Interpreting the spatial distribution of quantitative variation as evidence for mechanisms of morphological change

Meredith Tamminga
University of Pennsylvania

Introduction

This paper argues that the spatial distribution of quantitative linguistic data can provide evidence for or against different accounts of language change. As a case study I investigate possible transitions between a morphologically rich determiner paradigm and a more syncretic one in Dutch. By linking expected intermediate stages of the change to explicit spatial predictions, I argue for an account based on grammar competition rather than morphological reanalysis or lexical diffusion. The results illustrate the promise of geographic evidence for illuminating significant diachronic questions.

Background

The conservative Dutch determiner system has a gender distinction between masculine, feminine, and neuter nouns marked on both definite and indefinite singular determiners, labeled /ənən/ (m.), /ənə/ (f.), and /ən/ (n.). In the innovative two gender system, characteristic of Standard Northern Dutch, the distinction is no longer marked on the indefinite determiners (with /ən/ used for all nouns).

I outline three possible accounts of the mechanism of change. The first, following proposals on verbal syncretism by, e.g., Aalberse 2007, is a multi-step process of morphological reanalysis, driven primarily by feature structure. The second is a probabilistic competition between a conservative and an innovative categorical grammar, the latter being selected increasingly often until it replaces the conservative one (Kroch 1989). The third account is lexical diffusion, whereby /ən/ as the indefinite determiner for all nouns regardless of gender spreads through the language one word at a time.

Methodology

The data come from the publically-available corpus of the Goeman-Taeldeman-Van Reenan Project (the basis of the Morphological Atlas of Dutch Dialects). I include 75,861 data points, with up to 386 phonetically-transcribed questionnaire items of the form *indefinite determiner + noun* from each of 276 locations in Dutch-speaking Belgium and the southern provinces of the Netherlands.

I use the percentage of /ən/ (%-/ən/) determiners to measure how advanced the determiner syncretism is in each location, and divide the geographical area into three regions based on their %-/ən/. The area with 52-74%-/ən/ is considered the 'transition zone' (the eastern half of the Dutch province of Brabant and a narrower band along the western Dutch-Belgian border.); higher and lower percentages define the 'innovative' and 'conservative' zones respectively.

My analysis focuses on masculine nouns, since masculine indefinite /ənən/ is (both phonologically and morphologically) the most distant from incoming /ən/. As a result, the determiners of masculine words provide the greatest potential for differentiation across multiple stages of change.

Predictions

For each of the three models of change outlined above, I make explicit what intermediate stages should be involved in the change, then look for geographic evidence for or against the existence of such stages as evidence of which model predicts them. For Romance pronouns, Heap argues that investigating the geographic transition between two different systems can contribute to understanding how that change took place (2000:41). He proposes that spatial patterns may reflect the residue or leading edge of change in progress, revealing intermediate stages as the change spreads geographically.

The morphological reanalysis proposal, given in a Distributed Morphology framework, posits an intermediate grammar with a two-way (common/neuter) indefinite determiner distinction, parallel to that of the definite determiners in the innovative system. The feature structure and the phonological input should conspire to yield the /ənə/ form. Thus, masculine words with /ənən/ in the conservative region and /ən/ in the innovative region should have /ənə/ in the geographic transition zone.

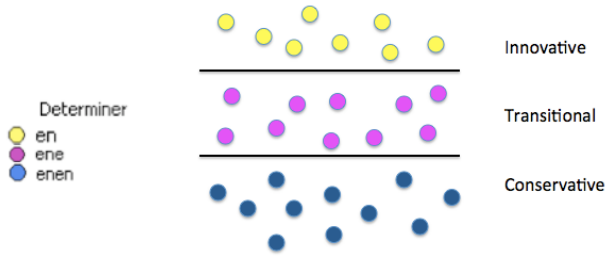

Figure 1. Schematic prediction from morphological reanalysis: masculine nouns take the indefinite determiner /ənə/ in the transition zone.

The competing grammars account predicts the maintenance of a consistent /ənə/:/ənən/ ratio as the overall rate of /ən/ increases, because the genders of the nouns in the conservative grammar should be the same whenever it is selected. At the same time we expect a random spatial distribution of /ən/ with masculine nouns in the transition zone, reflecting the probabilistic nature of the grammar selection.

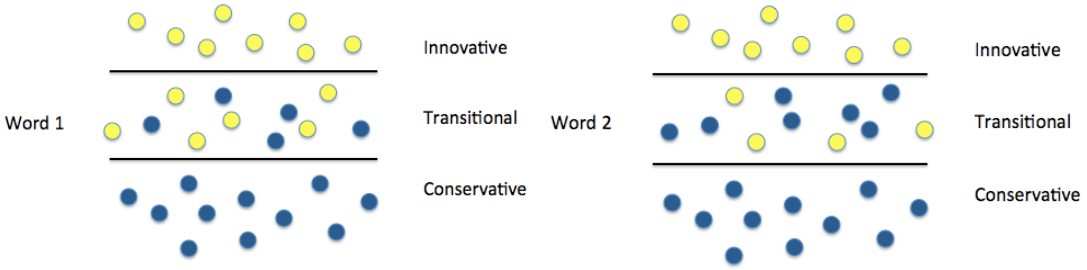

Figure 2. Schematic (partial) prediction from competing grammars: masculine nouns take /ənən/ or /ən/ in a spatially-random way in the transition zone.

Finally, lexical diffusion predicts coherent conservative and innovative areas for each word, but non-coincident lexical isoglosses.

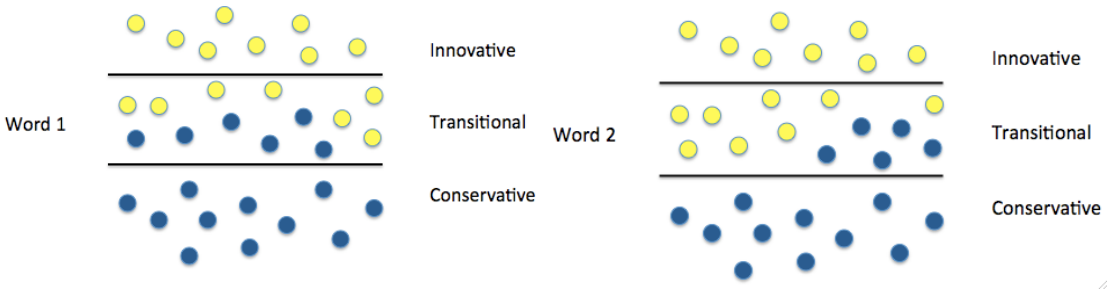

Figure 3. Schematic prediction from lexical diffusion: masculine nouns show contiguous but unique distributions between /ənən/ and /ən/, with isoglosses in the transition zone.

Results

I present multiple maps of two types: overall quantitative patterns, and individual lexical items. To compare the predictions of morphological reanalysis and competing grammars I map [(# /ənə/ tokens) ÷ (# /ənə/ tokens + # /ənən/ tokens)] at each location. Under morphological reanalysis assumptions this metric should peak to near 100% through the transition zone as /ənə/ and /ənən/ collapse into /ənə/, but under competing grammars assumptions it should remain consistent across all three zones. With only a small scattering of localized peaks and an otherwise homogeneous distribution, shown in Figure 4, this quantitative evidence is most consistent with competing grammars.

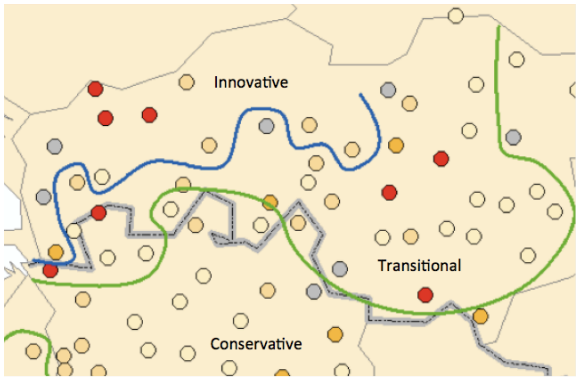

Figure 4. Percent /ənə/ tokens out of all non-/ən/ tokens. Red dots are highest percent; grey dots indicate zero denominator (100% /ən/).

Turning to individual lexical items, I compare the questionnaire responses for ten coronal-stop-initial masculine nouns on ten separate maps. It is immediately clear that the spatial distribution of the determiners conforms to the predictions of competing grammars, as illustrated in Figure 5.

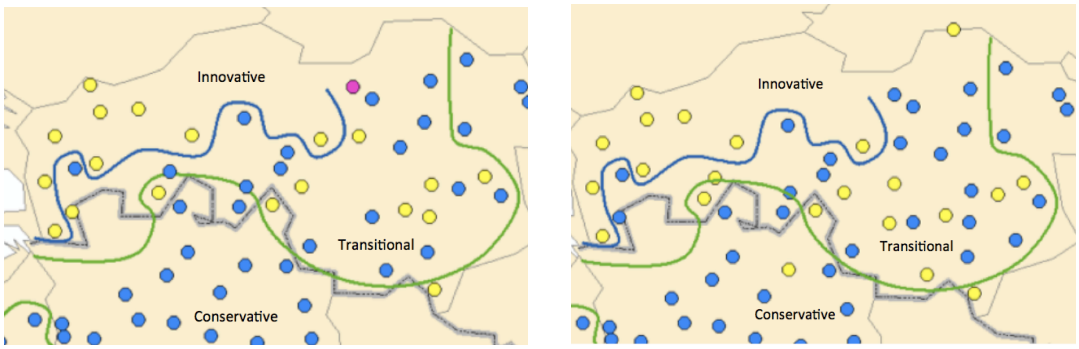

Figure 5. Determiner forms for *doek* 'cloth' (left) and *duim* 'thumb' (right).

The spatial data, both on the aggregated scale and the individual lexical scale, strongly support an analysis based on a grammar competition model of change, under which speakers choose probabilistically between a fully conservative grammar (which yields /ənən/ for a masculine noun) and a fully innovative grammar (which yields /ən/ for a masculine noun).

Discussion

These results illustrate the utility of converting theories of language change into concrete spatial predictions based on the identification of intermediate stages in a change. No evidence for lexical diffusion is observed, as all ten lexical items show a mix of conservative and non-conservative forms in the transition area. Additionally, we see no spatial reflexes of a transition through an intermediate stage where masculine nouns consistently have /ənə/, as predicted by the morphological reanalysis account. Instead, compelling geographic evidence exists for the competing grammars mechanism of morphological change. Pursuing this type of approach in other cases promises to bring dialectological data to bear as a useful source of evidence on questions of theoretical relevance.

References

Aalberse, S. 2007. The typology of syncretisms and the status of feature structure. Verbal paradigms across 355 Dutch dialects. *Morphology* 17: 109-149.

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

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

Van den Berg, B.L. 2003. *Phonology & Morphology of Dutch & Frisian Dialects in 1.1 million transcriptions. Goeman-Taeldeman-Van Reenen project 1980-1995*. Meertens Instituut Electronic Publications in Linguistics 3, Amsterdam: Meertens Instituut.