Try to account for this: the *probeer/ proberen* construction continuum in Afrikaans and Dutch

Cora Pots & Liesbeth Augustinus

KU Leuven / FWO-Vlaanderen

AGW2, Ghent - 6th October 2018

Outline

- Introduction

- Goals

- Methodology

- Results

- Conclusions & outlook

INTRODUCTION

Introduction

Afrikaans

- *dat Theresa (ge)probeer het **om te** dans.*

- *dat Theresa (ge)probeer dans het.*

  'that Theresa has tried to dance.'

- *dat Theresa dit (ge)probeer het.*
  'that Theresa has tried this.'

Introduction

Dutch

- *dat Leen heeft* geprobeerd */ * proberen* **om te** *dansen.*

- *dat Leen heeft* geprobeerd */ proberen* **te** *dansen.*

- *dat Leen heeft ** geprobeerd */ proberen dansen.*

  'that Leen has tried to dance.'

Introduction

Dutch

- *dat Leen heeft* *geprobeerd* / * *proberen* **om te** *dansen.*

- *dat Leen heeft* *geprobeerd* / *proberen* **te** *dansen.*   **IPP-effect**

- *dat Leen heeft* * *geprobeerd* / *proberen* *dansen.*

  'that Leen has tried to dance.'

- Cf. *dat Leen het heeft geprobeerd / *proberen.*
     'that Leen has tried it.'

Introduction

Dutch

- *dat Leen heeft geprobeerd / * proberen* **om te** *dansen.*   NL/VL

- *dat Leen heeft geprobeerd / proberen* **te** *dansen.*   NL/VL

- *dat Leen heeft * geprobeerd / proberen dansen.*   *NL/VL

'that Leen has tried to dance.'

Goals

- Investigate variation in *probeer/proberen* paradigm
  - Variation without difference in meaning
  - Special status of the verb (most verbs opt for one type of infinitival complement)

- Investigate relevant parameters
  - Infinitival complement (*om te*, *te* or bare infinitive)
  - Type and length of the object
  - For Dutch: NL vs VL, verb form, position of the object

- Ultimate goal: find out why Afrikaans *probeer* is more restricted than Dutch *proberen*

Methodology

Corpus investigation

- Afrikaans: Korpusportaal
  - 150M words
  - Written Afrikaans
  - Different genres

- Dutch: OpenSoNaR
  - 500M words
  - Written Dutch
  - Different genres

SoNaR-500	%
VL	78.24
NL	15.3
NL/VL	6.46

Queries

- Instances of Afrikaans *(ge)probeer* and Dutch *proberen/ geprobeerd* selected by a temporal auxiliary

- Followed by an *om te* infinitive, a *te* infinitive or a bare infinitive

- In main and embedded clauses (introduced by *dat* 'that')

- Dutch data: only consider data labelled for region (VL versus NL)

AFRIKAANS

Results *probeer*

	probeer	*geprobeer*	Total
Om te infinitive	6 228 (47.07%)	145 (95.39%)	6 373
Te infinitive	0	0	0
Bare infinitive	7 004 (52.93%)	7 (4.61%)	7 011
Total	**13 232 (100%)**	**152 (100%)**	**13 384**

- Results: manually filtered
- > 200 hits: random sample (200 hits) checked, results extrapolated to estimate frequencies
- 14.9% of the constructions with lemma *probeer*

Results *probeer*

- 14.9% of the constructions with lemma *probeer*

- Analogy with unambiguous verbs taking a bare infinitive (e.g. *kom* 'come', *bly* 'continue') → *probeer* = INF in those constructions

- Analogy with other verbs taking an (*om*) *te* infinitive (e.g. *begin* 'begin', *help* 'help') → *te* infinitive replaced by an *om te* infinitive

- *geprobeer*: colloquial form? Text genre?
  → future work

- Size and type of the object → regression analysis

Regression analysis *probeer*

Variable	Value
Object present	yes/no
Object type	bare, <u>def</u>, indef, pron, proper, clause, none
Object number	<u>sg</u>/pl/none
Object size (# syllables)	[0 -∞]
V2	yes/no
Om present	yes/no

Regression analysis *probeer*

Variable	Value
Object present	yes/no
Object type	bare, <u>def</u>, indef, pron, proper, clause, none
Object number	<u>sg</u>/pl/none
Object size (# syllables)	[0 -∞]
V2	yes/no
Om present	yes/no

… terwyl paramedici probeer het   om <u>hulle</u> te stabiliseer.
… while paramedics try   have for them to stabilize
'… while paramedics have tried to stabilize them.'

Regression analysis *probeer*

Variable	Estimate	*p*-value
Object present	0.25	2.24e-06
Object type: clause pron	-0.82 -0.16	7.48e-06 0.00122
Object number		not significant
Object size (# syllables)	0.03	5.93e-06
V2		not significant

DUTCH

Results *proberen*

	NL	VL	Total
Om te infinitive	414 (19.82%)	2103 (44.42%)	2517 (36.89%)
Te infinitive	1674 (80.13%)	2459 (51.94%)	4133 (60.57%)
Bare infinitive	1 (0.04%)	172 (3.63%)	173 (2.54%)
Total	**2089 (99.99%)**	**4734 (99.99%)**	**6823 (100%)**

- Results: manually filtered
- > 200 hits: random sample (200 hits) checked, results extrapolated to estimate frequencies
- 4% of the occurrences with [lemma="proberen"] in both the NL and VL part of the corpus

Results *proberen*

Goodness of fit:
does the *(om) (te)* infinitive occur significantly more in Belgian Dutch or Netherlandic Dutch?

om te	NL	VL	Total
Observed	414	2103	2517
Expected	412	2105	2517
te	NL	VL	Total
Observed	1674	2459	4133
Expected	676	3457	4133
inf	NL	VL	Total
Observed	1	172	173
Expected	28	145	173

not significant
($P = 0.914$; $p=\leq 0.05$)

significant
($P < 0.001$; $p=\leq 0.05$)

significant
($P < 0.001$; $p=\leq 0.05$)

Results *proberen*

	NL		VL	
	proberen	*geprobeerd*	*proberen*	*geprobeerd*
Om te infinitive	0	228	0	346
Te infinitive	64 (21.41%)	235 (78.59%)	120 (43.32%)	157 (56.68%)
Bare infinitive	1	0	172	0
Total	65	463	292	503

- Results: manually filtered
- Only checked hits considered (no estimations)

CONCLUSIONS & OUTLOOK

Comparison

	Afrikaans	Dutch
Om te infinitive	Form of *probeer*: ambiguous between INF and PP	• Form of *proberen*: PP • No significant difference NL - VL
Te infinitive	unattested	• Form of *proberen*: PP/INF • More PP in NL; more INF in VL • Significantly more frequent in NL • Internal variation (region, position of the object, verb form)
Bare infinitive	Form of *probeer*: ambiguous, but probably INF	• Form of *proberen*: INF • Only in VL

Conclusion

- Afrikaans *(ge)probeer:* selects bare V or *om te* V
  → more rigid than Dutch

    NL > VL > AF

- Influential factors: presence of object, type of object (clause & pronoun), number of syllables

Future work

- Investigate intralingual variables, such as:
  - Influence of negation
  - … (suggestions? ☺)

- Investigate interlingual variables, such as:
  - Register
  - Genre
  - Dialectal differences in Dutch (SAND)

- Comparison with other verbs allowing different complementation patterns (e.g. AF *begin* 'begin', *aanhou* 'continue'; DU *beginnen* 'begin', *durven* 'dare')

Thank you!

Questions or suggestions?

liesbeth.augustinus@kuleuven.be

cora.pots@kuleuven.be

Results *proberen te*

- *dat Leen heeft geprobeerd / * proberen **om** <u>een boek</u> **te** lezen.*

- *dat Leen heeft geprobeerd / [%] proberen <u>een boek</u> **te** lezen.*

- *dat Leen <u>een boek</u> heeft [%] geprobeerd / proberen **te** lezen.*

- *dat Leens <u>een boek</u> heeft * geprobeerd / proberen te lezen.*

  'that Leen has tried to read that book.'

Results *proberen te*

Position of object: terminology

- *dat Leen heeft geprobeerd een boek te lezen.*   PP + obj
  extraposition

- *dat Leen een boek heeft proberen te lezen.*   Obj + INF
  IPP

- *dat Leen een boek heeft $^{%}$ geprobeerd te lezen.*   Obj + PP
  3rd construction

- *dat Leen heeft $^{%}$ proberen een boek te lezen.*   INF + obj
  IPP + cluster interruption

'that Leen has tried to read that book.'

Results *proberen te*

	NL	VL	Total
PP + obj (extraposition)	212 (70.90%)	157 (56.68%)	369 (64.06%)
obj + INF (IPP)	64 (21.40%)	120 (43.32%)	184 (31.94%)
obj + PP (3rd construction)	23 (7.69%)	0	23 (3.99%)
Total	**299 (99.99%)**	**277 (100%)**	**576 (99.99%)**

- Only checked hits with an object considered (576/1323)
- 3rd construction: only in NL, but in general uncommon
  → interaction between position of object and morphological form of *proberen*
- Factors that trigger variation, especially in NL?
  → future work