Asymmetrical phoneme–grapheme mapping of coronal plosives in Dutch*

Anneke Neijt and Robert Schreuder
CLS and NICI, Radboud University Nijmegen

The distinction between deep and shallow orthographies is a central issue in studies of alphabetic writing. This paper aims to contribute to the debate on the relative merits of these by investigating how the coronal plosives [d] and [t] map onto the corresponding letters d and t. It turns out that an interesting asymmetry exists in word medial position in Dutch: both experienced users and children learning to write Dutch prefer ds over ts. Several mutually not always exclusive alternative explanations are provided. The fact that the d-preference is sensitive to phonological and morphological distinctions and may be influenced by speech rate, suggests that a deep orthographic representation might be the better option.

1. Introduction

Spoken language is a rich and complex system of forms, with the potential to express phonetic detail and shades of meaning which cannot be represented in alphabetic writing. By necessity, written forms are but approximations, abstractions which most often underdetermine the spoken form and sometimes overrepresent it.

All spelling systems must deal with these limitations in some way, and must choose between representations that follow the spoken forms as closely as possible, or more abstract forms that represent certain features only partly or indirectly, on a case by case basis, if not as a matter of principle. We will elaborate on the distinction between deep and shallow orthographies — underdetermination and overrepresentation — in the conclusion. This paper aims to contribute to the debate by investigating the curious ways in which the coronal plosives [d] and [t] map onto the corresponding letters d and t. The mapping seems to be straightforward in words such as dak ‘roof’ and tak ‘branch’, but turns out to be less straightforward in words such as kader ‘cadre’ and kater ‘male cat’, or wanden ‘walls’ and wanten ‘mittens’.

ISSN 1387–6732 / e-ISSN 1570–6001 © John Benjamins Publishing Company
It is relevant to distinguish word medial position from word onset position, because in recognizing words, language users rely on information at the beginning of words rather than on information in the middle or at the end of words. See for details e.g. Marslen-Wilson (1987) who explains these results in terms of the ‘Cohort Model’. In this model, words are selected from the mental lexicon on the basis of ever diminishing sets of candidates, the ‘cohorts’. Information from word onsets forms the first step in restricting the cohort, and information from the nucleus, the rime, and later syllables comes next in the selection process. Thus, the Cohort Model explains why longer words can be recognized before the whole word has been pronounced. But it also predicts precision of articulation at word onset positions, and greater sloppiness in medial and final positions.1

In line with the Cohort Model, learning to write begins with discriminating segmental oppositions in the onsets of words. In word initial contexts, the mapping is straightforward, as shown by the English pairs dab – tab and dagger – tagger, or Dutch dak ‘roof’ – tak ‘branch’ and deken ‘blanket’ – teken ‘sign’. In word medial position, however, the mapping of coronal plosives onto d and t is not so clear-cut, as is demonstrated in the acquisition by English speaking children of the spelling of words like city, dirty and even sometimes. With some regularity, such children will substitute d for t, most readily in city, and less often in dirty, on account of the related form dirt, in which the [t] is easily distinguished. Least likely to be misspelt is sometimes, because in this word the [t] is not rendered as a voiced rather than voiceless so-called ‘flap’, but fully pronounced (cf. Treiman, Cassar & Zukowski (1994)).

Dutch shows similarly vague contexts for t and d. Learning to spell d and t begins with pairs like dak ‘roof’ and tak ‘branch’, not wanden ‘walls’ and wanten ‘gloves’. The latter are neither as numerous nor as prototypical, nor as precisely pronounced and hence less informative than pairs with onset variation. In fact, choosing between d and t is so notoriously difficult in this language, that it is a matter requiring special attention in classrooms. Here are some of the reasons why (these examples are standard colloquial Dutch, not substandard):

- Dutch has d-t-allophony in function words (1a)
- Dutch has d-t-allophony in coda position (1b)
- Dutch has d-t-allophony in the past tense suffix of verbs (1c)

(1) a. [dəsmoj] – Dat is mooi! ‘That’s great!’
   [istat] – Is dat? ‘What is that?’

 b. [hant, hando] – hand, handen ‘hand, hands’

 c. [bloza, blozdɑ] – blozen, bloosde ‘to blush’
   [plasɑ, plasta] – plassen, plaste ‘to pee, to urinate’
The allophony in coda position is a matter of final devoicing, which operates across the board in Dutch. Since all word final *ds* are pronounced [t] anyway, the spelling of the singular form *hand* (1b) might equally well have been *hant*, following the phonology. But it so happens that the more abstract form echoing the morphological make-up of nouns is part of the Dutch standard spelling — *hand* in the case of 'hand'. Here, the constant representation of morphemes, pronunciation permitting, takes precedence over the phonological strategy of consistently representing the same sounds by the same letters.

Whereas the spelling of singular nouns is a matter of a clear choice, spelling of Dutch weak (i.e., regular) past tense forms, formed by adding the suffix *[da]* or *[ta]* to the verb stem, presents real problems (Ernestus & Baayen 2001). For one thing, in these medial contexts it is by no means always clear what coronal plosive is pronounced and should be written, just as with English *city* and *dirty*. For instance, pronunciation gives hardly a clue as to why the past tense of *krabben*, 'to scratch' should be spelled *krabde* rather than *krapte*. The distinction between *t* and *d* in this context is explicitly taught at length at primary school, with dubious results (Ernestus & Mak 2004).

In certain loan verbs in which pronunciation vacillates between [t] and [d] even though the spelling is fixed, both past tense forms are considered correct. Thus, from *leasen* 'to lease' we get *leasede* next to *leasete*; *golfen* 'to play golf' yields both *golfde* and *golfte*; and *bridgen* 'to play bridge' gives rise to both *bridgede* and *bridgte* (and the non-standard forms *bridgde* and *bridgte*). Even greater leniency obtains in certain native verbs, which may have two parallel spellings with concomitant past tenses. Examples are *sausen*/*sauzen* 'to pickle/to whitewash' and *plonsen*/*plonzen* 'to splash' with past tenses *sausde*/*sauste* and *plonsde*/*plonste*.

Not surprisingly then, Dutch verb spelling is especially difficult to learn, and this aspect of orthography has held a prominent position in research on spelling during the past half century as a consequence (Van de Velde 1956, Van Heuven 1978, Zuidema 1988). It turns out that even highly educated writers make mistakes in choosing *d*, *t* or *dt* (a special case outside the scope of this article) in the many homophonous contexts of Dutch. There are traces of storage, since errors are more frequent with rare forms (Assink 1983, Verhoeven 1985, Sandra et al. 1999, Frisson & Sandra 2002), and the distribution of *d* and *t* in the middle of words is in large part determined by analogy (Ernestus & Baayen 2003). One could suggest that, given this state of affairs in Dutch orthography, better concepts of teaching verb spelling should be designed. However, despite several efforts to improve Dutch spelling education, the spelling of Dutch verbs remains notoriously difficult. See Van de Velde 1956, Assink 1983, and Zuidema 1988.
2. The trouble with *ds* and *ts*: Three experiments

The above sketch of the predicament of Dutch learners prompts the question of what spelling system would be easiest and least error-prone for Dutch past tense suffixes: the existing one with its phonologically based forms, which represents the word medial coronal obstruent as either *t* or *d*, a surface representation, or an alternative with a single, more abstract ‘deep’ form, rendering all word medial plosives as *t*, or all word medial plosives as *d*, regardless of their phonetics. That question can be addressed, when we gain insight into what difficulties children encounter in choosing *t* or *d* in non-initial positions, and whether and if so, how, the morphological and phonological context is involved.

To do so, we investigated the use of *d* and *t* in the middle of words and followed by schwa, in both children learning to read and write and experienced readers/writers, i.e. adolescents and adults.

We collected past tenses of verb forms and words that hitherto have not been investigated, i.e., nouns and numerals with a phonological structure that resembles the phonological structure of such verbs. For instance, *zesde* ‘sixth’, and *dikte* ‘thick’, in which -*de* and -*te* are presumably pronounced in the same way as the past tense suffixes -*de* and -*te* in the verbal forms *reisde* ‘travelled’, and *pikte* ‘stole’.

We added plural nouns such as *woord* ‘words’, and *vuisten* ‘fists’ that resemble verb forms like *hoorden* ‘heard’, and *ruisten* ‘rustled’. The difference is that in these nouns the consonants [d] and [t] belong to the stem, not to a suffix. That way we explored the possibility that spelling segments of stems of words is easier or more difficult than spelling segments of suffixes.

This resulted in three experiments. Two of them aimed to invoke spelling choices and errors from primary school children learning to read and write. These setups could not be used with adolescents and adults, since experienced writers simply make too few spontaneous errors. Therefore, a third experiment eliciting judgments about specific errors was designed especially for experienced writers.

2.1 Experiment 1: Children in primary school — verbs

Above, we observed that Dutch adults meet difficulties with the choice between -*te* and -*de* when they form past tenses of loan verbs such as *leasde/leaste*. In this first experiment we investigate the performance of children with *d* and *t* in the past tense suffix of native verbs.

*Participants.* Forty-one Dutch children in the first year of primary school (Dutch group 3, 6–7 years of age). The experiment was carried out in March, after half a year of spelling education.
Materials. Forty verbs in the past tense together with pictures of the actions they signified were presented to the children (cf. (2) below). Only frequently used words, familiar to primary school children, were selected. (See Appendix A for the materials.)

(2) a. 20 verbs with -de as past tense suffix, such as belde 'called'.
   b. 20 verbs with -te as past tense suffix, such as blafte 'barked'.

Every item was presented on a separate page. The children thus could concentrate on one word at a time and would be minimally influenced by choices for earlier words. In the upper left and right corner of each page the letters d and t were printed, and in the printed word itself, there was a space where one of these should go. A picture was printed above each test word, to eliminate any doubt as to the identity of the word-with-a-hole-in-it. For instance, teken…e (draw) was illustrated with a picture of a pencil drawing a picture, and blaf…e (barked) came with dogs. The test items were gathered into booklets in two randomized versions, presented in reversed order as well.

Information on the frequency of d and t in the relevant contexts was gathered by summarizing the type and token frequencies in Celex for all words that form rimes with the test items, i.e. all words ending in -de(n) and -te(n) with similar preceding rimes. For instance, for the d-contexts of belde we calculated the sum of the frequencies of words such as schelden, schelde, vermelde and vermelden, all ending in [εldәn] or [εldә], and for its t-contexts we calculated the sum of the frequencies of words such as smelten, Kelten and stelten, ending in [εltәn] or [εltә]. In the set of verbs written with d, the d-contexts outnumber the t-contexts (type frequencies d 5354, and t 116; token frequencies d 340094, and t 14025). In the set of verbs written with t, no words occur that are spelled with d (type frequencies d 0, and t 1802; token frequencies d 0, and t 169363).

Procedure. Children were instructed to indicate the missing letter by circling either the d or the t printed at the top of the page. The instructor told the children that they had to choose a letter, even when they were not sure about their choice. Furthermore, she told the children that the results would be used for research, and that this exercise was not a test.

Results and discussion. The children found the task difficult, but not impossible. There were no indications that they did not understand the words printed and illustrated by the pictures. The error percentages are presented in Table 1.

This table shows that the choice between d and t in the middle of words is by no means obvious for the children after seven months of spelling instruction. Interestingly, as for English (Treiman e.a. 1994), we find more errors in words written with t than in words written with d. Although the spelling of verbs has been
investigated intensively, none of the investigations we mentioned above noticed this asymmetry.

2.2 Experiment 2: Children in primary school — nouns and numerals

The asymmetry of the phoneme–grapheme mapping, henceforth called ‘d-preference’, will be investigated in the remainder of this paper. Our second experiment investigates whether the observed asymmetrical pattern of errors is present as well in words other than verbs. We expect to find a similar preference, and perhaps a role of morphology.

Participants. Two groups of children took part in the experiment. Twenty-five Dutch children in the first year of primary school (Dutch group 3, 6–7 years of age) and 22 children in the second year (7–8 years of age). The experiment was carried out at the end of the school year, in early June.

Materials. In this experiment we contrasted words with d or t in the stem with words in which d or t is part of the suffix, cf. (3). When available, the form without final n was selected. This final n is usually not pronounced, and there is no reason to assume that the presence or absence of this letter could influence the results. Only frequently used words were selected, that would be familiar to primary school children. (Appendix B contains the full set of test items.)

(3) a. 10 nouns such as borden ‘plates’, with d in the stem (bord + plural suffix -en)
   b. 10 nouns and numerals such as liefde ‘love’ with d in the suffix (lief ‘dear’ + nominalizing suffix -de) and tiende ‘tenth’ (tien + ordinal suffix -de).
   c. 10 nouns such as harten ‘hearts’, with t in the stem (hart + plural suffix -en).
   d. 10 nouns such as ziekte ‘illness’, with t in the suffix (ziek ‘ill’ + nominalizing suffix -te).

As in Experiment 1, every item was presented on a separate page with a picture illustrating the word. The test items were gathered into booklets in four semi-randomized versions. The type and token frequencies of the nouns were calculated as before, e.g., as the number of, and the sum of the frequencies of all words ending...
in the same prefinal rime, followed by -de(n) or -te(n). For instance, for borden we calculated the sum of the frequencies of words such as worden, orde, morden, but also words ending in -te(n) such as sporten.

In the set of nouns that should be written with a $d$, the number of $d$-contexts is larger than the number of $t$-contexts (summed type frequency $d$ 8739, and $t$ 1002, summed token frequency $d$ 882255, and $t$ 88265). In the set of nouns that should be written with $t$, the type counts with contexts containing $t$ slightly outnumber the contexts with $d$. With respect to token frequencies, however, the number of $d$-contexts is larger once more (summed type frequency $d$ 1605, and $t$ 1653, summed token frequency $d$ 276835, and $t$ 140744). Based on these figures, one might expect a $d$-bias in the performance of children.

Procedure. The procedure was identical to that of Experiment 1.

Results and discussion. The error patterns are presented in Table 2.

Table 2. Experiment 2: Percentages of errors (variance between brackets)

<table>
<thead>
<tr>
<th>Condition</th>
<th>6–7 years</th>
<th>7–8 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3a) $d &gt; t$ in stem (borden)</td>
<td>10% (3)</td>
<td>1% (2)</td>
</tr>
<tr>
<td>(3b) $d &gt; t$ in suffix (liefde)</td>
<td>10% (2)</td>
<td>2% (1)</td>
</tr>
<tr>
<td>(3c) $t &gt; d$ in stem (harten)</td>
<td>30% (5)</td>
<td>7% (3)</td>
</tr>
<tr>
<td>(3d) $t &gt; d$ in suffix (ziekte)</td>
<td>26% (2)</td>
<td>12% (2)</td>
</tr>
</tbody>
</table>

Interestingly, the number of errors seems to be not related to morphology.$^6$ Highly significant, however, is the difference between words with $d$ and words with $t$, the difference between (3a) and (3b) on the one hand, and (3c) and (3d) on the other. The younger children made only 10% errors in both sets of words that should be spelled with $d$, and 30% or 26% errors in words that should be spelled with $t$. The older children made only 1% or 2% in words that should be spelled with $d$, and 7% or 12% errors in words that should be spelled with $t$.$^7$

Notice that the error percentages are lower in this set of nouns and numerals than in the set of verbs investigated in Experiment 1.$^8$ This may be due to the fact that the children of Experiment 1 were tested in March, thus after only 7 months of spelling instruction, whereas the children of Experiment 2 were tested in June, after 10 months of training. At the end of this paper, we will suggest an alternative explanation for this finding.

The findings in these two experiments are (a) an overall pattern of $d$-preferences, and (b) more errors in verbs than in nouns. Are these related to spelling acquisition in particular or are these aspects relevant too for experienced writers and readers of Dutch? Our third experiment investigates whether these patterns also occur in judgments on spelling errors by adolescents and adults.
2.3 Experiment 3: Judgments of adolescents and adults

Since most people eventually learn to write forms like those used in the previous experiments correctly almost without fail, a different kind of task was needed to investigate the possible existence of a $d$-preference in experienced users, and the relevance of word class distinctions. We designed a judgment task in which participants were asked to indicate the severity of an error on a seven-point scale.

Participants. Fifty-six adolescents (12–13 years of age, first year of secondary school) and 45 adults with a higher education (15–71 years of age).

Materials. 40 nouns and 40 verbs formed the basis of the test, each ending in a suffix [tә] / [dә], cf. (4).

(4) a. 20 Ns in -de(n): aarde ‘earth’, handen ‘hands’, etc.
   b. 20 Ns in -te(n): ziekte ‘illness’, olifanten ‘elephants’, etc.
   c. 20 Vs in -de: belde ‘called’, bloosde ‘blushed’, etc.
   d. 20 Vs in -te: blafte ‘barked’, bluste ‘extinguished’, etc.

As before, we selected singular verbs and as many singular nominal forms as possible, to avoid the silent final $n$ of the plural suffix.

The set of verbal forms, listed in Appendix B, is similar to that of Experiment 2, but not the set of nouns. We selected $d$-nouns with a $d$-bias in the type and token frequencies ($d$-nouns type frequency with $d$ 4841 $>$ $t$ 2191; token frequencies $d$ 1009642 $>$ $t$ 206840) as in Experiment 2. But among the $t$-nouns we selected those with a $t$-bias ($t$-nouns type frequency with $t$ 3004 $>$ $d$ 1910; token frequencies $t$ 471845 $>$ $d$ 166315). In this way, we tried to exclude frequency as a possible explanation for the $d$-preferences. (See Appendix C for the materials.)

Procedure. The words were presented to the participants in the wrong spelling, written with $t$ in stead of $d$ (belte, blooste, aarte, hanten, etc.) or written with $d$ in stead of $t$ (blafde, bluste, olifanden, ziekde, etc.). Participants were asked to indicate the severity of the error on a seven-point scale running from “a severe error” (1) to “not a very severe error” (7). First all the nouns were presented, and only then the verbs, in a separate list, so that participants were clear about the category of each word. Each list was used in four different, semi-randomized versions.

Results and discussion. One adolescent skipped over a whole page of verbs. The data on verbs from this participant were excluded from analysis. There were no reasons to exclude data from other participants. As Table 3 shows, again there was a $d$-preference.

The average judgments for words wrongly spelled with $d$ are higher (3.88) than the average judgments for words wrongly spelled with $t$ (3.13). The difference
is significant in the four sets of data of Table 3. The data show two further points of interest. First, adolescents are less strict in their judgments than adults. Second, errors in nouns are considered worse than errors in verbs. We will provide a tentative explanation for the latter finding at the end of the following section.

It is tempting to assume a common ground for the errors by children and the judgments of spelling errors by experienced users of Dutch. We calculated the correlations between the errors and the judgments for the words that were used in all experiments, i.e. 40 verbs and 31 nouns. The correlation (r) between the data of children and adolescents is 0.68 ($t(70) = 41.41, p < 0.001$). Between children and adults the correlation is 0.54 ($t(70) = 33.49, p < 0.001$), and between adolescents and adults the correlation is 0.88 ($t(70) = 7.10, p < 0.001$). These correlations suggest that there might indeed be a common explanation.

3. Towards explaining $d$-preferences

There are several, not always mutually exclusive, ways to explain these results. First, consider the possibility of a bias in the selection of test items. These were not chosen on linguistic grounds, but for extra-linguistic reasons such as frequency and familiarity (they needed to be known by first graders), and the availability of pictures that illustrated their meanings. One could argue that homography and homophony play a role, since only one noun written with $d$, *zwaarden* ‘swords’, had a homophone as counterpart with $t$ (*zwaarte* ‘heaviness, weight’), but no less than six $t$-nouns had $d$-counterparts. Thus, *zwaarte* has *zwaarden, leegde* ‘emptied’, *harten* ‘hearts’ has *harde* ‘hard’, and so on. To test whether this affected our results, we eliminated such items, and found that the $d$-preference was still present in the remaining data set.

Next, analogy might be the cause of $d$-preferences, if the phonological context in the test items were to favor $d$ over $t$. That might happen if the $d$-contexts outnumber the $t$-contexts. In experiments 1 and 2, this is indeed the case. However, the set of nouns in experiment 3 was equally balanced, and still a $d$-preference was found. Apart from conceding that it would be surprising if analogy, which exerts its influence throughout the language system, played no role at all in our data, we leave this issue for further study.
Yet another explanation could be hypercorrection. Especially through final devoicing, [t] is quite often spelled as d in Dutch, cf. the examples in (1b). The reverse, [d] written as t, is systematically absent. This might cause a d-preference by children learning to write Dutch. The instances of ts that teachers correct into ds might lead children to choose d when in doubt. Could hypercorrection be a permanent factor that explains the behavior of experienced readers as well? Little is known about hypercorrection in general, so we have to leave this issue for further investigation.

The explanation we will pursue in the remainder of this paper, without denying the possibility that above explanations may contribute as well, is lenition (Katamba 1989:103–105), a form of ease of articulation known from language change (Latin pater > Dutch vader). Ease of articulation may explain the d-preference as follows. The ds and ts in our test items form the link, or part of the link between the sonority peaks of two syllables. Articulation of [d] is easier in this context than articulation of [t], because voiced obstruents are more sonorant than voiceless ones. Careful articulation will not be necessary in the contexts investigated, because the difference between t and d is most often not distinctive (cf. the Cohort Model mentioned in the introduction). In sum, one may expect a loose style in which /t/ sounds like [d] in the middle of words. A first indication that this explanation might be on the right track is provided in Ernestus (2000:232), who noticed that voiceless stops were significantly more often classified by Dutch adults as voiced, than voiced stops are classified as voiceless.

If lenition is one of the explaining factors, we expect differences due to phonological context. The d-preference would be present, or more conspicuously present after sonorants (vowels and sonorant consonants), which form the context for lenition. It would be absent, or less conspicuously present after obstruents, which form the context for strengthening. We investigated this assumption by pooling the test items according to the sonority of the segment in front of d and t. The error percentages of the children (younger groups) are presented in Table 4, the judgments by adolescents and adults in Table 5.

In sonorant contexts, we find d-preferences in all groups of participants. In obstruent contexts, the d-preference is present only in the group of children, and we find a t-preference rather than a d-preference in the two groups of experienced writers. The differences found in sonorant contexts versus obstruent contexts show that articulation is relevant. This aspect will, of course, be reflected in analogy, but analogy itself cannot explain why we find this effect, and not, for instance, the reverse. In other words, the differences expressed by Table 4 and Table 5 form evidence for the relevance of lenition in present-day Dutch. We leave for further research the question why the effect of lenition is context-dependent for adolescents and adults, and not for children.
Additional evidence for the relevance of lenition in the mapping of /t/ and /d/ onto the letters t and d can be derived from the fact that the production and perception of voice are influenced by timing, cf. Borden, Harris & Raphael (2003:172–3): “Longer voice onset times, extended periods of aspiration, and longer closure durations cue /p,t,k/, the voiceless stops. Short voice onset times, little or no aspiration, and brief closure cue the voiced stops, /b,d,g/.” This aspect may at least partly explain the [t]-preference found in the production of Dutch infants (Van der Feest 2007:72). It might be argued that, if toddlers are slow speakers, it will be easier for them to produce a voiceless obstruent in word medial position. Experienced language users speak faster, and therefore they will produce a voiced obstruent more often. At first sight, one might have been inclined to think that it is contradictory to find that infants use [t] more often in word medial position and that young children write d more often. Both phenomena, however, may reflect the influence of speed on articulation.

Tempo might also explain the difference between nouns and verbs. Nouns are used more often in accented position (Streefkerk 2002:50), which causes longer durations (Sluijter 1995). Therefore, one would expect lenition of [t] on average to be stronger in verbs than in nouns. This may explain why fewer errors occur in nouns than in verbs (Table 2 and Table 1, respectively) and why such errors are frowned upon more in nouns than in verbs (Table 3).

### 4. Conclusion

The issues discussed in this paper are pertinent to the abstractness debate in spelling, which is traditionally concerned with orthographic depth and identifying relevant levels of representation disregarding the context of the mapping of phonemes to graphemes (Haas, 1976, Sproat 2000, for Dutch Neijt 2001 and Sproat...
Instead of striving for a general abstract level, we can concentrate on determining the level of abstractness that forms the best fit of the relation between sounds and letters in a given context.

For coronal plosives, the options vary. Using both $d$ and $t$ is the obvious choice in positions where the distinction of the sounds $[d]$ and $[t]$ is salient and where the alphabetic distinction prevents homonymy. The letter $d$ seems to be the optimal choice for English past tenses. Although the English past tense convention, with its single -ed, grossly underdetermines the pronunciations of these forms, which vary between $[d]$, $[t]$ and $[\text{id}]$ (cf. lived, stopped, and patted), it seems not to pose any difficulty to either readers or writers, be they experienced or mere beginners. If so, it might be useful to follow the English example more often when a change of spelling conventions is taken into consideration or when an alphabetic writing system is designed for one of the many languages without a written standard. Since the phoneme–grapheme mapping is context sensitive, a more abstract form of representation might be the better option.

Notes

* This research benefited from comments on an earlier version and discussions on $d$s and $t$s in phonology and orthography with Mirjam Ernestus, Susanne van der Feest, Paula Fikkert, Judith Hanssen, Vincent van Heuven, and Rik Smits. Special thanks go to Kim van de Hulsbeek, and Monique van den Broeck, for carrying out the experiments.

1. This approach presupposes a slightly different view on distinctiveness in languages than the traditional structuralist framework in which phonemic oppositions are defined irrespective of their position. Rather than an all-or-nothing distinction, the distinctive value of a sound in a word depends on its position. In a similar vein, letters will be more distinctive as they disambiguate larger sets of words, and conversely. In other words, letters matter less as we move to the right.

2. In the experiments reported here, $d$ was at the left hand side of the page, and $t$ at the right hand side. In later experiments (to be reported) we varied the position of these letters. It turned out that the position of the letters was of no influence at all.

3. The Celex database (Baayen, Piepenbrock, and Gulikers, 1995) contains a corpus of 42 million Dutch words. This corpus is based on written texts with phonematic transcriptions that took alphabetic writing as point of departure.

4. The orthographic forms express phonological forms with or without $[n]$, since the $[n]$ in word final position is often not pronounced.

5. T-test by items, two-sample assuming equal variances, two-tailed: $t(38) = 5.64$, $p < 0.001$, by subjects, two-sample for means, two-tailed: $t(40) = 3.03$, $p < 0.005$. All t-tests in this paper are two-sample assuming equal variances in the analyses by items, and paired two-sample for means in the analyses by subjects. All $p$-values are two-tailed.
6. In the younger group, 20% (the mean of (3a) and (3c), e.g. 10 and 30% in Table 2) errors occur when d or t are part of the stem, and 18% (the mean of (3b) and (3d), 10 and 26% in Table 2) when they are part of the suffix. In the older group, 4% errors occur when d or t are part of the stem, and 7% when they are part of the suffix (similar mean of (3a) plus (3c), and mean of (3b) plus (3d)). These differences are not significant. In the younger group, item analysis t(38) = 0.32, p > 0.3, and subject analysis t(24) = 0.67, p > 0.2. Similarly in the older group, item analysis t(38) = 0.66, p = .5, and subject analysis t(21) = 1.96, p = 0.06.

7. D-preferences in the younger group: item analysis t(38) = 4.28, p < 0.01, subject analysis t(24) = 3.89, p < 0.01, and in the older group: item analysis, t(38) = 2.39, p = 0.02; subject analysis, t(21) = 4.04, p < 0.01.

8. In the younger group, mean 38% errors in verbs and mean 19% errors in nouns, item analysis t(78) = 5.50, p < 0.01, subject analysis t(64) = 6.13, p < 0.01.

9. D-preferences. Adolescents, item analyses, nouns: t(38) = 4.70, p < 0.01, verbs: t(38) = 3.37, p < 0.01. Adults, item analyses, nouns: t(38) = 3.25, p < 0.01, verbs: t(38) = 2.69, p = 0.01. Adolescents, subject analyses, nouns: t(55) = 9.20, p < 0.01, verbs: t(54) = 7.41, p < 0.01. Adults, subject analyses, nouns: t(44) = 4.63, p < 0.01, verbs: t(44) = 7.82, p < 0.01.

10. The difference between nouns and verbs. Adolescents mean 3.30 versus 4.08, the difference is significant by items t(78) = 4.47, p < 0.01 and by subjects t(55) = 6.04, p < 0.01; adults mean 2.94 versus 3.72, significant by items t(78) = 4.10, p < 0.01 and by subjects t(44) = 6.48, p < 0.01.

11. 25% errors in words with t, and 10% errors in words with d; the difference is significant by items, t(29) = 4.00, p < 0.01, and by subjects (t(24) = 3.08, p = 0.01.

12. D-preferences in sonorant contexts. Children: item analysis t(34) = 3.07, p < 0.01, subject analysis t(63) = 3.03, p < 0.01. Adolescents: item analysis t(41) = 4.73, p < 0.01, subject analysis t(55) = 9.64, p < 0.01. Adults: item analysis t(41) = 2.79, p < 0.01 subject analysis t(44) = 6.21, p < 0.01.

13. D-preferences in obstruent contexts. Children: item analysis: t(42) = 3.10, p < 0.01, subject analysis: t(78) = 3.80, p < 0.01. Adolescents: item analysis t(35) = 0.99, p = 0.3, subject analysis t(55) = 2.40, p = 0.02. Adults: item analysis t(35) = 1.44, p = 0.08, subject analysis t(44) = 4.59, p < 0.01.

References


All rights reserved


Asymmetrical phoneme–grapheme mapping of coronal plosives in Dutch


Authors’ address

Radboud University
Department of Dutch language and culture
P.O. Box 9103
6500 HD Nijmegen
The Netherlands

Appendix A. Data used in Experiment 1 and 3.

d-verbs: belde (phoned); bloosde (blushed); bouwde (built); brandde (burnt); filmde (filmed); glansde (shone); hoorde (heard); kleefde (stuck); knikkerde (played marbles); krabde (scratched); leefde (lived); reisde (travelled); schudde (shook); speelde (played); spuugde (spat); straalde (beamed); tekende (draw); veegde (wiped); voelde (felt); zwaaiide (waved).
t-verb: blafte (barked); bluste (extinguished); boekte (posted); danste (danced); drupte (dripped); fietste (rode a bicycle); gaapte (gaped); juichte (cheered); klapte (clapped); kookte (cooked); lachte (laughed); maakte (made); piepte (squeaked); plukte (picked); praatte (talked); stootte (bumped); surfte (surfed); tenniste (played tennis); viste (fished); werkte (worked).

Appendix B. Data used in Experiment 2.

Nouns with -d in their stem: borden (plates); handen (hands); hemden (underhirts); honden (dogs); hoofden (heads); maanden (months); videobanden (videotapes); vrienden (friends); woorden (words); zwaarden (words).

Nouns and numerals with -d in their suffix: derde (third); elfde (eleventh); liefde (love); negende (ninth); tiende (tenth); twaalfde (twelfth); vierde (fourth); vijfde (fifth); zesde (sixth); zevende (seventh).

Nouns with -t in their stem: olifanten (elephants); planten (plants); harten (hearts); poorten (gates); herten (deers); vuisten (fists); kasten (closets); fotolijsten (photo frames); schriften (work-books); suikerklonten (sugar cubes).

Nouns with -t in their suffix: breedte (breadth); diepte (abyss); dikte (corpulence); drukte (crowd); hoogte (height); leegte (emptiness); lengte (length); stilte (silence); ziekte (illness); zwaarte (weight).

Appendix C. Data used in Experiment 3.

Nouns with -de(n): aarde (earth); borden (plates); handen (hands); hemden (undershirts); honden (dogs); hoofden (heads); kade (quay); legende (legend); liefde (love); maanden (months);
methode (method); orde (order); salade (salad); schade (damage); videobanden (videotapes); vrede (peace); vrienden (friends); woorden (words); zijde (flank); zwaarden (swords).

Nouns with -te(n): breedte (breadth); diepte (abyss); dikte (corpulence); drukte (crowd); fotolijsten (photo frames); harten (harts); herten (deers); hoogte (height); kasten (closets); leegte (emptiness); lengte (length); olifanten (elephants); planten (plants); poorten (gateways); schriften (work-books); stilte (silence); suikerklonten (sugar cubes); vuisten (fists); ziekte (illness); zwaarte (weight).