

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
REKENAFDELING

MR 89

Mechanische vorming van het
geschreven verkleinwoord

door

H. Brandt Corstius



juli 1967

BIBLIOTHEEK MATHEMATISCH CENTRUM
AMSTERDAM

The Mathematical Centre at Amsterdam, founded the 11th of February, 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.) and the Central Organization for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

§1. Het probleem

Hoe vormt men in het Nederlands van een gespeld woord het gespelde verkleinwoord? Is deze vorming te mechaniseren, in casu door een rekenautomaat te bepalen? Het nut van zo'n mechanische bepaling is tweërlei:

1. Theoretisch. Voor een mechanische bepaling is een exacte formulering van de regels voor het verkleinwoord noodzakelijk. De bestaande grammatika's geven die niet. De rekenautomaat verschaft ons de best mogelijke controle op de geldigheid van de regels.
2. Practisch. De "Woordenlijst van de Nederlandse taal" [1] (het "groene boekje") geeft thans het verkleinwoord "als dat gebruikelijk is en als het anders wordt gespeld dan door toevoeging van -je aan het grondwoord". Vergroot men het aantal van zulke regels, dan hoeft men van minder woorden de afwijkingen daarvan expliciet op te geven. De omvang van de woordenlijst in het geheugen van de rekenautomaat kan daardoor geringer worden.

Dezelfde twee doeleinden worden gediend door een mechanische bepaling van de meervoudsvorming en van de werkwoordsvormen.

Ons probleem zou van belang kunnen zijn bij mechanische vertaling uit een andere taal in het Nederlands. Andersom zou, bij vertaling uit het Nederlands, een algoritme voor het verkrijgen van het niet-verkleinde woord uit het verkleinwoord nuttig kunnen zijn. Dit is enerzijds eenvoudiger dan ons probleem, anderzijds is direct in te zien dat het in veel gevallen onoplosbaar is: de woorden "buur" en "buurt", "diploma" en "diplomaat", "logé" en "logée" vormen hetzelfde verkleinwoord. Dat, omgekeerd, een woord twee verschillende verkleinwoorden heeft, afhankelijk van zijn betekenis, is zeldzaam: "pad", "gat" en "filmster" zijn hiervan voorbeelden.

Ons probleem blijft beperkt tot de geschreven vorm van het woord. Morfonologisch liggen de problemen anders [2]. Ook vragen we ons niet af of van een woord het verkleinwoord gevormd kan worden. In dit opzicht geeft de Woordenlijst ook niet altijd uitsluitel: staat achter een zelfstandig naamwoord geen verkleinwoord dan kan dit betekenen dat het niet voorkomt, of dat het wordt gevormd door er "je" achter te plaatsen. Achter "rancune" wordt niets vermeld waaruit men - op grond van taalregels die de afwijkende vorm "rancunetje" doen verwachten - kan afleiden dat het verklein-

woord niet gebruikelijk werd geacht. Onze oplossing is in de vorm van een programma in ALGOL 60 [3] dat een woord inleest en het verkleinwoord (of de verkleinwoorden) ervan produceert.

In §2 geven we in grote lijnen de regels van de verkleinwoordvorming. Voor de exacte formulering zie men echter het programma in §3, dat vanaf de label LEES WOORD van taalkundig belang is. In §4 wordt de conclusie getrokken uit de resultaten van het programma.

§2. De regels

De regels worden gegeven als transformaties waarbij steeds links van de pijl het rechteruiteinde van het te transformeren woord staat, dat veranderd wordt tot wat rechts van de pijl staat. Groepen van letters worden met hoofdletters als volgt aangeduid:

$V = \{a, o, u\}$.

$W = \{\text{klinker}\}$. Aangenomen wordt dat deze klinkers, die uit 2 of 3 lettertekens kunnen bestaan, tevoren in het woord samengenomen zijn, evenals de bigrammen "ng" en "qu".

$L = \{l, m, n, r\}$.

$C = \{\text{medeklinker}\}$.

De regels worden van boven naar beneden doorlopen tot de eerste toepasbare.

1. $V \rightarrow VVtje$ (beide V's duiden dezelfde letter aan)
opa $\rightarrow$ opaatje, radio $\rightarrow$ radiootje, menu $\rightarrow$ menuutje
2. $\acute{e} \rightarrow eetje$ caf  $\rightarrow$ cafeetje
3. $y \rightarrow y'tje$ lolly $\rightarrow$ lolly'tje
4. $i \rightarrow ietje$ ski $\rightarrow$ skietje
5. $W \rightarrow Wtje$ kastanje $\rightarrow$ kastanjetje, aai $\rightarrow$ aaitje
6. Voor een lijst van woorden (blad, vat, glas....):
aC $\rightarrow$ aaCje blad $\rightarrow$ blaadje

Op dezelfde wijze wordt bij enkele woorden, als "schip", "augur" en "depot", de laatste klinker veranderd in resp.: "ee", "uu" en "oo".

$$7. \begin{Bmatrix} V \\ i \\ E \end{Bmatrix} \times \begin{Bmatrix} L \\ ng \\ b \\ g \end{Bmatrix} \rightarrow \begin{Bmatrix} V \\ i \\ E \end{Bmatrix} \times \begin{Bmatrix} LL \\ ng \\ bb \\ gg \end{Bmatrix} etje \quad \text{lam} \rightarrow \text{lammetje}$$

waarin het $\times$ -teken aangeeft dat de keuze uit de 2 accoladeparen on-

afhankelijk is, en de E de letter e voorstelt mits deze niet stom is. De woorden op "b" of "g" moeten 1-lettergrepig zijn. Deze hebben vaak een nevenvorm op "je": "wiggetje" en "wigje". De woorden op "ing" van meer dan een lettergreep krijgen "inkje": "koninkje".

8. m → mpje bloem → bloempje
9. $\begin{Bmatrix} L \\ w \end{Bmatrix} \rightarrow \begin{Bmatrix} L \\ w \end{Bmatrix} tje$ leeuw → leeuwtje, hiel → hieltje
10. C → Cje klad → kladje, oorlog → oorlogje.

Deze 10 regels geven een onvolledige beschrijving, getuige de volgende verkleinwoorden:

l'etje, machientje, radertje, parachuutje, leerlingetje, jongetje, clubje, kippetje, dineetje, fetisje, crapaudtje, aveetje, stedeke.

Wel wordt nu duidelijk welke onze moeilijkheden zijn:

1. het herkennen van de stomme e. Alleen voor 1-lettergrepige woorden kan men aan de geschreven vorm zien of een e stom is. De woorden "meester" en "zeester" tonen dat dit in andere gevallen niet kan.
2. de lijsten van woorden met klinkerverandering als "glaasje", "lootje" en "scheepje". Hier is duidelijk verband met de meervoudsvorming in die zin dat we nooit klinkerverandering in het verkleinwoord vinden zonder klinkerverandering in het meervoud. Andersom gaat dit verband niet op: "verdragje", "gebedje", "hofje", ondanks "verdragen", "gebeden" en "hoven". De woorden "depot", "jabot" etc. krijgen in de geschreven vorm ook klinkerverandering maar dit heeft een geheel andere oorzaak. (zie §4.4)
3. kennis over de klemtoon is nodig voor de beslissing ing → ingetje of ing → inkje.
4. een aantal woorden heeft meer dan een verkleinwoord:
vlootje, vlooitje; dagje, daagje; heggetje, hegje; machinetje, machientje.
Van "rad" bestaan zelfs drie verkleinwoorden: radje, raadje en radertje.

Het meervoud van het verkleinwoord wordt gevormd door "s" achter het verkleinwoord te plaatsen, behoudens in een aantal gevallen waar behalve deze regelmatige vorming ook een vorming bestaat die van het (onregelmatige) meervoud van het onverkleinde woord afkomstig is: eiertjes, lammertjes, rundertjes, kalvertjes, kindertjes, beendertjes, bladertjes,

hoendertjes, daakjes en leedjes. Dit verschijnsel behandelen we verder niet.

Het verkleinen van bijvoeglijke naamwoorden (type "blauwtjes" en type "geeltje" [4]) wordt door dezelfde regels geregeerd: "smalletjes", "zoetjes".

Bij onderzoekingen als deze zou een retrograad woordenboek een grote hulp zijn. Zo'n woordenboek waarin de woorden van achter naar voren gealfabetiseerd zijn is mechanisch eenvoudig te maken als eenmaal de normale woordenlijst in voor een machine leesbare vorm is gebracht.

§3. Het programma

In het nu volgende ALGOL-60 programma is geprobeerd het taalkundige gedeelte zo goed mogelijk te isoleren. Van de aanvang tot de label LEESWOORD op pag. 7 hoeft men alleen de commentaren te lezen, daarna volgt een formulering van de regels die veel exacter is dan die in §2. Voor de codering van de letters, die niet essentieel is, verwijzen we naar [5]. De opdracht "zet (119)" bewerkstelligt een nieuwe regel.

Pagina's 5 en 6 bevatten declaraties en procedures die het mogelijk maken dat de rest van het programma voor de menselijke lezer te begrijpen is. Pagina 7 behandelt het comprimeren van klinkers als "aai" en de medeklinkers "qu" en "ng". Hier worden ook de verkleinwoorden van de 27 letternamen gevormd. Pagina 8 geeft de verkleinwoordvorming van de woorden die op een klinker eindigen. Pagina 9 produceert een aantal nevenvormen van woorden op een klinker eindigend, en een aantal bijzondere gevallen van woorden die op een medeklinker eindigen. Pagina 10 tenslotte behandelt de verkleinwoordvorming van de woorden die op een medeklinker eindigen.

begin comment Van de aangeboden woorden wordt het verkleinwoord gevormd;

integer a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t,
u, v, w, x, y, z, ij, ng, qu, aa, ee, oo, eu, ie, oe, ou, accent,
apostrof, komma, vraagteken, woordscheider, voorlaatste, laatste,
aantal, ww, uitzondering, sytal;
integer array W[1:50], C[1:37,1:37];

procedure lees woord;

comment de letters van het te behandelen woord worden in array W gezet en ter controle uitgevoerd. Na het laatste woord geeft een vraagteken het eind van de bewerking aan. aantal geeft het aantal letters in het woord aan;

begin integer ii, symbool;

integer procedure in;

begin integer sy;

sy:= RESYM; if sy = vraagteken then EXIT;

in:= if sy = accent ∨ sy = apostrof then sy else if sy > 62
then woordscheider else if sy > 35 then sy - 27 else sy

end in;

zet(119); ii:= 0;

L: symbool:= in; if symbool ≠ woordscheider then

begin zet(symbool); ii:= ii + 1; if ii > 50 then EXIT;

W[ii]:= symbool; goto L

end

else if ii = 0 then goto L else

begin aantal:= ii; laatste:= W[aantal]

end woord gelezen

end lees woord;

procedure geef constanten waarden;

begin for i:= 1 step 1 until 37 do

for j:= 1 step 1 until 37 do C[i,j]:= 0; a:= 10; b:= 11; c:= 12;

d:= 13; e:= 14; f:= 15; g:= 16; h:= 17; i:= 18; j:= 19; k:= 20;

l:= 21; m:= 22; n:= 23; o:= 24; p:= 25; q:= 26; r:= 27; s:= 28;

```

t:= 29; u:= 30; v:= 31; w:= 32; x:= 33; y:= 34; z:= 35; qu:= 36;
ng:= 37; ij:= 2; aa:= 3; ee:= 4; oo:= 5; eu:= 6; ie:= 7; oe:= 8;
ou:= 9;
C[aa,i]:= C[aa,u]:= C[aa,y]:= C[ee,u]:= C[oo,i]:= C[oo,y]:=
C[a,e]:= C[a,i]:= C[a,u]:= C[a,y]:= C[e,i]:= C[e,y]:= C[u,i]:=
C[u,u]:= C[u,y]:= C[o,i]:= C[o,y]:= C[o,j]:= C[eu,i]:=
C[ie,u]:= C[oe,i]:= C[oe,u]:= C[oe,y]:= C[ou,i]:= 1; C[a,a]:= aa;
C[e,e]:= ee; C[o,o]:= oo; C[e,u]:= eu; C[i,e]:= ie; C[o,e]:= oe;
C[o,u]:= ou; C[i,j]:= ij; C[q,u]:= qu; C[n,g]:= ng;
accent:= 126; apostrof:= 120; komma:= 87; vraagteken:= 122;
woordscheider:= 100
end geef constanten waarden;

integer procedure w3(v,l,o); integer v, l, o;
w3:= (v - 9) × (26 ⌈ 2) + (l - 9) × 26 + o - 9;
integer procedure w4(g,l,a,s); integer g, l, a, s;
w4:= (g - 9) × (26 ⌈ 3) + w3(l, a, s);
integer procedure w5(s,c,h,i,p); integer s, c, h, i, p;
w5:= (s - 9) × (26 ⌈ 4) + w4(c, h, i, p);
boolean procedure klinker(letter); integer letter;
klinker:= aeiou(letter) ∨ letter = y ∨ letter < 10;
boolean procedure aeiou(letter); integer letter;
aeiou:= letter=a ∨ letter=e ∨ letter=i ∨ letter=o ∨ letter=u;
boolean procedure lmnr(letter); integer letter;
lmnr:= letter = l ∨ letter = m ∨ letter = n ∨ letter = r;
procedure zet(letter); integer letter; PUSYM(letter);
procedure woord;
comment aantal letters van W wordt uitgevoerd,voorafgegaan door
een komma;
begin integer ii;
    zet(komma);
    for ii:= 1 step 1 until aantal do zet(W[ii])
end woord;
procedure je;
begin zet(j); zet(e)
end je;

```

procedure comprimeer ng qu en klinkers;

comment met behulp van de informatie in C worden de meerletterige klinkers, de qu en de ng samengenomen. De laatste en de voorlaatste letter van het gecomprimeerde woord worden bepaald. Het getal ww karakteriseert de laatste vijf letters van het woord. syltal wordt gelijk aan het aantal lettergrepen;

begin integer letter, volgende, wi, ci;

integer array CW[0:50];

wi:= ci:= 0; letter:= x;

comprimeer: wi:= wi + 1; if wi > aantal then goto gecomprimeerd;

volgende:= W[wi];

if volgende = accent $\vee$ volgende = apostrof then goto comprimeer;

if C[letter, volgende] = 0 then

begin CW[ci]:= letter; ci:= ci + 1; letter:= volgende end

else letter:= C[letter, volgende]; goto comprimeer;

gecomprimeerd: CW[ci]:= laatste:= letter;

voorlaatste:= if ci > 1 then CW[ci - 1] else 0; syltal:= 0;

for wi:= 1 step 1 until ci do

begin if klinker(CW[wi]) then syltal:= syltal + 1 end;

ww:= SUM(wi, aantal - 4, aantal,

$(26 \wedge (\text{aantal} - \text{wi})) \times (\text{if } \text{wi} < 1 \text{ then } 0 \text{ else } (W[\text{wi}] - 9))$)

end comprimeer ng qu en klinkers;

geef constanten waarden;

comment Hier begint het taalkundig interessante gedeelte;

LEES WOORD: lees woord;

bepaal verkleinwoord: if aantal = 1 then

begin woord; zet(apostrof); if lmnr(laatste) then zet(e);

if $\neg$ (laatste = f $\vee$ laatste = s $\vee$ laatste = x $\vee$ laatste = z)

then zet(t); je

end letternamen

else

begin comprimeer ng qu en klinkers; if aantal=2 $\wedge$ laatste=ij then

begin woord; zet(apostrof); zet(t); je

end ij

```
else if klinker(laatste) then
begin if aeiou(laatste)  $\wedge$  (laatste = e  $\neg$  W[aantal - 1] = accent) then
  begin if laatste = e then
    begin aantal:= aantal - 1; W[aantal]:= e
    end cafe;
    woord; zet(if laatste = i then e else laatste); zet(t)
  end a i o u en geaccentueerde e
else
  begin if laatste = ee  $\wedge$ 
    W[if aantal > 3 then aantal-2 else 1] = accent then
    begin aantal:= aantal - 1; W[aantal - 1]:= e end employee;
    if laatste = e then
      begin if ww = w3(a,v,e)  $\vee$  ww = w5(m,p,o,r,e)  $\vee$ 
        ww = w5(i,m,i,l,e) then
          begin aantal:=aantal+1; W[aantal]:=e; laatste:=ee
          end ave,extempore,facsimile;
          if ww = w5(t,o,i,r,e)  $\vee$  ww = w5(l,e,i,n,e) then
            begin aantal:=aantal-1; goto bepaal verkleinwoord
            end directoire,kleine;
          if ww = w5(o,n,a,d,e)  $\vee$  ww = w5(c,h,u,t,e)  $\vee$ 
            ww = w5(e,c,u,r,e) then
              begin aantal:=aantal-1; laatste:=voorlaatste;
              voorlaatste:=W[aantal-1]; goto klinkerverandering
            end karbonade,parachute,sinecure;
          if ww = w5(n,e,t,t,e) then
            begin aantal:=aantal-2; goto bepaalverkleinwoord end brunette;
          if ww = w5(e,r,i,n,e)  $\vee$  ww = w5(r,r,i,n,e)  $\vee$  ww = w5(a,f,i,n,e) then
            begin W[aantal-1]:=e;W[aantal]:=n;goto bepaal verkleinwoord
            end pelerine,terriner,serafine
          end woorden waar slot-e vervalt;
        woord; if  $\neg$ (voorlaatste = qu  $\wedge$  laatste = e) then
          begin if laatste = y then zet(apostrof); zet(t) end
        end andere klinkers;
      je;
```

```
if ww = w3(v,l,o) then
  begin aantal:= aantal + 2; W[aantal - 1]:= o; W[aantal]:= i;
    goto bepaal verkleinwoord
  end vlooitje;
if laatste=e then
  begin if ww = w5(c,h,i,n,e) then
    begin W[aantal - 1]:= e; W[aantal]:=n; goto bepaal verkleinwoord
    end machientje;
    if ww = w5(r,e,t,t,e) then
      begin aantal:=aantal-2; goto bepaal verkleinwoord end operetje;
      if ww = w5(r,i,s,t,e) ∨ ww = w5(l,o,e,z,e) then
        begin if voorlaatste = z then W[aantal-1]:=s;
          aantal:= aantal - 1; goto bepaal verkleinwoord
        end koriste,bloeze
      end nevenvormen met vervallen van slot-e;
    end woord eindigend op klinker
  else
    begin for uitzondering:= w3(r,a,d), w3(d,a,g), w3(l,o,t),
      w3(p,a,d), w4(s,p,i,t), w4(s,t,a,f), w3(g,a,t) do
      begin if ww = uitzondering then
        begin woord; je; if ww = w3(r,a,d) then
          begin woord; zet(e); zet(r); zet(t); je
          end radertje;
          goto klinkerverandering
        end
      end padje en paadje;
      for uitzondering:= w4(b,l,a,d), w4(g,l,a,s), w3(v,a,t),
        w4(s,t,a,g), w4(s,l,a,g), w5(s,c,h,i,p), w5(a,u,g,u,r),
        w5(t,e,l,o,t), w5(l,e,t,o,t), w5(d,e,p,o,t), w5(j,a,b,o,t),
        w5(e,r,r,o,t) do
        begin if ww = uitzondering then goto klinkerverandering
        end blaadje;
      for uitzondering:= w5(d,i,n,e,r), w5(o,u,p,e,r), w5(e,u,n,e,r)
      do
        begin if ww = uitzondering then
          begin W[aantal]:= e; goto bepaal verkleinwoord end
        end dineetje;
```

```

if ((aeiou(voorlaatste)  $\wedge$  (laatste = b  $\vee$  laatste = g  $\vee$  laatste
= ng  $\vee$  lmnr(laatste))  $\wedge$  (voorlaatste = e  $\neg$  (syltal = 1))))  $\vee$ 
ww = w3(k,i,p)  $\vee$  ww = w3(p,o,p) then
begin if voorlaatste = i  $\wedge$  laatste = ng  $\wedge$  (syltal  $\neq$  1) then
    begin W[aantal]:= k; woord; je;
    end harinkje
    else
    begin if  $\neg$ ((laatste=g  $\vee$  laatste=b)  $\wedge$  (syltal  $\neq$  1))  $\wedge$ 
         $\neg$ (ww = w4(c,l,u,b)) then
        begin woord; if laatste  $\neq$  ng then zet(laatste); zet(e);
        zet(t); je
        end wiggetje;
        if laatste = b  $\vee$  laatste = g  $\vee$  ww = w3(p,o,p) then
        begin woord; je
        end wigje
        end
    end lammetje
    else if laatste = m then
    begin woord; zet(p); je
    end bloempje
    else if lmnr(laatste)  $\vee$  laatste = w then
    begin if ww = w5(o,n,g,e,n) then aantal:= aantal - 1;
    woord;zet(t); je
    end leeuwtje
    else
    begin woord; if laatste = j then zet(e) else je
    end woordje
    end woord eindigend op medeklinker
    end gecomprimeerde woord;
    goto LEES WOORD;

klinkerverandering: aantal:= aantal + 1;
if voorlaatste = i then W[aantal - 2]:= voorlaatste:= e;
W[aantal - 1]:= voorlaatste; W[aantal]:= laatste;
goto bepaal verkleinwoord
end

```

buur, buurtje	verdrag, verdragje
buurt, buurtje	gebed, gebedje
diploma, diplomaatje	hof, hofje
diplomaat, diplomaatje	jabot, jabootje
loge, logeetje	dag, dagje, daagje
loge, logeetje	heg, heggetje, hegje
pad, padje, paadje	ei, eitje
gat, gatje, gaatje	blauw, blauwtje
filmster, filmstertje	geel, geeltje
rancune, rancunetje	smal, smalletje
opa, opaatje	zoet, zoetje
radio, radiootje	abc, abcje
menu, menuetje	c, c'tje
cafe, cafeetje	stoomschip, stoomscheepje
lolly, lolly'tje	staatslot, staatslotje
ski, skietje	houtvlot, houtvlotje
kastanje, kastanjetje	aquarel, aquareltje
aai, aaitje	label, labeltje
blad, blaadje	oefening, oefeninkje
vat, vaatje	cahier, cahiertje
glas, glaasje	passepartout, passepartoutje
schip, scheepje	bourgeois, bourgeoisje
augur, auguurtje	entremets, entremetsje
depot, depootje	entredoux, entredouxje
wig, wiggetje, wigje	madam, madammetje
koning, koninkje	man, mannetje
lam, lammetje	moe, moetje
bloem, bloempje	schrede, schredetje
leeuw, leeuwte	ij, ij'tje
hiel, hielte	claque, claqueje
klad, kladje	koriste, koristetje, koristje
oorlog, oorlogje	pad, padje, paadje
l, l'etje	spit, spitje, speetje
karbonade, karbonaadje	staf, stafje, staafje
machine, machinetje, machientje	gat, gatje, gaatje
rad, radje, radertje, raadje	stag, staagje
parachute, parachuutje	pop, poppetje, popje
leerling, leerlinkje	extempore, extemporeetje
jongen, jongetje	facsimile, facsimileetje
club, clubje	directoire, directoirtje
kip, kippetje	kleine, kleintje
diner, dineetje	sinecure, sinecuurtje
fetisj, fetisje	brunette, brunetje
crapaud, crapaudje	pelerine, pelerientje
ave, aveetje	terraine, terrientje
stede, stedetje	serafine, serafientje
meester, meestertje	operette, operettetje, operetje
zeester, zeestertje	matelot, matelootje
lot, lotje, lootje	pierrot, pierrootje

4. Conclusie

Op pagina 11 is het resultaat van het programma afgedrukt zoals dat door de Electrologica X8 van het Mathematisch Centrum in 10 seconden werd geproduceerd voor de in dit rapport genoemde woorden. Voor de 24 woorden rechts bovenaan de bladzijden van de woordenlijst, waarvan een verkleinwoord wordt opgegeven, was het resultaat:

bloeze, bloezetje, bloesje
communiqué, communiqueetje
druiper, druipertje
flambeeuw, flambeeuwtje
fundering, funderinkje
groen, groentje
knuppel, knuppeltje
molen, molentje
pantomime, pantomimetje
patroon, patroontje
piedestal, pedestalletje
plager, plagertje
plastron, plastronnetje
plooi, plooitje
poeder, poedertje
richel, richeltje
schee, scheetje
smeel, smeeltje
verkoping, verkopinkje
vleugel, vleugeltje
vlo, vlootje, vlooitje
worm, wormpje
zanger, zangertje
zin, zinnetje

De verkeerd geproduceerde verkleinwoorden vallen uiteen in zes groepen:

1. Afkortingen

Is een afkorting van punten voorzien, dan kan het verkleinwoord mechanisch verkregen worden, doordat de verkleinwoorden van de 27 lëtternamen geprogrammeerd werden: "a.b.c" krijgt dan "a.b.c'tje".

Maar als deze punten ontbreken wordt een verkeerd verkleinwoord verkregen. Daar de Woordenlijst geen afkortingen opneemt vonden we daarin alleen "abc" dat door het programma "abcje" als verkleinwoord krijgt.

2. Samenstellingen

Doordat in de uitzonderingenlijsten niet alle samenstellingen worden opgenomen, krijgen samenstellingen waarvan het laatste lid een onregelmatig verkleinwoord heeft soms een verkeerd verkleinwoord. Een woord als "stoomschip" krijgt wel "stoomscheepje" daar alleen de laatste 5 letters van het woord bekeken worden. Deze oplossing zou echter voor b.v. "staatslot" niet voldoen, daar dan, op grond van "lot" → "lootje", "houtvlot" → "houtvlootje" als verkleinwoord zou krijgen. De groep van samengestelde woorden geeft voor de meervoudsvorming en de werkwoordvervoeging soortgelijke moeilijkheden.

3. Nederlandse woorden waar kennis over de uitspraak vereist is.

Dit geldt met name voor:

- a. het herkennen van de stomme e in de laatste lettergreep, en:
- b. kennis over de klemtoon bij een laatste lettergreep op ing.

Alleen over 1-lettergreepige woorden kon beslist worden. Bij meer-lettergreepige woorden werd, op grond van verwachtingen over de frequenties, beslist dat de e in de laatste lettergreep stom is (zodat "zeester" een verkeerd verkleinwoord krijgt) en dat de uitgang -ing tot inkje wordt (zodat "leerling" een verkeerd verkleinwoord krijgt).

Dit veroorzaakt foute verkleinwoorden bij de volgende woorden uit de Woordenlijst:

- a. aquarel, ulevel, label (ook "labeltje" is toegestaan) en 12 twee-lettergreepige woorden op "el", die ten onrechte de uitgang "eltje" kregen.
- b. oefening, opening, tekening, verandering, vergadering, versnapering en 11 woorden op "ling" die ten onrechte de uitgang "inkje" kregen.

4. Vreemde woorden

De (Franse) woorden die in de uitspraak op een klinker, maar in de spelling op een medeklinker eindigen geven moeilijkheden. Woorden als

"depot" en "jabot" worden behandeld als "lot". Woorden als "cahier" en "passe-partout" gaan toevallig goed. Bij andere woorden volgt de Woordenlijst geen consequente gedraglijn: "bourgeois" krijgt "bourgeoistje", maar "entremets" wordt verkleind tot "entremetje", terwijl voor "entre-deux" zowel "entre-deuxtje" als "entre-deutje" wordt opgegeven. Voor nog 5 Franse woorden geeft het programma verkeerde verkleinvormingen om deze reden.

Eveneens worden niet gevormd: "dessintje", "plafonnetje", "souvenir-tje", "framepje" en "stenciltje".

Wel zorgt het programma voor "dineetje", "dejeuneetje" en "soupeetje".

5. Verkleinwoorden op "ke"

De Woordenlijst geeft: "boorlingske", "kersouwke" en "stedeke"; en als nevenworm naast de normale: "madammeke", "manneke" en "moeke". Deze worden niet voortgebracht.

6. rariteiten als "schreetje" en "bloemetje"

Deze 6 groepen tellen tezamen 54 woorden. Van ongeveer 2700 woorden geeft de Woordenlijst het verkleinwoord op, zodat ongeveer 98% van de gebruikelijke verkleinwoorden, niet gevormd door "je" achter het woord te zetten, automatisch wordt voortgebracht.

Voor werkelijke toepassing voor de automatische vorming van verkleinwoorden in de "Woordenlijst van de Nederlandse taal" zal het programma vereenvoudigd kunnen worden, en deze uitzonderingen kunnen dan apart in de lijst opgenomen worden.

Voor de woorden met klinkerverandering ligt een kenmerking in verband met de meervoudsvorming voor de hand.

Ook is een versnelling van het programma mogelijk: ter wille van de duidelijkheid voor de lezer werd met de factor tijd in het geheel geen rekening gehouden.

Wat betreft de beide in §1 genoemde doeleinden kan nu geconcludeerd worden:

1. Een exacte formulering van de verkleinwoordvorming is in de notatie van §2 haast onmogelijk, maar is in ALGOL 60 goed op te schrijven, met

het voordeel van rigoureuze controle door de rekenautomaat. Een parafrase van het programma in "gewone taal" zou weinig zinvol, want moeilijk leesbaar zijn. Ook voor een taalkundige kan een algemene programmeertaal als ALGOL 60 dienst doen om een gecompliceerd systeem te beschrijven. Niet alle onderdelen van ALGOL zijn hierbij gebruikt, de voornaamste nuttige elementen zijn:

de begin-end-nesting, de if-then-else-constructie, het procedure-mechanisme en de for-statement.

2. Van slechts een honderdtal woorden zal een afwijkend verkleinwoord opgenomen hoeven te worden (n.l. de 43 in het programma genoemde woorden en de 54 hierboven genoemde bijzondere woorden).

Dank ben ik verschuldigd aan de heren Meertens en Geurts (die t.z.t. een soortgelijk programma voor de meervoudsvorming zullen publiceren).

- [1] Woordenlijst van de Nederlandse taal,
's-Gravenhage, 1954 (herdruk januari 1956).
- [2] A. Cohen, Het Nederlands diminutiefsuffix: een morfonologische proeve,
Nieuwe Taalgids 51 (1958), 40-45.
- [3] P. Naur e.a., Revised Report on the Algorithmic Language ALGOL 60,
Copenhagen 1962.
- [4] H. Schultink, De morfologische valentie van het ongelede adjectief
in modern Nederlands, Van Goor Zonen, Den Haag, 1962.
- [5] F.E.J. Kruseman Aretz, Het M.C.-ALGOL 60-systeem voor de X8.
Voorlopige programmeurshandleiding, Mathematisch Centrum MR 81, 1966.

