

The Locus of Spelling Variation in Dutch First Names

Doreen Gerritzen (Meertens Instituut Amsterdam), Anneke Neijt (Dutch Department Nijmegen) and Marc van Oostendorp (Meertens Instituut Amsterdam)

First names in the Netherlands provide us with a rich source of spelling variation, hence with a window on the linguistic rule system that underlies Dutch orthography. In this paper, we demonstrate how a combination of corpus searching and a small-scale experiment sheds light on a well-known generalisation on Dutch orthography: spelling adaptations tend to start out at the righthand periphery of the word. Furthermore, we demonstrate how ‘purely’ orthographic principles interact with principles inherent to the process of naming (such as naming a child after a relative or some other person, or choosing a spelling variant for ‘aesthetic’ reasons) in the choice of a spelling for a given name. In this paper, two separate strands of research are therefor brought together. First, this concerns research done within the field of name scholarship on modern namegiving to children in the Netherlands. In the second place this concerns a tradition of research on the linguistic principles of spelling of (loan)words in Dutch.

The reasons why names are such a rich source of spelling variation are threefold. First, other than for ‘normal’ words there is no official spelling for (first) names in Dutch. Even though there demonstrably are spelling conventions for names, parents are free in principle to write the name of their children as they like. Secondly, using an ‘irregular’ spelling seems to be one way to give an ‘original’ name to one’s child. Thirdly, there has been quite an influx of ‘new’ names in the Netherlands during the past few decades. All in all, we can observe a kind of near-anarchy in the space of first names that is an ideal testing bed for research on spelling variation.

Recent investigations into the spelling system of Dutch have been directed towards the spelling differences between native and non-native words (cf. Nunn 1998). Most scholars seem to agree that the relevant labels ‘native’ versus ‘non-native’ should not be interpreted in an etymological way, but rather purely formally. Roughly speaking, stems containing exactly one full vowel (with possibly additional schwas in other syllables) count as ‘native’, even if they are etymologically allochtonous (e.g. *kelder* ‘cellar’ < Latin *cellarius*). On the other hand, words with more than one full vowel count as non-native, even if they do not occur in any other language (e.g. *havo*, a type of secondary school; the word is originally an acronym).

The spelling rules for the two types of words are slightly different. Most interesting for us is the process of adaptation of non-native forms to native ones. One of the regularities in this adaptation (already noted by Te Winkel, 1865:194, and related to the adaptation of the pronunciation) is that letter symbols in the final syllable of a word tend to get adapted to the Dutch spelling system sooner than those in previous syllables. For instance, the sound [i] has been ‘normalised’ to <ie> in the third syllable of <politie> ‘police’, but not in the second one; nor did the adaptation take place in <politieel> ‘police, Adjective’, where <i> is not in the final syllable.

The question now arises whether we find the same regularity in the spelling of first names. The interesting cases here are those names in which there

is more than one locus of potential language variation. An example of this is the name which is pronounced in Dutch approximately as [ɑnzəlɪk]. A purely etymological spelling (rendering the original French) would be <Angélique>; a purely phonological spelling would probably be something like <Anzjeliek>. If we ignore the diacritic acute accent for the sake of simplicity, we thus have two loci of potential variation: the sound [ʒ] and the final syllable [ɪk].

We can now try to find spelling variants of this name in our corpus (the SVB corpus named after the Dutch 'Social Security Bank' which collected the material). This corpus contains the names of all the children born in the period 1996-1998 for which children's allowance was applied. The database contains 46,672 different names given to 602,111 children. A search in this corpus of the name [ɑnzəlɪk] reveals that there are three variants (again disregarding diacritics): <Angelique>, <Angeliek> and <Ansjelique>.

The first of these three forms is completely faithful to the French original, the second one has been adapted as we would expect it based on the rules of Dutch spelling, i.e. the final syllable takes the form found in native words. The third form, however, poses an interesting puzzle, because it seems to have been adapted in the middle but not at the end. Interestingly, however, it did not adapt in the exactly right way, phonetically: it has <sj> rather than <zj>. (Although a caveat is in order here, since some dialects of Dutch do not distinguish between voiced and voiceless fricatives.) Now it turns out that the combination <Ansje> is in itself a Dutch name, a diminutive of Anna which used to be very frequent, present in the corpus. It therefore seems that special naming conventions (e.g. naming a child after somebody else, or combining two names that are considered fitting for other reasons) 'overrule' the usual conventions of spelling in this case.

In our talk we present data from other names in the SVB database in which we find spelling variation in more than one locus. It is shown that most of the preferences that become apparent in studying the database can be understood as a result of the interaction between rather well-known linguistic principles of spelling (such as the rule that spelling adaptation should start at the end of the word or that the spelling of a word follows either the native or the non-native convention, but does not mix them) on the one hand and naming conventions on the other.

Using a written database in order to investigate spelling variations turns out to have some methodological problems attached to it, however. Most important among these is that it is often unclear whether a given variant should be attributed to spelling variation 'proper' or should rather be seen as a result of phonological variants. An example of this is the name which could be spelled as <Caroline> or <Carolien> (as well as with an initial <K>). Even though these two spelling variants seem to be often pronounced in the same way — with final [i] — there is also a possibility of pronouncing the first form as [i]. In that case the difference between the two names no longer is one purely of spelling.

In order to circumvent these and other problems, we have conducted an experiment, in which pairs of names with two loci of potential variation such as <{C/K}arol{ine/ien}> or <An{g/sj}el{ique/iek}> were read aloud to a number of subjects. At the same time, two spelling variants were shown to the subjects (e.g. <Caroline>, <Carolien>; <Angelique>, <Ansjeliek>). The subjects were asked to

express a preference for one of the two names ('Which name would you give to your own child?') on a five-point scale. In our talk, we present the results of both the database search and the experiment and discuss their implications for the theory of spelling, and the special case of spelling of first names.

References

Nunn, A. 1998. *Dutch Orthography*, The Hague.

Te Winkel, L.A. 1865. *Grondbeginselen der Nederlandsche spelling*, Leiden.