Amsterdam University Press

TAAL & TONGVAL 74.1 (2022) 10-54
HTTPS://DOI.ORG/10.5117/TET2022.1.002.PHEI

komen 'come' + Verb of Movement

Diatopic and Semantic Variation in Spoken Varieties of Dutch

Jeffrey Pheiff
Institut für Germanistik Universität Bern
jeffrey.pheiff@germ.unibe.ch

Lea Schäfer
Philipps-Universität Marburg
lea.schaefer@staff.uni-marburg.de

Abstract

Periphrastic constructions with COME have primarily been grammaticalized to express tense in Indo-European languages (Devos & van der Wal 2014). In the Germanic language group, COME has not undergone grammaticalization to the same degree that related GO has. Nevertheless, this verb has acquired some special functions when used in combination with other elements. One of them concerns the combination of COME with a motion verb. In Standard Dutch, the choice of the morphological form (INF/ PTCP) of the movement verb in this construction is variable (Haeseryn et al. 1997): *De agent kwam de straat ingefietst*.PTCP /*infietsen*.INF 'The police officer came cycling into the street'. This contribution investigates this special construction in terms of diatopic and register variation as well as from a semantic-functional perspective. We performed an experiment in which we tested for geographic and semantic factors. The results show that the distribution of the variants is not regionally conditioned contrary to our expectations. Instead, the infinitive variant is the preferred variant across all regions in regional Dutch. We then discuss the results for the semantic factors that we systematically integrated into the test conditions, i.e. lexical semantics and path and manner as has been previously proposed in the literature (Ebeling 2006, Honselaar 2010, Beliën 2016). The results of a regression analysis do not conform to expectations. We reflect on the results and propose an alternative hypothesis, based on Schäfer (2020), proposing that

the infinitive variant is the result of a stalled grammaticalization process, in which *komen* is – or better was – on its way toward becoming a future auxiliary. Future work will have to test this hypothesis.

Keywords: Dutch dialects, grammaticalization, language change, motion verbs, TAM (tense–aspect–mood), *komen* 'come' with past participle or infinitive

1 Introduction[1]

Periphrastic constructions with COME have primarily been grammaticalized to express tense in Indo-European languages (Devos & van der Wal 2014). In the Germanic language group, COME has not undergone grammaticalization to the same degree that related GO has. Nevertheless, this verb has acquired some special functions. One of them concerns the combination of COME with a motion verb in which the verb COME conveys an agent's motion, while a motion verb describes how the agent moves, i.e. its manner of motion. Typologically, West Germanic languages such as German, Dutch, and Yiddish show a degree of morpho-syntactic variation as the examples in 1–3 show (see also Schäfer 2020). In some languages, we find that speakers can only make use of either a single variant, either the past participle as in Standard German or the infinitive with *tsu* 'to' as in sources from written Yiddish. Yet even within a single language, we find variation as the example from Standard Dutch illustrates where speakers can make use of a motion verb in both its infinitive and past participle form.

1. Dutch: *De agent kwam de straat ingefietst*.PTCP/*infietsen*.INF
2. German: *Der Agent kam in die Straße geradelt*.PTCP/**radeln*.INF
3. Yiddish: *der agent komt tsu forn*.INF/**gefornPTCP in gas arayn mit a rover*
   'The police officer came cycling into the street'

1 We would like to thank our informants for taking time to complete our questionnaires. Furthermore, our thanks go out to Michael Cysouw (Marburg) for advising and supporting us with the statistical analysis as well as Sophie Ellsäßer (Osnabrück), Marina Frank (Oldenburg), Simon Kasper (Marburg), and the anonymous reviewers for their comments on previous drafts of this paper. Further, special thanks go to the editor of this journal, Gunther De Vogelaer (Münster), for his invaluable input in putting the final touches on this paper. Lastly, we would like to express our gratitude to participants of the Sociolinguistics Circle (2018), TABU Dag (2018), and ICLaVE 10 (2019) for constructively discussing this research with us.

In the following, we will focus on the construction with an infinitive (INF) and past participle (PTCP). In Standard Dutch, there are other variants of the construction such as with a present participle (PPTCP) (4) and a gerund (GRD) (5), the latter of which still occurs in some southeastern Dutch dialects (e.g. Weijnen 1966:313–314). The variant with a present participle will receive some attention in the context of the discussion of the dialectological evidence; however, for more on both of these variants, see Schäfer (2020). Following Beliën (2016), excluded from our discussion will be cases that have a "sequential" reading, and, in which a past participle is not possible (6).

4. *De agent kwam de straat infietsend*.PPTCP
   'The police officer came cycling into the street'
5. *Loupentere*.GRD *kaom ich häom tege* (Barbiers et al. 2006)
   'I met him while walking'
6. *Aan het diner kwam Paula naast mij zitten*.INF (Beliën 2016:21)
   'At dinner, Paula came and sat next to me'

Instead of pursuing a typological and diachronic perspective (as in Schäfer 2020), we would instead like to investigate the construction in terms of diatopic and register variation as well as from a semantic-functional perspective. Research has shown that there is variation on the level of Standard Dutch, manifesting itself in the form of regional preferences (Haeseryn et al. 1997). However, as yet, there have not been any empirical studies showing that these regional preferences exist, neither for local varieties of Dutch (i.e. dialects) nor for registers of Dutch closer to the standard language. Following Patocka (1993: 409), in light of the fact that regionally conditioned (morpho)syntactic variants can be expected in higher speech levels and varieties, we expect to find a similar distribution of variants in the dialects as in varieties situated closer to Standard Dutch. In addition to regional preferences, the choice of variants is claimed to be in part semantically conditioned (Ebeling 2006, Honselaar 2010, Beliën 2016); however, claims in this regard are based on small samples of qualitatively judged evidence and have not yet been systematically investigated, neither with a corpus study nor with a questionnaire. From these preliminary considerations, we can derive our preliminary research questions:

I. Is the construction COME + Motion Verb regionally conditioned in the dialects of Dutch?
II. Do varieties and speech levels closer to the standard language show a similar regional distribution as the dialects?
III. Are there semantic constraints that are responsible for structuring the variation as previously claimed in the literature (e.g. Beliën 2016, Ebeling 2006, Honselaar 2010)?

This contribution is structured as follows. In Section 2, we will give an overview of the variation in Standard Dutch. Then, in Section 3, we will discuss syntactic and semantic constraints previously discussed in the literature. In Section 4, we will then review typological, diachronic, and dialectological evidence. The discussion of the typological and diachronic evidence is based on the literature, and the typological discussion will focus exclusively on variation within the Germanic language grouping. The dialectological evidence that we present and discuss is novel. This evidence will give a preliminary answer to the question regarding the construction's distribution in the dialects.[2] To this end, we will examine data from four different series of data. This data is heterogeneous in nature from a methodological and historical standpoint. In this way, a diachronic component is also introduced into the analysis of the dialect data. It turns out that despite methodological differences used to collect the data and its varying historical provenience, the regional distribution of the construction COME + Motion Verb not only shows clear regional preferences for the morphological form of the motion verb, but that the distribution has remained stable in the dialects over the last ca. 100 years. The dialectological data that we present in Section 4 is of importance for two reasons. First, it provides extensive *empirical* evidence for the assumed regional distribution claimed throughout the literature. Second, it helps in setting the scene for the subsequent online survey that is the topic of Section 5. There, we present our methodology and results of a survey that we conducted to test the influence of geography and semantic factors on the choice of the morphological form of the movement verb. In Section 6, we will then present the results of a study we performed on registers of Dutch situated more closely to the standard variety[3], which not only tested for the regional distribution of the variants, but also systematically took semantic factors into account. After interpreting our results, we round off our contribution with a short summary in Section 7.

2 With Auer (2005), the term dialect is a relational concept, meaning that without a standard there is no dialect (of course, see the criticism of this viewpoint in Weiß 2009). The term dialect is used to refer to areal variability in language exclusively. Further, it is not only limited "to the 'base dialects', i.e. the most ancient, rural, conservative dialects, but will be used such as to include regional and urban varieties with a larger geographical reach as well" (Auer 2005:7–8). Between this pole dialect and another pole standard language, there is room for additional levels in between, resulting in different types of speech repertoires.

3 We will return to this point in Section 5 when discussing the registers that we attempted to capture in this investigation.

2 COME + Motion Verb in Standard Dutch

Haeseryn et al. (1997) describe the conditions under which the verb *komen* 'come' can be combined with a verb that specifies 'de manier van komen' ('the way of coming') in Standard Dutch. Generally speaking, this additional verb of movement can either occur as an infinitive or a past participle in three different circumstances. First, it can be formed with verbs of movement with a prefix on the verb indicating the direction as in examples 7–8.[4]

7. *Zonder kloppen kwam hij de kamer binnengelopen*.PTCP/*binnenlopen*.INF
   'Without knocking, he came walking into the room'
8. *Iedere morgen komt ze hier voorbijgefietst*.PTCP/*voorbijfietsen*.INF
   'She comes riding by here every morning'

Second, it can be formed with verbs of movement with a complement indicating direction such as in examples 9–10.[5] Here, the complement is an adverbial in the form of a prepositional phrase and there is no prefix on the verb.

9. *Als de baby maar even huilde, kwam vader naar boven gesneld*.PTCP/*snellen*.INF
   'As the baby began to cry, father came running upstairs'
10. *En daarvoor komt hij nou viermaal in de week naar Nijmegen gereden*.PTCP/*rijden*.INF
    'And he now comes [driving] to Nijmegen four times a week for it'

Third, it can be formed with verbs of movement combined with the prefix *aan* 'to'[6] such as *aanlopen/snellen/fietsen/rijden* 'run/hurry/bike/ride up to' as in example 11 as well as with other verbs that indicate the manner of coming by stating what the person who is coming is doing

4 Originally:"werkwoorden van beweging die samengesteld zijn met een bijwoord van richting" (Haeseryn et al. 1997).

5 Originally:"werkwoorden van beweging vergezeld van een bepaling van richting" (Haeseryn et al. 1997).

6 In Dutch, there is variation with regard to whether the separable prefix occurs adjacently to the motion verb in verb clusters. There is considerable intra- and interspeaker variation in this regard (van Usen, Haeseryn & Fickert 2013, Barbiers et al. 2008: Map: 2.3.2.2). Since we did not test for particle splits, we will not discuss them any further. Since some respondents produced them in their responses to our survey questions, we will return to particle splits briefly in Section 6.3.

while coming, e.g. *aanfluiten*/*mopperen* 'whistle/grumble up to' as in examples 12.[7]

11. *Van alle kanten kwamen mensen aangelopen*.PTCP/*aanlopen*.INF
    'People came running from all sides'
12. *Daar komt mijn neef Nurks weer aangemopperd*.PTCP/*aanmopperen*.INF
    'Here comes my nephew Nurks grumbling again'

In this context, there are also individual preferences for the use of the participle or infinitive variant. Investigating the Standard Dutch of Heerlen with a corpus of spontaneous speech, Cornips (2002) finds the infinitive variant and participle variant in the first two uses identified by Haeseryn et al. (1997) without any apparent difference in meaning. With regard to a use such as in 12, Cornips (2002:4) does find a preference for the infinitive form. Such individual speaker variation suggests that a linguistic change is ongoing; however, at the same time, the existence of competing formal variants may also suggest that there is a concomitant semantic differentiation for certain speakers. In the next section, we will discuss semantic and syntactic constraints on the use of the infinitive and participle variants of the *komen*-construction that have been previously identified in the research literature.

3 Semantic and Syntactic Constraints

3.1 Semantic Constraints

Haeseryn et al. (1997) do not hint at any semantic differences in meaning between the two forms. However, Ebeling (2006), Honselaar (2010), and Beliën (2016, 2017) suggest that the two constructions have very subtle semantic differences, i.e. that the occurrence of either variant may possibly be determined by semantic factors.

Ebeling (2006:409) argues that the INF is a verbal form not specified for tense or mood; the PTCP on the other hand is characterized by 'bisituationality' (Ebeling 2006:389, 392), i.e. it signals the existence of two situations, a process and its 'reminiscence', i.e. the resulting situation. The PTCP focuses on the final states, the situation at/from the moment of arrival, and the INF on the preceding phase, the situation of approaching.

7 Originally: "werkwoorden die gecombineerd zijn met het bijwoord *aan*" and "andere werkwoorden die 'de manier van komen' aanduiden door aan te geven wat degene die komt tijdens het komen doet" (Haeseryn et al. 1997)

Honselaar (2010:5–8), following Ebeling (2006), attempts to show a semantic difference between the INF and PTCP forms. Honselaar (2010) argues that the verb *komen* designates a state of affairs characterized by a motion towards a certain point (the deictic center). This is in contrast to the verb *gaan* 'go', which designates a state of affairs characterized by a motion away from the deictic center. Furthermore, *komen* neither specifies the manner of the motion nor the successful completion of the motion, i.e. *komen* does not specify whether the intended endpoint was reached. Lastly, the directional adverbial or prefix indicates the direction of the motion, with or without any specification of the endpoint or the route. Under these conditions, it becomes clear, according to Honselaar, that both the INF and the PTCP are equipped to specify the character of the 'coming'. That is to say that both verbal forms conform to the situation; they differ only with regard to their specification of the route vs. endpoint. Assuming that directional motions have a path and an endpoint that presuppose one another, then it is reasonable to expect that the INF and PTCP can be used interchangeably without any substantial practical communicative consequences. Honselaar demonstrates this with the sentence in (7), in which *de kamer* 'the room' acts as the deictic center into which *hij* 'he' walks. If the INF is used, then the speaker focuses on the event of arrival; if on the other hand the speaker uses the PTCP, then he places the focus on the arrival, i.e. the resulting state of being in the room.

Following Ebeling (2006) and Honselaar (2010), Beliën (2016, 2017) proposes in terms of cognitive grammar that 'the two variants in Dutch offer subtly different "construals" of the motion event: they differ in the way these events are conceptualized. At the highest level of semantic composition, where *komen* combines with the manner of motion verb + the directional phrase (or particle), either variant designates a deictic motion event that is unfolding' (Beliën 2016:228). However, the difference between the two variants PTCP and INF, is that while 'both variants describe an unfolding, unidirectional motion event towards a contextually construable vantage point, the variant with the past participle highlights the *end* of a process, while the infinitive variant does not' (Beliën 2016:30, italics in original).

To this end, Beliën (2016:24–30) reports on some small-scale corpus-based evidence to support her hypothesis. First, she searched the archives of the Dutch national newspaper *De Volkskrant* for instances of *aangerend*.PTCP 'ran (up) to' and *aanrennen*.INF 'run (up) to' with *komen*, for which she

Table 1: Results (Beliën 2016:28–30)

Contexts / Variants	Past participle	Infinitive
uit het ei gekropen/kruipen	84	8
de straat ingereden/inrijden	82	148

could find seven and eleven instances, respectively.[8] Beliën qualitatively compared the contexts in which both variants occurred. She found that the PTCP variant *aangerend* 'ran (up) to' has a 'sense of completion' that is absent with the INF variant *aanrennen*. Further, the PTCP variant is often accompanied by a description of what happens next, while the INF variant is accompanied by a description of what is happening at the same time. In the second instance, she compares Google hits for *uit het ei gekropen*.PTCP 'crept out of the egg' and *uit het ei kruipen*.INF 'creep out of the egg' on the one hand, as well as for *de straat ingereden*.PTCP 'ridden into the street' and *de straat inrijden*.INF 'ride into the street' on the other hand. Since the motion event in the former only involves a short path because the agent only has to move from the inside to the outside of the egg to complete it, the PTCP should be the dominant variant. Since the latter motion event is characterized by a potentially longer path, making it amenable to an interpretation focusing on the final state or not, speakers have more of a choice in their selection of variants depending on their perspective, thus we should expect more variation (Beliën 2016:28–30). Table 1 summarizes her data. It turns out that her expectations pan out. The PTCP is dominant in the context with a short path, while there is more variation in the latter context where either perspective is possible.

It can thus be assumed that the use of PTCP or INF depends on how a movement behaves in the *event frame*. According to Talmy (2000:257–309), there are five types of event frames: paths, causal chains, cycles, participant interactions, and interrelationships. Some aspects of an event frame can be highlighted, *windowed*, and others can be backgrounded, or *gapped* in his terminology:

> [T]he coherent referent situation with respect to which the windowing must take place is an *event frame*, the portions that are foregrounded by inclusion are *windowed*, and the portions that are backgrounded by exclusion are *gapped* (Talmy 2000:257; italics in the original).

8 In general, the COME + motion verb construction is of very low frequency in literary and spoken corpora of Dutch, German, Yiddish, and Afrikaans (Schäfer 2020).

With respect to Beliën (2016), it seems that the INF is not possible if the motion event is windowed, but rather for motion that is happening outside the windowed event or in the background. The Dutch differentiation between using *komen* + motion verb with PTCP or INF would therefore be a fine-tuned and syntactically denoted description of what is figure and what is ground.

3.2 Syntactic Constraints

In addition to the semantic constraints discussed in Section 3.1, there are also certain morpho-syntactic constraints that have been discussed in the literature for Standard Dutch and registers closer to the standard language. The following discussion is based largely on Cornips (2002); the sentences and the corresponding judgements also stem from Cornips (2002).

First, there is the IPP effect (*infinitivus pro participio*). IPP is the designation for the phenomenon whereby an infinitive takes the place of a participle form. In Standard Dutch, IPP occurs in the perfect tense of complex verb clusters with a modal verb. Cornips (2002) and Haeseryn et al. (1997) claim that there is an IPP effect for complex verb clusters with *komen* and a motion verb. Cornips (2002) presents the following data claiming that the complement of *komen* must obligatorily occur in the infinitive form as in 14 and 15; in these two latter sentences, the past participle variant is unacceptable.

13. *dat Jan is komen aanlopen* [1–2.INF–3.INF]
14. **dat Jan is komen aangelopen* [1–2.INF–3.PTCP]
15. **dat Jan is gekomen aan*(*ge*)*lopen* [1–2.PTCP–3.PTCP/INF]
    'that Jan has come running'

We would like to note, however, the existence of sentences such as 16, testifying to the uncertain status of this constraint. We suggest that future studies could investigate this syntactic factor more thoroughly, which we were unable to incorporate into our present study.

16. *De bestuurster van de Audi, een veertiger uit Lettelingen, verklaarde verrast te zijn geweest door de bijzonder hoge snelheid waarmee de Porsche is komen aangereden* [1–2.INF –3.PTCP].
    'The driver of the Audi, a forty-something year-old from Lettelingen, explained having been surprised by the especially high speed with which the Porsche came riding.'

Second, in the event that *komen* appears with an additional verb of movement in a subordinate clause, then finite *komen* precedes the infinitive form of verb of movement in the right periphery of the sentence: only the sequence [1–2] is acceptable, while the sequence [2–1] is unacceptable

(Cornips 2002). If finite *komen* occurs together with a past participle in a subordinate clause, then Cornips (2002) finds that the sequence 1–2 is preferred over the sequence 2–1; however, Cornips notes that speakers found both variants to be questionable, and that there was a number of speakers who preferred the sequence 2–1 over the sequence 1–2.[9] Therefore, there appears to be considerable variation in this regard as well.

17. *Ik weet zeker dat Jan daar komt aanlopen* [1–2]
18. **Ik weet zeker dat Jan daar aanlopen komt* [2–1]
19. *?Ik weet zeker dat Jan daar komt aangelopen* [1–2]
20. *??Ik weet zeker dat Jan daar aangelopen komt* [2–1]
    'I know for sure that Jan will come running there'

Third, in imperative sentences, only the infinitive variant is possible (Cornips 2002). If highlighting the end of a movement demands the past participle, as Honselaar (2010: 321) and Beliën (2016) claim, then we should expect the past participle to occur with the imperative, which clearly insists on the end and the goal of the motion; however, this is not the case.

21. **Kom hier langsgelopen*.PTCP
22. *Kom hier langslopen*.INF
    'Come this way!'

A fourth and final special syntactic property of the construction *komen* with a verb of motion is the presence of a separable prefix (see Section 6.3). Although empirical data shows that the prefix is rarely missing, it is considered optional (Haeseryn et al. 1997). This is especially true in comparison to other Germanic languages (Schäfer 2020:182), and it applies to both infinitive and past participle variants. Thus, the role of the prefix is irrelevant for our survey.

In sum, there are a few syntactic constraints, which motivate the use of the infinitive variant. While the influence of the IPP-effect is unclear, at least subordinating clauses and the imperative appear to lead to a preference for the infinitive variant. Having reviewed the existing literature, we would now like to turn our attention to (micro)typological, diachronic, and dialect variation that the construction shows.

9 Cornips (2002:5–6) describes the judgements in 16 and 17 as "erg verdeeld" 'very split' (our translation). She neither provides the original judgement data nor does she specify whom she asked, but states that she gave the sentences to ca. ten colleagues.

4 Typological, Diachronic, and Dialectological Evidence

In what follows, we will briefly discuss the phenomenon in closely related Germanic languages, which have similar constructions. We will then turn to the extant literature with regard to older stages of Dutch. Lastly, we will present new evidence from four sources of dialect data, showing that the spatial distribution of the variants in the dialects of Dutch has remained stable over the course of the last 100 years despite the fact that the datasets are heterogenous in terms of data collection methods. Further, with a dataset from 1977, we can show based on two tasks which differ semantically, that the predictions discussed above in Section 3.1, are not borne out in the data analysis. Finally, we can quantify the acceptability of the variants spatially, showing that there are indeed national and regional differences in the dialects as expected.

4.1 Evidence from Related Germanic Languages

In many Indo-European languages, the verb COME has become an auxiliary for future tense (Devos & van der Wal 2014; Fleischmann 1982:79). In the Germanic languages, this has happened in Swedish but not in Continental West Germanic languages. In addition to the grammaticalization to a future auxiliary, German varieties also use *kommen* as a copula for inchoatives and as a passive auxiliary (Nübling 2006; Schäfer 2020). Similarly, our Dutch construction may be a reflex of a grammaticalization of COME. All Continental West Germanic varieties have developed periphrastic constructions with COME and a movement verb in their recent language history. For example, High and Low German dialects allow formations with infinitive, past participle, and *zu/te*-infinitive to compete with each other, while Yiddish varieties have strictly grammaticalized the formation with *tsu*-infinitive, and Afrikaans kept the 'old' form with past participle (Schäfer 2020). In all cases, the (*to*)-infinitive is the innovative form.

Schäfer (2020) proposes that this variation was triggered by the grammaticalization of COME to a future auxiliary, as happened in Swedish (Hilpert 2008:54–69; 125–131). In the last 500 years, the Swedish future construction with COME reduced from *kommer till att* + INF 'come to to' to *kommer att* + INF 'come to' down to *kommer* + INF 'come' in present-day spoken Swedish (Hilpert 2008:127).

The grammaticalization of COME + motion verb has either come to completion in some varieties like Yiddish, or it has not. From the construction with movement verbs, the pattern would have to spread to other verbs, as has partly happened with stative verbs or acoustic

verbs in German varieties (Schäfer 2020). The Dutch situation is unique because speakers have a choice between two variants (Cornips 2002, Beliën 2016). Whether speakers of German or Afrikaans also vary, has yet to be shown.

In particular, given the fact that the variation that we find in the Dutch-speaking area is part of a larger continuum of dialects that extends across the German-Dutch state border (e.g. Auer & Hinskens 1996:16–17;), it may be interesting to examine more closely how the phenomenon behaves in neighboring dialect varieties of German. A cursory glimpse into extant sources suggests that German dialects might exhibit similar diatopic patterns.[10] For example, in Low German, we can find the infinitive variants with and without *to* and the past participle variants (23–25); in Central (26–27) and Upper German (28–29), we find the past participle and infinitive with *to*.

23. *wii bråchn hör ja Zåtədach ååmt weer un dau kuam S. dår anloup*n.INF (Reershemius 2004:144)
    'We brought her back Saturday night, and then S. came running'
24. *he küəmt to lôpen*.INF (Woeste 1882:148)
    'He comes walking'
25. *hei kömmt angelaupen*.PTCP (Westfälisches Wörterbuch:1046)
    'He comes walking (up to me)'
26. *he kom ze regge*.INF (Rheinisches Wörterbuch Vol. 4:1151)
    'He comes riding'
27. *e koum mat de Päerd gerannt*.PTCP (Luxemburgisches Wörterbuch Vol. 4, Col. 42a–44a)
    'He comes running with the horse'
28. *er kommt gfahre*.PTCP (Badisches Wörterbuch Bd. 3:213)
    'He comes riding'
29. *jetz kummen sie ze fahren*.INF (Elsässisches Wörterbuch Bd. 2, Sp. 888a)
    'Now they come riding'

10 This list is in no way exhaustive and does not include all the potentially occurring variants in the dialects; rather it simply serves to illustrate that there appears to be considerable variation in the regional varieties of German (Schäfer 2020:45–46 provides additional examples). Diatopic variation has also been observed for texts handed down from the Early New High German period. The verb COME with a past participle was widely used in this period; however, COME also occurs with an infinitive in texts primarily from West Central German, West Upper German, and Low German areas. In East Central German, COME with a past participle is more common (Schöndorf 1991:24). The fact that the variant with the past participle has prevailed in Standard German may thus be influenced by the East Central German dialects (see Lameli 2013:234).

4.2 Diachronic Evidence

Historically, the PTCP variant is older than the INF variant (Hirao 1965, Van der Horst 2008a:910). Weijnen (1971:110) gives some qualitative evidence for the occurrence of the construction *komen* 'come' + motion verb in Middle Dutch. According to Weijnen, it was also possible to use a present participle, a past participle, or an infinitive with the verb *komen* 'come' as in examples 30–32.

30. *doe cam ic gaende*.PPTCP *in een valeie* (Weijnen 1971:110)
    'then I came into a valley'
31. *doe sach tfolc dat eene duve quam gevlogen*.PTCP (Weijnen 1971:110)
    'then the people saw that a dove came flying'
32. *quam die moerdenaer lopen*.INF *totten monic* (Weijnen 1971:110)
    'the murderer came running to the monk'

Our investigation is focused on constructions with INF and PTCP since we suspect that they express aspectual expressive possibilities of movement. Formations with PPTCP, on the other hand, which are documented in all Germanic languages from an early stage on (cf. Schäfer 2020:150), are less relevant in this context since they serve to express the manner or the way of movement instead of the action of movement.

Following Van der Horst (2008a:910), there are already examples of *komen* 'come' with an INF in the 13th century; however, these combinations had a final interpretation. It was not until the 15th century that *komen* + INF also acquired the manner interpretation and began to become more frequent to the detriment of the PTCP variant. The frequency of the PTCP variant decreases, and the INF variant increases up through the 17th century (Van der Horst 2008a:1186). This development is spurred on by a syntactic change, in which *komen* 'come' is used more often with *te* 'to', and then still later with *om te* 'in order to' to express finality. The existence of two competing syntactic variants led to a semantic specialization of the PTCP variant, and thereby to an increase in the use of directional adverbials or prefixes, which, according to Van der Horst (2008a:910) became obligatory in the 18th century:

> Door de oppositie infinitief versus voltooid deelwoord kreeg die met volt. dw. (door de betekenis van het volt.dw. als zodanig) al gauw een perfectiever interpretatie dan die met infinitief. […] De perfectievere interpretaties van *komen* + volt.dw. leiden ertoe dat we die combinatie meer en meer gepaard zien gaan van een richtingsbepaling […] bij werkwoorden van beweging […] (Van der Horst 2008a:910).

> Because of the opposition infinitive versus past participle, the one with the past participle (because of the meaning of the past participle as such) soon got a more perfective interpretation than the one with the infinitive. […] The more perfective interpretations of *komen* + past participle result in us seeing that combination occurring more and more often together with a directional adverbial […] with motion verbs […] (Van der Horst 2008a:910, our translation).

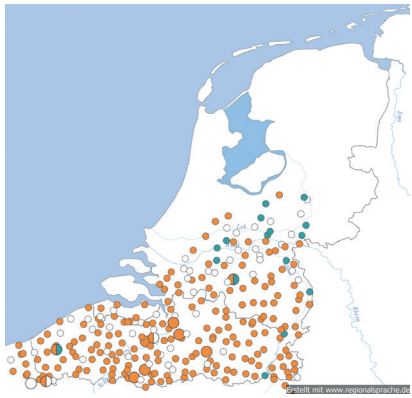
Dal (1954:493) claims that in Middle Dutch *komen* occurred with both infinitives and past participles synonymously, presenting two examples:

33. *Die quam ghevaren.*PTCP (Dal 1954:493)
    'that one came riding'
34. *Die goore, die daer ut quam varen.*INF (Dal 1954:493)
    'The smells that came from out of there'

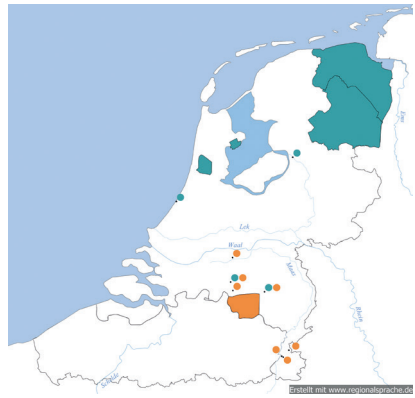
Vogel (2005) and Dal (1954) argue that the past participle variant has been losing ground in spoken Standard Dutch; however, Dal (1954:493) hedges by stating that the past participle only occurs 'in einzelnen speziellen Fällen' ('in certain special cases'), but she does not provide any examples. Similarly, Van der Horst (2008b:1783) claims that the PTCP variant has not completely disappeared, although it has decreased in frequency. While Beliën (2016), and the results of our own study for that matter (Section 5), shows that the past participle variant is still common in present-day Dutch, this discussion certainly highlights the need for a representative, diachronic corpus study to finally settle the matter as to the frequency of the constructional variants in the history of Dutch. To summarize, the infinitive variant has become more and more common throughout the history of Dutch, but it has not completely supplanted the past participle variant because the past participle variant allegedly acquired a more 'special' meaning, viz. a more perfective interpretation, reflecting the old meaning of the prefix <ge->.

4.3 Evidence from Dutch Dialects

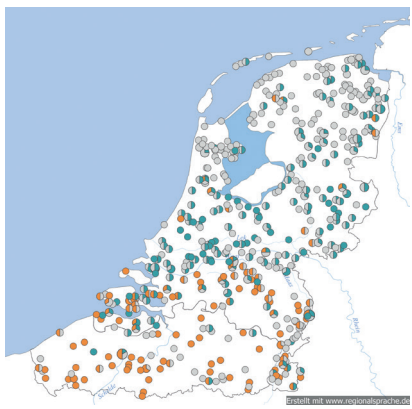
Let us turn to the differences in the spoken dialects. It has been previously claimed that on the level of the standard language, there are slight regional, country-specific preferences in the use of the past participle or the infinitive in constructions with COME and a motion verb (Haeseryn et al. 1997). In the south of the Dutch-speaking area (i.e. Belgium and the southern part of the Netherlands), the variant with past participle is preferred, while in the north of the Dutch-speaking area (i.e. the north of the Netherlands), particularly in the west of the Netherlands, there is a preference for the variant with the infinitive



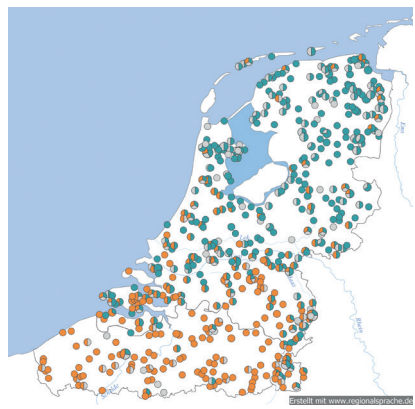
Map 1: Willems (1885)



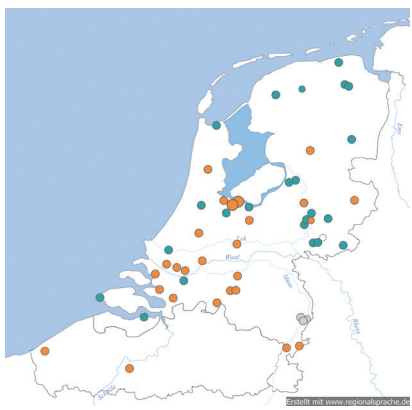
Map 2: Dialect grammars and dictionaries (1886–2011)



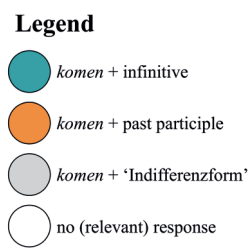
Map 3: Vragenlijst No. 52 (1977): Sentence 12/6



Map 4: Vragenlijst No. 52 (1977): Sentence 12/7



Map 5: SAND (2000–2003)



(Haeseryn et al. 1997).[11] However, to be sure, the variation is not solely regionally based. There are also idiolectal preferences, i.e. speaker-related variation, for the use of the participle or infinitive variant (Cornips 2002). The dialects of Dutch are claimed to show similar regional preferences in that the use of the infinitive variant increases and the use of the past participle decreases toward the north, and, on the other hand, the use of the past participle increases and the use of infinitive decreases toward the south; however, empirical evidence for the whole of the Dutch-speaking area is missing.

To examine the variation in the dialects, we consulted four sources of dialect data. These sources include competence data from dialect grammars and dictionaries, acceptability data from a written questionnaire distributed by the Meertens Instituut in the 1970s as well as translation data that was collected with a written questionnaire in 1885 by Pieter Willems and by the fieldworkers of the SAND project in the early 2000s. The data will be presented in chronological order.[12] The acceptability data and translation data are presented in the form of pie-chart maps since there is often more than one response for a given location. The reason for this is either that there are multiple responses from one respondent, or that there are multiple questionnaires for one location. The size of the chart is directly related to the number of responses from a location, i.e. the bigger the symbol, the more responses, and *vice versa*. Maps 1–5 show the results of the analyses. The results for each map will be discussed in turn in the following sub-sections. In what follows, when we make reference to a particular town's questionnaire, we will cite the Kloekecode, in addition to the town's name, in parentheses. Kloekecodes are a common tool used in the Dutch dialectological tradition that act as unique identifiers for towns (Kloeke & Grootaers 1926).

4.3.1 Enquête Willems (1885)

Pieter Willems distributed questionnaires in the 1880s with lexical, syntactic, and phonological questions for about 15,554 items (Goossens 1989:9). The questionnaire was distributed in the southern Netherlands, Belgium, France, Germany, and even in Luxemburg. The entirety of Belgium and

11 Overdiep (1937:§ 183) further noted a preference for the infinitive variant in the northern dialects of Dutch:De verbinding van *komen* met het partic[ipium] [...] blijkt nòch in Katw[ijk] nòch in Gron[ingen] en Dre[nthe] gebruikelijk; wel die met den infinitief (tracking in original). The combination of *komen* with a participle [...] does not appear to be common in neither Katwijk nor in Groningen and Drenthe; though the one with the infinitive (our translation).

12 The data that we collected will be available online via GitHub for interested readers. We created the maps using the mapping application REDE SprachGIS available at Regionalsprache.de.

Dutch-speaking France were covered, while only the southern half of the Netherlands was surveyed.[13] The German questionnaires are restricted to the West Central German dialect grouping, with Low Franconian dialects and Ripuarian dialects as well as their respective transition zones and two Mosel Franconian questionnaires (see the dialect classification after Wiesinger 1983). In the questionnaires, one of the tasks asks speakers to provide a translation of the following Standard Dutch sentence into their local dialect, containing the relevant COME + motion verb. In the task, the motion verb is presented in the form of the past participle.

35. *hij komt aangeloopen*.PTCP[14]
'he comes walking [up towards us]'

In total, we were able to analyze the responses from 344 questionnaires, taking into account not only Belgium (172), French Flanders (13), and the Netherlands (110), but also the neighboring German (59) and Luxemburgish (1) varieties. In 75 cases, there was no available data. Often the reason for this was either because the subject did not translate the sentence, or because the relevant pages were missing from the corpus for one reason or another. In four instances, the subject simply did not translate the verb of motion *angelopen* such as in the translation from Strijen (K124p) in South Holland: hij komt dăĕr ân 'he arrives'.

In total, there are 271 translations available for analysis. Subjects translated the sentence using the variant with the past participle in roughly 93.7% of the translations (n=254/271), while in approximately 5.9% of the translations that respondents used the variant with the infinitive (n=16/228). If we exclude the German and Luxemburgish data, which only show the past participle variant, there are 225 translations with the infinitive occurring in approximately 7.1% of the cases (n=16/225) and the past participle in approximately 92.4% of the cases (n=208/225). There was one instance of an infinitive occurring with the verbal prefix *te* 'to' in West Flanders. Map 1 shows the results for the Dutch-speaking area. Table 2 gives an example response and the absolute and relative frequency of each variant in the dataset.

13 For many dialects, especially in the south of the Dutch-speaking language area, the Willems questionnaire from 1885 is one of the oldest sources of dialect data, giving one of the first documentations of these dialects. The informants were all men (with one exception); they were typically younger, having been born in the 1850s and 1860s, and of the upper middle class (Goossens 1989:10–11).

14 The <oo> is the former orthographical convention for long vowels in open syllables. Today, it would be written <aangelopen>.

Table 2: Total number of variants with an example (Willems 1885)

Variant	Example	N/225	Percentage	N/271	Percentage
komen PTCP	hẹi kòmp āngẹ̄loôpẹ̄ (E192p)	208/225	92.4%	254/271	93.7%
komen INF	dāōr komtie āōnlōape (K099p)	16/225	7.1%	16/271	5.9%
komen te INF	hij komt te loope (H112p)	1/225	0.4%	1/271	0.4%

The high occurrence of the participle variant is no doubt also partially a result of the fact that subjects may have been more predisposed to using it because of influence from the template sentence. Seiler (2010: 522) has previously discussed that subjects tend to use the same construction offered in a stimulus if the construction is facultative in their own dialect, a methodological weakness of translation tasks. On the other hand, if subjects diverge from the template variant (in this case with the past participle), then this is arguably strong evidence for the occurrence of a variant in a particular region. As Map 1 shows, the two variants have a particularly coherent areal distribution. While the participle form occurs everywhere in the south of the Dutch-speaking area and in the neighboring West Central German area, the infinitive variant increases towards the north. This coherent areal distribution suggests that the participle variant may indeed be the prevailing variant in the southern dialects, otherwise, we might expect there to be the occasional infinitive variant interspersed throughout the participle variants.

4.3.2 Dialect Grammars and Dictionaries: 1886–2011

We consulted 56 dialect dictionaries, available online in the eWND (s. Van den Heuvel et al. 2015), as well as an additional 23 dialect grammars. There are a few methodological caveats when using dialect grammars and dialect dictionaries (see also Fischer 2019). First, the dictionaries and grammars span nearly a century ranging from publication in 1886 to the early 2000s meaning that the data were collected at varying points in time. Therefore, the resulting areal distribution may reflect varying synchronic states, which requires the data to be corroborated by other extant materials. Second, the dictionaries and grammars vary in their depiction of the linguistic facts because they may claim to be representative of one location, a collection of locations, or a larger region. Third, dialect grammars were traditionally written with an eye for the diachronic development of sounds from a proto-system. Generally, morphological and syntactic phenomena were not discussed in the grammars. This means that the presentation of any information on these phenomena is

largely a result of luck, a point which also holds for the dictionaries. In our analysis, we were able to uncover 51 attestations of the construction COME with a motion verb.[15] Map 2 shows the areal distribution of areas and towns where we found attestations of the infinitive and past participle variants. The map allows for little interpretation, but it confirms the north-south divide of this phenomenon in the Dutch dialects. Two examples of the attestations from the dictionaries are given in examples 36–37.

36. *hae kumt aangeloupe* (Limburg, Roebroek 1886)
    'he comes running'
37. *hai kwam tou 't stro oet kroepen* (Groningen, Ter Laan 1953)
    'he came crawling out from the straw'

4.3.3 Vragenlijst No. 52 (1977)

The Dialectological Department from the *Instituut voor Dialectologie, Volkskunde en Naamkunde* (forerunner to the Meertens Instituut in Amsterdam) distributed their 52nd survey in 1977. In total, 708 respondents took part in the survey from 597 different locations distributed throughout the Netherlands and Flanders (Belgium). The second half of the survey contained many tasks which were analyzed by Marinel Gerritsen in her *Atlas van de Nederlandse Dialectsyntaxis* (Gerritsen 1991).

This questionnaire includes three tasks that are relevant for the present study. In tasks 12/6 and 12/7, respondents are presented with the variants in sentence form, and are then instructed to indicate whether the variant occurs in the local dialect by circling either *ja* 'yes' or *nee* 'no'. Then, respondents should provide their best judgement as to which of the three variants they find to be the *meest gewoon* 'most normal' in the local dialect.[16]

15 The bibliographic information for the dialect grammars and dictionaries in which we were able to find information about the construction is given in the references.

16 79 speakers chose the INF as the most natural construction, while 83 speakers chose the PTCP variant and 307 speakers the PPTCP variant, respectively. Since the question remained unanswered in 252 cases, meaning that not all survey participants chose a most normal variant, we decided against providing a map of the answers to this question. In the event that speakers only indicated one variant as acceptable, we could assume that this variant is also the variant that they find "the most normal". However, in the event that speakers chose two variants, we can no longer make out after the fact whether speakers simply overlooked this question or if they found both variants to be equally "normal". Also, there is sometimes a difference between the possible variants and the "most normal" variants; for example, participants noted only A as possible, but marked B as the most normal. In addition, speakers had the opportunity to write down a variant if they did not find the presented response possibilities to be acceptable. Sometimes, this variant does not correspond to "the most normal" variant. For example, a speaker may have used a participle in their response but marked the infinitive variant as the most normal. What's more, participants sometimes chose multiple variants as "the most normal".

Table 3: Summary of the total number of variants (Vragenlijst No. 52)

Sentence (12/6)	**N/708**	**Percentage**
Jan komt niet met de bus, hij komt gelopen.PTCP	129/708	18.2%
Jan komt niet met de bus, hij komt lopend.PPTCP	412/708	58.2%
Jan komt niet met de bus, hij komt lopen.INF	198/708	28%
'Jan is not coming with the bus, he's coming walking'		
Sentence (12/7)	**N/708**	**Percentage**
De agent kwam de straat in gefietst.PTCP	292/708	41.2%
De agent kwam fietsend de straat in. PPTCP	206/708	29.1%
De agent kwam de straat in fietsen.INF	410/708	57.9%
'The police officer came cycling into the street'		

Finally, respondents should provide a translation of the variant that they marked as being the 'most normal'. The variants are shown below in Table 3 with their absolute and relative frequencies.[17]

We will concentrate solely on the responses to the questions about which of the variants are *acceptable* or *possible* in a dialect. While sentence 12/6 determines the manner of movement without a combination with a directional phrase or a particle, sentence 12/7 has a telic focus on the arrival in the street. Therefore, 12/7 gives a context with a manner of motion verb with a directional phrase. According to Beliën (2016), the past participle should be expected in 12/7, especially in Hollandic varieties. This prediction is not borne out in the data analysis. Since 12/7 has a telic focus, we should expect, with Beliën (2016), more responses with the past participle than with the infinitive; instead, we find more attestations of the infinitive in this

It is obvious that the question and the methodology shows weaknesses and pitfalls. The other reason for not considering this question was, of course, that we were interested in the potential *variation*, i.e. in all of the variants that speakers might use.

17 While the responses to these questions did not make their way into the *Atlas van de Nederlandse Dialectsyntaxis* (Gerritsen 1991), we did uncover five map drafts which give an overview of the phenomenon and are available over *De Kaartenbank*. Maps 22724, 22725, and 22727 deal with task 12/6 and maps 22726 and 22728 deal with task 12/7. While we could have made use of the maps in their form, we decided to remap the data for two reasons. First, we decided that we would be able to provide a more accurate presentation of the data if we examined the raw data once more from scratch. Second, we wanted to be able to make the data available for future analyses. It turns out that it was worthwhile to reexamine the raw data. We were able to uncover and transcribe participants' responses to a third question regarding semantic differences in the two constructions. We also found that the maps available over *De Kaartenbank* do not contain all of the raw data, thus, in addition to being more accurate, our maps are also more detailed than the original hand-drawn maps.

Table 4: Distribution of the variants in each country

Variant	Belgium (n=86)	Netherlands (n=653)	Variant	Belgium (n=143)	Netherlands (n=765)
gelopen.PTCP	57% (49)	12.3% (80)	*gefietst*.PTCP	76.2% (109)	23.9% (183)
lopend.PPTCP	36% (31)	58.3% (381)	*fietsend*.PPTCP	18.2% (26)	23.5% (180)
lopen.INF	7% (6)	29.4% (192)	*fietsen* .INF	5.6% (8)	52.5% (402)

context. If we compare only the occurrence of the INF and PTCP variants in both tasks, there is no significant difference between the two (χ^2=0.34, df=1, p=.56). Instead, the infinitive variant is the dominant variant in both tasks. The areal distribution of the responses to both tasks can be found in Map 3 and Map 4, respectively.

If we examine the occurrence of the variants by countries, we find that in both contexts, the INF variant is preferred over the PTCP variant in the Netherlands, while the PTCP variant is preferred over the INF in Belgium. If we include the PPTCP variant, the difference is significant for both contexts (12/6: χ^2=108.13, df=2, ***p<.0001; 12/7: χ^2=162.83, df=2, ***p<.0001). Similarly, if we only compare the acceptability of the INF to PTCP, the difference is also significant in both contexts (12/6: χ^2=65.74, df=1, ***p<.0001; 12/7: χ^2=151.14, df=1, ***p<.0001). If we examine the breakdown of the variants by province, we can observe that precisely those three Dutch provinces that directly border Belgium, i.e. North Brabant, Zeeland, and Limburg (NL), have the highest rates of the PTCP variant in both contexts; however, the rate is not as high as in Belgium itself. There is a dramatic drop in the acceptability of the PTCP variant between these three provinces and the other Dutch provinces. We can say, however, that with the exception of Utrecht in the context of task 12/6 and West Flanders in the context of task 12/7, both variants occur in each province, even though they occur with differing frequencies. Our use of administrative borders, i.e., provinces, to classify the dialect variants was simply a practical heuristic device. We are well aware that dialect borders and province borders do not necessarily coincide. We also quantified how many locations found more than one variant acceptable. For 12/6, of the 473 locations for which data is available, 94.5% (n=447) of locations have only one documented variant, while 5.5% of the locations have at least two documented variants. For task 12/7, of the 524 locations for which data is available, 78.2% (n=410) accepted only one variant, while 21.8% (n=114) accepted more than one variant. Thus, there is indeed intra-local variation to a degree, but for the most part, most locations have only one documented variant in each context.

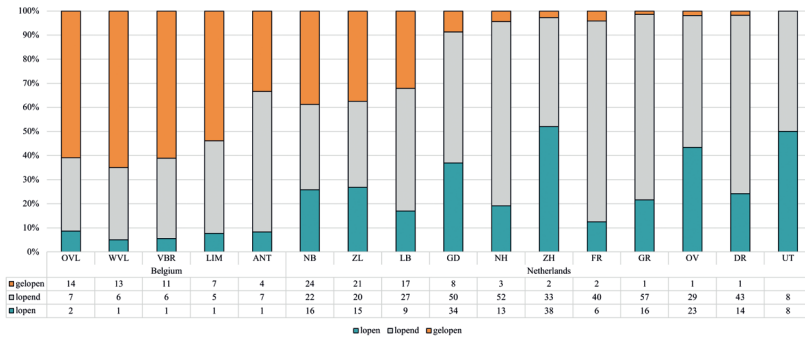


Figure 1: Distribution of the variants by country and province (12/6)[18]

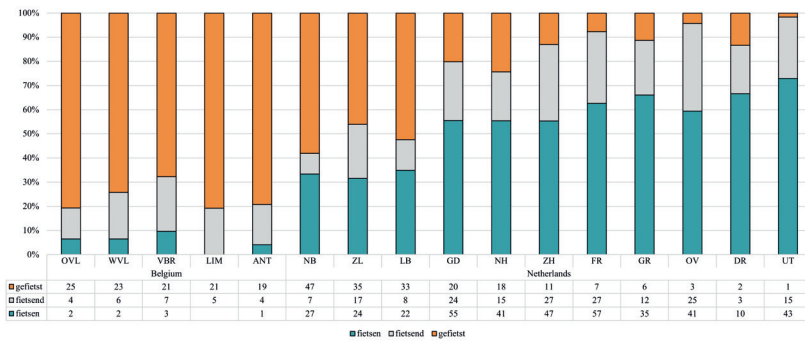


Figure 2: Distribution of the variants by country and province (12/7)

In a third task, respondents were asked to provide a short explanation as to what the difference in meaning is between the construction with an infinitive and past participle, in the event that both variants were deemed acceptable.[19] While the majority of respondents skipped over the question

18 For reasons of space, we made use abbreviations for the provinces in Figure 1 and Figure 2. For Belgium, they include the following abbreviations: OVL = East Flanders, WVL = West Flanders, VBR = Flemish Brabant, LIM = Limburg, and ANT = Antwerp. For the Netherlands, the abbreviations are NB = North Brabant, ZL = Zealand, LB = Limburg, GD = Gelderland, NH = North Holland, ZH = South Holland, FR = Friesland, GR = Groningen, OV = Overijssel, DR = Drenthe, and UT = Utrecht.

19 The question reads as follows in the questionnaire:

Als uw dialect zowel de constructie met 'lopen' of 'fietsen' als die met 'gelopen' of 'gefietst' kent, wat is dan het verschil in betekenis? Hieronder graag uw toelichting.

If your dialect has both the construction with 'lopen' or 'fietsen' as well as the one with 'gelopen' or 'gefietst', what is the difference in meaning then? Please write your explanation below (our translation).

for one reason or another, some respondents did provide short answers, which give some clues about the construction. To give an impression of the responses, we will present and briefly discuss a few of them. Table 5 contains an overview of some relevant metadata with a transliteration of the speakers' responses together with a translation into English. While speakers from Hellevoetsluis (I021p) and Waubach (Q117a) both point out that the infinitive variant seems to indicate that the action is still happening, and the past participle variant seems to indicate that an action is completed, another speaker from Waubach (Q117a) seems to contradict this point by remarking that the past participle indicates an ongoing action. The speaker from Rothem (Q099p) on the other hand points out that the past participle highlights the manner of the movement. There are also syntactic remarks. For example, a speaker from Berg (Q103p) notes that the infinitive is used when there is also an adverbial prepositional phrase. Finally, there are also speakers who note that there is no difference. For example, the speakers from Geleen (Q021p) and Nieuwaal (K116a) state that both variants do not differ in any way. Further, the speaker from Geleen (Q021p) hedges his response and indicates that the infinitive variant connotes that the action is more 'surprising', which may point to an ingressive meaning. From this discussion, we conclude that we find a partial confirmation of the hypothesis with regard to the occurrence of the past participle versus the infinitive form. However, the responses also contain some contradictory remarks as well as indications that there is no difference between the two variants, highlighting the necessity to perform a more systematic investigation, which we will discuss in Section 5. It is striking that speakers do not attribute a clearly defined uniform semantic function to the construction; instead, they remain generally vague.

4.3.4 DynaSAND (2000–2003)

Lastly, we consulted the DynaSAND corpus (Barbiers et al. 2006) for which data were collected during the field work for the *Syntactische Atlas van de Nederlandse Dialecten* (SAND) at the beginning of the 21st century (Barbiers et al. 2005, 2006, 2008, see also Barbiers 2005). The SAND contains data from 267 locations 'collected in oral and telephone interviews and in a postal survey' (Barbiers et al. 2007: 54).

Our analysis is based on test sentence no. 310, which was collected in the course of the postal survey. The postal survey was the pretest for the later large-scale project. The questionnaire 'was sent to the informants network of the Meertens Institute. These informants were not controlled for social variables' (Barbiers & Bennis 2007: 61) and also the geographical distribution

Table 5: Informant responses to differences in meaning

Location	Description
Hellevoetsluis (I021p South Holland)	*lopen* of *fietsen* is men nog bezig; *gelopen* of *gefietst* is gebeurd 'one is still busy with *lopen* or *fietsen*; *gelopen* or *gefietst* has happened'
Waubach (Q117a NL Limburg)	met 'ge' benadrukt men 't punt of tijdstip v. aankomst; met inf à benadrukt men de handeling 'with 'ge' one emphasizes the point or moment of arrival; with inf[initive] à one emphasizes the action'
Waubach (Q117a NL Limburg)	wordt er gezegd; *gelopen* of *gefietst* dan is de handeling nog aan de gang 'if it is said; *gelopen* or *gefietst* then the action is still happening'
Rothem (Q099p NL Limburg)	gĕfiets en geloupĕ wijst meer op de manier van voortbewegen 'gĕfiets and geloupĕ (=PTCP) rather highlights the manner of moving forward'
Berg (Q103p NL Limburg)	met'n voorzetsel – hier b.v. *in* – wordt de infinitief gebruikt 'with a preposition – here for example *in* – the infinitive is used'
Geleen (Q021p NL Limburg)	de eerste en de tweede betekenen 't zelfde, 't eerste is meer verrassend 'the first (=INF) and the second (=PTCP) mean the same thing; the first one is more surprising'
Nieuwaal (K116a Gelderland)	Beide spreekwijzen worden wel gebruikt. Een bepaald onderscheid is er niet 'Both ways of saying it are used. There is no particular difference'

was not controlled for. That is the reason as to why some places have more than one informant and other regions have none at all.

In this questionnaire, the subjects were asked to translate the sentence in 38 into their dialect. The stimulus forms in the written questionnaires were presented '(...) in Standard Dutch although some questions accommodate to a specific geographical distribution of the variant by inclusion of *specific local features*...' (Cornips & Jongenburger 2001: 56; our italics). Test sentence 310 is an example of the researchers incorporating a specific dialect variant into the otherwise Standard Dutch template sentence.

38. *zij kwamen aan te gewandelen*
    'they came walking'

In total, there are 106 translations of the sentence from 53 different locations, which are mostly in the Netherlands with the exception of two locations in

Table 6: Summary of the total number of variants with an example (DynaSAND test sentence 310)

Variant	Example	N/106	Percentage
kwam PTCP	*ze kwamme aongewandeld* (K039p)	55/106	51.9%
kwam INF	*zie kwam'n d'r an lopen* (of: *kuier'n*) (G052p)	42/106	39.6%
kwam te INF	*sie kwame an te wandele* (E004p)	3/106	2.8%
kwam om te INF[20]	*zi-j kwammen um te wandelen* (G278p)	1/106	0.9%
*kwam te ge-*INF	*zeej kwaeme aan te gelaupe* (L295p)	5/106	4.7%

Belgium. There are two translations from every location with the exception of four locations with only one translation, and two locations with four translations. There is not a single location in which speakers use more than one variant despite the fact that there are multiple questionnaires for some locations. The analysis of the translations revealed five different variants. Table 6 shows an example translation for each variant as well as their absolute and relative frequencies.

Map 5 shows the areal distribution of the variants based on the data from this translation task in the dialects of Dutch. While there are five variants in total, the infinitive and participle variants are both the relevant and prevailing constructions for this analysis. The variant *kwam te* INF has been included with the INF variants, while the *om te* INF variant is displayed in white. The exclusively Limburgish variant *kwam te ge-*INF[21] has been rendered with grey circles. The map shows three things. First, the infinitive variants are more prevalent in the dialects in the north, while the participle variants are more prevalent in the south (though to be sure there are only two data points in Belgium). Second, this observation is nothing more than a tendency. The map shows that infinitive variants occur as far south as Zeeland, and that the participle variant occurs as far north as Drenthe. Third, competing variants often occur in neighboring villages. This most often seems to occur in the center of the Dutch-speaking area, particularly around the great rivers.

20 This variant is only listed in the table since it occurs in the translations of this sentence. However, we believe that this variant is not an instantiation of the construction we are interested in, but rather has a final or purposive reading.

21 Such *ge-* prefixed infinitives are rare in this context, but they appear in several historical and modern West Germanic varieties (cf. Schallert 2014: 287–289). Dal (1954) speaks of them as "Indifferenzformen" (Forms of indifference).

4.3.5 Summary

In this section, we were able to show multiple things. First, we were able to show *empirically* that the morphological variants of the motion verb used in the *komen*-construction show a north-south distribution in the base dialects, i.e. the infinitive is more common in the north, while the past participle prevails in the south. Seemingly, there is a transition zone around the large rivers. Second, with the necessary caveat that these datasets are subject to limitations, they present a cohesive image, suggesting that this construction has remained stable in the dialects for the last 100 years. In light of the fact that regionally conditioned (morpho)syntactic variants can be expected in registers closer to the standard (Patocka 1993), this evidence would appear to support, but does not confirm, Haeseryn's (1997) claims that speakers of registers of Dutch closer to the standard language also show these preferences.

These facts confront us with a further question. If speakers of a given dialect have a choice between two morphological variants in a construction, is this choice due to a semantic distinction between the infinitive and the past participle, which has been suggested for the standard language? While the discussion of responses to the surveys from 1977 shed some light on this matter, it remains elusive for the time being. In light of practical issues involved in the acquisition of dialect speakers for a linguistic study, this question will have to be addressed in another follow-up study.

Furthermore, as noted, the dialect data does not provide evidence for higher registers. Since we have empirical evidence supporting our expectations regarding the geographical distribution of the morphological variants in the *komen*-construction as well as theoretical expectations as to which factors may motivate the choice beyond geography, we developed an online survey.

5 The Design of the Online Survey

To test the influence of semantic and geographic factors, we developed a short online survey using *SoSci Survey* (Leiner 2019). We asked test subjects to construct sentences with a set of words that have to be imbedded into a given context to get them to produce a relevant construction without allowing them to produce too many irrelevant variants. The task type has been successfully applied in other research projects (Fleischer, Kasper & Lenz 2012). To guarantee that no serialization effects occur in the answers, the order and number of elements that informants use for all tasks in our survey was varied. We created an online survey consisting of ten randomized puzzle-judgement tasks in order to avoid such serialization effects. In order

to be able to participate in the survey, participants had to be native speakers of Dutch (for more information on the method of our survey, see sections 5.2–5.5).

5.1 Register

Following Möller & Elspaß (2014), we tried to capture registers of speech that speakers use in everyday communication (=*Alltagssprache*), the totality of speech forms that are used 'in the social and functional ("nearness") domain of private life, of spontaneous conversation with friends, relatives or acquaintances or also in informal exchanges with strangers from the same town [...]'.[22] In our questionnaires, we asked speakers to respond to our questions in a form which they would use when speaking to a familiar person. In this way, we capture registers of speech which may potentially range from base dialect to standard Dutch. Since in many parts of the Dutch-speaking area, speakers no longer speak the traditional dialects (Auer 2005), depending on such factors as geography, socio-demographic factors like age and social class, as well as functional domains such as the speech setting (Willemyns 1997), we reckon with a high degree of responses which correspond to a register closer to the standard language.

5.2 Participants

In total, 121 native speakers of Dutch took part in the survey. The respondents ranged in age from 16 to 90 years old with the median age being 43. At least 74 women and 48 men took part in the survey (two respondents did not indicate their biological sex). 92 of the subjects indicated having no knowledge of linguistics, while 32 had at least an introductory course to linguistics. Eight of the respondents reported having completed only secondary education (=*middelbaar onderwijs*), 24 had completed a higher professional education (=*hoger beroepsonderwijs*), and 92 had completed a university education (=*universitair onderwijs*).

5.3 Tasks

At the beginning of each task, participants were reminded to imagine that they were speaking with a person whom they are well acquainted with. Each task consisted of two parts. The first part consisted of a sentence completion task. Survey participants were shown an image and then an incomplete

22 Originally: "[...] im sozialen und funktionalen (‚Nähe'-)Bereich des Privaten, des spontanen Gesprächs unter Freunden, Verwandten oder Bekannten oder auch im informellen Austausch unter nicht näher Bekannten aus demselben Ort [...]" (Elspaß & Möller 2014: 122).

sentence directly below it. They were asked to describe the image and to make use of the words in parentheses to complete the sentence. When the participants had finished completing the sentence using the available words, they could then proceed to the next page where they were asked about the movement they had just described in the sentence. They then had to give their subjective judgement ranging from whether the event will happen in the future or whether it is already complete. There were ten such tasks. In addition, we also included five tasks on the use of the future with *zullen* 'shall' and *gaan* 'go' as well as two tasks on ingressive constructions such as *Er komt een dronken student op de bank* (*?te*) *zitten* ('A drunken student comes to sit down on the bench').[23] We have provided an example of one of the tasks on the use of the *komen* 'come' + motion verb construction in Figure 3 showing the puzzle task (left) and the subjective judgement task (right).

5.4 Conditions

The ten test sentences were constructed using six different lexical verbs: *rijden* 'drive, ride', *vliegen* 'fly', *zwemmen* 'swim', *kruipen* 'crawl, creep', *fietsen* 'bike, cycle', and *rennen* 'run'. The verbs *kruipen* 'crawl, creep' and *rennen* 'run' were each used in two different test sentences, *vliegen* 'fly' in three. The sentences were constructed in such a way as to test the effects of motion event (= windowed, i.e. internal motion (manner) or gapped, i.e. external motion (path) after Talmy 1985, 1991, 2000) and lexical semantics, or *Aktionsart* (= egressive, ingressive, progressive) on the realization of the infinitive or past participle variants. Table 7 summarizes the set of conditions used in the survey experiment. The puzzle-task sentences can be found in Table 9 in the Appendix.

The variable *motion event* stems from Talmy (1985). He characterizes a motion event, which is a situation containing movement or the maintenance of a stationary location, as consisting of 'one object (the "Figure") moving or located with respect to another object (the referente-object or "Ground")' (Talmy 1985:61). In addition to the components 'Figure' and 'Ground', borrowed from Gestalt Psychology,[24] the motion event also consists

23 For more information of the relation between future and ingressive COME with our construction, see Schäfer (2020).

24 "The terms 'Figure' and 'Ground' are taken from Gestalt psychology but we give them a distinct semantic interpretation here: the Figure is a moving or conceptually movable object whose path or site is at issue; the Ground is a referente-frame [sic!], or a referente-point [sic!] stationary within a referente-frame [sic!], with respect to which the Figure's path or site is characterized" (Talmy 1985:61).

7. Beschrijf a.u.b. het beeld. Gelieve gebruik te maken van de woorden tussen haakjes om de zin er onder af te ronden.

Stel u voor dat u met een vertrouwde persoon spreekt.

Lina is een begaafde amateurfotograaf. Vandaag wil ze een nieuw macro-objectief uitproberen. Als ze het mos op een tak wil fotograferen, _____
[komen] [een vogel] [vliegen]

8. De beschreven beweging...

- ...zal gebeuren
- ...staat op het punt te beginnen
- ...is net begonnen
- ...is in volle gang
- ...zal binnenkort eindigen
- ...is voltooid
- geen van bovenstaande is van toepassing

Figure 3: Question no. 48 from the questionnaire

Table 7: Summary of conditions

ID	Lexical Verb	Meaning	Motion Event	*Aktionsart*
01	*rijden*	'drive, ride'	windowed	Egressive
06	*vliegen*	'fly'	windowed	Egressive
07	*zwemmen*	'swim'	gapped	Ingressive
09	*kruipen*	'crawl'	gapped	Ingressive
10	*fietsen*	'drive, ride'	gapped	Progressive
11	*rennen*	'run'	gapped	Ingressive
49	*vliegen*	'fly'	gapped	Ingressive
48	*vliegen*	'fly'	gapped	Ingressive
05	*rennen*	'run'	windowed	Ingressive
04	*kruipen*	'crawl'	windowed	Egressive

of the motion itself as well as the path that the 'Figure' has followed: 'The "Path" [...] is the course followed or site occupied by the Figure object with respect to the Ground object. "Motion" [...] refers to the presence *per se* in the event of motion or location (only these two motion states are structurally distinguished by language)' (Talmy 1985:61). In addition to these 'internal components', the motion event can also have 'external components': 'Manner' and 'Cause' (Talmy 1985:61). Since there are different uses of ingressive in the research literature, we define *ingressive* as emphasing the sudden occurrence of a state (Flämig 1965:5). On the other hand, *egressive* designates the end of a situation.

6 The Results of Our Online Survey

We will now present the results of our online questionnaire study. Figure 5 shows the distribution of the participants' responses according to the task number. The bars in the chart are ordered according to decreasing frequency of the infinitive variant and increasing frequency of the participle variant. We exclude the present participle variant in the following. Recall that according to Beliën (2016:30) 'the variant with the past participle highlights the end of a process, while the infinitive variant does not'. We want to check whether this claim holds for registers of spoken Dutch that we attempted to capture with our questionnaire (Section 5.1). If this claim holds, then we expect that speakers prefer the past participle variant for egressive contexts, i.e. test sentences 06, 10, and 04, and that speakers prefer the infinitive variant in non-egressive contexts. Figure 4 shows that our study cannot corroborate

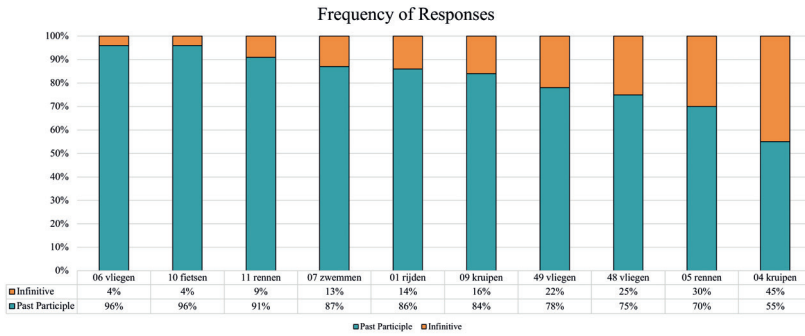


Figure 4: Distribution of participants' responses by task

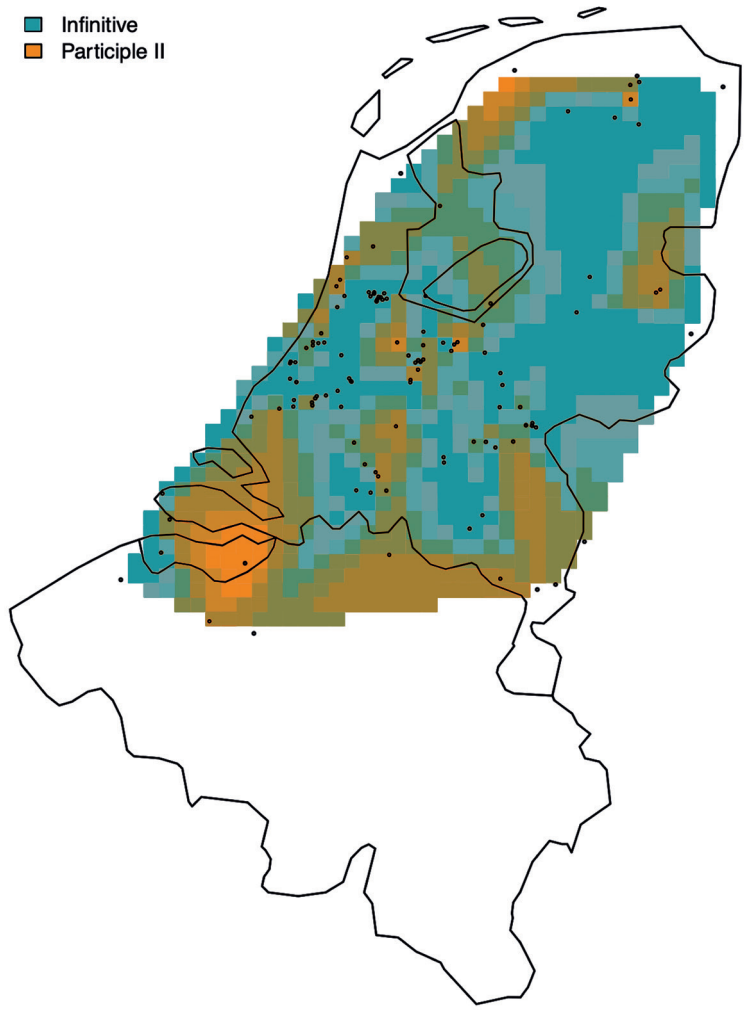
this claim.[25] In fact, in test sentences 06 and 10, the infinitive variant prevails. We could only find this effect of the past participle occurring with telic processes in test sentence 04; however, while we do find the past participle variant in this sentence the most often, the infinitive variant is still the most dominant even in this case. In fact, with regard to test sentence 04, this result is based on fewer responses than the others since we excluded uses of the present participle. Nearly half of the participants used the present participle to highlight this special, unusual kind of movement as in (39). In the next section, we will turn to the effect of geography to check whether geography plays a role in determining the choice of variants.

39. […] *dus kwam zij kruipend*.PPTCP *over de finish*
'[…] thus, she came crawling over the finish line'

6.1 Regional Impact

Map 6 shows the locations which survey participants hailed from as black dots. The results are presented in the form of a heat map. The results for the dialect analyses presented above would suggest that there is a north-south distinction with regard to the distribution of the two variants. We can see that in registers of Dutch situated more closely to the standard language, geography no longer plays such an important role in determining the choice of variants (also see the results of the regression model in Figure 5). Instead, we can see that the infinitive variant, which formerly primarily occurred in

25 Figure 4 does not include irrelevant responses, i.e. responses without the target construction, nor does it include the present participle variant. Table 10 in the Appendix contains the absolute frequency of each variant for each task in the questionnaire. By not including irrelevant responses and responses with a present participle, the absolute frequency of the variants for task 04 is comparatively lower than the others.



Map 6: The spatial distribution of the infinitive and past participle as a heat map (n=1018 responses from 121 questionnaires)

northern dialects, has established itself as the prevailing variant in higher varieties and speech levels in most parts of the language area. Nevertheless, in the far north there are still some remnants of the past participle.

6.2 PTCP vs. INF: Semantic Constraints

Thus far, the results of our study suggest that neither egressive semantics nor geography necessarily steers the apparent variation that we have been able to find. Moreover, the lexical verb itself does not play a role. In addition to

```
Call:
glm(formula = inf == "INF" ~ manner + ingressive + area +
verbs,
    family = binomial, data = vars)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7070   0.2695   0.4325   0.6272   1.6397

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      3.510964   0.814148   4.312 1.61e-05 ***
mannerTRUE      -1.522116   0.403426  -3.773 0.000161 ***
ingressiveTRUE  -3.528381   0.641926  -5.497 3.87e-08 ***
(and others, not significant)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 761.29  on 817  degrees of freedom
Residual deviance: 657.46  on 797  degrees of freedom
  (422 observations deleted due to missingness)
AIC: 699.46

Number of Fisher Scoring iterations: 14
```

Figure 5: R-code and result of the regression analysis

the independent variable AKTIONSART with the levels egressive, ingressive, and progressive semantics, we also tested for the independent variable MOTION EVENT (*manner TRUE*) with the levels gapped and windowed as well as the independent variables AREA and VERBS. For AKTIONSART we only tested here if ±INGRESSIVE applies (*ingressive TRUE*). To this end, we can use a binominal model. We fit a logistic regression using the gml R-function for a generalized linear model to our tested sentences (see Figure 5).[26] We were able to find two effects. First, windowed, or external, movement (Tschander 1999) disfavors the presence of the infinitive variant and, at the same time, favors the presence of the past participle variant. Second, ingressive semantics disfavors the use of an infinitive, and favors the use

26 We also made an attempt to estimate the performance of the logistic regression. There is no general agreement as to the metrics to be used for this, so we will provide several so-called pseudo- R^2 values: Cragg-Uhler (Nagelkerke) R^2=0.197, McFadden R^2=0.136, McFadden Adj R^2 =0.081, Efron/Lave R^2=0.131.

of a past participle, once more confirming the result discussed in Section 6, which did not conform to the expectations set out in Beliën (2016:30) with regard to written Standard Dutch. There were no effects for the AREA and VERB. Code and results of the regression analysis are given in Figure 5.

Furthermore, the question arises as to whether there is an "objective" way to determine the *Aktionsart* in these sentences. Therefore, in addition to the *objective* categorization we also collected participants' judgements on the semantics of each sentence. Directly after each task, participants were asked to indicate when the movement would take place, i.e. whether it will begin, is taking place, or has already come to completion. The results for this judgement test are given in Table 8. Ingressive semantics should disfavor the infinitive and favor the past participle. Test sentences 11, 07, 09, 48, 49, and 05 are semantically ingressive. However, participants' subjective judgements localizing the test sentences, have little to do with the "objective" semantics of the sentence. Participants judge test sentences 11, 09, 48, 49, and 05 to be more progressive in nature, and they judge test sentence 07 to be rather egressive. Test sentences 06, 01, and 04 are "objectively" egressive, yet, participants judged 06 as having something akin to ingressive and progressive semantics, and 01 to show rather progressive sentences. 04 was the only sentence that participants subjectively judged to show egressive semantics.

Interestingly enough, however, it turns out that there is a slight positive correlation between the use of past participles on the one hand (percentages in Figure 4) and participants' subjective judgements on the other hand (average values in Table 8): Pearson correlation coefficient 0.42 (=medium effect), R2 0.18, p=0.23.[27] This result could be interpreted to mean that if a speaker judges a verbal action as being egressive, then there is a concomitant increase in the use of the past participle, leading to a decrease in the use of the infinitive form of the verb. Thus, in subjective terms, the infinitive appears to be more associated with ingressive semantics, while the past

27 One potential rebuttal toward our use of this subjective evidence might be that the respondents simply either randomly evaluated the completeness of the event, or they themselves might not have properly understood the task. We would object that the respondents' responses are too consistent to allow for this conclusion. More problematic would be an objection to the inclusion of the pictures in the tasks. Is it possible that the respondents did not judge the sentence, but rather simply judged the picture. After comparing the results to questions 48 and 49, which have an image of bird(s) in a resting position and in motion, respectively, it turns out that speakers judged the infinitive and participle variants to be acceptable at roughly equal frequencies in both tasks (Q48:75% vs 25%, Q49:78% vs. 22%, respectively). At least here, the pictures did not lead to any differences.

Table 8: Participants' subjective judgements (1:max. ingressive – 6:max. egressive)

Test sentence ID	Average Value	Standard Deviation
06 *vliegen* 'fly'	2.67	1.41
10 *fietsen* 'bike, cycle'	4.1	0.91
11 *rennen* 'run'	3.61	0.78
07 *zwemmen* 'swim'	4.64	1.49
01 *rijden* 'drive, ride'	3.69	1.03
09 *kruipen* 'crawl, creep'	3.76	0.7
49 *vliegen* 'fly'	3.6	0.9
48 *vliegen* 'fly'	3.5	0.91
05 *rennen* 'run'	3.19	0.61
04 *kruipen* 'crawl, creep'	5.02	1.15

participle is more associated with egressive semantics. However, Pearson's R is only reliable for larger sample sizes (>30). If we check Kendall's coefficient, which is especially suited for smaller groups (<10), there is no effect at all (τ -0.045 2-sided p=0.92), meaning that there is no correlation between the past participle variant and egressive semantics. For this reason, we reject the possibility that the subjective data support the hypothesis that the past participle is associated with egressive semantics.

In light of this mismatch between the results of the regression analysis and participants' subjective judgements, we also reject this analysis in favor of a syntactic analysis, which we will further explore in Section 6.3.

6.3 INF VS. PTCP: Syntactic Constraints

A particular syntactic restriction on the Dutch *komen*-construction is that it requires the addition of a directional particle or phrase (Schäfer 2020, Beliën 2016: 20). The puzzle tasks deliberately omitted particles because we wanted to test how necessary they are in Dutch varieties (in German, by contrast, they are not necessary). However, the responses from speakers always contain an added particle, despite not being deliberately included in the stimuli in the puzzle task (compare the stimuli in Figure 3, for example). Therefore, the particles seem to be important for the Dutch construction (for an explanation of the mandatory particle in the Dutch structure see Schäfer 2020: 180–182).[28] If we examine the results of our experiment more closely, we are unable to find any evidence for *particle splits* when speakers

28 For more information on the function of particles in COME-constructions in the Germanic languages, see Schäfer (2020).

use a past participle; yet they seem to occur quite naturally when speakers use an infinitive. This is in line with the empirical research literature, which shows that in the event that a particle split occurs, then it tends to overwhelmingly occur with infinitives rather than past participles (40–41) (Haeseryn 1990: 87, Pauwels 1970: 99–100, cf. van Usen, Haeseryn & Fikkert 2013: 106, 111), despite a general tendency for speakers of Dutch to not split particle verbs (van Usen, Haeseryn & Fikkert 2013). The fact that particle splits tend to occur with infinitives rather than participles can be interpreted as an indication that *komen* 'come' may be on its way to grammaticalizing to a modal or auxiliary verb. This kind of grammaticalization channel from a full verb to a modal or auxiliary verb is typologically frequent (e.g. Lehmann 2015:39) and has also been observed with COME + motion verb constructions in West Germanic languages (Schäfer 2020).

40. ?/*[…] *aan komt gevlogen*.PTCP
41. […] *aan komt vliegen*.INF
    '[…] comes flying'

There are verbs other than *komen* 'come' for which particle splits can be interpreted as being a consequence of grammaticalization. We also tested the use of the future auxiliaries *zullen* 'shall.INF' and *gaan* 'go.INF' in our experiment, albeit for another purpose. The task results show that particle split occurs quite often with these two verbs. We therefore hypothesize that in analogy to the future auxiliaries, particle split is only possible when *komen* 'come' acts as an auxiliary verb, thus requiring the motion verb to occur as an infinitive (first status after Bech 1955). This would be in line with the alleged historical transition from the use of a past participle to an infinitive. We assume that this transition started in the dialects in the north and then spread to the south but has not yet spread to all of the dialects there. However, in higher registers of present-day Dutch, this historical transition from a past participle to an infinitive has almost reached an endpoint.

In essence, we claim that the transition from past participle to infinitive is a result of a change in the *komen* 'come' from a full verb to a modal or auxiliary verb, in which it has undergone a loss of semantic content. As such, it qualifies as a classic example of grammaticalization. We will explore these thoughts further in the next section.

6.4 Grammaticalization

Grammaticalization refers to the development of a linguistic sign from a lexical to a grammatical form or from a grammatical to an even more grammatical form. With recourse to Hopper (1991) who presents five

principles to identify grammaticalization in progress while it is still in its 'incipient stages', we rely on his Principle of Divergence that 'refers to the fact that when a lexical form undergoes grammaticization, for example to an auxiliary, clitic or affix, the original form may remain an autonomous lexical element and undergo the same changes as any other lexical items' (Hopper 1991:24). Divergence can be understood as a special type of Layering. While the latter involves different degrees of grammaticalization of different lexical forms in a similar functional domain, the former involves different degrees of grammaticization of the same lexical item (Hopper 1991:24). Thus, in the case of the Dutch verb *komen* 'come', we would be dealing with one lexical item that has become increasingly grammaticized in one context, but not in another context. In this case, the grammaticalization of *komen* 'come' specifically involves its use in conjunction with a lexical verb.

Evidence from Germanic languages shows that COME can take on a more grammatical function. For example, in varieties of Dutch, it has been observed that *komen* 'come' is undergoing semantic loss in the sense that it can be used as a copula verb in some varieties of Dutch (Heinsius 1928). In Germanic languages, more generally, we know that COME has come to be used as a future auxiliary in Swedish and as a passive auxiliary in Bavarian and Alemannic dialects (Nübling 2006; Schäfer 2020). These facts lead us to the hypothesis, initially introduced by Schäfer (2020), that the construction with *komen* 'come' in Dutch is the result of a grammaticalization process, in which *komen* is – or better was – on its way toward becoming a future auxiliary similar to *gaan* 'go'. As discussed in Section 6.3, the verb *komen* 'come' seems to behave like an auxiliary in the Dutch varieties investigated, meaning it has similarly lost semantic content. This can be interpreted as a preliminary step to a possible grammaticalization towards an auxiliary for (future) tense.

The data from the 19th and 20th centuries allowed for glimpses into a current process of language change, showing that there is a leveling process (=*Ausgleichsprozess*) going on in varieties of Dutch, which may have two possible stages. First, the morpho-syntactic variation that appears in Dutch dialects can be interpreted as 'historical junk' (Lass 1990) from a grammaticalization process to a future auxiliary. Second, the construction with *komen* 'come' might potentially undergo in some dialects further functional extensions towards special expressions for aspect, mode, or even tense such as for progressive expressions, as Beliën (2016) suggests. At the moment, however, lacking any evidence, we can only speculate about such a functional extension.

7 Summary

The historical development of the construction discussed here as well as the areal variation that we find in the Dutch dialects of the 19th and 20th centuries show that the formation with an infinitive is the innovative form. This suggests a tendency towards a grammaticalization of *komen* 'come' towards an auxiliary; perhaps towards an auxiliary for future tense in analogy to *gaan* 'go'.

Extant studies suggest that a semantic differentiation is responsible for the synchronic variation between the infinitive and past participle variants. The past participle is suspected of denoting telicity. This explanation implies a specialization of the 'old' variant, viz. the past participle, but it does not explain the spread of the innovative variant, viz. the infinitive. More importantly, our survey results do not provide evidence that the choice of a morphological variant is semantically determined (Section 6). Nor did the analysis of the dialect data show this result (specifically Section 4.3.3). The results of our regression analysis showed that windowed (i.e. internal movement) and ingressive semantics both disfavor the use of an infinitive, and favor the use of a past participle. This allows for the hypothesis that the infinitive expresses a more progressive and egressive (i.e. more *future-ish*) interpretation than the past participle. This in turn fits with syntactic considerations that *komen* 'come' in our construction has taken a grammaticalization step towards becoming a future auxiliary: the infinitive is diachronically the more innovative form and synchronically the form with a stronger tendency to express tense.

What else is responsible for the variation if semantics is not the only influencing factor? Of course, geography plays a role (speakers from the south in varieties closer and further from Standard Dutch chose the past participle variant more often than speakers from the north). However, it is not the geographic variation that needs to be explained, but rather the speaker-specific variation. One possible explanation would be assuming an incomplete or stagnated grammaticalization of COME as an auxiliary verb, as Schäfer (2020) suggests for West Germanic varieties. The morpho-syntax tries to balance this grammatical *grey area* in structures with IPP, imperative, and a directional adverb by increasingly favoring the infinitive. But there are still contexts that elude morpho-syntactic conditioning. At this point, we cannot yet say what determines the choice for the infinitive or the participle; nonetheless, we do have reason to suspect that the construction of the type *komen* 'come' + motion verb is or was involved in a

process of grammaticalization. These observations would make a diachronic investigation of periphrastic expressions with *komen* 'come' in the history of Dutch a worthwhile undertaking.

References

Research Literature

Auer, Peter (2005): 'Europe's Sociolinguistic Unity, or: A Typology of European Dialect/Standard Constellations'. In: Nicole Delbecque, Johan van der Auwera, & Dirk Geeraerts (eds.): *Perspectives on Variation*. Berlin: Mouton de Gruyter, 7–42. (Trends in Linguistics. 163).

Auer, Peter & Frans Hinskens (1996): 'The Convergence and Divergence of Dialects in Europe. New and not so New Developments in an Old Area'. In: Sociolinguistica, 10, 1–30.

Barbiers, Sjef (2005): 'De Syntactische Atlas van de Nederlandse Dialecten'. In: Taal & Tongval, 18, 17–40.

Barbiers, Sjef, Hans Bennis, Gunther de Vogelaer, Magda Devos, & Margreet van der Ham (2005): *Syntactische Atlas van de Nederlandse Dialecten. Deel I.* Amsterdam: Amsterdam University Press.

Barbiers, Sjef et al. (2006): *Dynamische Syntactische Atlas van de Nederlandse Dialecten* (DynaSAND). Amsterdam: Meertens Instituut. URL: http://www.meertens.knaw.nl/sand, accessed on April 11, 2019.

Barbiers, Sjef, Johan van der Auwera, Hans Bennis, Eefje Boef, Gunther de Vogelaer, & Margreet van der Ham (2008): *Syntactische Atlas van de Nederlandse Dialecten. Deel II.* Amsterdam: Amsterdam University Press.

Barbiers, Sjef, Leonie Cornips, & Jan Pieter Kunst (2007). The Syntactic Atlas of the Dutch Dialects (SAND): A corpus of elicited speech and text as an online dynamic atlas. In: Beal J.C., Corrigan K.P., Moisl H.L. (eds): Creating and Digitizing Language Corpora. Palgrave Macmillan, London, 54–90.

Barbiers, Sjef & Hans Bennis (2007): The Syntactic Atlas of the Dutch Dialects: A discussion of choices in the SAND project. Nordlyd 34, 53–72.

Bech, Gunnar (1955): *Studien über das deutsche verbum infinitum.* Volume I. Køpbenhavn: Ejnar Munksgaard.

Beliën, Maaike (2016): 'Exploring Semantic Differences in Syntactic Variation: Dutch *komen* 'come' with a Past Participle or an Infinitive'. In: Anne Bannink, & Wim Honselaar (eds.): *From Variation to Iconicity: Festschrift for Olga Fischer on the Occasion of her 65th Birthday.* Amsterdam: Pegasus, 17–32.

Beliën, Maaike (2017): 'Auxiliary choice with particle verbs of motion in Dutch'. In. Acta Linguistica Hafniensia 49(2), 212–231.

Bybee, Joan (2009): 'Grammaticization. Implications for a Theory of Language'. In: Jiansheng Guo, Eelena Lieven, Nanchy Budwig, Susan Ervin-Tripp, Keiko Nakamura, & Seyda Özçaliskan (eds.): *Crosslinguistic Approaches to the Psychology of Language. Research in the Tradition of Isaac Slobin.* Hove: Psychology Press, 345–355.

Cornips, Leonie (2002): 'Een vreemde eend in het rijtje. Over het aspectueel hulpwerkwoord 'komen''. In: *In verband met jan Luif: 29 Variaties op een thema door vrienden en collega's bij het afscheid van Jan Luif.* URL: http://cf.hum.uva.nl/poldernederlands/backup_luif/luif/cornips.html, accessed on January 7, 2020.

Cornips, Leonie & Willy Jongenburger (2001): 'Elicitation Techniques in a Dutch Syntactic Dialect Atlas Project'. In: Ton van der Wouden & Hans Broekhuis (eds.): *Linguistics in the Netherlands 2001.* Amsterdam: John Benjamins Publishing Company, 53–63. (AVT Publications. 18).

Corpus Dialectmateriaal Pieter Willems = Corpus Dialectmateriaal Pieter Willems. Centrum voor Teksteditie en Bronnenstudie. URL: http://bouwstoffen.kantl.be/CPWNL/#search, accessed on January 7, 2020.

Dal, Ingrid (1954): 'Indifferenzformen in der deutschen Syntax. Betrachtungen zur Fügung *ich kam gegangen*'. In: Norsk Tidsskrift for Sprogvidenskap, 17, 489–497.

De Kaartenbank = Kruijsen, Joep, & Nicoline van der Sijs (eds.) (2016): Meertens Kaartenbank. URL: www.meertens.knaw.nl/kaartenbank/, accessed on May 11, 2020.

Devos, Maud, & Jenneke van der Wal (eds.) (2014): *COME and GO off the beaten grammaticalization path*. Berlin: De Gruyter.

Ebeling, Carl Lodewijk (2006): *Semiotaxis. Over theoretische en Nederlandse syntaxis*. Amsterdam: Amsterdam University Press.

eWND = Nicoline van der Sijs (samensteller) (2015-): eWND. URL: ewnd.ivdnt.org. Gehost door het Instituut voor de Nederlandse Taal.

Fischer, Hanna (2019): 'Dialektgrammatiken als Datenquelle? Zum Wert der dialektgrammatischen Literatur für dialektvergleichende Studien'. In: Sebastian Kürschner, Mechthild Habermann, & Peter O. Müller (eds.): *Methodik moderner Dialektforschung: Erhebung, Aufbereitung in und Auswertung von Daten am Beispiel des Oberdeutschen*. Hildesheim, Zürich & New York: Olms, 309–333. (Germanistische Linguistik. 241–243).

Fläming, Walter (1965): 'Zur Funktion des Verbs III. Aktionsart und Aktionalität'. In: Deutsch als Fremdsprache Bd. 2.2, 4–12.

Fleischer, Jürg, Simon Kasper, & Alexandra N. Lenz (2012): 'Die Erhebung syntaktischer Phänomene durch die indirekte Methode: Ergebnisse und Erfahrungen

aus dem Forschungsprojekt *Syntax hessischer Dialekte* (SyHD)'. In: Zeitschrift für Dialektologie und Linguistik, 79(1), 2–42.

Fleischmann, Suzanne (1982): *The Future in Thought and Language. Diachronic Evidence from Romance*. Cambridge: University Press. (Cambridge Studies in Linguistics. 36).

Gerritsen, Marinel (1991): *Atlas van de Nederlandse Dialectsyntacis* (*AND*). Amsterdam: P.J. Meertens Instituut.

Goossens, Jan (1989): 'Pieter Willems en zijn dialectenquête'. In: Taal & Tongval. Honderd Jaar Enquête Willems, 5–15.

Grootaers, Ludovic, & Gesinnius G. Kloeke (1926): *Handleiding bij het Noord- en Zuid-Nederlandsch Dialectonderzoek*. 's-Gravenhage: Martinus Nijhoff.

Haeseryn, Walter (1990): Syntactische normen in het Nederlands. Een empirisch onderzoek naar volgordevariatie in de werkwoordelijk eindgroep. Nijmegen: Katholieke Universiteit Nijmegen.

Haeseryn, Walter et al. (ed.) (1997): *Algemene Nederlandse Spraakkunst*. Groningen: Martinus Nijhoff.

Heine, Bernd, & Tania Kuteva (2004): *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Heinsius, Jacobus (1928): Een eigenaardig gebruik van het ww. komen. In: Tijdschrift voor Nederlandse Taal- en Letterkunde, 47, 26–32.

Hilpert, Martin (2008): *Germanic Future Constructions: A Usage-based Approach to Language Change.* Amsterdam & Philadelphia: John Benjamins.

Hirao, Kuzo (1965): 'Fügungen des Typs 'kam gefahren' im Deutschen'. In: Beiträge zur Geschichte der deutschen Sprache und Literatur, 87, 204–226.

Honselaar, Wim (2010): 'The Infinitive or the Past Participle with the Verb *komen* in Dutch: Variation or Opposition?' In: Jolanta Chojak, Tomasz Korpysz, & Krystyna Waszakowa (eds.): Człowiek. Słowo. Świat. Warsaw: Wydawnictwa Uniwersitetu Warszawskiego, 316–325.

Hopper, Paul (1991): 'On some Principles of Grammaticalization.' In: Elizabeth Closs Traugott & Bernd Heine (eds.): *Approaches to Grammaticalization*. Volume I. Focus on Theoretical and Methodological Issues. Amsterdam & Philadelphia: John Benjamins Publishing Company, 17–35. (*Typological Studies in Language* 19.1).

Lameli, Alfred (2013): Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexi-tät der Dialekte in Deutschland. Berlin / New York: De Gruyter. (Linguistik – Impulse & Ten-denzen. 54).

Lass, Roger (1990): 'How to Do Things with Junk: Exaptation in Language Evolution'. In: Linguistics, 26, 79–102.

Leiner, D. J. (2019). SoSci Survey (Version 2.5.00-i1142) [Computer software]. Available at http://www.soscisurvey.com, accessed on January 7, 2020.

Lehmann, Christian (2015 [1995]): *Thoughts on grammaticalization.* 3rd edition. Berlin: Language Science Press (Classics in Linguistics, 1).

Möller, Robert & Stephan Elspaß (2014): 'Zur Erhebung und kartographischen Darstellung von Daten zur deutschen Alltagssprache online: Möglichkeiten und Grenzen'. In: Fabio Tosques (ed.): *20 Jahre digitale Sprachgeographie.* Berlin: Humboldt-Universität. Institut für Romanistik, 121–131.

Nübling, Damaris (2006): 'Auf Umwegen zum Passivauxiliar – Die Grammatikalisierungspfade von GEBEN, WERDEN, KOMMEN und BLEIBEN im Luxemburgischen, Deutschen und Schwedischen'. In: Claudine Moulin, & Damaris Nübling (eds.): *Perspektiven einer linguistischen Luxemburgistik. Studien zu Synchronie und Diachronie.* Heidelberg: Winter, 171–202.

Overdiep, G.S. (1937): *Stilistische grammatica van het moderne Nederlandsch.* Twolle: W.E.J. Tjeenk-Willink.

Patocka, Franz (1993): 'Zu Problemen der Erhebung und Kartierung syntaktischer Erscheinun-gen'. In: Wolfang Viereck (ed.): *Historische Dialektologie und Sprachwandel. Sprachatlanten und Wörterbücher.* Band 2. Verhandlungen des Internationalen Dialektologenkongresses. Bamberg 29.7–4.8.1990. Stuttgart: Franz Steiner Verlag, 400–417. (Zeitschrift für Dialektologie und Linguistik. 75).

Pauwels, J.L. (1970): Statistisch onderzoek van de Nederlandse zinsbouw. In: De Nieuwe Taalgids, 63, 93–100.

REDE SprachGIS = Jürgen Erich Schmidt, Joachim Herrgen, & Roland Kehrein (eds.) (2008–): Regionalsprache.de (REDE). Forschungsplattform zu den modernen Regionalsprachen des Deutschen. Teil 6: REDE SprachGIS – Das forschungszentrierte sprachgeographische Informationssystem von Regionalsprache.de. Edited by Dennis Bock, Robert Engsterhold and Slawomir Messner. In collaboration with Hanna Fischer, Brigitte Ganswindt, Simon Kasper, Juliane Limper, Mark Pennay, Jeffrey Pheiff, Tillmann Pistor, Christoph Purschke, Philipp Spang and Anna Wolańska. Marburg: Forschungszentrum Deutscher Sprachatlas. Available at https://regionalsprache.de/SprachGIS/Map.aspx, accessed on May 11, 2020.

Schäfer, Lea (2020): '*kommen* und Bewegungsverb in westgermanischen Varietäten'. In: Julia Hertel, & Augustin Speyer (eds)*: Syntax aus Saarbrücker Sicht 3. Beiträge der SaRDiS-Tagung zur Dialektsyntax*. Stuttgart: Steiner, 145–190. (Zeitschrift für Dialektologie und Linguistik Beihefte.).

Schallert, Oliver (2014): Zur Syntax der Ersatzinfinitivkonstruktion: Typologie und Variation. Tübingen: Stauffenburg.

Schöndorf, Kurt Erich (1991): '*kommen* mit Infinitiv, Partizip Präsens oder Partizip Präteritum. Zu Varianten einer Prädikatsfügung im Frühneuhochdeutschen'. In: John Ole Askedal, Gisela Brandt, & Kurt Erich Schöndorf (eds.): Osloer und

Rostocker Studien zur deutschen Sprach- und Literaturwissenschaft. Oslo, 11–26. (Osloer Beiträge zur Germanistik. 13).

Seiler, Guido (2010): 'Investigating Language in Space: Questionnaire and Interview'. In: Peter Auer, & Jürgen Erich Schmidt (eds.): *Language and Space: An International Handbook of Linguistic Variation*. Volume I: Theories and Methods. Berlin & New York: De Gruyter, 512–527. (Handbooks of Linguistics and Communication Science. 30.1).

Talmy, Leonard (1985): 'Lexicalization Patterns: Semantic Structure in Lexical Forms'. In: T. Shopen (ed.): *Language Typology and Syntactic Description.* Volume 3: Grammatical Categories and the Lexicon. Cambridge: Cambridge University Press, 57–149.

Talmy, Leonard (1991): 'Path to Realization: A Typology of Event Conflation'. In: *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Grammar of Event Structure*, 480–519.

Talmy, Leonard (2000): *Toward a Cognitive Semantics.* Volume 1, Concept Structuring Systems. Cambridge: MIT Press.

Tschander, Ladina (1999): 'Bewegung und Bewegungsverben'. In: Proceedings der *KogWiss*, 99, 25–30.

Van den Heuvel, Henk, Eric Sanders, & Nicoline van der Sijs (2015): 'Curation of Dutch Regional Dictionaries'. In: Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.): *Proceedings of the Tenth International Conference on Language Resources and Evaluation.* Paris: European Language Resources Association (ELRA), 3249–3255.

Van der Horst, Joop (2008a): *Geschiedenis van de Nederlandse Syntaxis*. Volume 1. Leuven: Universitaire Pers Leuven.

Van der Horst, Joop (2008b): *Geschiedenis van de Nederlandse Syntaxis*. Volume 2. Leuven: Universitaire Pers Leuven.

Van Usen, Karin, Walter Haeseryn, & Paula Fikkert (2013): 'Weet jij wanneer je *op* gaat *splitsen*? Onderzoek naar de voorkeur voor het bijeenhouden of splitsen van partikelwerkwoorden in de bijzin'. In: Taal & Tongval, 64, 103–127.

Vogel, Petra (2005): 'Neue Überlegungen zu den Fügungen des Typus *sie kamen gelaufen*'. In: Zeitschrift für Germanistische Linguistik, 33(1), 57–77.

Weijnen, A.A. (1966): *Nederlandse Dialectkunde.* Tweede Druk. Assen: Van Gorcum & Comp.

Weijnen, A.A. (1971): *Schets van de Geschiedenis van de Nederlandse Syntaxis*. Assen: Van Gorcum & Comp.

Weiß, Helmut (2009): 'How to Define Dialect and Language – A Proposal for Further Discussion'. In: Linguistische Berichte, 219, 251–270.

Wiesinger, Peter (1983): 'Die Einteilung der deutschen Dialekte'. In: Werner Besch, Ulrich Knoop, Wolfgang Putschke, & Herbert Ernst Wiegang (eds.): *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. 2. Halbband. Berlin & New York, 807–900.

Willemyns, Roland (1997): 'Dialektverlust im niederländischen Sprachraum'. In: Zeitschrift für Dialektologie und Linguistik, 64(2), 129–154.

Dialect Grammars and Dialect Dictionaries

Boekenoogen, G.J. (1897 [1971]): *De Zaanse Volkstaal*. Deel II: Zaans Idioticon-Aanvullingen. Reprint. Zaandijk: Drukkerij J. Heijnis.

De Bont, A.P. (1962): *Dialekt van Kempenland. Meer in het bijzonder d'Oerse taol*. Deel 1. Klank- en Vormleer en enige Syntaktische Bijzonderheden. Assen: Van Gorcum & Comp.

De Gast, Cees (2011): *'t boekske van de Aolburgse taol.* Wijk & Aalburg: Stichting behoud Aalburgs dialect.

De Laat, Gerard (1992): *Zoo prôte wèij in Nuejne mi mekaâr.* Nuenen: Heemkundekring De Drijehornick.

Fien, A., J.B. Keuter, A. van der Line-van der Weerd, P. van Mierlo & K. van 't Oever (2000): *Woordenboek van de Kamper Taal*. Kampen: Stichting IJsselacademie.

Ganderheyden, A.A. (1897): *Groningana: Supplement op H. Molema's Woordenboek der Groningsche Volkstaal.* Winsum: J.C. Mekel.

Karsten, G. (1934): *Het Dialect van Drechterland II*. Purmerend: J. Muusses.

Molema, H. (1889): *Proeve van een woordenboek der Drentsche volkstaal in de 19e eeuw.* Groningen: Manuscript.

Molema, H. (1895): *Woordenboek der Groningsche Volkstaal in de 19e eeuw.* Groningen: Manuscript.

Naaijkens, J. (1992): *Dè's Biks – Verklarende Dialectwoordenlijst*. Hilvarenbeek: De Hilverbode.

Overdiep, G.S. (1940): *De Volkstaal van Katwijk aan Zee.* Antwerpen: Standaard-Boekhandel.

Roebroek, J.L.H. (1886): *Dialect van Beek-Elsloo*. Beek: Manuscript.

Schelberg, P.J.G (1986): *Woordenboek van het Sittards dialect*. Amsterdam: Rodopi.

Sterenborg, Will, & Ed Schilders (2010–2011) *Woordenboek van de Tilburgse Taal* (WTT). Tilburg: Stichting Cultureel Brabant.

Ter Laan, K. (1953): *Proeve van een Groninger spraakkunst*. Winschoten: J.D. van der Veen.

Appendix

Table 9: Puzzle tasks with semantic conditions

ID	Puzzle-task sentence	Motion event	*Aktionsart*
01	Eindelijk, met een vertraging van 40 minuten, … [rijden] [het station] [komen] [de metro]	windowed	Egressive
04	Stefanie Goodman kon de laatste paar meter van de St. Louis Marathon niet lopen, dus … [kruipen] [komen] [zij]	windowed	Egressive
05	Eindelijk opent het starthek en … [komen] [de paarden] [rennen]	windowed	Ingressive
06	Je weet dat de zomer eindelijk voorbij is als … [het park] [de eerste trekvogels] [vliegen]	windowed	Egressive
07	Uw buurman vertelt u over zijn laatste vakantie, een cruisevaart. Hij was al heel lang niet meer op zee en had er een beetje schrik voor. Maar hij is de angst snel vergeten, want toen het schip van wal stak, … [dolfijnen] [zwemmen] [komen]	gapped	Ingressive
09	Om de slakkenplaag het hoofd te bieden, heeft u vallen in de tuin gezet, maar de pest neemt niet af. Terwijl u een dag in de tuin zit te lezen, kijkt u toe hoe de slakken … [kruipen] [de potjes] [komen]	gapped	Ingressive
10	Anne is naar Amsterdam gereden om te winkelen. Ze wil net de straat oversteken, als haar buurmeisjes … [komen] [voorbij] [fietsen] [zonder haar te zien]	gapped	Progressive
11	U bent een wandeling gaan doen. U verstijft van schrik als er … [een hond] [rennen] [komen]	gapped	Ingressive
49	Lina is een begaafde amateurfotograaf. Vandaag wil ze een nieuw macro-objectief uitproberen. Als ze het mos op een tak wil fotograferen, … [komen] [een vogel] [vliegen]	gapped	Ingressive
48	Nog terwijl heer Pieters vogelvoer in het vogelhuisje strooit … [vliegen] [komen] [op de omliggende takken] [de eerste vogels]	gapped	Ingressive

Table 10: Absolute frequency of each variant for each task

Variant / Task	06	10	11	07	01	09	49	48	05	04
Infinitive	105	110	107	96	94	88	91	78	66	21
Past Participle	4	4	11	14	15	17	26	26	28	17
Present Participle	0	0	0	0	0	4	0	0	1	61
Total	109	114	118	110	109	109	117	104	95	99