**Projectbeschrijving Boontje**

**Doel van het project**

**Fases in het project**

1. *Corpusvoorbereiding*
2. *Testen stijlkenmerken Boon*
   1. **Verschillende soorten testen oplijsten.**
      1. Duidelijke voorkeur voor typische dialectwoorden uit het Oost-Vlaams. Voorbeelden hiervan zijn *klodde*, *gremel*, *frak*, en *tatjespap*.
      2. Woorden met een archaïsche klank (maar in het Vlaams vaak nog gangbaar)  
         Voorbeelden zijn *immer*, *aanschijn* en *proper*.
      3. Gebruik van genitief s  
         bijvoorbeeld: *des wolfs stok*
      4. Genitief in tijdsaanduidingen  
         Bijvoorbeeld: ‘*sanderendaags* en ‘*snoenens*
      5. Vaker gebruik van participia praesentis  
         Bijvoorbeeld : *En dat alles eraan toegevoegd hebbende.*
      6. Vaker verbinden van infinitieven met beginnen  
         Bijvoorbeeld: *ze waren beginnen te redetwisten*
      7. Vaker het hulpwerkwoord *zullen* door *gaan* of *moeten* vervangen.  
         Bijvoorbeeld: *Ze gingen eraan bedrogen worden* en *Moest ik Ondine niet zijn*
      8. Verbinding van lidwoord met daaropvolgend zelfstandig naamwoord  
         Bijvoorbeeld: ‘tminnesterie, ‘tkasteel, ‘twitpapier
      9. Verbinding van lidwoord en werkwoord  
         Bijvoorbeeld: ‘tis en ist’
      10. Verbinding van persoonlijk voornaamwoord en werkwoord  
          Bijvoorbeeld: ‘kzal
      11. Vocaalverzwakking  
          Bijvoorbeeld: *paster* in plaats van *pastoor*
      12. Verdwijnen van de h aan woordbegin  
          Bijvoorbeeld: *allo* voor *hallo*
      13. Gebruik van hoofdletters: Gewoonlijk alleen aan begin van zin en bij woord(delen) die nadruk krijgen.   
          Bijvoorbeeld: *allEEn, het is DAT wat ….*
      14. Opvallend gebruik van het woord *dingen*.
      15. Overvloedig gebruik van tussenvoegsels en uitroepen  
          Bijvoorbeeld: *milledju, bogot, bah, hewel, ei, godomme, tsja, a là.*
      16. Veelvuldig gebruik van vaste uitdrukkingen of zegswijzen en spreekwoorden.
          1. Veelvuldig gebruik van de vaste verbinding  
             Bijvoorbeeld: *‘tis te zeggen* of *plus daarbij of + daarbij* of *en plus daarbij* of *+ daarbij en daarenboven* of *t.t.z.*
          2. Verbasterde spreekwoorden en gezegden  
             Bijvoorbeeld: *Van een pink een arm maken* of *het vroos kasseistenen* of *iemand bij zijn pietje hebben.* Ook stukken van spreekwoorden gecombineerd tot verbasterde vorm.
      17. Woordgrappen (en klanknabootsingen)  
          Bijvoorbeeld: juniversiteit, medicasijnen, tsikkelogisch, pasternoster, alderijligen.
      18. Veelvuldig gebruik van rijm, zowel eindrijm als alliteratie  
          Bijvoorbeeld: rijm in zin: … en zag dat allen, zeer ontdaan, de koning en zijn nieuwe nobeline staarden aan … want hijzelf was nog zeer onlangs een nieuw huwelijk aangegaan.  
          Bijvoorbeeld: *fladderende vlinder, dorre droge ogen, hij spuwt zijn spog.*
      19. Veelvuldig gebruik van de zelfcorrectie  
          Bijvoorbeeld: *ik was benieuwd wat die gek over pannekoeken… neen, over kunst… zou te zeggen hebben*Ook vaak aangekondigd via: *tis te zeggen* of *t.t.z.*
      20. Veelvuldig gebruik van enumeratie  
          Bijvoorbeeld: *steelt en bedriegt en hoereert en boeleert, zoals het ….*
      21. Discussie over gebruik epitheta en gebruik adjectieven. Nagaan of er veel epitheta zijn en voor de rest weinig adjectieven.  
          Bijvoorbeeld: epitheta: de schone vrouw Lucette, de kantieke schoolmeester, Johan Janssens, dichter en dagbladschrijver, de oude Bultharkas, jeannine de voorlopige, de schele malvine
   2. **Per test aanduiden wat nodig is. + erbij oplijsten welke secundaire bronnen nodig zijn voor de testen**
      1. Duidelijke voorkeur voor typische dialectwoorden uit het Oost-Vlaams.   
         Externe bron: dialectwoordenboek (www.vlaamswoordenboek.be)  
         Te onderzoeken: Hoe binnenhalen?  
           
         Geen tagging nodig
      2. Woorden met een archaïsche klank (maar in het Vlaams vaak nog gangbaar)  
         Voorbeelden zijn *immer*, *aanschijn* en *proper*.  
           
         Externe bron: Van Dale met omschrijving archaîsch?  
           
         Te onderzoeken: Hoe binnenhalen? (wel 3 dagen gratis online proberen)

Eventueel te combineren met tagging

Laagste prioriteit.

* + 1. Gebruik van genitief s  
       bijvoorbeeld: *des wolfs stok*  
       🡪 De officiële regel is dat de bezits-s in principe aan een naam vast geschreven wordt: Piets fiets, Mariekes vriendin, Oscars tante, etc.  
       3 regels: a. Alleen als de s de uitspraak van de voorafgaande klank zou veranderen, staat er een apostrof voor de s. Dat is het geval als de laatste lettergreep op [één enkele a, i, o, u of y](https://onzetaal.nl/taaladvies/bezits-s-algemene-regels#enkele klinker) eindigt, of op een [enkele e die als [ee]](https://onzetaal.nl/taaladvies/bezits-s-algemene-regels#enkele klinker) klinkt. Bijvoorbeeld: Anna's auto, Johnny's baardje en Zoë's baan.  
       b. Na een [dubbele klinker of een tweeklank](https://onzetaal.nl/taaladvies/bezits-s-algemene-regels#dubbele klinker) komt geen apostrof (Coetzees roman, Lidewijs gedicht). c. Na een naam die op een [hoorbare sisklank](https://onzetaal.nl/taaladvies/bezits-s-algemene-regels#sisklank) eindigt, komt wel een apostrof maar geen extra s (Bas' column).

Geen externe bron nodig.

Eventueel te combineren met tagging

5 gevallen te checken:

a. aantal keer *des* op totaal aantal woorden. Uitsluiten van *des* *te* en van delen van een woord als *paddenstoel.*  
b. woorden eindigend op ‘s

c. woorden eindigend op s’

d. woorden eindigend op ‘

e. woorden eindigend op s

Telkens meerdere mogelijkheden bij b tem e.

Te onderzoeken: hoe verwarring met quotes vermijden?

Prioriteit middelmatig.

* + 1. Genitief in tijdsaanduidingen  
       Bijvoorbeeld: ‘*sanderendaags* en ‘*snoenens*  
       🡪 te herkennen aan ‘s

Zie oefening iii. Te combineren met tagging.

* + 1. Vaker gebruik van participia praesentis  
       Bijvoorbeeld : *En dat alles eraan toegevoegd hebbende.*   
       Vervoegd via infinitief + d of de  
       Kan ook gebruikt worden als bijvoeglijk naamwoord of zelfstandig naamwoord.

Tagging handig maar niet noodzakelijk

Screenen op infititieven = Screenen op woorduitgang: minstens 2 letters + en + d of de. OF woordeinde (er mogen letters voor staan) *zijn*, *gaan*, *staan* of *slaan* + d of de

prioriteit hoog zonder tagging

* + 1. Vaker verbinden van infinitieven met beginnen  
       Bijvoorbeeld: *ze waren beginnen te redetwisten*

Tagging niet nodig.

Te bundelen met oefening V, namelijk herkennen van infinitieven.  
Te bundelen met werkterrein binnen een zin: oefening VII en XIII.

Screenen op ‘beginnen’ of ‘beginnen te’ + infinitief in dezelfde zin. (infinitief = minstens 2 letters, waarvan laatste medeklinker is + ‘en’ OF zijn, woordeinde van gaan, staan, slaan, doen of zien in dezelfde zin.

Prioriteit hoog.

* + 1. Vaker het hulpwerkwoord *zullen* door *gaan* of *moeten* vervangen.  
       Bijvoorbeeld: *Ze gingen eraan bedrogen worden* en *Moest ik Ondine niet zijn*Te bundelen met V en VI, nl. zoeken op infinitieven.   
         
       Screenen op ga, gaat, gaan, ging, gingen + infinitief in dezelfde zin OF moet, moeten, moest, moesten + infinitief in dezelfde zin.   
         
       Prioriteit hoog.
    2. Verbinding van lidwoord met daaropvolgend zelfstandig naamwoord  
       Bijvoorbeeld: ‘tminnesterie, ‘tkasteel, ‘twitpapier  
       Geen tagging nodig.  
         
       Te screenen op ’t of ‘d.   
       Samen te nemen met IX 🡪 bijkomende screening op t’  
       Samen te nemen met X 🡪 bijkomende screening op ‘kPrioriteit hoog
    3. Verbinding van lidwoord en werkwoord  
       Bijvoorbeeld: ‘tis en ist’Zie Viii
    4. Verbinding van persoonlijk voornaamwoord en werkwoord  
       Bijvoorbeeld: ‘kzal  
       Zie Viii
    5. Vocaalverzwakking  
       Bijvoorbeeld: *paster* in plaats van *pastoor*

Woorden zoeken in alle corpora die identiek zijn, behalve de laatste klinker(s) van het woord.   
  
Tagging kan helpen om woorden in dezelfde categorie te zoeken.

Prioriteit laag.

* + 1. Verdwijnen van de h aan woordbegin  
       Bijvoorbeeld: *allo* voor *hallo*  
         
       Woorden zoeken in alle corpora die identiek zijn, behalve de eerste letter   
       *h* die weggevallen is.  
         
       Tagging kan helpen om woorden in dezelfde categorie te zoeken.  
         
       Prioriteit laag.  
         
       Te bundelen met XI
    2. Gebruik van hoofdletters: Gewoonlijk alleen aan begin van zin en bij woord(delen) die nadruk krijgen.   
       Bijvoorbeeld: *allEEn, het is DAT wat ….*  
         
       Geen tagging nodig.  
         
       Nagaan of er ergens anders dan aan het begin van de zin hoofdletters worden gebruikt. Indien ja, nagaan of het meer dan 1 opeenvolgende letters zijn.   
         
       Begin van een zin herkennen na 1 punt + spatie of na 1 of meerdere witregel(s).  
         
       Hier moet gekeken worden binnen de begrenzing van een zin, net zoals in oefening VI en VII.  
       Prioriteit hoog.
    3. Opvallend gebruik van het woord *dingen*  
         
       Geen tagging nodig.  
         
       Te herkennen door aantallen te tellen en te vergelijken met andere corpora, in verhouding tot totaal aantal woorden.   
       Ook eventueel betrekken op aantal zinnen met/zonder ‘dingen’.  
         
       Hoogste prioriteit
    4. Overvloedig gebruik van tussenvoegsels en uitroepen  
       Bijvoorbeeld: *milledju, bogot, bah, hewel, ei, godomme, tsja, a là.*  
       Tagging om tussenvoegsels te herkennen?  
         
       Te vergelijken met lijst met tussenvoegsels of uitroepen?  
         
       Eventueel de voorbeelden letterlijk nakijken en tellen.  
         
       Lage prioriteit
    5. Veelvuldig gebruik van vaste verbinding

Veelvuldig gebruik van de vaste verbinding  
Bijvoorbeeld: *‘tis te zeggen* of *plus daarbij of + daarbij* of *en plus daarbij* of *+ daarbij en daarenboven* of *t.t.z.*  
Geen tagging nodig  
  
Letterlijk controleren en tellen en vergelijken met andere corpora.  
  
Hoge prioriteit

* + 1. Veelvuldig gebruik van verbasterde spreekwoorden en gezegden  
       Bijvoorbeeld: *Van een pink een arm maken* of *het vroos kasseistenen* of *iemand bij zijn pietje hebben*   
         
       Geen tagging nodig.  
         
       Externe bron nodig: wikipagina met spreekwoorden of dbnl pdf met spreekwoorden.   
         
       Te checken met woordenboek van spreekwoorden. Ook stukken van het spreekwoord laten meetellen om de verbasterde vormen te detecteren.  
         
       Te bekijken binnen de begrenzing van 1 zin. Zie ook oefening VI en VII en XIII.  
         
       Middelmatige prioriteit.
    2. Woordgrappen (en klanknabootsingen)  
       Bijvoorbeeld: juniversiteit, medicasijnen, tsikkelogisch, pasternoster, alderijligen.  
         
       Te herkennen door te screenen op woorden die veel letters gemeen hebben, maar niet allemaal.  
         
       Geen tagging nodig, eventueel wel handig indien gescreend wordt op woorden uit andere corpora.

Zeer lage prioriteit.

* + 1. Veelvuldig gebruik van rijm: eindrijm   
       Bijvoorbeeld: rijm in zin: … en zag dat allen, zeer ontdaan, de koning en zijn nieuwe nobeline staarden aan … want hijzelf was nog zeer onlangs een nieuw huwelijk aangegaan.  
       Bijvoorbeeld: *fladderende vlinder, dorre droge ogen, hij spuwt zijn spog.*Geen tagging nodig.   
         
       Binnen bereik van 1 zin blijven. Zie ook oefening VI en VII en XIII en XVI en XX.  
       Zoeken naar lettercombinaties aan einde (minstens 2 letters) van het woord in elkaars buurt. (bijvoorbeeld maximum 1 woord ertussen toelaten).   
         
       Opletten voor gelijkaardige klanken met verschillende schrijfwijzen *(au en ou, d en t, ei en ij*) en voor klankhebbend en klankloos gecombineerd.  
         
       Prioriteit middelmatig
    2. Veelvuldig gebruik van rijm: alliteratie  
         
       Geen tagging nodig.   
         
       Binnen bereik van 1 zin blijven. Zie ook oefening VI en VII en XIII en XVI en XIX.  
         
       Zoeken naar lettercombinaties aan begin (1 letter volstaat) van het woord in elkaars buurt. (bijvoorbeeld maximum 1 woord ertussen toelaten).   
         
       Opletten voor gelijkaardige klanken met verschillende schrijfwijzen *(au en ou, ei en ij*) en voor klankhebbend en klankloos gecombineerd.  
         
       Prioriteit middelmatig
    3. Veelvuldig gebruik van de zelfcorrectie  
       Bijvoorbeeld: *ik was benieuwd wat die gek over pannekoeken… neen, over kunst… zou te zeggen hebben*

Geen tagging nodig.  
Letterlijk zoeken via *neen* en via *tis te zeggen* of *t.t.z.*Hoogste prioriteit wegens eenvoud.

* + 1. Veelvuldig gebruik van enumeratie  
       Bijvoorbeeld: *steelt en bedriegt en hoereert en boeleert, zoals het ….*

Geen tagging nodig.  
  
Letterlijk zoeken op gebruik van *en* en het aantal tellen op het totaal aantal woorden.

Hoogste prioriteit.

Ook te tellen binnen het bereik van 1 zin 🡪 gemiddelde per zin.  
Ook te tellen als aantal zinnen waarin *en* als woord (dus niet binnen een woord) voorkomt.

* + 1. Discussie over gebruik epitheta en gebruik adjectieven. Nagaan of er veel epitheta zijn en voor de rest weinig adjectieven.  
       Bijvoorbeeld: epitheta: de schone vrouw Lucette, de kantieke schoolmeester, Johan Janssens, dichter en dagbladschrijver, de oude Bultharkas, jeannine de voorlopige, de schele malvine  
         
       Kan zonder tagging.   
         
       Te testen als woorden die meerdere keren in vaste volgorde na elkaar voorkomen. Uitsluitel van lidwoorden.   
       Zie ook oefening xvi. 🡪 nagaan of bepaalde woorden vaker na elkaar voorkomen. Beginnen bij hoogst mogelijke dat meer dan eens voorkomt (bv. vaste verbinding van 5 woorden), dan vaste verbinding van 4 woorden enz. Samen te bekijken met oefening op vaste uitdrukkingen.

Indien tagging, dan testen op vast voorkomen van adjectief + substantief.   
Binnen de begrenzing van 1 zin bekijken.   
  
Bijvoorbeeld: weinig adjectieven  
  
Tagging nodig.   
  
Uitsluiting epitheta en vaste verbinding.  
Middelmatige prioriteit.

* 1. Testen groeperen

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Woord-niveau | Zins-niveau | Tekst-niveau | Tag nodig | Tag mag | Externe bron | Combi? |
| I | Dialect | X |  |  |  | X | X |  |
| II | Archaisch | X |  |  |  |  |  |  |
| III | Des + genitief | X |  |  |  |  |  |  |
| IV | Genitief tijd | X |  |  | X |  |  |  |
| V | Participia praesentis | X |  |  |  | X |  | Voor VI en VII |
| VI | Beginnen hulpww |  | X |  |  |  |  | V en VI |
| VII | Gaan of moeten hulpww |  | X |  |  |  |  | V en VI |
| VIII | Clisis lidw + zelfs nw | X |  |  |  | X |  | VIII, IX en X |
| IX | Clisis lidw + werkw | X |  |  |  | X |  | VIII, IX en X |
| X | Clisis pers vnw+ww | X |  |  |  | X |  | VIII, IX en X |
| XI | Vocaalver-zwakking | X |  |  |  |  |  | XI en WII |
| XII | Geen h aan begin | X |  |  |  |  |  | XI en XII |
| XIII | hoofdletter |  | X |  |  |  |  |  |
| XIV | Dingen | X |  |  |  |  |  |  |
| XV | tussenwerpsels |  | X |  | X |  |  |  |
| XVI | Vaste verbinding |  | X |  |  |  |  | Na XXIII |
| XVII | Spreekwoorden ook verbasterd |  | X |  |  |  | X | Na XVI en XXIII |
| XVIII | Woordgrappen | x |  |  | X |  |  |  |
| XIX | Eindrijm |  | X |  |  |  |  | XIX en XX |
| XX | Alliteratie |  | X |  |  |  |  | XIX en XX |
| XXI | Zelfcorrectie |  | X |  |  |  |  | Zie XVI en XXIII |
| XXII | Enumeratie |  | X |  |  |  |  |  |
| XXIII | Epitheta |  | X |  |  | X |  | XVI en XVII |

* 1. Testen prioretiseren  
     Zeker nodig:
     1. Telling aantal woorden per corpus
     2. Telling aantal zinnen per corpus
     3. Infinitieven definiëren
     4. Telling + lijst opstellen van vaste combinaties op zinsniveau

Beginnen met 3 gemakkelijkste:  
Oef 14: dingen  
Oef 22: enumeratie  
Oef 5 : participia praesentis

* 1. Secundaire bronnen binnenhalen
  2. Tagging uitvoeren
  3. Per test beschrijven welke output in welke vorm ik wil (lijst, percentage …)
  4. Evaluatie per test inrichten

1. *Presentatie resultaten*
   1. Hoe moet het resultaat eruit zien?
   2. Resultaten visualiseren
   3. Omschrijvingen bij resultaten?
2. *Eventueel later: Testen stijlkenmerken Volgelingen*
3. *Eventueel later: Verwerking Volgelingen in presentatie resultaten*