N-deletion in reading style

Hans Van de Velde and Roeland van Hout
Université Libre de Bruxelles / Nijmegen University/CLS

1. Introduction

In standard Dutch, syllable final /n/ can be dropped after a schwa at the end of a morpheme (Booij 1995: 141). Exceptions to this rule are the article *een* and — according to Koefoed (1979) and Booij (1995: 140) — verbal stems (e.g., *ik teken* 'I draw'). In most descriptive and prescriptive studies on standard Dutch pronunciation it is claimed that the deletion of (n) is more common than its realization, and that (n) is most often realized in prevocalic position. Nevertheless, in the introduction of the official Dutch orthography guide (Woordenlijst 1995: 17), it is stated that in standard speech (n) is realized. It is also claimed that n-deletion is a postlexical process (Koefoed 1979; Hinskens 1992: 336; Booij 1995), but Van de Velde and Van Hout (1998) and to a lesser extent also Van Oss and Gussenhoven (1984) raise serious doubts about this claim, which furthermore appears to be contradictory to the fact that (n) at the end of a verbal stem cannot be deleted. Van Oss and Gussenhoven (1984) and Van de Velde and Van Hout (1998) also showed that the process of n-deletion is much more variable than commonly expected. The large inter-speaker differences turn out not to be exclusively quantitative, but also qualitative in nature. On the basis of an analysis of spontaneous speech data and the behaviour of (n) before the right hand environment (Vowel, Consonant, Pause) four types of speakers can be distinguished: non-realisers (C = V = P = zero), liaisoners (V > (P = C)), deleters (C < (V = P)) and pausers (P > (V = C)).

However, it appears to be impossible to test the impact of all relevant linguistic factors in n-deletion on the basis of spontaneous speech corpora. Such an analysis is hampered by three problems. First, there is the frequency problem, i.e. the unequal distribution of a linguistic variable over linguistic contexts, as a result of which some crucial contexts hardly occur. E.g., monomorphemic words only account for about 10% of all words ending on (n) in our speech corpora, and monomorphemic finite verb forms (1st person singular) do not occur at all. Second, there is the co-occurrence problem, i.e. the entanglement of linguistic

factors in such a way that specific combinations of these factors occur rather (in)frequently. E.g., in Dutch finite forms seldomly occur before a pause, contrary to infinitives and past participles. Third, there is the interaction problem, i.e. the existence of groups of speakers showing different patterns of variation. As mentioned above, four types of speakers can be distinguished in n-deletion. Consequently, it is not possible to construct a global community grammar.

To test the impact of some of the factors that might play a role in the realization of (n) we decided to construct a quasi-experimental study. In this study the speakers are not randomly assigned to controlled conditions, but are real subjects representing differences in region, age and sex (Section 3). The linguistic factors are systematically distributed over stimuli / items. The design of the stimuli is sketched in Section 4. With such a controlled reading experiment some of the questions we were confronted with in our studies of (n) on the basis of spontaneous speech corpora might be answered and the frequency and co-occurrence problems can be coped with. The interaction problem however wil not be tackled in this study. First, it would require a high amount of observations per speaker. Second, the focus of this analysis is on those factors we were unable to investigate adequately on the basis of spontaneous speech.

2. Research questions

With this study we want to get insight in some internal and external factors ruling the realization of (n) in standard Dutch. We will focus on the following research questions:

1. Are there contexts for obligatory realization of (n)? Unsystematic observations raise serious doubts about the claim of obligatory realization of (n) at the end of a verbal stem.
2. Is n-deletion a postlexical process? Koefoed (1979) and Booij (1995) claim that there is no difference between monomorphemic and polymorphemic words but in Van de Velde and Van Hout (1998) it is shown that (n) is realised more frequently in monomorphemic than in polymorphemic words. Van de Velde (1996: 153) found that in prevocalic position (n) is realized more often in finite than in infinite verb forms.
3. What is the influence of the external factors region, sex and age on the realization of (n)? Insight in these external factors, and their interaction, might provide insight in the prestige linked with the realization or deletion of (n) in standard Dutch. The realization of /n/ after a schwa is an important criterion for the division of dialect areas (Daan and Blok 1967) and language users of the Northern part of the Netherlands are often stigmatised by their realization of (n). Van de Velde (1996) showed that there is no change in progress, but a pattern of age grading might be

present as it is often claimed that older people realize the (n) more than younger people do. However, this claim has never been supported by quantitative data.

3. Subjects

The subjects are 80 Dutch language teachers from the Netherlands, stratified for region (4), sex (2) and age (2), as can be seen in Table 1. All subjects are participating in a Flemish-Dutch research project on the pronunciation of standard Dutch (Van Hout et al. 1999). We opted for Dutch language teachers for several reasons. First, they are professional language users speaking standard Dutch on a daily basis. Second, as instructors of the standard language they play an important normative role. Van de Velde and Houtermans (1999) showed that they are also accepted as a normative authority. Third, their speech is expected to show more variation than that of broadcasters, whose speech is used in most other studies of variation and change in standard Dutch pronunciation.

Table 1. The corpus of Dutch language teachers, stratified for region, sex and age ($N=80$)

	Male		**Female**	
	Young	**Middle**	**Young**	**Middle**
Randstad	5	5	5	5
Middle	5	5	5	5
North	5	5	5	5
South	5	5	5	5

Subjects were selected at schools in middle-sized cities in four regions in the Netherlands: 1. the Randstad, i.e. the economic and cultural centre of the Netherlands, which also appears to be the core area for ongoing changes in the standard language; 2. Middle, i.e. an intermediate zone in the South of the province of Gelderland, along the borders of the Great Rivers; 3. North, a peripheral area in Groningen and the North of Drenthe; 4. South, a second peripheral area, i.e. Limburg. Subjects are currently living in the region, must have lived there before their 8th birthday, and have been living there for at least eight years before their 18th birthday. Two age groups are distinguished: young (between 22 and 40) and middle (between 45 and 60). For sex, a biological distinction between male and female is made. For more information about the research design, see Van Hout et al. (1999).

4. Method

We distinguished five categories of words ending on (n) in our reading task: monomorphemic nouns (mono N), monomorphemic verbal stems (mono V), polymorphemic finite verb forms (poly F), polymorphemic infinitives (poly I) and spatial adverbs/prepositions (spatial A/P). The research question about obligatory (n) realization contexts can be answered by the results found in the mono category, especially mono V. Whether n-deletion is a postlexical process, the second research question, can be tested by analyzing the differences between monomorphemic and polymorphemic words (mono vs. poly), between poly I and poly F, and between mono N and mono V. The spatial A/P category will also provide us with this type of information. Two other variables were systematically included in the research design, i.e. context (right hand environment, five different contexts for each word category) and lexical item (five different words in each word category). To be able to control these (nuisance) variables and to reduce the number of stimuli, a Latin square design was used. The differences between mono N and mono V, and between poly F and poly I are studied with the same lexical items in exactly the same contexts. An overview of the lexical items and the contexts is presented in Table 2. Differences between lexical items will give us additional information on the role of the lexicon in the process of n-deletion.

A fully crossed design would have given 125 combinations of category, lexical item and context ($5 \times 5 \times 5$) for each subject. With a Latin square design for the combinations of lexical item and context, the number of items per subject can be reduced to 25 in such a way that each subject gets each item and each context, but only in five specific combinations. As each cell of our corpus (Table 1) contains five informants, all 125 items are present in each cell. It means that there are five versions of the reading task, each consisting of 25 carrier sentences containing the items to be tested. The 25 sentences containing (n) were mixed with sentences focussing on other phonological variables. The sentences were split in two separate reading tasks in such a way that lexical items did not occur twice in the same task. In the interview there was a gap of about 20 minutes between the two reading tasks.

The sentences were individually presented on the screen of a laptop computer by means of a PowerPoint presentation. The speech of the subject was recorded on digital audiotape with a portable TASCAM DA-P1 recorder and an AKG C420 headset microphone. All interviews were administered by a young, Dutch male interviewer, who speaks modern standard Dutch without a regional accent. The recordings were digitalized on computer and downsampled to 16 kHz (16 bits). For the auditory analysis each sentence was saved as a separate soundfile. Two trained transcribers, one being a Flemish native speaker of Dutch, the other a native speaker of French with a limited knowledge of Dutch, made a consensus transcription. In cases of doubt (0.5%) a third trained transcriber (a native speaker of Dutch from

Table 2. Overview of the variables word category, lexical item and context

	Monomorphemic		Polymorphemic		Spatial A/P
	Noun	Verb	Finite	Infinite	
lexical item	baken 'beacon'		bijten 'bite'		beneden 'below'
	keten 'chain'		buigen 'bend'		binnen 'inside'
	teken 'sign'		leven 'live'		boven 'above'
	toren 'tower'		lopen 'run'		buiten 'outside'
	zegen 'bless(ing)'		tekenen 'draw'		midden 'in the middle of'
context	pause		pause		pause
	vowel		vowel		vowel
	consonant		consonant		consonant
	clitic		vowel other clause		vowel within word
	schwa		consonant other clause		schwa

Table 3. Frequency distribution of the 10 variants of (n) (2000 observations)

	N	%	+/− n
schwa	687	34.35	−
[I]	2	0.10	−
Ø (deletion of syllable)	34	1.70	−
schwa + n	1246	62.30	+
syllabic n	4	0.20	+
nasal schwa	10	0.50	+
n	2	0.10	+
schwa +m	13	0.65	+
syllabic m	1	0.05	+
schwa + t	1	0.05	+

the Netherlands) was consulted. Ten variants had to be distinguished (see Table 3). For the analysis we recoded the variants into a binomial variable, absence (−) or presence (+) of a nasal (or alveolar in the case of schwa + t) feature.

5. Results

As can be seen from Table 3, realizations other than schwa or schwa + n are rare: the eight other variants account for 3.35% of the realizations. In comparison with spontaneous speech data, the amount of (n) realization is high. Although figures are not directly comparable — due to differences in the distribution of the final (n) — the amount of (n) realization (63.8%) in our data is much higher than in

spontaneous speech (12.6% in Van de Velde (1996), 3.7% in Van Oss and Gussenhoven (1984)), but comparable to reading style in newscasts (56.2% in Van Oss and Gussenhoven (1984) and 40.3% in Van de Velde and Van Hout (1998)). It is clear that orthography and carefulness which is typical of reading style, play a role in the realization of (n). Despite the high amount of monitoring in our reading task, complete syllables were reduced (1.7%); hypercorrections (insertion of /n/ after a schwa which is not followed by /n/) were very rare. It should be noted that the realizations containing [m] are not the result of (regressive) assimilation of place, and that syllabic (n) realizations, which are a marker of the speech of people from the North-East of the Netherlands, are extremely rare in our corpus.

The analysis is split in three parts: monomorphemic words (mono N vs. mono V), polymorphemic words (poly F vs. poly I) and spatial A/P's. In the analyses of variance, sex, age and region are between-subjects factors; word class (N vs. V) for the monomorphemic words and finiteness (F vs. I) for the polymorphemic words are within-subjects factors. The results of the analyses of variance are summarized in Table 4. All main effects and those interaction effects which are significant in one of the three analyses are included. The mean scores for sex and region are presented in Table 5 and those for the interaction effect sex by age in Table 6. The mean scores for word class and finiteness can be found in Table 7.

Table 4. Summary of the analyses of variance for monomorphemic, polymorphemic and spatial words

	Monomorphemic	**Polymorphemic**	**Spatial A/P**
Between-subjects factors			
Sex	n.s.	*	*
Age	n.s.	n.s.	n.s.
Region	n.s.	*	*
Sex by age	**	*	n.s.
	N vs V	**F vs I**	
Within-subjects factors			
Linguistic factor (= LF)	***	*	
LF by region	*	n.s.	
LF by sex by age	**	n.s.	

All main effects are included plus those interaction effects which are significant in one of the three analyses (n.s. = not significant at .05 level; $*p<.05$; $**p<.01$; $***p<.001$; when the analysis does not apply the cell is left blank).

As can be seen from Table 4, the realization of (n) in monomorphemic, polymorphemic and spatial A/P words is influenced by different factors. For none of these types of words is there a significant main effect for age. Sex and region

Table 5. Mean scores for the three word types, split up by sex and region

		Monomorphemic	Polymorphemic	Spatial A/P
Sex	Male	88.00	54.00	60.50
	Female	84.50	39.25	46.00
Region	Randstad	86.50	45.50	50.00
	Middle	89.50	32.00	47.00
	North	85.00	60.00	73.00
	South	84.00	49.00	43.00

factors are significant for polymorphemic and spatial A/P words (see Table 5), but not for monomorphemic words. As expected, (n) is realized more often in the North than in the other regions (the region factor is significant twice; see mean scores in Table 6). The interaction effect of the linguistic factor for the monomorphemic words and region (LF by region) seems to be the result of a ceiling effect. The amount of realization of (n) is very high in all regions for mono V; only in word class mono N do differences between regions show up, again people from the North having the highest amount of realization. Overall, there is a significant difference between monomorphemic nouns and verbs: (n) is realized more in verbal stems than in nouns (Table 7). However, the claim that (n) at the end of a verbal stem cannot be deleted is contradicted: in 6% of the cases is (n) not realized, which shows that realization of /n/ at the end of a verbal stem is not categorical. Furthermore, a significant difference is found between finite and infinite verb forms. We will return to this difference when we compare all five word categories.

Table 6. Mean scores for sex by age

	Monomorphemic		Polymorphemic		Spatial A/P	
	Young	Middle	Young	Middle	Young	Middle
Male	92.50	83.50	61.00	47.00	63.00	58.00
Female	78.00	91.00	30.50	48.00	39.00	53.00

The absence of a main age effect is in line with the real time observations: Van de Velde (1996) showed that n-deletion is not a change in progress in standard Dutch. There is however an interesting and very systematic sex by age effect (see Table 6). The interaction is significant for monomorphemic and polymorphemic words (with a main effect of sex for polymorphemic words), but the pattern is also present in the spatial A/P's (with a main effect for sex). In general, young women delete the (n) more than young men. Young men are even those who realize it the most. Women

Table 7. Mean scores for the linguistic factors word class and finiteness

Monomorphemic	**Noun**	**Verb**
	78.50	94.00
Polymorphemic	**Finite**	**Infinite**
	50.75	42.50

having a higher rate of n-deletion than men probably seems to reflect the fact that the deletion of (n) has higher level of prestige than its realization, as women commonly use more prestige variants than men (Chambers 1994: 102). It could imply that (n) is becoming part of a change, the sex differential pattern being a portent.

The differences between the five word categories seem to be obvious. However, as context 4 and 5 vary between word categories (see Table 2), the real distinctions can be blurred. Therefore, only the first three contexts (pause, vowel, consonant) were selected for making comparisons between all five categories (n = 1200). The design is less balanced by this reduction but we will not investigate here the extra-linguistic factors sex, age and area. The scores of 80 informants are available to make the category comparisons and this will balance out the context reduction. An analysis of variance (repeated measures) shows a significant category effect ($F = 60.901$, $df = 4,316$, $p = .000$). By applying paired t-tests in combination with the Bonferroni procedure post-hoc comparisons between the contexts can be made. The results are given in Table 8.

Table 8. Post-hoc comparisons between the five categories

Category	**Mono N**	**Mono V**	**Poly F**	**Poly I**	**Spatial A/P**
Mean	75.41	90.88	51.68	38.95	43.76
Mono N		*	*	*	*
Mono V	*		*	*	*
Poly F	*	*		*	n.s.
Poly I	*	*	*		n.s.
Spatial	*	*	n.s.	n.s.	

Paired t-tests, Bonferroni procedure; *$p < .05$; n.s. = not significant

In comparing Tables 7 and 8, a small decrease in the mean scores is observed for mono N, mono V, poly F and poly I. Also for the spatial words the reduction of the contexts results in a decrease in the mean score (53.25 vs. 43.76), which can be explained by the vowel-like nature of contexts 4 and 5. Importantly, significant differences are found where we found them before: mono N vs. mono V, and poly

F vs. poly I. The fact, however, that there is a higher frequency of realization of (n) in poly F than in poly I is contradictary to the observations in Van de Velde (1996: 153). Anyway, the difference contradicts (n) deletion as a purely postlexical process. There is also a straightforward morpheme effect: the mono categories differ significantly from the poly categories. This again contradicts pure postlexicality. The spatial A/P's, although being monomorphemic words, show more quantitative similarities with the polymorphemic words, having a high rate of n-deletion. The spatial A/P's behave differently from another perspective as well. Table 9 contains the realization percentages for the individual lexical items. The words within the same category behave rather homogeneously, except for the finite form of *buigen* and the spatial A/P's. There are much more cases of realization of (n) in *boven* and *buiten* than in *beneden*, *binnen* and *midden*. We have no explanation for these lexical differences.

Table 9. Realization of (n) split up by word category and item

Mono N	**%**	**Mono V**	**%**	**Poly F**	**%**	**Poly I**	**%**	**Spatial**	**%**
baken	85.0	baken	91.3	bijten	45.0	bijten	37.5	beneden	46.3
keten	81.3	keten	93.8	buigen	67.5	buigen	48.8	binnen	42.5
teken	82.5	teken	97.5	leven	52.5	leven	36.3	boven	71.3
toren	70.0	toren	91.3	lopen	42.5	lopen	43.8	buiten	62.5
zegen	73.8	zegen	96.3	tekenen	46.3	tekenen	46.3	midden	43.8

We cannot give a straightforward explanation for the observed differences between the categories, but they are probably due to differences in lexical storage. Monomorphemic nouns and particularly verbal stems seem to be stored including the final /n/. On these words a deletion rule is applied. Spatial A/P's (*beneden*, *binnen*, *midden*, but not *boven* and *buiten*) and infinitives, however, are possibly stored without the final /n/. In these cases an insertion rule is applied to generate a nasal. But the picture is definitely more complex. Baayen et al. (1997) showed that plurals of nouns can be generated by a rule or can be stored separately. Hence, for plural formation on *-en* there are four possible points of departure: a stored plural form with final /n/, a stored plural form without final /n/, a rule that adds schwa to the singular form and a rule that adds schwa + n to the singular form. However, an argument in favour of a model in which final /n/ is stored is the fact that hypercorrection (the insertion of a nasal in contexts where /n/ is not present) is rare.

6. Conclusions

In this quasi-experimental study of n-deletion in standard Dutch it is shown that n-deletion is bound to some external factors. First, there is a regional difference: (n) is realised more in the North than in the other parts of the Netherlands. Second, there is an interaction between age and sex, with young women deleting the most and young men realizing the most. This may point out that deletion of (n) has more prestige than realization, but further research on different speech styles is necessary to confirm this assumption. Age is not a significant factor, pointing out that n-deletion is not a change in progress. This claim is supported by real time data (Van de Velde 1996). The sex by age interaction however could imply that the process of n-deletion is in the incipient phase of change, towards more deletion of (n).

In this paper we also showed that /n/ after a schwa at the end of a verbal stem can be dropped in standard Dutch. So the rule of n-deletion can be expanded to all occurrences of /n/ after a schwa at the end of a morpheme. It is also obvious that n-deletion is not only a postlexical process. Systematic differences between monomorphemic and polymorphemic words are observed, as well as between monomorphemic verbs and nouns, and finite and infinite verb forms. Spatial A/P's seem to behave differently and point out a clear lexical effect.

Further research is necessary to explain n-deletion in standard Dutch. A valid model will probably be very complicated. In such a model variation and differences in lexical storage would be found. In storage the degree of literacy of the speaker probably plays an important role. Furthermore, speakers may have different types of rules for n-insertion, n-liaison, n-deletion and/or n-realization. They differ in the extent to which one or other rule dominates, as is shown by our typology of speakers (Van de Velde and Van Hout 1998). Finally, we think that most speakers have (a form with) an underlying /n/ for most of the words. If this is not the case, the amount of hypercorrection would be much higher.

References

Baayen, R.H., T. Dijkstra and R. Schreuder (1997) 'Singulars and Plurals in Dutch: Evidence for a Parallel Dual-Route Model'. *Journal of Memory and Language* 37: 94–117.

Booij, G.E. (1995) *The Phonology of Dutch.* Clarendon Press, Oxford.

Chambers, J.K. (1994) *Sociolinguistic Theory. Linguistic Variation and its Social Significance.* Blackwell, Oxford.

Daan, J. and D.P. Blok (1967) *Van Randstad tot Landrand. Toelichting bij de kaart: dialecten en naamkunde.* B.V. Noord-Hollandse Uitgevers Maatschappij, Amsterdam.

Hinskens, F. (1992) *Dialect Levelling in Limburg. Structural and Sociolinguistic Aspects.* Dissertation, Nijmegen.

Koefoed, G. A. T. (1979) 'Paradigmatische invloeden op fonetische processen'. In T. Hoekstra and H. v. d. Hulst (eds.), *Morfologie in Nederland*. Glot Special, Leiden, 51–72.

Van de Velde, H. (1996) *Variatie en verandering in het gesproken Standaard-Nederlands (1935–1993)*. Dissertation, Nijmegen.

Van de Velde, H. and M. Houtermans (1999) 'Vlamingen en Nederlanders over de uitspraak van nieuwslezers'. In E. Huls and B. Weltens (eds.), *Artikelen van de Derde Sociolinguïstische Conferentie*. Eburon, Delft, 451–462.

Van de Velde, H. and R. van Hout (1998) 'Dangerous aggregations. A case study of Dutch (n) deletion'. In C. Paradis et al. (eds.), *Papers in Sociolinguistics*. Nuits Blanche éditeur, Québec, 137–147.

Van Hout, R., G. De Schutter, E. De Crom, W. Huinck, H. Kloots and H. Van de Velde (1999) 'De uitspraak van het Standaard-Nederlands. Variatie en varianten in Vlaanderen en Nederland'. In E. Huls and B. Weltens (eds.), *Artikelen van de Derde Sociolinguïstische Conferentie*. Eburon, Delft, 183–196.

Van Oss, F. and C. Gussenhoven (1984) 'De Nederlandse slot-n in het nieuws'. *Gramma* 8(1): 37–45.

Woordenlijst (1995) *Woordenlijst Nederlandse taal. Samengesteld door het Instituut voor Nederlandse Lexicologie in opdracht van de Nederlandse Taalunie*. Sdu/Standaard, Den Haag/Antwerpen.