Journal of Linguistic Geography (2013), **1**, 115–124. © Cambridge University Press 2013
doi:10.1017/jlg.2013.10

ORIGINAL RESEARCH

Phonology and morphology in Dutch indefinite determiner syncretism: Spatial and quantitative perspectives

Meredith Tamminga*

University of Pennsylvania, Department of Linguistics, Williams Hall, Philadelphia

This paper uses dialect data to disentangle the contributions of phonology and morphology to the emergence of gender syncretism in the Dutch determiner paradigm. Quantitative and spatio-statistical analyses are used to identify an inverse relationship between phonological erosion and adoption of the innovative syncretic system, counter to expectation. That inverse relationship is shown to obscure the parallel development of the determiners in masculine and feminine contexts, leading to the suggestion that the syncretism results from a single morphological change triggered by phonological variability.

1. Introduction

Syncretism, the partial neutralization of surface inflectional distinctions relative to their underlying morphosyntactic features, is a common phenomenon at the phonology–morphology interface. Because of its location at this interface, tracing the diachronic development of syncretism offers the promise of elucidating the interplay between phonology and morphology in language change. In this study I use evidence from a transitional dialect area to investigate the emergence of gender syncretism in Dutch determiners, a case which implicates both phonological and morphological variation and change.

Baerman, Brown & Corbett (2005) outline two widely-recognized pathways to syncretism: Regular phonological change and morphosyntactic change. The former mechanism is at work when uniform sound change that erodes word-final segments leaves inflectional homophony in its wake. Baerman et al. cite as an example the collapse of the nominative and accusative singular in the first declension from Classical Latin to Vulgar Latin. When word-final /m/ was lost due to regular phonological change, so too was the distinction between *luna* ('moon'-NOM-SG) and *lunam* ('moon'-ACC-SG). Similarly, Barðdal and Kulikov (2009) point to the loss of word-final vowels in Middle Arabic as the reason for the loss of the three-case system marked with *–u* (NOM-SG), *–i* (GEN-SG), and *–a* (ACC-SG). The requisite observation of overlap between the segments that are lost through sound change and the segments that crucially distinguish morphosyntactic inflections, though, does not hold for all instances of syncretism. In cases like the loss of the dative–locative distinction in the development of Ancient Greek from Proto-Indo-European, which cannot be accounted for by any concomitant sound changes, Baerman et al. attribute the syncretism to "some fundamental reanalysis of the system of morphosyntactic oppositions" (2005:6).

Baerman et al. (2005) acknowledge, though, that it can in practice be difficult to distinguish between syncretism resulting from phonological change and syncretism resulting from morphological change. A major reason for this difficulty is that analyses of the development of syncretism are often based on static observations of long-completed phonological and morphological changes, in some cases even observations derived from reconstructed morphosyntactic paradigms. In contrast, in the case study I undertake here of the loss of gender marking on indefinite determiners in Dutch, both the conservative (non-syncretic) system and the innovative (syncretic) system are still used by living speakers in neighboring dialect regions. This case study thus presents the possibility of finding evidence for the diachronic pathway of the change in the variable synchronic patterns of the regional transition zone, namely the Dutch province of North Brabant.

At first glance, it would be entirely reasonable to hypothesize that the phonological process of word-final segment erosion is the driving force behind the emergence of gender syncretism in Dutch determiners. The transition from the indefinite determiner triad of masculine /ənən/, feminine /ənə/, and neuter /ən/ to a system with a single indefinite determiner /ən/ could easily be related to the phonological processes of word-final schwa-deletion and /n/-deletion known to be active in Dutch (De Wulf & Taeldeman, 2001). In this paper, however, I take the position that a phonological erosion account is not supported by the synchronic quantitative and geographic evidence.

*Adderss for correspondence: Meredith Tamminga, University of Pennsylvania, Department of Linguistics, Williams Hall, 255 S. 46th St. Philadelphia, PA 19104; 814-321-7062
(Email: tamminga@ling.upenn.edu)

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

Rather than corresponding with the ultimate loss of final segments, the emergence of the innovative syncretic system is most advanced in exactly the locations where the deletion of final schwa is at a minimum.

In the following sections, I develop a quantitative argument in two parts. First, I show that the innovative form /ən/ used with masculine nouns is not phonologically derived from the conservative masculine determiner /ənən/. This allows for the estimation of the underlying rates at which speakers select the innovative determiner with masculine nouns, which turn out to show significant spatial conditioning across North Brabant. Second, I estimate rates of schwa-deletion for individual locations and then use those rates to estimate the rates at which speakers in those locations select the innovative determiner /ən/ with feminine nouns. When schwa-deletion rates are taken into account, the feminine determiners turn out to follow the same spatial pattern as the masculine ones.

I raise the possibility that the role of phonology in the development of syncretism may, in this case, be one of shaping the learners' hypotheses rather than one of erasing phonological material. A drop in schwa-deletion rates may make the reanalysis of /ənə/ as /ən/ more likely and thus serve as a catalyst for the emergence of gender syncretism in the determiner system. This role for phonology in the development of syncretism is distinct from, but could in principle interact with, the erosion of final segments. The relationship between phonological and morphological change in the development of syncretism is thus simultaneously complicated and clarified. Although the unexpected apparent independence of gender syncretism from phonological homophony mid-change calls into question the analysis of previous changes for which we can only observe the final outcome, identifying phonological variability as a potential triggering factor for reanalysis is a step towards pinning down the as-yet ill-defined nature of morphosyntactic change.

2. Data and methods

2.1 Dutch dialect data

The data for this study come from the Goeman—Taeldeman—Van Reenen Project (GTRP) (Van den Berg, 2003). The fieldwork results from this project form the basis of the *Morphological Atlas of Dutch Dialects, Volumes I and II* (*Morfologische Atlas van de Nederlandse Dialecten, Deel I en II*, or MAND) (De Schutter, Van den Berg, Goeman & De Jong, 2005; Goeman, Van Oostendorp, Van Reenen, Kornwinder, Van den Berg & Van Reenen, 2008) as well as the *Phonological Atlas of Dutch Dialects, Volumes I-IV* (*Fonologische Atlas van de Nederlandse Dialecten*, FAND) (Goossens, Taeldeman & Verleyen, 1998–2005). Only a fraction of the data that was collected is actually mapped in MAND and FAND, however. The full dataset is made available both online and on a CD-ROM (Van den Berg, 2003); it includes both the fine phonetic transcriptions and a simplified version using a modified alphabetic spelling system. For the purposes of this study the simplified version proved to be sufficient.

The main method of data elicitation was the administration of an 1,876-item questionnaire that was completed orally by the interviewees and recorded by the GTRP fieldworkers, then transcribed phonetically by trained phoneticians (see Goeman, 1999:ch. 2–3 for discussion of the validity and reliability of the transcribed data). This fieldwork took place between 1980 and 1995. To ensure smooth geographic coverage, a hexagonal grid was overlaid on a map of all of the provinces of the Netherlands as well as the Dutch-speaking (that is, northern) provinces of Belgium and French Flanders. Within each hexagon, the questionnaire was administered to one subject who self-identified as a daily speaker of the local dialect (as opposed to Standard Dutch). Extra speakers were added in known transition zones. The resulting sample included 688 total locations, of which 613 were transcribed; this paper focuses on the 51 locations in the Dutch province of North Brabant, each of which is represented by a single speaker from that location. Over 70% of the subjects completing the questionnaire were male, with a mean age of 61.7 (Van Oostendorp, forthcoming:3). Unsurprisingly given this demographic and the fact that most of the locations are rural, the speakers in the sample also have low levels of educational attainment. This particular demographic profile, the classic dialectological NORM (*nonmobile, older, rural male* [Chambers & Trudgill, 1998:29]), is advantageous for the current study because we expect these speakers to use their local vernaculars with a minimal degree of interference from the formal standard varieties of higher education and professional occupations.

The province of North Brabant, commonly called Brabant, shares its southern border with the Belgian province of Antwerp. It was selected for close investigation because of its status as a religious, historical, and most importantly linguistic transitional area between the Netherlands and Belgium. Unlike much of the rest of the Netherlands and despite many attempts by the Dutch to impose Protestantism, Brabant is majority Catholic. The religious background is just one reflection of a long shared history with Belgium going back to the pre-17th-century Duchy

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

of Brabant. Even the border, which in places carves up towns, farms, and even restaurants between the two nations, bears the imprint of their convoluted history. The results from Tamminga (2012) indicate that, of the locations included in the GTRP fieldwork, the greatest mixture of conservative and innovative determiners is to be found within Brabant's borders, with dialects to the north primarily using a single indefinite determiner and dialects to the south largely maintaining the traditional tripartite distinction.

2.2 *Statistical analysis and mapping*

The premise of this paper is to use dialect data, which demonstrate the geographic transition between two different grammatical systems, and to shed light on the pathway that the change from one system to the other might follow diachronically. My approach to interpreting spatial linguistic data takes as its foundation what Heap calls "the hypothesis that successive stages of a diachronic change can present interesting analogies with a geolinguistic continuum" (2000:43, translation mine). This hypothesis rests on the observation that geographical patterns, especially gradient or nested ones, may be the residue of change in progress, revealing intermediate stages of linguistic interest; as Labov, Ash & Boberg put it, "the diffusion of a change outward typically shows the ordering of successive stages as a series of concentric rings around the originating center, with the initial changes diffused most widely" (2006:4). This, then, is a spatial analogue to the apparent time construct whereby speakers of different ages are taken to represent successive stages in a change in progress. Like the apparent time construct, the spatial reflections of language change have underpinned many studies of language change, particularly regarding the development of mergers (see, *inter alia*, Herzog, 1969; Bailey, Wikle, Tillery & Sand, 1993) and chain shifts (see, *inter alia*, Moulton, 1962; Labov, 1994).

Given quantitative measures of linguistic variables at geographically-distributed points, we can apply the tools of modern spatial statistics, in particular spatial autocorrelation, to facilitate the identification of significant patterns in dialect data. Spatial autocorrelation is tested first globally, then locally. In section 4 I begin with the Moran's *I* test of global spatial autocorrelation (Moran, 1948; Odland, 1988). This is a statistical test of whether the proportions of a variant at each location are spatially distributed in a significantly non-random way. The test statistic, *I*, ranges in value from negative to positive 1, with 0 indicating random spatial distribution. Positive values indicate a tendency towards clustering of similar values, while negative values indicate a tendency towards dispersion of similar values. Moran's I is calculated from the following equation (Odland, 1988):

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

In this equation, N is the number of locations, x is the proportion at the subscripted location, and $\bar{x}$ is the mean across all locations. The value w_{ij} is a weighting function used to relate locations i and j in terms of distance. I use a reciprocal distance weighting function so that the influence of any pair decreases with distance between the locations. Grieve, Speelman & Geeraerts (2013) recommend this weighting function for dialect analyses where relevant dialect boundaries might be quite close together, as it strongly privileges pairs of locations that are near each other.

If global spatial autocorrelation is established, it is reasonable to proceed with a test for local spatial autocorrelation,[1] which asks whether each location is part of a significant high or low value cluster with its neighbors (Ord & Getis, 2001; Grieve et al., 2011). The Getis-Ord G_i* statistic is a *z*-score that is negative for locations that are part of low-value clusters and positive for locations that are part of high-value clusters. Getis-Ord G_i* is calculated with the following equation (Ord & Getis, 1995):

$$G_i^* = \frac{\sum_j w_{ij}x_j - \bar{x}\sum_j w_{ij}}{\sqrt{\frac{\sum_j x^2{}_j}{N} - \bar{x}^2 \sqrt{N\sum_j w^2{}_{ij} - \frac{\left(\sum_j w_{ij}\right)^2}{N}} - 1}}$$

All statistical analyses were conducted using R (R Core Team, 2012). The spatial statistics and maps discussed in section 4 made use of the following R packages: *sp: Classes and methods for spatial data in R* (Pebesma & Bivand, 2005); *spdep: Spatial dependence: weighting schemes, statistics and models* (Bivand et al. 2012); *RANN: Fast Nearest Neighbour Search* (Kemp & Jefferis, 2011); and *maps: Draw Geographical Maps* (Becker, Wilks, Brownrigg & Minka, 2012).

3. Dutch morphology and phonology

Standard Dutch has two genders for nouns: common and neuter. Earlier forms of Dutch, however, had three genders: masculine, feminine, and neuter. In the history of Dutch, the masculine and feminine forms fell together into what is now called the common gender. Some dialects of Dutch, particularly the Flemish varieties spoken in northern Belgium and the Brabantian dialect of the southern Netherlands, still have a tripartite gender system.

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

Table 1. *Standard Dutch system for singular definite and indefinite determiners*

	Common	Neuter
Indefinite	ən	ən
Definite	də	(h)ət

Table 2. *Standard Dutch system for singular definite and indefinite adjective inflection*

	Common	Neuter
Indefinite	-ə	-∅
Definite	-ə	-ə

Table 3. *Conservative Dutch system for singular definite and indefinite determiners*

	Masculine	Feminine	Neuter
Indefinite	ənən	ənə	ən
Definite	dən	də	hət

Table 4. *Conservative Dutch system for singular definite and indefinite adjective inflection*

	Masculine	Feminine	Neuter
Indefinite	-ən	-ə	-∅
Definite	-ən	-ə	-∅

Gender is not marked on the noun in Dutch, nor are Dutch determiners declined for case. Gender is seen on determiners and certain adnominal adjectives. The adnominal gender agreement system of Standard Dutch is outlined in Table 1 for determiners and Table 2 for adjectives.[2]

Most varieties with a robust tripartite gender system maintain unique indefinite determiner forms for all three genders. While different dialects have different surface forms for the indefinite determiners, what remains of the conservative system in North Brabant conforms to the system reported in MAND, as presented in Table 3 for determiners and Table 4 for adjectives (De Schutter et al., 2005:35).

Syncretism is "the situation where a single inflectional form corresponds to multiple morphosyntactic feature values" (Baerman, 2006:363). The situation in the Dutch determiner system, then, can be described as syncretic insofar as the features that we have evidence for from one part of the paradigm are neutralized in their phonological form in another part of the paradigm. In the standard modern system, we can see from the definite determiners that Dutch still has two gender specifications, but there are not two corresponding inflectional forms of the indefinite determiner. In this sense the innovative system is more syncretic than the conservative form, which has three inflectional forms of the determiner in both the definite and indefinite contexts, to match up with the three grammatical genders. The question of when and how the loss of the masculine–feminine distinction in favor of the common gender took place is related to the issues discussed here, but cannot be dealt with without evidence from the definite determiners. As the GTRP questionnaire did not elicit definite forms, I limit my analysis to consideration of the stages of change in the configuration of the indefinite determiners.

The /n/-final determiner forms, like other Dutch words ending in /-ən/, are subject to phonologically-conditioned /n/-deletion. Deletion of /n/ in non-adnominal positions, such as in verbs and nouns, is generally high. In the determiner cases it seems that the phonological environments are essentially categorical, with deletion being the default and retention occurring predictably in a limited environment. De Wulf & Taeldeman report that the situation throughout the province of Brabant is that "in adnominal words 'gender-*n*' remains (most) often preserved before a following word with an initial *t-* (e.g. *tak*), *-d* (e.g. *dag*), sometimes also before *b-*… regularly also before *h-* and occasionally before *r-*" (2001:11, translation mine). In the data for this study, the rate of /n/-retention before /r/ is below 4%, indicating that it is solidly part of the /n/-deleting set of environments, while /b, d, h/ and all the vowels comprise the /n/-preserving context for the majority of the speakers. The /t/-initial words are excluded from the analyses below because they consistently cause robustly variable (rather than categorical) /n/-deletion.

The other relevant phonological process to consider is schwa-deletion. Schwa-deletion is not conditioned in any categorical way by the following phonological environment, but is sensitive to the preceding phonological environment. Taeldeman states that final schwa deletes when the stem contains either a long vowel or a schwa-consonant sequence and the final consonant of the stem is a sonorant (1980). He describes instances where this rule fails to apply as "scaled somewhere between the extremes 'impossible' and 'common usage'" (Taeldeman, 1980:235). I take this to indicate that the rule is variable, which is consistent with the

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

data here. Notice that the indefinite determiner /ənə/ fits the structural description of the rule and is thus vulnerable to schwa-deletion.

Given the phonological facts discussed here, one clear possibility is that the change to a syncretic determiner system takes place through phonological erosion of word-final segments in two stages. First, the loss of /n/ on both definite /dən/ and indefinite /ənən/ could lead to an intermediate state of the language where there is a two-way gender distinction realized in both the definite and indefinite parts of the paradigm. This might be followed by the further loss of schwa on indefinite /ənə/ causing the complete syncretism of the indefinites. The following section, however, will be dedicated to a demonstration that the quantitative and spatial evidence from the GTRP data is not consistent with such an account.

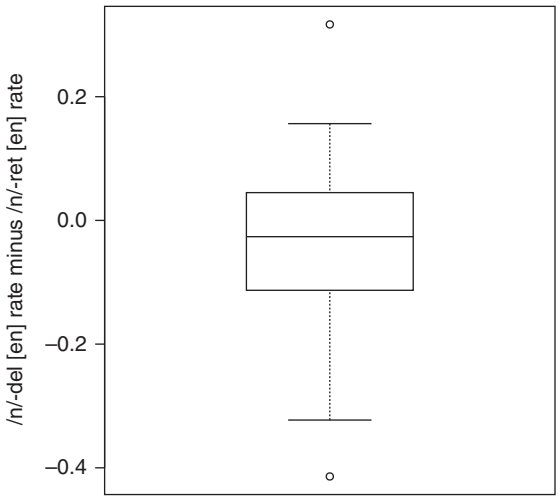


Figure 1. Boxplot of individual speaker differences between masculine [ən] rates in /n/-deleting and /n/-retaining phonological contexts.

4. Analysis

4.1 Masculine indefinite determiners

I begin by showing quantitatively that for masculine indefinite determiners, there is no pathway from the conservative underlying /ənən/ to surface [ən] via phonological processes. Between /n/-deletion and schwa-deletion, it might be possible to end up with a surface [ən] form after starting with an underlying /ənən/. To be more explicit, /n/-deletion might feed schwa-deletion so that a derivation could start with /ənən/, delete the final /n/ to get [ənə], and then delete the newly-vulnerable schwa to get [ən]. But if these hypothetical surface [ən] forms deriving from underlying /ənən/ exist, could they be disentangled from surface [ən] forms deriving from underlying /ən/? Because of the homophony of [ən] from underlying /ən/ and [ən] derived by deletion, it would be difficult if not impossible to distinguish them from a strictly theoretical viewpoint. The existence of derived [ən] in masculine environments, however, can be refuted quantitatively, relieving us of the problem of distinguishing between the hypothetical two types of surface [ən].

For a majority of the speakers surveyed in Brabant, the deletion of /n/ is a categorical process in the conservative masculine determiners; there is no [ənən] in the /n/-deleting environment and no [ənə] in the /n/-preserving environment (before /b/, /d/, /h/, and vowels as described in section 3). For 23 of the 51 speakers there is scattered variability, mostly in the direction of small amounts of /n/-deletion in the /n/-preserving environment. However, the question to be answered here is not what the exact conditions on /n/-deletion are but rather whether /n/-deletion feeds schwa-deletion. The quantitatively simplest path to answering this question is to examine only the speakers with consistent /n/-deletion behavior.

The surface form [ən] is attested in both the /n/-deleting and /n/-preserving environments.[3] In the /n/-preserving environment, the never-deleted final /n/ protects the second /ə/ in /ənən/ from deletion. Instances of surface [ən] in the /n/-preserving context can thus be confidently attributed to underlying /ən/, the innovative form. The rate of [ən] in this environment can be taken as the true rate at which the innovative underlying form /ən/ is selected. This rate can then be compared to the frequency with which [ən] is used in the /n/-deleting contexts. If forms in the /n/-deleting environment are further vulnerable to variable schwa-deletion, yielding surface [ən], the surface rate of [ən] in this context should be greater than in the /n/-preserving context. A proportion of the [ən]s equal to the rate of [ən] in the /n/-retaining context would result from underlying /ən/ in the /n/-deleting environment, and the extra ones could be attributed to schwa-deletion.

Figure 1, however, suggests that these extra instances of [ən] do not exist. This boxplot shows the by-speaker differences in the rates of surface [ən] between the /n/-deleting and /n/-retaining environments for the consistent /n/-deleters (five speakers who use exclusively innovative forms are also excluded). A Wilcoxon signed rank test fails to reject the null hypothesis that the differences have a mean of zero ($p = 0.23$). I conclude that there is no evidence for surface masculine [ən] forms derived from underlying /ənən/ via phonological erosion, and, therefore, that /n/-deletion does not feed schwa-deletion. This allows us to take the overall surface rate of [ən] across both

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

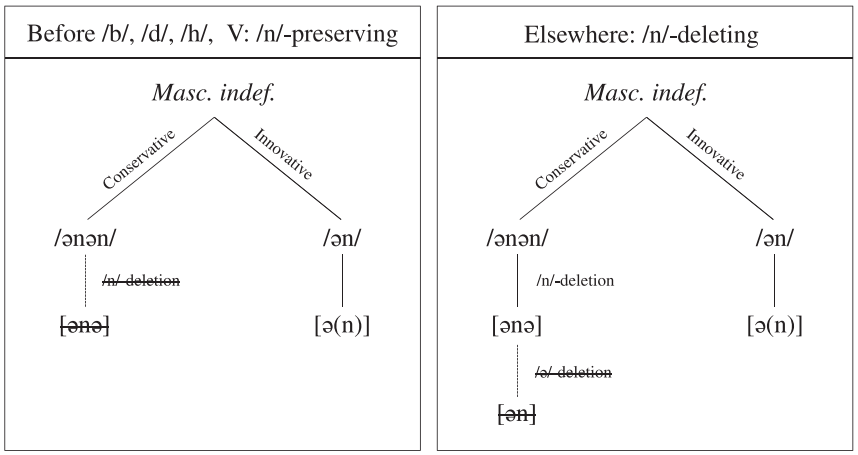
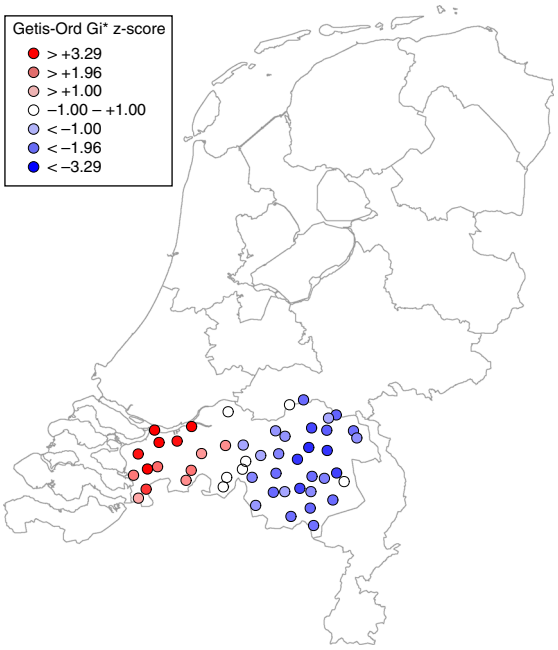


Figure 2. Mapping masculine determiners from underlying to surface forms in /n/-deleting and /n/-retaining phonological contexts.

contexts as the true underlying rate of the innovative form /ən/ with masculine nouns. Figure 2 illustrates how selection of an underlyingly innovative or conservative form maps to a unique surface form within each phonological context.

I now turn back to the full dataset. The mean masculine innovative rate calculated across the means for the 51 speakers in the sample is 33%. This mean obscures the fact that the individual speaker means range from a minimum speaker mean of 1% to a maximum of 100%. A spatial statistical analysis of the data, following Grieve, Speelman & Geeraerts (2011), reveals that these differences fall into a striking pattern. The Moran's I test produces evidence for significant global spatial autocorrelation of the rate of [ən] with masculine nouns ($I = 0.085$, $p < 0.0001$). I thus proceed to calculate the Getis-Ord G_i* scores for each location. In this case, low-G_i*-value clusters are ones where [ən] is used at low rates and high-value clusters are ones where [ən] is used at high rates. These scores, for the use of [ən] with masculine nouns, are mapped in Map 1.

The red dots on the map indicate locations that are similar to their neighbors in having high rates of /ən/, while blue dots indicate locations that are similar to their neighbors in having low rates of /ən/. The white dots are intermediate and may be considered transitional. We can see from Map 1 that the eastern half of the province is characterized by low use of the innovative /ən/ while the western half (and particularly the northwestern corner) is characterized by high use of innovative /ən/. In the middle of this east-to-west progression there seems to be an area that is transitional in nature, indicated by pale and white dots on the map. This kind of spatial patterning is consistent with a change in progress that has spread gradually through a transition zone. I suggest that the increasing use of the innovative form with masculine nouns is spreading across Brabant from (north)west to east; if so, then it is naturally most advanced in the area where it originated.



Map 1. Local spatial autocorrelation of /ən/ rates with masculine indefinite determiners.

4.2 Feminine indefinite determiners

Unlike the masculine indefinite determiners, the feminine ones show consistently high rates of surface [ən] across all speakers. Feminine surface [ən] rates range from 37–100% across speakers, but 46 of the 51 speakers have rates greater than 80%; the Moran's I test for global spatial autocorrelation finds no evidence

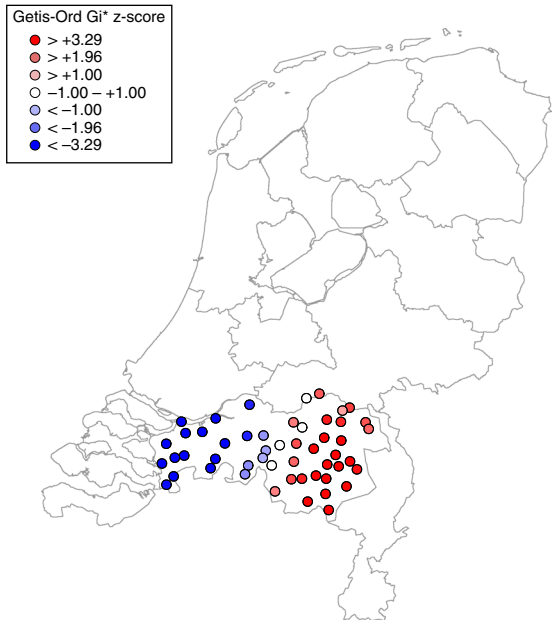
https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

for spatial clustering ($I = -0.021$, $p = 0.51$). While in the masculine case I showed that the surface [ən] forms could be uniquely attributed to the innovative determiner, in the feminine case we have every reason to believe that variable schwa-deletion can produce surface [ən] forms from underlying /ənə/. The surface rate of [ən] is not meaningful, then, because [ən] derived from conservative /ənə/ is indistinguishable from the morphologically innovative form. Since we are interested in the progression of the change towards the innovative form, we need to know what proportion of surface [ən] forms are truly /ən/ underlyingly so we can calculate the rate of use of the innovative determiner with feminine nouns.

One way to estimate this proportion is by first estimating the schwa-deletion rate. To estimate the schwa-deletion rate, I turn to an environment where the choice between innovative and conservative does not produce variability: the feminine adnominal adjectives. Because innovative and conservative agreement on feminine adjectives is the same, /-ə/, any observed variation can be attributed to phonological deletion. An examination of the data shows that the set of adjectives that allow schwa-deletion is larger than would be expected from Taeldeman's (1980) description, with 25 out of the 28 feminine adjective–noun items showing variation. Rather than hewing closely to the original phonological description counter to the attested variation, I include all adjectives that show variability across the province and exclude only the adjectives that show categorical schwa retention. The overall rate of schwa-deletion within the variable set of 25 adjectives is 38%. Once again, though, there is strong spatial conditioning of the rate of schwa-deletion across the individual locations in Brabant. The Moran's I test provides evidence for global spatial autocorrelation ($I = 0.242$, $p < 0.0001$) and the map of the local spatial autocorrelation G_i* scores is in Map 2.

We see in Map 2 that the western half of the province, which as we saw in section 4.1 is characterized by elevated use of the innovative masculine determiner, is also characterized by *lower* rates of schwa-deletion than the eastern half. The narrow transitional zone for schwa-deletion bisects the province along a similar midline as the transitional zone for use of the innovative masculine determiner. There is a significant negative Pearson's product-moment correlation of −0.58 between rates of the innovative masculine determiner and rates of schwa-deletion in feminine adjectives ($p < 0.0001$), further suggesting an inverse relationship between innovative determiners and schwa-deletion.

The observation that the rate of surface [ən] as a determiner with feminine nouns is stable across Brabant is misleading, then, because it obscures a

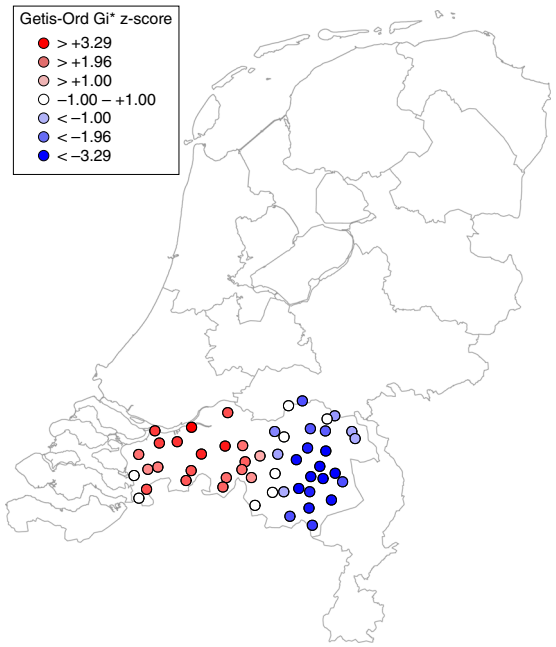


Map 2. Local spatial autocorrelation of schwa-deletion in feminine adjectives.

spatial pattern in schwa-deletion that implies an inverse spatial pattern in the use of innovative determiners with feminine nouns. If there is less schwa-deletion in a given area but the overall rate of surface [ən] stays the same, there must be more underlying /ən/ to make up the difference. The feminine determiners, like the masculine ones, must actually show higher rates of /ən/ in the western half of the province. We are now in a position to estimate the true rate of selection of the innovative form with feminine determiners.

I make the assumption here that schwa-deletion has a uniform rate across the adnominal domain, with schwa-deletion rates on adjectives being a reasonable estimate of schwa-deletion rates on determiners. For each location, the surface rate of [ənə] can be calculated directly from the determiner data. The surface [ənə] rate can then be combined with the estimated rate of adjectival schwa-deletion for that location in order to estimate the proportion of surface [ən] derived from underlying /ənə/ by schwa-deletion. The rate of surface [ən] generated by underlying /ən/ is then simply the remaining portion of the data, which is also the rate of selection of the innovative grammar. In other words, the schwa-deletion rate can be used to factor out the surface [ən] forms derived from underlying /ənə/ in order to estimate how often innovative /ən/ is selected. For example, consider a hypothetical speaker with 21 instances of [ənə] with feminine nouns, 79 instances of [ən] with feminine nouns, and a schwa-deletion rate of 25% based on the feminine adjectives. Those 21 [ənə]

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press



Map 3. Spatial autocorrelation of /ən/ rates with feminine indefinite determiners.

tokens are the 75% of underlying conservative determiners that "survived" schwa-deletion, meaning that there were 28 instances of the conservative determiner. The speaker's rate of selection of the innovative determiner, then, is not 79% but 72%. When this calculation is performed within each speaker's data, the Pearson's product-moment correlation between estimated by-speaker rates of feminine innovative /ən/ and schwa-deletion is even stronger than in the masculine context at -0.71 ($p < 0.0001$).

As expected, there is significant global spatial autocorrelation for this measure of the innovative feminine determiners according to the Moran's I test ($I = 0.206$, $p < 0.0001$). Map 3 shows that the normalized local spatial autocorrelation statistics are high in the west (mostly the innovative form) and lower in the east (more of the conservative form). These estimates for feminine determiners thus follow the spatial pattern of the masculine determiners closely. I split the locations into the eastern and western halves of the province by the median longitude and calculated the mean of the masculine and feminine innovative determiner rates for each to give an estimate of the central tendencies in the innovative and conservative areas. These means are given in Table 5. It is clear from the table, though, that it is not possible to completely unify the masculine and feminine contexts because the proportions of /ən/ are very different. A Wilcoxon signed rank test of the by-speaker differences between the masculine and feminine innovative determiner rates strongly rejects the null hypothesis that the difference is zero, with $p < 0.0001$. There is not a single rate at which speakers select the innovative determiner regardless of gender, but there is a good correlation between the masculine and feminine innovative determiner rates, with the Pearson's product-moment correlation equal to 0.64 ($p < 0.0001$).

Table 5. *Mean rates of /ən/ in the innovative (western) and conservative (eastern) halves of Brabant*

	Masculine	Feminine
West (innovative)	0.45 (0.39)	0.88 (0.12)
East (conservative)	0.20 (0.20)	0.69 (0.13)

5. Discussion

The finding that the change of the masculine indefinite article from /ənən/ to /ən/ is not phonologically mediated is inconsistent with the possibility that the syncretism in Dutch determiners results straightforwardly from the loss of final /n/ and /ə/. Although the qualitative description of the phonological processes at play would have suggested a pathway for any given determiner instance from /ənən/ to [ən] with [ənə] in the middle, the quantitative facts indicate that it is not possible to get from the conservative to the innovative masculine form by purely phonological means. Nonetheless, the use of the innovative form /ən/ increases progressively right up to near-completion in the more innovative area of Brabant, disregarding phonological contexts completely. In the feminine contexts, a plausible hypothesis was that variable schwa-deletion might play a role in eroding the feminine forms. Upon closer examination, though, it turns out that the use of the innovative form continues to increase while the rate of schwa-deletion actually *decreases*. The process of change implied by phonological-erosion accounts of syncretism, where sound change goes to completion and leaves homophony in its wake, is not supported here. Neither do we see evidence of an intermediate stage where the masculine and feminine have collapsed into /ənə/ prior to the complete loss of gender marking on indefinite determiners. Instead, it appears that the innovative form /ən/ is spreading across the province in both the masculine and feminine contexts in tandem.

There are several possibilities regarding the origins and diffusion of such a change that are compatible with the data presented here. One possibility is that the change in Brabant simply reflects convergence to Standard Dutch, although this does not in itself

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

provide an explanation for the geographic distribution of the innovative forms. Further inquiry into the demographics, language attitudes, and exposure of these areas to the standard variety would be required to assess the viability of this possibility. A second possibility is that within Brabant, the change is induced by contact with the Hollandic and Zeelandic dialects that border it to the west. This would be consistent with the direction of the observed geographic patterns, with the innovative form being borrowed first in the west and then spreading eastward, but does not explain the apparent inverse relationship between schwa-deletion and use of the innovative form. I suggest that high rates of schwa-deletion, as found in eastern Brabant, provide learners with robust evidence for the relationship between the underlying feminine determiner /ənə/ and its surface reflex [ən]. With an abundance of evidence for a variable phonological process that can relate /ənə/ to [ən], it is reasonable for a learner to simply assume that the two forms are related in that way. On the other hand, if the learner is exposed to a variety where schwa-deletion is rare, that learner might be more likely to attribute a surface form [ən] to a morphological difference, reanalyzing that token as an instance of an underlying form /ən/. Somewhat counterintuitively, a lower rate of schwa-deletion could prompt morphological reanalysis within the feminine context. This proposal might be integrated with a borrowing account if the phonological processes already at play in any given location are seen as relevant to the spread of innovations. Low rates of schwa-deletion, as in the western half of the province, might facilitate adoption of the neighboring syncretic system, while high rates could slow its spread eastward.

Explanations appealing to the role of schwa-deletion in the analysis of feminine forms do not explain the appearance of /ən/ in the masculine context. I suggest that the feminine indefinite determiner is the weak point in the system where the incoming indefinite determiner form /ən/ is innovated or borrowed. From the feminine context it may spread, perhaps even immediately, to the masculine context under the influence of analogy, phonological ambiguity, or feature structural pressures. The parallel patterns of local spatial autocorrelation across the masculine and feminine forms are strongly suggestive of a unified change where the gradual adoption of the innovative form proceeds as a matter of morphological competition with the conservative form irrespective of gender context. This suggestion resembles the Constant Rate Hypothesis proposed by Kroch (1989) for syntactic changes. As in the case studies where the Constant Rate Hypothesis was developed, the proportion of the incoming form is different across different contexts. Real time diachronic evidence would be needed to confirm that the rates of change are the same because it is not yet clear how to map geographic distances onto temporal differences. In other words, although this study rests on the assumption that spatial distributions may reflect historical changes, the relationship between geography and diachrony is likely to be sufficiently non-linear to make the mathematics of demonstrating constant rates difficult.

6. Conclusion

This case study of the development of syncretism in the Dutch indefinite determiners marshaled both quantitative and spatial evidence to argue for a complex relationship between phonology and morphology. Although an explanation where sound change leads straightforwardly to homophony seemed plausible prior to closer scrutiny, the evidence for such an explanation was not found. Instead there is evidence for a unified morphological change towards the innovative form in all contexts and a possible triggering role of phonology that runs counter to the expected direction of phonological influence. It would not have been possible to draw these conclusions from the quantitative surface distributions of the various determiner forms; a view of language variation patterns that might otherwise appear chaotic is brought into focus through the postulation of linguistic structure and processes at multiple levels.

Acknowledgements

I would like to thank Bill Labov, Dave Embick, Charles Yang, Kobey Shwayder, Hilary Prichard, Akiva Bacovcin, Joe Fruehwald, Daniel Ezra Johnson, Kyle Gorman, and Tom Ruette for their helpful discussion and technical assistance throughout the development of this project. I am also indebted to the Meertens Institute and everyone associated with the Goeman—Taeldeman—Van Reenan Project for making their data available, as well as to two anonymous reviewers for their comments.

Notes

[1] An anonymous reviewer points out that the relationship between global and local spatial autocorrelation is not entirely straightforward; the reader is referred to Goeman (1999:ch.5) for some discussion of the relevant issues.

[2] For all of the nasal-final determiners and adjectival inflections discussed here, some dialects show nasal place assimilation. I abstract away from such assimilation throughout, treating forms like [əm boon] ('a bean') as equivalent to [ən boon].

[3] The surface form [ə] also appears; /n/-deletion appears to apply variably to /ən/ rather than categorically. I treat [ən] and [ə] together as instances of [ən].

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press

References

Baerman, Matthew. 2006. Syncretism. In K. Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., 363–366. Amsterdam: Elsevier.

Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand. 1993. Some patterns of linguistic diffusion. *Language Variation and Change* 5(3):359–390.

Barðdal, Johanna & Leonid Kulikov. 2009. Case in decline. In A. Malchukov & A. Spencer (eds), *The Oxford handbook of case*, chapter 30. Oxford: Oxford University Press.

Becker, Richard A., Allan R. Wilks (original S code), Ray Brownrigg (R version) & Thomas P. Minka (enhancements). 2012. *maps: Draw Geographical Maps.* R package version 2.2-6. http://CRAN.R-project.org/package=maps

Berg, B.L. van den. 2003. *Phonology and morphology of Dutch and Frisian dialects in 1.1 million transcriptions*. Goeman-Taeldeman-Van Reenen project 1980–1995, Meertens Instituut Electronic Publications in Linguistics 3. Amsterdam: Meertens Instituut. CD-ROM.

Bivand, Roger, with contributions by Micah Altman, Luc Anselin, Renato Assunção, Olaf Berke, Andrew Bernat, Guillaume Blanchet, Eric Blankmeyer, Marilia Carvalho, Bjarke Christensen, Yongwan Chun, Carsten Dormann, Stéphane Dray, Rein Halbersma, Elias Krainski, Pierre Legendre, Nicholas Lewin-Koh, Hongfei Li, Jielai Ma, Giovanni Millo, Werner Mueller, Hisaji Ono, Pedro Peres-Neto, Gianfranco Piras, Markus Reder, Michael Tiefelsdorf & Danlin Yu. 2012. *spdep: Spatial dependence: weighting schemes, statistics and models*. R package version 0.5-51. http://CRAN.R-project.org/package=spdep

Chambers, Jack & Peter Trudgill. 1998. *Dialectology*, 2nd edn. Cambridge: Cambridge University Press.

Goeman, Ton. 1999. *T-deletie in Nederlandse dialecten: kwantitatieve analyse van structurele, ruimtelijke en temporele variete*. Amsterdam: Holland Academic Graphics Vol. 26.

Goeman, Ton, Marc van Oostendorp, Pieter van Reenen, Oele Koornwinder, Boudewijn van den Berg & Anke van Reenen. 2008. *MAND: Morphological atlas of the Dutch dialects: Volume II*. Amsterdam: Amsterdam University Press.

Goossens, Jan, Johan Taeldeman & Geert Verleyen. 1998–2005. *Fonologische atlas van de Nederlandse dialecten Volumes I-IV*. Gent: Koninklijke Academie voor Nederlandse Taal en Letterkunde.

Grieve, Jack, Dirk Speelman & Dirk Geeraerts. 2011. A statistical method for the identification and aggregation of regional linguistic variation. *Language Variation and Change* 23:193–221.

Grieve, Jack, Dirk Speelman & Dirk Geeraerts. 2013. A multivariate spatial analysis of vowel formants in American English. *Journal of Linguistic Geography* 1:31–51.

Heap, David. 2000. *La variation grammaticale en géolinguistique: les pronoms sujets en roman central*. München: Lincom Europa.

Herzog, M. 1969. Yiddish in the Ukraine: Isoglosses and historical inferences. *The Field of Yiddish* 3:58–81.

Kemp, Samuel E. & Gregory Jefferis. 2011. *RANN: fast nearest neighbour search*. R package version 2.1.3. http://CRAN.R-project.org/package=RANN

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.

Labov, William. 1994. *Principles of linguistic change. Vol. 1. Internal factors*. Oxford: Blackwell.

Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin/New York: Mouton de Gruyter.

Moran, Patrick. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society*, Series B 37:243–251.

Moulton, William G. 1962. Dialect geography and the concept of phonological space. *Word* 18:23–32.

Oostendorp, Marc van. Forthcoming. Phonological and phonetic databases at the Meertens Institute. Draft chapter for a handbook.

Odland, John. 1988. *Spatial autocorrelation*. Beverly Hills/ London/New Delhi: Sage Publications.

Ord, J. Keith & Arthur Getis. 2001. Testing for local spatial autocorrelation in the presence of global autocorrelation. *Journal of Regional Science* 41(3):411–432.

Pebesma, Edzer J. & Roger S. Bivand 2005. *sp: Classes and methods for spatial data in R*. R News 5(2). http://cran.r-project.org/doc/Rnews/

Schutter, Georges de, Boudewijn van den Berg, Ton Goeman & Thera de Jong. 2005. *MAND: Morphological atlas of the Dutch dialects: Volume I*. Amsterdam: Amsterdam University Press.

Taeldeman, Johan. 1980. Inflectional aspects of adjectives in the dialects of Dutchspeaking Belgium. In W. Zonnerveld et al. (eds), *Studies in Dutch Phonology*, volume 4 of *Dutch Studies*, 265–292. Den Haag: Martinus Nijhoff.

Tamminga, Meredith. 2012. Interpreting the spatial distribution of quantitative variation as evidence for mechanisms of morphological change. In *Pre-Proceedings of the International Conference Linguistic Evidence 2012: Empirical, Theoretical, and Computational Perspectives*, 95–99.

Wulf, Chris & Johan Taeldeman. 2001. Apocope en insertie van -n na sjwa in de zuidelijke nederlandse dialecten: Conditionering en geografie. *Taal en Tongval Themanummer* 14:7–51.

https://doi.org/10.1017/jlg.2013.10 Published online by Cambridge University Press