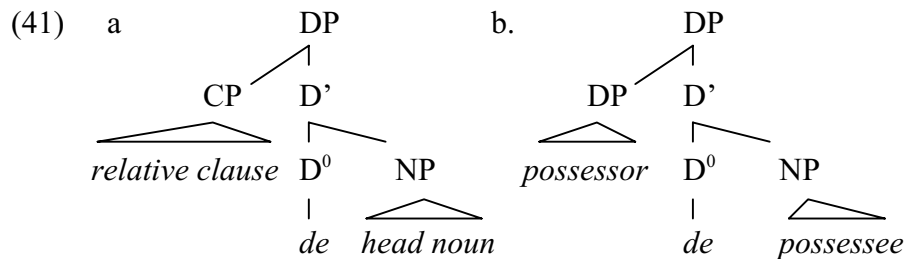
- (40) a. wo de shu b. wo zhe shu
 I DE book I this book
 ‘my book’ ‘my book’

This alternation suggests that there is indeed some special kind of link between the marking of a genitive relation and demonstratives – the element which currently marks the genitive relation *de* is arguably developed from an earlier demonstrative *zhi* and the same genitive relation can also be marked with an element which elsewhere still is a clear demonstrative.

Ultimately then despite earlier remarks that the categorial identity of *de* might be rather difficult to determine, further investigation reveals a range of evidence pointing to the conclusion that *de* has in fact developed from an earlier demonstrative and that it is consequently more than likely that *de* continues to instantiate the D⁰ position in DPs. Relative clauses and genitive constructions can therefore be suggested to have underlying representations such as those in (41):^{4 5}

⁴ Note that in Simpson (1997) and Wu (to appear) a rather more complex analysis of relative clause forms is offered in which the relative clause raises to its surface position from a Kaynean (1994) base. Note also that the structure in (41a) is essentially parallel to the structure which Murasugi (1997) proposes for Japanese.

⁵ Suggesting that *de* occurs in D⁰ we assume that demonstratives such *zhei/nei* ‘this/that’ will occur in other positions. There is abundant evidence that determiners in D⁰ do co-occur with demonstratives in DPs in many languages and that the demonstratives *in these cases* (see here also footnote 11) must be located in other non-D⁰ positions. Simpson & Wu (in preparation) suggest that the ‘doubling’ of determiners and demonstratives occurs when a determiner in D⁰ optionally selects for an emphatic focus projection and that the demonstrative is in these instances base-generated in the specifier of the focus-phrase.



Turning finally to consider Japanese *no*, the categorial status of this element is again rather difficult to determine, and Murasugi (1991) notes that *no* has been variously referred to as (a) a genitive case marker in examples such as (42), (b) a pronoun in instances such as (43), and (c) even a complementizer in Hoji's (1990) analysis of cleft-like occurrences as in (44):

- (42) Taroo no hon
 Taroo NO book
 'Taroo's book'
- (43) Hanako-wa shiroi-no-o katta
 Hanako-TOP white-NO-O bought
 'Hanako bought the white one.'
- (44) Yamada-ga atta no wa Russell da
 Yamada-NOM met NO TOP Russell be
 'It was Russell that Yamada met.'

In addition to such classifications, *no* is also commonly referred to a-categorially as a nominalizer, particularly in cases such as (45):

- (45) Hanako-wa Taroo-ga kekkon shita no-o shirimasen
 Hanako-TOP Taroo-GA wedding did NO ACC not-know
 'Hanako doesn't know that Taroo got married.'

Concerning the 'pronominal' cases of the type seen in (43), Murasugi (1991) notes that "It is widely assumed that the *no* in these examples is an N and is in the head position of the NP...Roughly speaking it is analagous to English *one*." (p.56). However, despite such a common view it is clear that *no* also shows patterns which distinguish it from other nouns and might seem to suggest that *no* is not in fact a noun/N⁰. As pointed out by Murasugi, unlike other nouns *no* cannot be modified by a demonstrative (in this sense being like Chinese *de* (28) and unlike Korean *kes* (24)):

- (46) *ano no wa omoshiroi
 that NO TOP interesting
 intended: 'That one is interesting.' (Murasugi 1991)

No can also not be modified by numerals as seen in (47):

- (47) *san bon no o katte kita
 3 CL NO ACC buying came
 intended: 'He bought three.' (Murasugi 1991)

On the other hand, citing work in Yuzawa (1944), Murasugi (1991) points out that *no* does allow for genitive modification in various dialects, which would appear to be a good argument in favour of analyzing *no* as a noun in the N⁰ head of NP:

- (48) kore-wa watashi-no-no de-wa-arimasen
 this-TOP I-NO-NO not-be
 ‘This one is not mine.’

No might then perhaps be taken to be nominal in the ‘formal noun’ sense that Korean *kes* is, like *kes* requiring obligatory modification of just a certain type. Murasugi notes that such a requirement in fact distinguishes *no* from other pronouns which do not require any modification and are free to stand alone in discourse:

- (49) Taroo-wa (hidoi) sore-o/(hidoi) no-o mita
 Taroo-TOP horrible it ACC horrible NO ACC saw
 ‘Taroo saw a horrible her/the horrible one.’

Elsewhere though, *no* seems to provide evidence of being of type D⁰ however. It has been noted above that genitive case elements are commonly analyzed as occurring in the D⁰ head of DP constituents and *no* clearly occurs in a genitive marking function. Such an aspect of its distribution should naturally lead one to assume that *no* is categorially a D⁰ element rather than an N⁰.

From a historical point of view, the origins of *no* are rather obscure. It is however well-known (see e.g. Shibatani 1990) that in classical Japanese *no* used to alternate with *ga* as a genitive-like marker linking two nominals as in (50).⁶:

- (50) wa no/ga sono ni
 I NO/GA garden in
 ‘in my garden’

Concerning the genitive *ga* here, more recently it has been speculated that this morpheme may have arisen from the demonstrative element *ka(-re)* (Lone Takeuchi p.c.). If this is indeed so, it would interestingly be another case of a demonstrative being used to encode a genitive possessor relation as in Chinese. Assuming that such a demonstrative would be of type D⁰ and that both *ga* and *no* occurred in the same position when signalling the genitive relation, this naturally leads to the further conclusion that genitive case *no* in classical Japanese would also have been of type D⁰. Given that the genitive function encoded by *no* has remained constant through to modern Japanese, this therefore suggests that *no* here still occurs in the D⁰-position, as indeed assumed in common analyses of genitive-marking.

Summarizing then, with Japanese *no* there is a variety of potentially conflicting evidence concerning its current categorial status. Some aspects of its patterning seem to support the widespread belief that that *no* is nominal and an N⁰, yet other clear restrictions such as the unavailability of demonstrative-modification (46) together with *no*’s occurrence as a genitive-marker appear to link *no* more with D⁰. Overall the broad picture which can be suggested to emerge after a fuller consideration of *de*, *no* and *kes* is that none of the paradigm appears to pattern either

⁶ Shibatani observes that the only apparent distinction between *no* and *ga* genitive-marking is that *ga* was used to mark elements which sometimes might be deemed to be less worthy of respect.

like a fully prototypical N^0 or D^0 , and that instead they seem to exhibit properties which variously associate them to a greater or lesser extent with both D^0 and N^0 types. In order to resolve this apparent conflict in categorial specification and capture what seems to be the ambivalent status of this set of elements, we will shortly develop the hypothesis that *de*, *no* and *kes* are currently in different stages of categorial re-analysis and that this ongoing grammaticalization is indeed directly responsible for the range of differences found in the *de/no/kes* paradigm. Before we proceed to this however, we first consider certain quite general problems raised by the phenomenon of grammaticalization and propose a formal approach to category change based on the simple ideas of movement and re-analysis.

3.2. Grammaticalization and the gradualness of language change

The process of grammaticalization is often taken to involve a certain loss of descriptive meaning or ‘bleaching’ of semantic content in a linguistic element and also frequently a change in syntactic category type. Such processes are generally described as being critically gradual and giving rise to elements which share the properties of both the original source and the ‘target’, with there being an incremental increase in those of the target end-point over time.

In typological descriptions such an incremental change in properties associated with two developmental points is often referred to by means of a *cline* with a feature matrix plotting the various properties which characterise the two discrete end-points of the developmental chain and how these properties may be realized in intermediate stages of development. A good example of this is Comrie’s (1981) treatment of numbers in Russian, which may be classified as being either more noun-like or more adjective-like according to how they pattern with regard to certain syntactic phenomena. When a variety of properties are compared, it can be seen that lower numbers are much more adjective-like, whereas further up the cline numbers are found to exhibit more-noun-like attributes. This is charted in the table in (51) below. For reasons of space we do not detail the specifics of the distributional properties (a-g), broadly these relate to differences in case, gender and number-marking; the interested reader is invited to refer to Comrie (1991, 107-110) for a fuller discussion:

(51) Numbers in Russian – Adjectives or Nouns?

Property	<i>odin</i>	<i>dva</i>	<i>tri</i>	<i>pjat</i>	<i>sto</i>	<i>tysjaca</i>	<i>million</i>
	1	2	3	5	100	1,000	1,000,000
(a)	A	N	N	N	N	N	N
(b)	A	N	N	N	N	N	N
(c)	A	(A)	N	N	N	N	N
(d)	A	A/(N)	A/(N)	N	N	N	N
(e)	A	A	A	A	(N)	N	N
(f)	A	A	A	A	A	N	N
(g)	A	A	A	A	A	A/N	N

In transformational-type approaches to language, the use of such clines and feature matrices is viewed as an unsatisfactory way of modelling linguistic change, as synchronically it is argued there should always be a fully discrete categorization of an element as being either of one type or another (e.g. either a noun or an adjective) and not, for example, 30% like an adjective and 70% like a noun as allowed in

cline/feature matrix approaches. Generally then, whereas the final endpoints of grammaticalization processes are relatively easy to describe, the intermediate stages where elements appear to have an unfixed and dual status prove to be more challenging as they involve a greater degree of complication.

In a Chomskyan-style phrase-structure grammar the assumption that lexical heads are dominated by an array of functional heads allows for a potential solution to the problems created by certain cline effects. Essentially it is possible to suggest that a set of properties which otherwise might be associated exclusively with either of the two endpoints in a grammaticalization chain alone may instead be spread over a number of functional heads occurring between the end-points of the cline, and that the presence or absence of certain functional-heads will correlate with the change in the properties attested. For example, *gerunds* present a classic case where verbal and nominal properties are considered to be mixed up together to a greater or lesser degree in different instances resulting in the classification of gerunds as being part noun and part verb-like. In instances where the English gerund is more noun-like and licenses genitive case on its thematic subject, it will only support adjectival not adverbial modification and license its object with *of*-insertion as in (52). Where the gerund behaves more like a verbal element it does license adverbial modification, bare accusative case on its object and a subject which is not marked with genitive *'s* as in (53):

(52) [Bill's careful/*carefully riding *(of) his bicycle] impressed us.

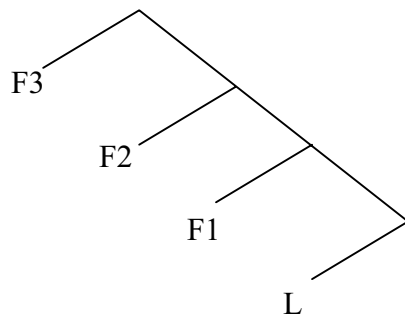
(53) [Bill carefully/*careful riding (*of) his bicycle] was a thing to behold.

Assuming case assignment and adjective/adverbial modification to be dependent on the presence of certain functional heads associated with either an N^0 -head or a V^0 -head it is possible to give a principled analysis of these and other 'intermediate' patterns found with gerunds with the suggestion that both an N^0 head and a V^0 head are present in gerund structures and that variation relates to the functional heads which these N^0 and V^0 heads project in any instance. Such an approach avoids the conclusion that a single lexical element (the gerund) is itself both verbal and nominal to a varying degree and allows for the ambivalent status of certain lexical elements to be formally attributed to the functional structure which may be projected in greater or lesser amounts.⁷

Here quite generally, we would now like to suggest that one rather simple way of understanding certain processes of grammaticalization is to suggest that various instances of category change may occur when a lexical element is re-analyzed as being base-generated in a higher position in the functional structure projected by the lexical head. Take, for example, a representative lexical head *L* which projects a functional structure consisting of three heads *F1-F3* as in (54). The head *L* here could be understood to be a verb which projects an array of agreement, aspectual and tense heads, or perhaps a noun projecting a set of nominal functional heads up to the DP-level:

⁷ See also Lapointe & Nielsen (1994) for an interesting treatment of the differences in type II and type III gerunds in Korean along these general lines. The availability of scrambling in the former but not the latter is taken to be directly related to the presence of an overtly instantiated tense projection in type II gerunds but not type III gerunds. Type II gerunds are therefore generally more clause-like/verbal than type III gerunds due to the presence of an additional functional category projected by the verb.

(54)



In such a structure the head positions F1-F3 might perhaps be occupied by discrete functional morphemes base-generated in these positions. However it is also possible that the lexical head L carries a set of (possibly null) affixes and raises up to these heads for licensing of the morphological features in the affixes. If such a head L raising up to F1-F3 then gradually becomes more associated with the functional properties of these higher heads and simultaneously undergoes a loss of descriptive content, this may ultimately cause the lexical head to be re-analyzed as being directly base-generated in the functional structure rather than originating in L. If F1-F3 are furthermore labelled as being distinct category types, then the end result is that the lexical head L will undergo a formal category change.

A concrete illustration of this process can be suggested to be the historical creation of the French negative morpheme *pas*. Originally negation in French was expressed just by a pre-verbal negative element *ne*. Later on this changed somewhat and the pre-posing of certain verbal objects to a non-canonical object position preceding the past participle came to signal *emphatic* negation (see Hopper & Traugott 1997, p.58). Among these emphatic re-inforcers of negation was the element *pas* literally meaning ‘a pace’ or ‘a step’ and occurring naturally as the object of a variety of verbs of motion such as ‘walk’, ‘run’ etc. The effect of emphatic negation with *pas* was very similar to that in the English example (56):

(55) Je ne suis pas alle
I NEG am step went
‘I didn’t go (at all).’

(56) I didn’t walk a STEP!

With other types of verbs different appropriate objects would be used, giving rise to forms similar to modern English emphatic negation in (57):

(57) I didn’t drink/touch a DROP!

From amongst the wide range of objects used as emphatic re-inforcers of negation in early French, *pas* later became much more common with negation and having lost its original purely literal meaning of ‘step’ *pas* came to be used with verbs which had no connection at all with walking or running or other actions involving ‘steps’ as potentially genuine objects. An important bi-product of this process of historical change is that *pas* then came to be used to express negation with transitive verbs which had discrete overt objects as in (58):

The possibility of explaining the clustering of properties and developmental changes with *pas* in this way thus arguably supports the suggestion that grammaticalization may indeed consist in a sequence of movement and subsequent re-analysis in a functional projection targeted by a lexical element. Here *pas* is originally a lexical object which begins to undergo raising to the pre-verbal SpecFocP when it signals emphatic negation. Significantly later on after continued association of *pas* with negation and focus it is re-analyzed as being only a functional element and as therefore being base-generated in the functional projection itself. Such re-analysis importantly frees up the object position in the VP and a genuine thematic object can then be base-generated in this lexical position, leading to the co-occurrence of *pas* with overt lexical objects. In the course of this re-analysis there are three significant stages. In stage 1 *pas* is a genuine lexical object of verbs of walking and running. In stage 2 *pas* is the object of these and other verbs but also functions as an emphatic re-enforcer of negation raising to the pre-participial focus position. During this part of the grammaticalization process *pas* is critically ambivalent in its status, occupying both the object-of-verb position and the focus-position and consequently having properties of both such positions. In stage 3 *pas* is fully re-analyzed and grammaticalized in the higher position no longer being an NP verbal object but categorially being simply a marker of emphatic negation;⁸ it therefore loses all the properties of its original source position and allows for a new element to occur in this position.

Elsewhere other common occurrences of grammaticalization also seem to significantly proceed in an ‘upward’ direction in syntactic structure. For example it is frequently found that lexical verbs undergo re-analysis into auxiliary-verb elements, which is clearly a development in which a lexical V^0 first raises up to a higher Infl-type head from V^0 and then is later simply base-generated in the higher functional head.⁹ In support of the movement-based approach to grammaticalization proposed here it can furthermore be noted that changes going in a ‘downward’ direction do not seem to take place. For example, whereas root modals which are canonically situated low down in the functional structure dominating VP commonly turn into structurally higher epistemic modals with sentential scope, the opposite type of change of epistemic modal into root modal is critically not found (see e.g. Bybee, Perkins & Pagliuca 1994). Importantly it would seem then that change and re-analysis consistently occurs in an upwards direction, mirroring the direction of syntactic movement and that grammaticalization may often involve the simple re-analysis of an element in a position it has been previously raised to.¹⁰

Having spent a little time attempting to justify a movement and re-analysis approach to at least certain occurrences of grammaticalization and having also suggested that the gradual nature of language change and the ambivalent status of elements undergoing change may critically involve changes relating to the functional structure dominating a lexical projection, we now turn back to the *de/no/kes* paradigm and argue that such a general approach to grammaticalization is able to capture both the similarities found with *de*, *no* and *kes* as well as their important differences.

⁸ Note that the emphatic force originally associated with the use of *pas* has now disappeared in modern French. This has been accompanied by the virtual disappearance of the higher marker of negation *ne* from colloquial French and a re-interpretation of *pas* as the primary marker of negation.

⁹ Such a change and re-analysis is indeed assumed to have occurred in the ‘creation’ of the modal verb series in English.

¹⁰ Note that since the initiation and development of the approach to grammaticalization detailed here, similar ideas have also been independently suggested and applied to rather different data in Roberts & Roussou (1999).

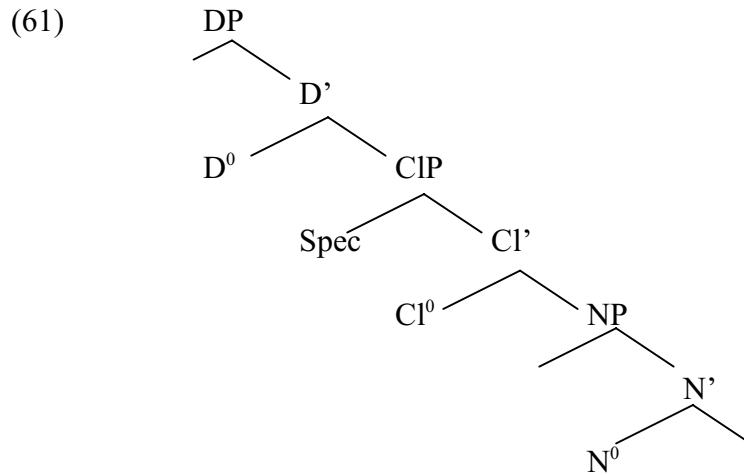
3.3. The grammaticalization of *de*, *no* and *kes*

Having pointed out in section 2 that *de*, *no* and *kes* show a tremendous similarity in their distribution, we have taken this as clear indication that they should *in some sense* all be assumed to be of the same basic linguistic type. Having also noted a number of differences in their fuller patterning, we will now suggest that these differences are critically due to *de*, *no* and *kes* instantiating different stages of an essentially unitary route of grammaticalization and development and show that the approach to grammaticalization sketched out in the preceding section allows for a successful formal model of the hypothesized changes which broadly captures both the similarities and the differences in the paradigm.

Reflecting on the potential categorial identity of *de/no/kes* in section 3.1, it seemed that there is evidence linking these elements to both the category N/noun and the category D/Determiner, though synchronically none of the *de/no/kes* set can be taken to instantiate a fully prototypical member of either N/D type. Individually, Chinese *de* was argued to have a clear association with type D, being most probably a former demonstrative bleached of its earlier definiteness and now fulfilling the genitive D-function in addition to its occurrence in other structures. Korean *kes* by way of contrast seemed to be more like a noun/an N, though again not prototypically so, being loosely referred to as a formal noun subject to a variety of restrictions which do not hold of regular nouns/Ns. Finally Japanese *no* was seen to show characteristics somewhere in the middle of *de* and *kes* and have properties linking *no* to both (formal) N and D types. It is this arguably cline-like property of the *de/no/kes* set which leads us to suggest now that *de/no/kes* simply lie at different developmental points in a single grammaticalization chain in which nouns/Ns over time become re-analyzed as D-type elements. Below we will basically argue that *kes* with its clearer noun/N-like characteristics occurs at a lower point in the re-analysis, *no* with its very mixed properties more ambivalently in the middle, and *de* at the endpoint of the grammaticalization as a D^0 element, these different statuses corresponding to the different syntactic behaviour found with *de*, *no* and *kes*.

Formally, in Chomskyan-inspired studies carried out on the internal syntax of NPs in recent years it has been suggested that NPs in fact project a variety of higher functional structure, just as VPs do in the clausal domain. In addition to the assumption that non-predicative referential NPs occur as DPs headed by a D^0 position encoding their potential referential value, a variety of other intervening functional projections have also been proposed, for example NumP hosting numerals, and CIP hosting classifiers in languages such as Chinese, Japanese and Korean. For present purposes and the representation of DPs in Chinese, Japanese and Korean, we will assume the presence of just one additional DP-internal functional projection (though there could in principle be others) and take this to be CIP, hosting a DP's numeral and classifier unit in its $C1^0$ head. (61) below therefore represents what is taken to be the underlying form of DPs in all three languages:¹¹

¹¹ Classifiers and their associated numerals do of course commonly occur following the DP/NP in surface forms in Japanese and Korean. We assume that this results from raising of the NP constituent in (61) to some higher Spec position, possibly SpecCIP.



Above in section 3.2 it was suggested that grammaticalization and the apparently gradual nature of certain language change is very often due to the changing relation of a lexical head to the functional structure it projects. Having assumed the minimal set of functional projections CIP and DP dominating the lexical head N^0 and its projection NP in (61), we will now attempt to show how such a basic structure allows for a simple account of the distributional differences between *de*, *no* and *kes*.

From the patterns examined here there are essentially four aspects of the *de/no/kes* paradigm which stand out as being significant and in need of some explanation. These are first noted in (62) and will subsequently be considered each in turn:

(62) **Pattern 1**

Japanese *no* and Korean *kes* do not occur in adult headed relative clauses.

Pattern 2

Korean *kes* does not occur as a genitive marker.

Pattern 3

Japanese *no* and Korean *kes* do not allow for modification by numerals and classifiers.

Pattern 4

Japanese *no* and Korean *kes* do occur in children's headed relative clauses.

Pattern 1 is very significant. If one considers the table of distribution in (23) repeated below for convenience, this aspect of the distribution of *no* and *kes* really stands out as being exceptional. Whereas there appears to be a strikingly similar patterning shared by *de*, *no* and *kes* in just about all other construction types, in this particular instance *no* and *kes* seem to go against the general trend towards uniformity in two obvious ways. First of all *no* and *kes* are clearly different from *de* here, which does occur in adult headed relative clauses. Secondly, *no* and *kes* have been observed to occur in other relative clause types such as children's headed relative clauses, headless relative clauses, internally-headed relative clauses and also cleft-type structures based on relative clause forms, so the absence of *no* and *kes* from adult headed relative clauses seems to be in strong contrast with the legitimate occurrence of the same elements in other similar environments in Japanese and Korean.

(23)

	relative clauses				pseudo-cleft	cleft/copula	genitive	clausal 'nominalizer' function
	adult headed	child headed	head-internal	headless				
Chinese	✓ (1)	✓ (1)	✓(children) (3)	✓ (2)	✓ (6)	✓ (5)	✓ (4)	
Japanese	Ψ (10)	✓ (13)	✓ (12)	✓ (11)	✓ (9)	✓ (8)	✓ (7)	✓ (14)
Korean	Ψ (15)	✓ (18)	✓ (17)	✓ (16)	✓ (19)	✓ (20)	Ψ (22)	✓ (21)

We would like to suggest that there is in fact a rather simple way of reconciling the fact that *no* and *kes* elsewhere pattern with *de* but in this one instance seem to be exceptional. What is really critical and identifies the adult headed relative clause case as being significantly different from the other constructions where *no* and *kes* legitimately occur in adult speech is that in all the other adult constructions, with the sole exception of *no*'s occurrence as a genitive marker, *no* and *kes* are final in the DP and are not followed by any overt nominal element/NP. This can be verified as true of head-internal and headless relative clauses, the cleft constructions and also clausal nominalizations formed with *no* and *kes*. Where a token of *no* or *kes* is however accompanied by a following head-noun/NP in adult speech as in examples (10) and (15) this is automatically unacceptable. Making use of this simple observation and linking it to the conclusion that *no* and *kes* are frequently assumed to have nominal properties allows one to begin to deconstruct the differences in the paradigm. Assuming a basic DP structure such as in (61) we suggest that a first important difference between *de* on the one hand and *no* and *kes* on the other is that *de* is an element base-generated in the D⁰-position whereas *no* and *kes* are inserted into the N⁰ position in the DP. Because *de* is directly inserted into D⁰ this allows for the N⁰ position to be overtly instantiated in Chinese and *de* regularly co-occurs with overt head-nouns in (headed) relative clauses. In Japanese and Korean however, as *no* and *kes* are base-generated in N⁰ this will block the occurrence of any other overt head-noun and *no* and *kes* will be restricted to occurring only in those constructions with no overt noun following *no/kes*. Such a simple assumption constitutes a first step in accounting for the wider paradigm and captures the basic differences between *de* and *no/kes* in headed relative clauses together with the fact that *no/kes* are fully acceptable in other relative clause structures (because in other adult relative clauses there is no overt N⁰ following *no* and *no* is itself base-generated in N⁰).

Turning now to the second pattern, that Korean *kes* does not occur as a genitive-like marker, this contrasts importantly with Japanese *no* and Chinese *de* which both occur regularly as markers of a genitive relation. In this possibility then Japanese *no* interestingly seems to pattern with Chinese *de* rather than Korean *kes*. We suggest that this fact about *no* indicates that while *no* can be assumed to be base-generated in N⁰ it also subsequently raises up to D⁰ during the course of the derivation. Assuming that a genitive relation is signalled by an element in the D⁰-position, we suggest that regular N⁰-to-D⁰ raising has furthermore allowed for *no* to be re-analyzed as also being base-generated directly in the D⁰ position. In other words, we suggest that in non-genitive constructions *no* is initially inserted into N⁰ and then raised up to D⁰, and that such movement to D⁰ has allowed for *no* to be re-analyzed as optionally

occurring also base-generated directly in D⁰ and instantiating just a genitive relation (and in these cases the N⁰ position will frequently be filled by a discrete nominal element, the possessee). Because Korean *kes* does not allow for interpretation as a genitive-marker in contrast to *de* and *no*, we suggest this indicates that *kes* simply does not raise up to the D⁰-position. This hypothesized difference between *no* and *kes* that the former raises to D⁰ whereas the latter does not also provides a potentially straightforward explanation of a further related difference, the fact that *kes* allows for modification by a demonstrative (example (24)) whereas *no* does not (example (46)). If demonstratives are assumed to normally occur in D⁰ (when there is no other overt D-element, see footnote 5 and footnote 11), the fact that they cannot occur with *no* is clearly explained if *no* also raises to occupy the D⁰-position; the legitimate possibility of *kes* co-occurring with a demonstrative is similarly explained on the assumption that *kes* does *not* raise up to the D⁰ position. Finally, it can be remarked that Chinese *de* which is always assumed to be base-generated in D⁰ naturally allows for use as a genitive marker and as expected does not allow for direct modification by a demonstrative (example (28)).¹²

The proposed differences between *de*, *no* and *kes* outlined so far can be usefully compared with the three principal stages in grammaticalization hypothesized to have occurred with French *pas*. Korean *kes* which might seem to be the least grammaticalized of the three elements in the paradigm is inserted into N⁰ the lexical head, this corresponding to stage 1 with *pas* where *pas* is inserted as the lexical object of verbs of walking/running. Japanese *no* in adult relative clauses is regularly inserted into N⁰ and then raises up to D⁰ thus having an ambivalent status and occupying two discrete positions during the course of the derivation, blocking use of these positions by other elements (overt head nouns and demonstratives). Here Japanese *no* is like *pas* in stage 2, *pas* being base-generated as the lexical object of a verb and then raising to a higher functional head to encode emphatic focus. In stage 2 *pas* like *no* has a dual status, instantiating both direct object and focus-specifier positions. Japanese genitive *no* which is taken to be fully grammaticalized and re-analyzed in D⁰ can be said to correspond to the final stage 3 of the grammaticalization process. In stage 3 *pas* is re-analyzed as being base-generated directly in the focus-specifier and therefore allows for new direct objects to be inserted in the object-of-verb position. Similarly *no* re-analyzed fully as genitive in D⁰ allows for overt nouns to be inserted

¹² Something additional has to be said about cases where *de* co-occurs with a demonstrative in a lower position in relative clause structures such as (i):

- (i) [wo mai]-de nei-ben shu
 I buy DE that-CL-book
 ‘that book I bought’

Here we believe that the demonstrative is optionally introduced into the structure in a lower emphatic focus-type position (actually rather similar to the emphatic re-inforcement with French negative *pas*). Such doubling of a D⁰ element and an emphatic demonstrative is found in a number of languages, e.g. Spanish:

- (ii) el hombre este
 the man this
 ‘this man’

In normal non-emphatic instances the demonstrative occurs simply in the initial D⁰-position:

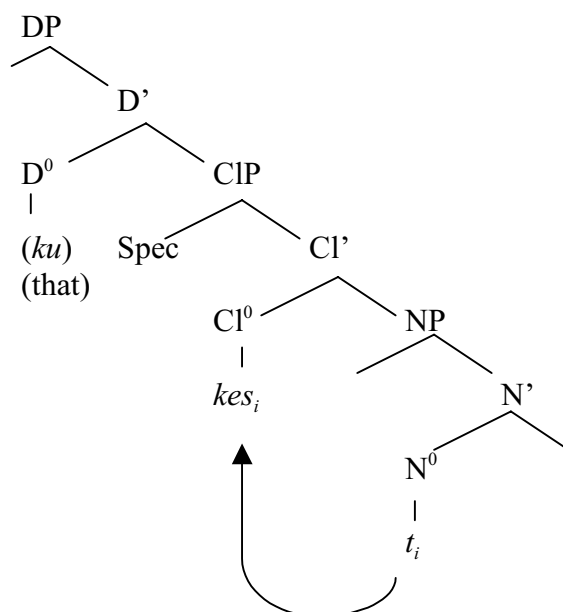
- (iii) este hombre
 this man

We therefore continue to assume that the normal position for demonstratives to be inserted into is indeed the D⁰-position, but that in certain instances it may be possible for a demonstrative to be inserted in a lower focus-position doubling an element already in D⁰. See Simpson & Wu (in preparation) for further discussion of this phenomenon.

in the N^0 position and encode possessives. Finally Chinese *de* can also be suggested to represent the same endpoint in the grammaticalization chain being currently directly base-generated in the higher functional head position.¹³

Continuing on to pattern 3 and the fact that Japanese *no* and Korean *kes* do not allow for modification by numerals and classifiers, this patterning actually does not indicate any immediate difference between *de*, *no* and *kes* as all three of these elements have been seen to disallow numeral-classifier modification (examples (27), (29) and (47)). However, the patterning can be argued to reveal more about the synchronic status of Korean *kes*. Earlier in section 3 it has been noted that there seems to be a clear intuition that *kes* is not just a regular noun and that various restrictions on its use indicate *kes* is undergoing some kind of grammaticalization as a ‘formal noun’. In the present approach, grammaticalization is taken to be (most commonly) a redefinition of the relation existing between a lexical element and the functional projections it is related to, and frequently this is argued to involve movement of a lexical element into higher positions in the functional superstructure projected. Here we would like to suggest that pattern 3 and the restriction that Korean *kes* cannot be modified by a classifier and numerals is a direct reflection of such grammaticalization and is due to movement of *kes* to the functional head position normally occupied by the classifier+numeral unit, CI^0 the head of CIP as illustrated below:

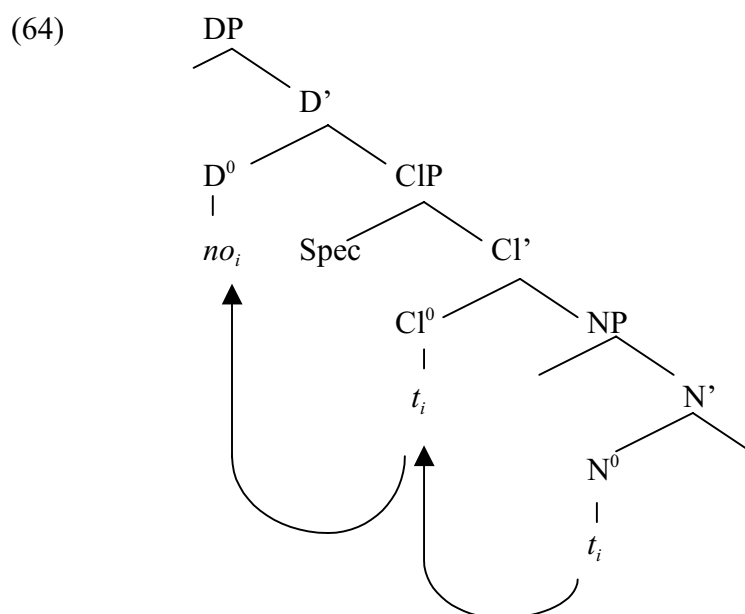
(63):



Because *kes* is grammaticalizing and raising up to the CI^0 position, this blocks insertion of any numeral-classifier pair and numerical/classifier modification of *kes* is simply unavailable as an option. As *kes* only raises as far as CI^0 however and not up to D^0 , it is still possible for *kes* to be modified by demonstratives inserted into D^0 unlike in Japanese where *no* raises to D^0 and therefore cannot co-occur modified by a simple demonstrative (example (46)). The impossibility of numerical-classifier modification with Japanese *no* is also fully expected and naturally accounted for in such an approach – as *no* is taken to head-raise from N^0 to D^0 it will have to raise

¹³ Due to a lack of evidence on how Chinese *zhi* originally came to be used as a demonstrative it is not possible to argue with any certainty that *de/zhi* necessarily grammaticalized in the D^0 -position after raising from a lower N^0 position. However, this is what the present approach might naturally expect.

through the intervening CI^0 position and this will only be possible if this position is empty as in (63). Consequently it will not be possible for any number-classifier unit to occur with *no*:^{14 15}



Finally there is pattern 4, which interestingly shows that Japanese and Korean children allow for *no* and *kes* to occur in *headed* relative clauses in addition to the other environments, unlike adult speakers (but like speakers of Chinese). Having argued for the structures and underlying derivations in (63) and (64), the account developed here now has available a very straightforward explanation of this fact. Quite simply we suggest that evidence from restrictions on the occurrence of *no* and *kes* first leads Japanese and Korean children to the adult analyses in (63) and (64) and then Japanese and Korean children independently take the natural grammaticalization cycle one stage further and allow for *no* and *kes* to be directly base-generated in CI^0 (or perhaps even D^0 in Japanese). Such a re-analysis will automatically allow for a discrete overt noun-head to be inserted into the N^0 position and so result in the occurrence of *no* and *kes* in relative clauses where *no* and *kes* are followed by an overt head-noun as in (13) and (18). Subsequently later on when children continue to fail to hear *kes* in adult headed relative clauses they can be suggested to re-adjust their assumptions about *no* and *kes* and re-analyze *no/kes* as instead being inserted into N^0

¹⁴ One might wonder what relation *kes* holds to the CI^0 head and why it should occur in such a position. Here we would like to note that functional heads are commonly assumed to be potentially associated with both positive and negative specifications, hence T^0/I^0 may be instantiated by a +finite tense element or alternatively T^0/I^0 can be occupied by the negative/opposite value –finite, in English this being realized by insertion of the infinitive marker ‘to’. If CI^0 is essentially the position in DPs where the individuation and counting of a noun is effected and made possible, we suggest that there may naturally be a negative/opposite specification and value associated with the CI^0 head ‘non-count’ and that it is just such a specification which *kes* instantiates in CI^0 (and which also attracts it to such a position).

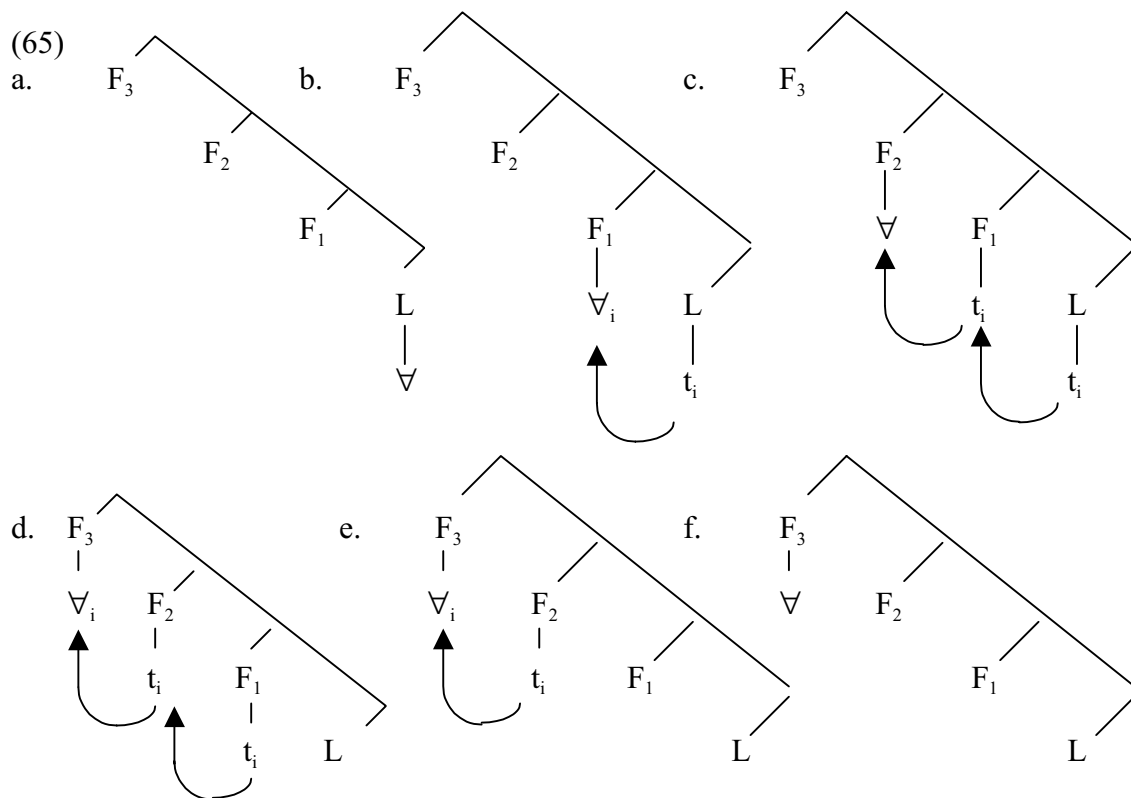
¹⁵ Supposing one were to suggest that numbers are inserted into SpecCIP rather than combined as a unit with the classifier in the CI^0 head, it might perhaps be expected that they would co-occur with *kes* when it is raised into CI^0 . However, against this it can be noted that it is rather natural to assume that the presence of a classifier is necessary for any individuation and counting and that if *kes* blocks the insertion of a classifier this will therefore also disallow the possibility of any numerical modification.

and then raised to CI^0 (or D^0), conforming to the patterns found in adult speech. During the pre-adult period though, the tendency for children to apparently analyze *no* and *kes* as purely functional elements and not necessarily as nouns in N^0 arguably further supports the view that *no* and *kes* are indeed already involved in a grammaticalization chain relating them to the functional structure above N^0 , and it is critically the established occurrence of N^0 -to- CI^0 raising which allows for children to “mis”-analyze *no/kes* as being base-generated in the functional structure dominating N^0 rather than in the N^0 position itself.

Ultimately then it has been possible to provide a principled explanation of all four of the significant patterns noted in (62) and also various other differences in the distribution of *de*, *no* and *kes* in a way which has furthermore established clear and consistent connections between the relevant properties. What the account developed here had to assume was a fairly minimal DP functional structure consisting of a determiner and a classifier position and the proposal in section 3.2 that grammaticalization is most commonly a change in the relation of a lexical element to the functional structure it projects. Considering *de*, *no* and *kes* we have basically argued that these elements represent different points in a natural grammaticalization cycle in which lexical nouns over time undergo re-analysis as instantiations of higher functional categories. Korean *kes* has been argued to be a noun/ N^0 which is beginning the grammaticalization cycle and raising to instantiate CI^0 as diagrammed in (63), this accounting for the unavailability of numeral-classifier modification, the temporary occurrence of *kes* in children’s headed relative clauses and the continued availability of demonstrative modification. Japanese *no* was argued to be regularly base-generated in N^0 and then raised through CI^0 to D^0 as in (64). This accounted for the unavailability of both numeral-classifier modification and demonstrative modification (vs. *kes*), and for the re-analysis of *no* fully in the D^0 position as a genitive marker permitting a following overt noun possessee (unlike Korean *kes* which does not raise to D^0 and so has not allowed for possible re-analysis as a genitive marker). The fact that both *kes* and (non-genitive) *no* are regularly base-generated in N^0 accounted for the major difference between *no/kes* and *de* that the former do not occur in adult headed relative clauses followed by a head-noun, even though they are found elsewhere in other relative clause types. Finally *de* was argued to be base-generated directly in D^0 accounting for its occurrence in all relative clauses and as a genitive marker. Although it is not possible to show that *de* reached such a position via movement and re-analysis from a lower lexical position, simple base-generation in the the D^0 -position does represent the natural endpoint of the grammaticalization cycle and one instantiated by genitive *no* which has allowed for full re-analysis in D^0 .

An important part of the present approach to grammaticalization and variation is the way that the apparently *gradual* nature of change is accommodated in the model. Essentially it has been suggested that surface variation occurs when the relations of a lexical element to the functional structure dominating it undergo some kind of change, with this frequently being argued to involve actual *movement* of the lexical element into higher positions in the functional structure. Critically such movement will result in a single element synchronically being associated with multiple syntactic positions and the properties which are associated with those positions. Concretely, the hypothesized raising of Japanese *no* through CI^0 to D^0 has the result that it cannot be modified by demonstratives and numeral-classifier pairs in the way that other nominal elements can and so *no* appears to be ‘less nominal’ than other elements of type N. Korean *kes* by way of contrast is assumed to raise only to CI^0 and so displays a patterning which gives it the appearance of being ‘more nominal’ than Japanese *no*

(as it can be simply modified by demonstratives), but ‘less nominal’ than regular nouns. Chinese *de* as the third member of the paradigm having such a similar wider distribution is argued to be base-generated in D^0 and because of this *de* displays a patterning which results in it appearing significantly less nominal than either *no* or *kes*. The apparent differences here between elements which seem nominal/noun-like in varying degrees is therefore a direct and simple result of the different relations they have to the same basic DP functional structure, and the gradualness of variation in a paradigm of similar elements results from the potential association of an element with more than a single structural position. Rather more abstractly, given for example a lexical head L and a functional structure consisting of three heads F_1 - F_3 , the grammaticalization of a lexical element ∇ originally of category L into an element of type F_3 via movement and re-analysis can proceed in a number of stages as illustrated in (65). In each of these developmental stages the element ∇ will have a different set of characteristics corresponding to the set of lexical and functional heads it instantiates during the course of the derivation, and each stage of development will result in ∇ seeming to be less and less like a prototypical element of type L . At the end of the cycle ∇ will be fully re-analyzed as being base-generated in F_3 and instantiate (only) this new categorial type. The apparently gradual nature of the change from category L to category F_3 is therefore suggested to be the simple result of a series of discrete smaller changes and stages in which a single lexical item may formally represent a number of adjacent functional head positions:

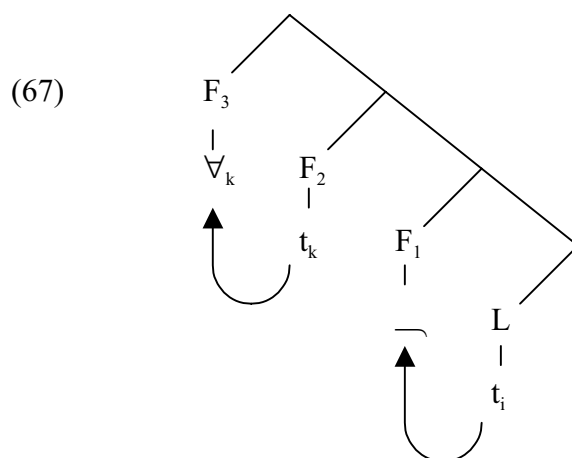


The changes in (65a-f) are tabulated in (66):

(66) Stage \forall instantiates and has properties of:¹⁶

- (a) L
- (b) L, F₁
- (c) L, F₁, F₂
- (d) F₁, F₂, F₃
- (e) F₂, F₃
- (f) F₃

Such a view of grammaticalization as sketched out here contrasts with more traditional treatments which do not assume any kind of functional structure associated with lexical elements and which have difficulty in capturing in any precise and insightful way the gradual and incremental route in which lexical elements become less prototypical instantiations of their lexical category. It should also be emphasized here that the functional heads posited as part of the present approach to grammaticalization are not just metaphorical points in an abstract grammaticalization chain but actually physical positions in syntactic structure which can duly be filled with discrete lexical elements. This strong commitment to the existence of a functional structure containing multiple positions now leads to an interesting prediction. Although we have here considered cases here in which a single formal noun incrementally undergoes successive stages of re-analysis to eventually become an instantiation of the higher D⁰ position, given the hypothetical existence of multiple functional positions above N⁰, one might also expect to find cases where there is actually more than just a single element undergoing grammaticalization in the same functional structure. Specifically, in cases such as (65e) where a nominal element \forall has grammaticalized as far as F_{2,3} one might expect to find instances where there is a second nominal element undergoing grammaticalization and raising from the N⁰ head L to the immediately higher head F₁ as diagrammed in (67). If grammaticalization critically occurs when an element raises to instantiate a higher position in the functional structure and if there are multiple positions in a single functional structure, then clearly there is the possibility that two elements might simultaneously occur grammaticalizing in the same functional string:



¹⁶ Note that wherever \forall does not occur in any of the heads L-F_{1,3} some other element will be inserted into this position. Thus in (d-f) a new lexical head will be base-generated in L, and in (a-c) there will be an independent head in F₃.

- (72) [zure egunero berandu hel-**tze**]-**a** parkaezina da
 you-Gen every.day late arrive-NZR-Det unforiveable be
 ‘Your arriving late every day is unforgiveable.’

Finally in Lilloet Salish there are again elements which have been labelled as determiners but which also have no fixed definiteness value, allowing for either a definite or indefinite interpretation, as illustrated in (73) and (74):

- (73) taxp-min-lkan [**ti**-puk-**a**] lkunsa LILLOET SALISH
 buy-1SG.NOM DET-book-DET today
 ‘I bought a/the book today.’

- (74) xaz **kw**-s-taxp-s [**kw**-puk] LILLOET SALISH
 not DET-NZR-buy-3rd.SG.GEN DET-book
 ‘She didn’t buy a book.’

These elements as in Basque co-occur with nominalizers. Consequently it might seem that once again there are two elements which have both undergone grammaticalization co-occurring in the same nominal functional structure, a ‘determiner’ element in D⁰ whose role is rather unclear and a second element grammaticalized into a nominalizer in a lower DP-internal position:

- (75) ama [**t(i)** – **s** – shiq – sw – **a**]
 good DET.NZR arrive 2SG.GEN DET
 ‘Its good that you came.’

- (76) xaz [**kw** –n –**s** shiq]
 not DET-1st.SG.GEN.NZR arrive
 ‘I didn’t arrive.’

Patterns of this kind then arguably illustrate what has been suggested to be the case with *de*, *no* and *kes* – *de* is argued to be a heavily bleached nominalizer which has undergone full re-analysis in the D⁰-position but encodes no definiteness value, whereas *no* and *kes* are suggested to be nominalizing elements in a less advanced state of grammaticalization occurring base-generated lower down in N⁰ and subsequently raised higher into the nominal functional structure. In Lotha one finds two nominalizers in various stages of grammaticalization able to co-occur and occupying different positions in the nominal functional structure, whereas in Basque and Salish there are fully grammaticalized elements in D⁰ with no apparent definiteness value being accompanied by formal nominalizers in lower DP-internal functional positions.

4. Concluding remarks

Summarizing now, this paper began by noting that the functional elements *de*, *no* and *kes* in Chinese, Japanese and Korean show a remarkable similarity in their patterning across a whole range of construction types suggesting that these elements should in some sense be assumed to be basically of the same type. The existence of certain other distributional differences however also indicated that the syntactic properties of *de*, *no* and *kes* are actually not entirely similar. The paper therefore set out to confront the challenge posed by the *de/no/kes* paradigm and construct an analysis which both

recognises the tremendous similarity in the patterning of these elements and at the same time accounts for the variation they exhibit in a principled way. Suggesting that the differences between *de*, *no* and *kes* may be due to these elements being at different stages of grammaticalisation, we introduced a general approach to grammaticalization which is able to model the apparently gradual nature of syntactic change in a formally precise way and then applied this to four salient patterns and differences in the *de/no/kes* paradigm. Assuming the existence of a fairly minimal functional structure internal to the DP in Chinese, Japanese and Korean, the approach was shown to allow for a fully integrated account which links the various differences in a coherent way and suggests that variation amongst the *de/no/kes* set is due to the existence of changes in the relations *de*, *no* and *kes* have to the same basic nominal structure. Critical in the analysis developed was the suggestion that elements undergoing categorial change may synchronically instantiate more than just a single position in a functional structure, and that it is this possibility combined with the occurrence of multiple positions in certain functional structures which results in appearances of gradual rather than catastrophic instantaneous change. Finally it was remarked that certain patterns of multiple grammaticalization found within a single nominal domain seem to resemble changes argued to be occurring with *de*, *no* and *kes* and strengthen the assumption that the functional structure supporting the ‘gradual’ steps of grammaticalization is indeed physically instantiated rather than being just a purely abstract notion.

References

- Bybee, Joan, Perkins, Revere & William Pagliuca. (1994) *The Evolution of Grammar*. University of Chicago Press, Chicago & London.
- Chiu, Bonnie. (1998) Relative Clauses in Child Chinese. manuscript, National Taiwan University.
- Comrie, Bernard. (1989) *Language universals and linguistic typology*. Blackwell, Oxford.
- Davis, Henry & Lisa Matthewson. (1998) Determiners, finiteness and the entity/event distinction. in *Proceedings of NELS 28*, 95-110. edited by N. Tamanji & K. Kusumoto, GLSA, U.Mass.
- Herring, Susan. (1991) Nominalization, relativization, and attribution in Lotha, Angami and Burmese. *Linguistics of the Tibeto-Burman Area 14:1*, 55-72.
- Hoji, Hajime. (1990) Sloppy identity in Japanese. manuscript, University of Southern California.
- Hopper, Paul & Elizabeth Traugott. (1997) *Grammaticalization*. Cambridge University Press
- Kayne, Richard. (1994) *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.
- Keenan, Edward. (1985) Relative Clauses. in: *Language Typology and Syntactic Description*. Vol 2: *Complex Constructions*, 141-170, ed: T.Shopen. Cambridge.
- Kitagawa, Chisato & Claudia Ross. 1982. Prenominal Modification in Chinese and Japanese. *Linguistic Analysis*, Vol 9:119-53
- Kuno, Susumo. (1973) *The structure of the Japanese language*. MIT Press, Cambridge, Mass.
- Lapointe, Steven & Sarah Nielsen. (1996) A reconsideration of type III gerunds. in: *Japanese/Korean Linguistics 5*, 305-320. edited by: N. Akatsuka, S. Iwasaki & S. Strauss. CSLI, Stanford.

- Matsushita, D. (1930) *Hyoojun nihonkoogo hoo*. Chubunkan, Tokyo.
- Mengzi yinde, Harvard-Yenching Institute Sinological Index Series, supplement number 17. Peiping: Yenching University Press, (1941)
- Murasugi, Keiko. (1991) Noun Phrases in Japanese and English: A study in syntax, learnability and acquisition. PhD dissertation. University of Connecticut.
- Murasugi, Keiko. (1997) An Antisymmetry Analysis of Japanese Relative Clauses, manuscript, Kinjo Gakuin University.
- Roberts, Ian & Anna Roussou. (1999) A Formal Approach to “Grammaticalisation”. manuscript, University of Stuttgart and University of Cyprus.
- Shibatani, M. (1996) *The languages of Japan*. Cambridge University Press.
- Simpson, Andrew. (1997) On the Status of Modifying DE and the Syntax of the Chinese DP. ms. London, to appear in: *On the Formal to Chinese Languages*. eds C.S.Liu & S.W.Tang, CSLI
- Whitman, John, Lee, Kwee-Ock & Lust Barbara. (1991) Continuity of the principles of Universal Grammar in first language acquisition: the issue of functional categories. in Proceedings on NELS 21, 383-398. edited by Tim Sherer. GSLA, Amherst.
- Williamson, Janis. (1987) An Indefiniteness Restriction for Relative Clauses in Lhakota. in: *The Representation of (In)definiteness*. eds E.Reuland & A.ter Meulen, MIT Press, Cambridge, Mass.
- Wu, Xiu-Zhi Zoe. in preparation. Functional categories in Mandarin Chinese. PhD dissertation, University of Southern California.
- Yuzawa, Y. (1944) *Gendaigohoo no shomondai*. Nihongo kyoiku Shinkookai. Benseisha, Tokyo.
- Zabala, Igone & Juan Carlos Odriozola. (1996) On the relation between DP and TP. The structure of Basque infinitivals. in: *Catalan Working Papers in Linguistics Vol 5, No. 2*, 231-281. Universitat Autònoma de Barcelona.
- Zhuangzi yinde, Harvard-Yenching Institute Sinological Index Series, supplement number 20. Peiping: Yenching University Press, (1947)
- Zuozhuan, Chunqiu jingzhuan yinde, Harvard-Yenching Institute Sinological Index Series, supplement no.11 Peiping: Yenching University Press, 1937.