The morphosyntactic structure of articles and pronouns in Dutch

Johan Rooryck
Leiden University - ULCL

Abstract[*]
This article argues for a radical morphosyntactic analysis of Dutch articles and pronouns. On the basis of Southern Dutch data, the indefinite article /œn/ is analyzed as a combination of Number /-n-/ and Gender /œ/. This analysis is extended to the definite articles /dœ/ and /hœt/, where /-D-/ is a marker for definiteness preceding or following Gender /œ/ via a syntactic movement operation. Agreement on demonstratives and adjectives is argued to follow a similar pattern. In all cases, movement is correlated with identity in feature specification type. Demonstratives such as /dit/ 'this', /dat/ 'that', and /hi:r/ 'here' and /da:r/ 'there', as well as question words such as /wat/ 'what' and /wa:r/ 'where', include a marker for proximate /-i-/ and distal /-a-/. The absence of question words containing proximate /-i-/ (such as */wi:r/ and */wit/) derives from the semantic incompatibility between Wh- /w-/ and . proximate /-i-/.

Keywords: morphosyntax, agreement, underspecification, gender, number, (in)definite, article, pronoun, proximate, distal.

1. The morphological structure of indefinite articles

In Romance languages, articles and pronouns are often morphologically complex, consisting of a morpheme indicating (in)definiteness and a morpheme marking gender. French definite –l- combines with –e/a to yields the definite articles and clitics *le* 'the_masc' and *la* the_fem'. The morpheme –l- also appears in the pronouns *il* 'he', *lui* 'him', and *elle* 'she' (e.g Rooryck 2001: Ch 8). Similar such analyses apply to indefinite and demonstrative articles and pronouns in e.g. Spanish (Harris 1991, Bernstein 1993).

In some varieties of Southern Dutch (SoD), indefinite articles are marked for gender differences which are conspicuously absent in Standard Dutch (SD).

(1) a. [nœ stul]   b. [œn straf]   c. [œ stul.œkœ]   (SoD)
        a.masc chair   a.fem punishment   a.neuter chair.diminutive

(2) a. [œn stul]   b. [œn straf]   c. [œn stul.tʃe]   (SD)
        'a chair'   'a punishment'   a chair.diminutive
                                                          'a small chair'

It is well known that diminutive morphology triggers neuter agreement in Dutch: the (1c-2c) examples thus illustrates gender shift with a diminutive, originally masculine, noun.

It is tempting to see the pattern in (1) as involving two morphemes, an indefinite marker /–n–/ and a gender marker /–œ–/. In this way, the morphological structure of Germanic languages would resemble Romance even more closely. It is relatively uncontroversial that the schwa morpheme corresponds to gender. Both varieties of Dutch also use this morpheme for adjectival gender agreement as in (3), although in a slightly different way to which I will come back in section 4.

(3)   a. [œn zwa:r.œ stul] b. [œn zwa:r.œ straf]   c. [œn zwa:r stul.tʃe]
            'a heavy.masc chair'   'a heavy.fem punishment' a heavy chair.diminutive
            'a heavy chair'   'a heavy punishment'   'a heavy small chair'

It the pattern in (1) indeed involves two separate morphemes, indefinite /–n–/ and gender /–œ–/, only the ordering of these morphemes determines masculine or feminine gender, with /n–œ/ masculine and /œ–n/ feminine. Neuter gender involves /œ/ alone, without the /–n–/ morpheme. This is in line with a language such as German, where neuter gender also involves less marking (*ein Buch* 'a book'*)* than masculine (*einer Mann* 'a man') or feminine (*eine Frau* 'a woman'). The comparison with German is also instructive in another regard. German gender markings do not involve morpheme ordering: gender markings are simply added to the right of the indefinite morpheme *ein*. By contrast, the ordering of morphemes in (1ab) is reminiscent of Swedish N to D movement (Delsing 1988, 1993; Taraldsen 1990) illustrated in (4). In these cases, the different ordering of N and D corresponds to a definite/ indefinite interpretation of the DP.[1]

(4)   a. en stul   b.   stulen   (Swedish)
           'a chair'   'the chair'
       c. ett bord   d.   bordet
           'a table'   'the table'

Although the interpretive results of morpheme ordering in (1) and (4) are quite different, the strategy employed seems to be identical: the ordering of morphemes correlates with the choice for a value of definiteness in (4) and gender in (1). Preposing the article in

---

[1] Bernstein (2001:550) argues that this phenomenon does not involve N to D of the kind attested in Romance languages, but rather definiteness agreement on the noun. Bernstein agues that N-to-D movement should uniformly move the noun past the adjective, as in Italian *casa nuova*. Swedish postposed determiners do not fit that pattern: adjectives occur prenominally and are themselves preceded by a copy of the determiner attached to a *d-* morpheme as in *d-en stora mann-en* 'the tall man'.
Bernstein's (2001) solution in terms of definiteness agreement has considerable disadvantages. It entails that the 'higher', preposed indefinite article *en* or *ett* in (4ac) is a D, while the lower postposed and definite occurrence of *en* or *ett* in (4bd) would be an expression of agreement. As a result, the morphological identity between the paradigms of preposed Ds and postposed agreement markers becomes a mere coincidence. Paradigms for agreement on the one hand and (in)definite articles on the other do not generally tend to be identical. Bernstein's solution therefore does not yield any new insight into the relation between definiteness and indefiniteness in Swedish. The analysis mentioned seems unlikely to be on the right track. I will therefore continue to assume that Swedish postposed determiners are a result of N-to-D movement. The 'double definiteness' cases as in *d-en stora mann-en* 'the tall man'need to be studied on a par with similar such cases in Modern Greek and Aromanian (Gr. *To puli (afto) to mikro* the bird (this) the small 'The/ this small bird') (Campos & Stavrou to appear , Campos to appear).

(4ac) correlates with an indefinite interpretation of the article, while preposing the gender marker in (1b) correlates with feminine gender. Similarly, postposing the article in (4bd) determines its interpretation as involving definiteness, while postposing the gender marker in (1a) determines its interpretation as feminine gender.

In the light of these facts, I would like to assume that indefinite/ gender morpheme ordering in (1) involves syntactic movement of one morpheme to the left of the other. The /–n–/ morpheme corresponds to the indefinite, represented here as a Num head. I also assume that Num is generated below Gender. This yields the representations in (5):

(5)   a. $[_{\text{Gender}}$ n–œ $[_{\text{Num}}$ n̶ $]]$ (masculine indefinite)
       b. $[_{\text{Gender}}$ œ $[_{\text{Num}}$ n $]]$ (feminine indefinite)
       c. $[_{\text{Gender}}$ 0–œ $[_{\text{Num}}$ θ $]]$ (neuter indefinite)

In (5c), I represent a zero allomorph of the indefinite which appears in the context of neuter nouns. I would like to argue that this zero allomorph moves to Gender in the same way as its overt masculine counterpart.

Evidence for the analysis in (5) comes from morphophonological changes of the indefinite article in some contexts. When a masculine noun starts with /b/, a process of assimilation will affect the ending of the indefinite in SoD, adding an /m/ coda to indefinite /nœ/. When the masculine noun starts with a vowel, assimilation adds an /n/ coda to /nœ/. Compare (1a) above with (6–7):

(6)   a.   [nœm bur]   b.   [nœm bal]   (SoD)
             'a farmer'   'a ball'
(7)   a.   [nœn ezœl]   b.   [nœn a:p]
             'a donkey'   'a monkey'

I would like to argue that it is not an accident that the feature [nasal] appears on the assimilating coda. In line with Chomsky's (1995) copy and deletion approach to traces, the trace of the /–n–/ morpheme in (5a) can be taken to be a true copy of it, containing all its information. A simple morphophonological rule can therefore be assumed, stipulating that part of the phonological information of the copy of /–n–/ is spelled out in the appropriate phonological context. The idea that a head is spelled out twice should not be surprising: there a number of cases reported in the literature where the heads of functional projections are subject to multiple spellout..[2] Note also that if the analysis presented here is rejected in favor of a purely phonological account, it is necessary to formulate a rule copying phonological information from the onset to the coda of the indefinite morpheme.

How can this analysis be related to the idea that the zero allomorph of /–n–/ moves to Gender in the case of neuter indefinites? Neuter indefinites display similar assimilation effects on their coda as masculine indefinites, as can be observed in (8), to be compared with (1c). In front of neuter nouns starting with /b/, assimilation is optional, while in front of neuter nouns starting with a vowel, assimilation is obligatory:

---

[2]   For other cases that can be interpreted as multiple spellout, see Van Koppen (2001) reinterpreting data from Kooiman (1958) and Wilde–van Buul (1943). For a theoretical discussion of the notion multiple spellout, see Nunes (2000).

(8)   a.   [œ(m) bal.œkœ]   b.   [œn apœl.ke]   (SoD)
             a ball.DIM   an apple.DIM
             'a small ball'   'a small apple'

In theoretical terms, the assimilation process also affects the copy of the zero allomorph in (5c), forcing it to spell out. Note that strictly speaking, this is not an argument for movement of the zero allomorph *per se*. It could very well be that the zero–allomorph stays in place in (5c), being spelled out *in situ*. I will however stick to the movement analysis for the sake of simplicity: on the analysis involving movement of the indefinite zero–allomorph, morphophonological spellout only affects copies of moved material.[3]

The Standard Dutch cases in (2) can now be analyzed along the same lines. In terms of the analysis in (5), it is clear that this variety of Dutch has simplified the complex system of SoD by extending the model of the feminine indefinite to both masculine and neuter indefinites.

2. Definite articles and adjective agreement

It is now tempting to extend this analysis to definite articles and adjective agreement. I will not make a distinction between SD and SoD in these cases, as the facts are identical in both varieties in the relevant respects. It is striking that the schwa morpheme marks masculine and feminine gender when it is found at the right of the definite article or the adjective. Witness (9ab), and (3a-c) repeated here:

(9)   a.   [d-œ stul]   b.   [d-œ straf]   c.   [œ-t stultʃœ]
             'the chair'   'the punishment'   'the small chair'

(3)   a.   [œn zwa:r.œ stul] b.   [œn zwa:r.œ straf]   c.   [œn zwa:r stul.tʃe]
             'a heavy.masc chair'   'a heavy.fem punishment' a heavy chair.diminutive
             'a heavy chair'   'a heavy punishment'   'a heavy small chair'

In the case of neuter articles, the schwa morpheme occurs in front of what I take to be the same /–D–/ morpheme as in (9ab). Apart from morpheme ordering, the only difference between (9ab) and (9c) is that /–D–/ has undergone the final devoicing typical of Dutch. In the case of neuter adjectives, the schwa morpheme is not spelled out.

In terms of an analysis along the lines of (5), I propose the configurations in (10) for the definite article:

(10)   a.   [$_{Gender}$ d–œ [$_D$ đ]] (masculine definite)
       b.   [$_{Gender}$ d–œ [$_D$ đ]] (feminine definite)
       c.   [$_{Gender}$ œ   [$_D$ t]] (neuter definite)

Comparing the structures of the indefinites in (5) with the configurations of the definite articles in (10), a paradox emerges. In both cases, movement of D or Num to Gender

---

³   Admittedly, it remains unclear why this spellout process is limited to the contexts specified.

occurs. However, with indefinite articles, Num moves to Gender if Gender is masculine or neuter, while definite articles involve movement of D to Gender in the case of masculine and feminine Gender. Masculine thus patterns with feminine for definite articles, and with neuter for indefinite articles. This might suggest at first sight that the analysis is on the wrong track, since the schwa morpheme does not seem to consistently mark a specific gender value. Note however that the morpheme does mark a combination of gender values in each case. How can this pattern be represented and made insightful?

It stands to reason to assume that the pattern observed has something to do with the internal representation of Gender *per se*. It is well known that Gender specifications are not just freely substitutable values of a paradigm with identical status. In systems involving two genders, masculine and feminine, as in French, masculine functions as a 'default' or 'underspecified' value, while feminine is marked and specified (cf. also Rooryck 2000). This can be seen in the use of the 'masculine' third person pronoun *il* 'he/ it' being used as an expletive (*il pleut* 'it rains', and in the use of masculine agreement in cases where a combination of masculine and feminine nouns require agreement *(Jean et Marie sont intelligent(*e)s* ' Jean and Marie are intelligent'). Similarly, Dutch neuter seems to function as a 'default' value, since the neuter value of the article functions as a pronominal expletive (*het regent* 'it rains'). This observation then argues for a binary representation of Gender values, with neuter separate from the gender subset masculine/ feminine, which I will designate informally as *Subgen* for lack of a better label.

Now let us take another look at the facts involving gender with definite and indefinite articles. Masculine gender patterns with feminine for definite articles, and with neuter for indefinite articles. It thus seems to have an intermediate status between 'strongly marked' feminine and 'unmarked' neuter. This might be taken to mean that within the gender subset uniting masculine and feminine, masculine is less marked than feminine. In a system of unary feature values making use of underspecification through an Attribute – Value system,[4] the necessary feature values could be represented as follows:

(11)   [Gen : [Subgen: fem]]   (feminine)
         [Gen : [Subgen:]]   (masculine)
         [Gen :   ]   (neuter)

Similar considerations apply to notions such as definite and indefinite: definiteness is consistently more 'marked' crosslinguistically than indefiniteness (e.g. in a system with a morphological opposition between 'bare' and 'non–bare' nouns, the 'bare' nouns will be indefinite). Again, this could be represented as follows, where Ref informally stands for Referentiality or whatever label would be used to refer to the set of definite and indefinite values.

(12)   [Ref : def](definite)
         [Ref :   ] (indefinite)

These feature representations can now be related to the analyses in (5) and (10). In (5ac), the representations for indefinites show that underspecification in the composition of Gender correlates with movement of an underspecified Referential element, Num. In (5b) the most specified Gender, feminine, fails to attract Num. By contrast, for definite articles in (10ab), a specification in Gender correlates with movement of D, the most specified Referential element. In (10c), the least specified Gender (neuter) fails to attract D. In other words, there is a correlation between (under)specification and (lack of) movement. Specified functional heads move to specified functional heads, underspecified functional heads move to underspecified functional heads. A difference in specification type results in lack of movement. The following generalization can be formulated:

(13)   In the morphosyntactic structure of Dutch determiners,
            Num and D move to Gender iff their specification type is identical.

The observation that masculine patterns with feminine for D movement and with neuter for Num movement is due to the fact that masculine Gender is at once specified and underspecified. Masculine is specified for Gender with a Subgen value, allowing it to pattern with (fully) specified feminine. It is also underspecified for Subgen, allowing it to pattern with neuter which is fully underspecified for Gender.

3. Demonstratives
In the previous section, it was argued that /d–/ and /–t/ in *de* 'the' and *het* 'the' are in fact allomorphs of the same underlying /–D–/ morpheme. This analysis now forces us to take a new look at the neuter demonstrative articles and pronouns *dit* 'this' and *dat* 'that'. I would like to argue that *dit* 'this' and *dat* 'that' are composed of three morphemes. The desirability of a morphologically compositional analysis should be clear from the fact that the interrogative/ indefinite pronoun *wat* 'what' is minimally different from *dat* 'that', with /w–/ 'wh–' substituting for /d–/ in the same way as in the locative PPs *daar* 'there' and *waar* 'where'. I will come back to these parallelisms in section 4.3.

The most obvious morphemes contained in *dit* 'this' and *dat* 'that' are proximate /–i–/ and distal /–a–/, which also occur in locative PPs such as *hier* 'here' and *daar* 'there'. I would like to argue that the [d–] and [–t] elements contained in *dit* 'this' and *dat* 'that' are two instances of the underlying morpheme /–D–/ which also occurs in *de* 'the' and *het* 'the'. In other words, the morphological structure of *dit* 'this' and *dat* 'that' is an instance of a phenomenon attested in Northern Germanic, Modern Greek and Aromanian, and is sometimes called *double definiteness*. Instances of double definiteness occur when a definite DP is modified prenominally by either an adjective or a PP.

(14) a. d.en här bok.en   b.   d.en   där   bok.en.   (Swedish)
          d.DET here book.DET   d.DET   there   book.DET
          'This book'   'That book'
      c. en stor bok   d.   d.en   stor.a   bok.en
          DET big book   d.DET   big.AGR   book.DET
          'a big book'   'the   big   book'

I submit that the morphosyntactic structure of the demonstrative Ds *dit* 'this' and *dat* 'that' exploits the same syntactic strategy as that present in Norwegian and Swedish definite DPs. The exact analysis of double definiteness constructions is unclear. Various proposals exist in the literature. Delsing (1988), Santelmann (1993) and Kester (1996) distinguish two different functional projections for the highest D and the lowest D, Delsing (1989, 1993) leaves the lowest D as an enclitic on N. I will not take position on this question here, noting that an analysis should be attempted in terms of multiple spellout of two positions of the chain created by movement of D.

The analysis of *dit* 'this' and *dat* 'that' as a case of double definiteness has an additional advantage. A comparison with interrogative/ indefinite *wat* 'what/ something', makes it clear that the highest /–D–/ spelled out as [d–], is responsible for the definite interpretation of the demonstrative. I would like to suggest that the contribution of the lowest /–D–/ spelled out as [–t] is to contribute to the marking of neuter Gender. Although /–D–/ is intrinsically a marker of definiteness, its position after /-œ-/ in the definite article *het* 'the' defines that combination as [neuter, definite]. I would like to suggest that the same is true in *dit* 'this' and *dat* 'that'. That is, the final [–t] on these morphemes is a definiteness marker whose final position in combination with a preceding, covert neuter Gender marker contributes neuter gender to these demonstratives. In other words, *dit* 'this' and *dat* 'that' covertly contain a fourth morpheme for Gender, besides two instances of /–D–/ and locative /–i–/ or /–a–/. This covert morpheme is marked for neuter Gender. The lowest definite /–D–/ does not move to this position by virtue of the generalization in (13).

Let us now turn our attention to masculine and feminine demonstratives *deze* 'this' and *die* 'that'. Again, language variation proves instructive. The contrast in (15) between SoD masculine/ feminine [dezœ] 'this' and neuter [des] 'this' shows that *deze* 'this' is bimorphemic, consisting of a demonstrative morpheme /des/ and a Gender morpheme /–œ/. This usage does not exist in SD, where it is replaced by *dit* 'this' (16). Importantly, the Gender inflection for /des/ is exactly parallel to that of /–D–/ and adjectives: masculine and feminine display Gender in final position, neuter does not.

(15) a. [dezœ stul]   b. [dezœ straf]   c. [des stul.œkœ] (SoD)
           this.masc chair   this.fem punishment   this.neuter chair.diminutive

(16) a. [dezœ stul]   b. [dezœ straf]   c. [dit stultʃe]   (SD)
           'this chair'   'this punishment'   'this small chair'

The facts for *die* 'that' are slightly different. First of all, *die* 'that' cannot be used for neuter nouns in both SD and SoD. Strikingly, however, Gender agreement in SoD does not conform to the pattern of /des/, /–D–/ and adjectives, but rather to the pattern displayed by indefinite articles in SoD. Specifically, Gender agreement does not show up in final position of articles in the context of feminine nouns. The contrast between (17a) and (17b) thus resembles that between (1a) and (1b):

(17) a. [dij.œ stul]   b. [di straf]   c. [da stul.œkœ]   (SoD)
           that.masc chair   that.fem punishment   that.neuter chair.diminutive

(18)   a.   [di stul]   b.   [di straf]   c.   [dat stultʃe]   (SD)
            'that chair'   'that punishment'   'that small chair'

The question then arises as to why this is the case. Ideally, this state of affairs should follow from the generalization stated in (13). In other words, the Gender difference between *deze* 'this' and *die* 'that' in (15) and (17) would follow from a difference in their specification type. If *die* 'that' is less specified in some way than *deze* 'this', then the facts in (15) through (18) would conform to the generalization in (13). Let us briefly see why this is the case in terms of the analysis presented above for D and Num.

In terms of (13), the most specified Gender, feminine, would fail to attract /di/, because /di/, like Num /–n–/, is an underspecified morpheme. The underspecified Gender masculine by contrast would attract the underspecified /di/. In the case of the combination /des/ + /–œ/ 'this', masculine and feminine /–œ/, being both specified for Subgen, would be specified enough to attract specified /des/. Note that the relevant feature specification of demonstratives should be one that is typical for the demonstrative morphemes targeted for attraction. In other words, the phi–features of demonstratives will be irrelevant, since demonstratives are not inherently defined in terms of phi–features. However, demonstrative *deze* 'this' and *die* 'that' are inherently different in that the former is proximate and the latter distal. In terms of a unary approach to feature specification making use of underspecification, this means that *die* 'that' is characterized by the absence of the feature proximate. Note that the reverse is not possible: absence of a distal feature does not result in the unspecified feature being specified as proximate without further ado. While [distal] is 'anywhere that is not here', [proximate] cannot be defined as 'anywhere that is not there'. In other words, the specification for *deze* 'this' and *die* 'that' is as follows:[5]

(19)   *deze* 'this'   [Location : proximate]
               die 'that'   [Location :   ]

The underspecification of [distal] can also be observed in the fact that distal, but not proximate demonstrative pronouns are typically used as relative and D- pronouns in relative clauses and left dislocation.

(20)   a.   Het boek dat/*dit ik gelezen heb
                'The book that/*this I read'
       b.   Deze jongen, die ken ik niet
                This boy, that know I not
                This boy, I don't know

In addition, the underspecification of [distal] provides an explanation for the absence of feature clashes. If the D-pronoun in (20b) were lexically specified for [distal], a feature

---

5   I have no explanation for why proximate /–i–/ and distal /–a–/ in *dit* 'this' and *dat* 'that', with the same demonstrative specification as *deze* 'this' and *die* 'that', do not likewise trigger agreement differences in the same way. I will assume this is due to interference of the *double definiteness* configuration.

clash should arise with the [proximate] specification on the nominal head provided by *deze* 'this'.[6]

The different specification of *deze* 'this' and *die* 'that' now allows for an explanation of their agreement patterns in terms of the generalization in (13). The most specified Gender, feminine, fails to attract /di/, because /di/ is underspecified for Dem. The partially underspecified Gender masculine, by contrast, does attract the underspecified /di/. In the case of the combination /des/ + /–œ/ yielding [dezœ] 'this', masculine and feminine /–œ/, being both (partially) specified for Subgen, are specified enough to attract specified /des/.

4. Some further issues
4.1. Adjectival agreement
I would like to extend the explanatory power of the generalization in (13) to the case of adjectival agreement in Dutch which was briefly mentioned in (4), repeated here.

(4)   a.   een zwar–e stoel   b.   een zwar–e straf   c.   een zwaar stoeltje
            'a heavy chair'   'a heavy punishment'   'a heavy small chair'

It has often been observed that the absence of gender with neuter nouns is only visible with indefinite articles. With definite articles and (in)definite plurals, /–œ/ reappears:

(21)   a.   het zwar–e stoeltje   b.   (vel–e/ die) zwar–e stoeltjes
            'The heavy small chair'   '(many/ those) heavy chairs'

The relevant generalization seems to be that the adjective remains uninflected only when DP is [–definite, +neuter, +singular] (e.g. Barbiers 1992, Kester 1996). Needless to say, this is not a nice stipulation, as it even casts doubt on the status of /–œ/ as a Gender morpheme. I would like to suggest that this observation can be derived from the generalization in (13). The inherent feature specification for adjectives is [+N, +V]. In terms of unary features and underspecification, the observation can be reformulated as follows: [–definite] can be translated as lack of definiteness, [+neuter] as lack of gender specification, and [+singular] as lack of number specification. Since in this case the adjective and the Agreement phrase have different feature specification types, no movement ensues. By contrast, as soon as one of the features of AGR is fully specified, movement to AGR occurs. Similar to the configuration for *de* 'the' in (10), the following structure can be proposed for adjectival agreement:

(22)   a.   [AGR   AGR   [A   zwaar]]   ([–definite, +neuter, +singular])
                    | |
                 [Gender:]   [N, V]
                 [Ref:]
                 [Num:]
       b.   [AGR   zwar–e   [A   ~~zwaar~~]]   (any specification on AGR)

---

[6]   Thanks to Sjef Barbiers for pointing this out to me.

The analysis in (22b) affords a further insight into the nature of elliptical DPs in Dutch. In Standaard Dutch, N-ellipsis is impossible if the adjective cannot be marked by /-œ/

(23)   a.   * Over boeken gesproken, ik heb [een leuk *en*] gelezen
               'Talking about books, I read a nice (one).'
        b.   Over films gesproken, ik heb [een leuke *en*] gezien
               'Talking about films, I saw a nice (one).'

Barbiers (1990) relates this observation to the requirement that an empty N (*en*) needs to be identified by a specified gender feature. Since *boek* 'book', the antecedent of the empy N is underspecified neuter, empty N in (23a) cannot be identified. A paradox remained, however. The /-œ-/ morpheme is not only used for masculine or feminine singular nouns, but also for plural nouns of any gender, while the /-œ-/ -less form occurs in a unique context, i.e. [neutral, singular]. The paradox then is that precisely the only form that is uniquely specified for gender (i.e. neuter) in the adjectival paradigm cannot identify the empty noun in (23a). That problem is lifted under this analysis. It is the total underspecification of the /-œ-/ -less adjectival form that is relevant with respect to the inability to identify the empty noun, and not its uniqueness in [neuter, singular] contexts.[7]

4.2. More on demonstrative *deze* 'this' and *die* 'that'
The morphological analysis of *deze* 'this' and *die* 'that' can be pushed one step farther. In the context of an bimorphemic analysis of the French nominative complementizer and interrogative pronoun *qui* 'who', Rooryck (2000: ch 8) argues that pronoun and demonstrative article *die* is likewise bimorphemic, as it alternates with the interrogative pronoun *wie* 'who' in the same way *dat* 'that' alternates with *wat* 'what'. Note that *ie*'he/it' in Dutch can also function as an enclitic third person pronoun as in (24):

(24)   a.   Heeft ie dat al gedaan?   b.   Is ie gebroken?
               Has he that already done   'Is it broken?'
               'Did he already do it?'

Rooryck (2000) argues that *–ie* is very similar to French *i(l)* 'he/it' in that it is underspecified for person, number and gender features. The fact that French *i(l)* 'he/it' and Dutch *–ie* can only be interpreted as 3rd person singular masculine is attributed to a 'default' interpretation of the Person, Number and Gender underspecification. There is a slight unexplained twist to this, however. In combination with /–D–/ and /w–/, Dutch *–ie* is compatible with both singular and plural interpretation, and occur in both masculine and feminine contexts whether they function as an article (17–18) or as a pronoun (25b) :

(25)   a.   Wie heeft/ hebben dat gedaan? Marie/ Jan
               'Who has/ have done that? Marie/ Jan
        b.   Die heeft/ hebben dat gedaan. (Marie/ Jan)
               'That one/ those has have done that. (Marie/ Jan)'

---

[7]   I owe this observation to Sjef Barbiers.

It seems then that the morphological combination of /–ie/ with /–D–/ and /w–/ prevents a 'default' interpretation of underspecified features from taking place. Rather, *die* 'that/those' and *wie* 'who' are radically unspecified for the phi features Person, Gender and Number.

Note that the distal feature of *die* 'that' (underspecified [loc:]) is not contributed by either /–D–/ or /–ie/. I would like to propose that *die* 'that' acquires this interpretation because it is inserted in the demonstrative paradigm to which *deze* 'this' belongs. Since *deze* 'this' is positively specified for [loc : prox], the inherently unspecified *die* 'that' can only take up the other value left in the demonstrative paradigm.

The features of the proximate article and pronoun *deze* 'this', are slightly different from its distal counterpart *die* 'that'. It is likewise underspecified for Person, triggering a 3rd person 'default' interpretation. It is also underspecified for Gender, at least in SoD, where it can appear with masculine, feminine and neuter nouns (cf. supra (17)). Contrary to *die* 'that', however, the Number feature of *deze* 'this' is specified for singular. This cannot be observed in its use as an article, as articles never receive plural agreement in Dutch, but its use as a pronoun clearly reveals that *deze* 'this' needs a plural ending to trigger plural agreement (plural in (26a) is just due to the coordination). This means that *deze* 'this' must be either inherently or by default singular.

(26)   a.   Dez.e   en   die   is/ zijn gekomen
             this.GEN   and   that.GEN   is/ are come
             'This one and that one has/ have come'
       b.   Dez.e.n   en   die   zijn/ *is gekomen
             this.GEN.PL   and   that.GEN.PL   is/ are come
             'This one and that one have come'

4.3. Where are *wit and *wier?

In section 3, I argued that *dit* 'this' and *dat* 'that' are composed of three morphemes: proximate /–i–/ and distal /–a–/, which also occur in locative PPs such as *hier* 'here' and *daar* 'there', and a double occurrence of the /–D–/ morpheme. *Dat* 'that' alternates with the interrogative/ indefinite pronoun *wat* 'what'. In this case, /w–/ 'wh–' substitutes for /d–/ in the same way as in the locative PPs *daar* 'there' and *waar* 'where', and the article/ pronoun *die* that' and *wie* 'who'. The paradigm combining proximate /–i–/ and distal /–a–/ with PP /–r/, D /–D–/ and Wh– /w–/, can be represented as follows:[8]

(27)	Def PP	Wh– PP	Def D	Wh– D
proximate	*hier*	**wier*	*dit*	**wit*
distal	*daar*	*waar*	*dat*	*wat*

In view of the combination pattern observed, it is surprising that interrogative /w–/ cannot combine with proximate /–i–/. The answer to this impossibility lies in the nature of the Wh– morpheme. By its very definition, [+Wh–] /w–/ implies that the referential range of

---

[8]   Note that the /–ie/ in *die* 'that' and *wie* 'who' is not the same as the one that enters the proximate/ distal i/a opposition (cf. supra).

the PP or D it attaches to is open, i.e. unknown to the person formulating the question.[9] This meaning specification is in contradiction with the demonstrative specification of /–i–/ as proximate, i.e. close to the speaker and therefore necessarily known to him or her. I therefore conclude that *wit* and *wier* are excluded as the result of a feature clash between two inherently incompatible meaning specifications.

It is less clear why Dutch does not feature *dier* 'here' instead of *hier* 'here'. It is unlikely that that /d–/ in *daar* 'there' indicates definiteness, since *hier* 'here' and *daar* 'there' are both equally definite. It might be the case that /d–/ in *daar* 'there' functions as a dummy morpheme without specification, a diachronic fossil from an older stage of Germanic. Note that in Swedish, and less transparently, English, the proximate/ distal distinction is still actively rendered by the opposition between /h–/ and /d–/, rather than through /–i–/ vs /–a–/:[10]

(28) a. h.är   b. d.är
          PROX.LOC/*DIR   DIST.LOC/*DIR
          'here'   'there'
      c. h.it   d. d.it
          PROX.DIR   DIST.DIR
          'hither'   'thither'

It is therefore likely that the *h/ d* alternation in Dutch *hier* 'here' and *daar* 'there' is a relic from a period when this opposition was still active.

5. Conclusion

A closer look at the morphological composition of Dutch articles and pronouns reveals intricate morphosyntactic patterns. It would be intriguing to extend the analysis to other cases, such as the combination of the distributive morpheme *elke* with range–indicating Wh– /w–/ in *welke* 'which (= wh+each?)' and comparative /z–/ in *zulke* such'. To do so is beyond the scope of this paper. I hope to have shown that a morphosyntactic approach to the internal structure of Dutch articles and pronouns, supplemented with a fully explicit representation of feature (under)specification, offers some insights into the otherwise curious interaction between articles, pronouns and agreement in Dutch.

---

[9]   It was suggested to me that *wit* could perfectly well be imagined as a D–linked Wh– element in the system. However, note that in non Wh- contexts, coreference with previously mentioned discourse elements is taken up by D–linked *dat*, not *dit*:
i.   Jan heeft een fout gemaakt en dat/*dit heb ik hem niet vergeven.
          'Jan made an error, and I didn't forgive him for that'
This makes it unlikely that *wit* could ever function as a D–linked Wh– element.
[10]   The locative/ directional opposition between *här* 'here' and *hit* 'hither' is fully active in Swedish: unlike English here (*come here!*), *här* 'here' cannot be used directionally (*komm hit/ *här*).

References

Barbiers, Sjef. 1990. Telwoorden, adjectieven en lege NPs. Unpublished master's thesis, Leiden University.

Barbiers, Sjef. 1992. Adjectives as auxiliaries of the noun phrase. Linguistics in the Netherlands. Reineke Bok-Bennema & Roeland van Hout (eds), Amsterdam: John Benjamins, 13-24.

Bernstein, Judy. 1993. The syntactic role of word markers in null nominal constructions. Probus 5, 5-38.

Bernstein, Judy. 2001. The DP hypothesis: identifying clausal properties in the nominal domain. The handbook of contemporary syntactic theory. Mark Baltin & Chris Collins (eds), London Blackwell, 536-561.

Cheng, Lisa & Johan Rooryck. 2001. Licensing Wh- in situ. *Syntax* 3, 1-19.

Campos, Hector & Melita Stavrou. (to appear) Polydefinite constructions in Modern Greek and in Aromanian. Topics in Balkan Sprachbund syntax and semantics. Olga Tomic (ed.) Amsterdam: Benjamins, 147-186.

Campos, Hector. (to appear). Noun modification, pseudo articles and last resort operations in Arvantovlaxika and in Romanian. Lingua.

Delsing, L.O. 1988. The Scandinavian Noun Phrase. Working papers in Scandinavian syntax 42. 57–79.

Delsing, L.O. 1993. The internal structure of noun phrases in the Scandinavian languages. Ph. D. dissertation, Lund University.

Harris, Jim. 1991. The exponence of gender in Spanish. Linguistic Inquiry 22, 27-62.

Kester, Ellen–Petra. 1996. The nature of adjectival inflection. Ph. D. dissertation, University of Utrecht, OTS.

Kooiman, K. 1958. Zuid–Hollands 'hoordiede', 'lachtiede'. In: De nieuwe taalgids 51, 240.

Nunes, Jairo. 2000. Linearization of Chains and phonetic realization of multiple copies. Lecture presented at the workshop on antisymmetry theory. Cortona.

Rooryck, Johan 2000. Configurations of sentential complementation. Perspectives from Romance languages. London: Routledge.

Santelmann, L. 1993. The distribution of double determiners in Swedish: den support in D°. Studia Linguistica 47.2. 154–176.

Scobbie, James. 1991. Attribute value phonology, Ph. D. dissertation: Edinburgh.

Taraldsen, Tarald. 1990. D–projections and N–projections in Norwegian. Grammar in Progress. Joan Mascaró and Marina Nespor, (eds) Dordrecht: Foris, 419–431.

Van Koppen, Marjo. 2001. Subject – Tense inversion. Ms. ULCL.

Wilde–van Buul, G. de. 1943. Het enclitische pronomen personale van de tweede en derde persoon singularis in het Rotterdams. Tijdschrift voor Nederlandse taal– en letterkunde 62, 290 – 301.