



Universiteit
Leiden
The Netherlands

The discovery of aspect: a heuristic parallel corpus study of ingressive, continuative and resumptive viewpoint aspect

Bogaards, M.P.M.

Citation

Bogaards, M. P. M. (2022). The discovery of aspect: a heuristic parallel corpus study of ingressive, continuative and resumptive viewpoint aspect. *Languages*, 7(3), 158.
doi:10.3390/languages7030158

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/3429585>

Note: To cite this publication please use the final published version (if applicable).

Article

The Discovery of Aspect: A Heuristic Parallel Corpus Study of Ingressive, Continuative and Resumptive Viewpoint Aspect

Maarten Bogaards 

Leiden University Centre for Linguistics (LUCL), Leiden University, 2300 RA Leiden, The Netherlands;
m.p.m.bogaards@hum.leidenuniv.nl

Abstract: Languages differ in how systematically and obligatorily they encode conceptual categories such as tense and aspect. By drawing on large parallel corpora, these differences can be exploited heuristically: expressive obligatoriness and the systematicity of a conceptual category in one language can function as a probe for other languages that do not (evidently) encode it. This study applies this method—called *heuristic translation mining* (HTM)—to viewpoint aspect in Mandarin (an aspect-oriented language) and Dutch (a non-aspect-oriented language). Specifically, it takes the Mandarin aspect markers 起来-*qilai* (“ingressive”) and 下去-*xiaqu* (“continuative”) and collects translation strategies for these markers from a corpus of five Mandarin novels and their Dutch translations. The outcomes are methodological, descriptive and theoretical in nature. Methodologically, it is shown how conceptual templates consisting of temporal boundaries and phases facilitate annotating specific types of viewpoint aspect consistently. Descriptively, the exercise indicates at which linguistic levels viewpoint aspect may be encoded in a non-aspect-oriented language. Theoretically, conducting an HTM analysis with several aspect markers at once makes it possible to quantify (non-)marking of conceptual content; it turns out that the types of viewpoint under study correspond to varying marking frequencies, which may correlate with conceptual complexity.



Citation: Bogaards, Maarten. 2022. The Discovery of Aspect: A Heuristic Parallel Corpus Study of Ingressive, Continuative and Resumptive Viewpoint Aspect. *Languages* 7: 158. <https://doi.org/10.3390/languages7030158>

Academic Editors: Henriëtte de Swart and Bert Le Bruyn

Received: 6 October 2021

Accepted: 17 June 2022

Published: 24 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: viewpoint aspect; ingressive aspect; continuative aspect; resumptive aspect; parallel corpus; Mandarin; Dutch

1. Introduction

Languages vary synchronically in the ways they encode conceptual categories such as tense, aspect and modality (i.e., [Comrie 1976](#); [Bybee et al. 1994](#); [Bhat 1999](#); [Binnick 2012](#)). Major axes of variation are the degrees of obligatoriness and systematicity of expression. A well-known construction illustrating this point is the English Progressive, which is systematic in its composition (i.e., *be V-ing*), cf. (1a), and obligatory for ongoing situations vis-à-vis the simple form of the verb, cf. (1b). Lower on these axes—at least when it comes to progressivity—is Dutch, which allows a simple present verb for an ongoing situation, cf. (2a), and has at least two progressive constructions: the prepositional progressive (*aan het V_{INF} zijn*) in (2b), and the posture progressive (*zitten te V_{INF}*) in (2c). Since (2a–c) may all translate to (1a), Dutch encodes progressive aspect in a less obligatory and systematic way than English does (cf. [Bertrand et al. 2022](#) for a similar juxtaposition of the English and Dutch present perfect constructions).

- (1) a. Rosie is eating a sandwich.
b. *Rosie eats a sandwich.
- (2) a. Rosie eet een broodje.
b. Rosie is een broodje aan het eten.
c. Rosie zit een broodje te eten.
'Rosie is eating a sandwich.'

Advances in the accessibility and searchability of large-scale parallel corpora (cf. [Barlow 2008](#); [Cysouw and Wälchli 2007](#); [Dahl 2007, 2014](#)) have made it possible to exploit this type of crosslin-

guistic variation *heuristically*, remedying a basic methodological problem: identifying the means of expression for a conceptual category in a language that does not (evidently) encode that category obligatorily or systematically. By working with parallel translation corpora of languages on opposite ends of these axes, the obligatoriness and systematicity of expression in one language can help to discover the equivalent means of expression in the other—using the source language, in other words, as a probe for the target language.

This particular application of parallel corpus methods—which builds upon corpus-based contrastive linguistics (CCL: i.a., [Aijmer et al. 1996](#); [Johansson 2007](#); [Granger and Lefer 2020](#)), Translation Mining (TM: i.a., [van derKlis et al. 2017, 2020, 2021](#); [de Swart et al. 2022](#)) and multiple parallel text analysis (MParT: i.a., [Lu and Verhagen 2016](#); [Lu et al. 2018](#); [Knotková and Lu 2020](#))—has been called *heuristic translation mining* (HTM: [Bogaards 2019a](#)).¹ HTM is a useful instrument in the linguist’s toolkit for both descriptive and theoretical questions. Descriptively, HTM may point the researcher in the direction of means of expression they were not aware of. And theoretically, HTM may contribute to understanding more precisely how languages vary in how they encode grammatical concepts, which may in turn inform general theories about those concepts by broadening their empirical basis.

The probing perspective HTM capitalizes on is not new; the idea has some precedent in CCL in particular. What distinguishes HTM from CCL, though, is that HTM specifically seeks out languages that are maximally divergent in terms of optionality and systematicity of expression of a given grammatical concept. This provides empirical input for theories working from the assumption that crosslinguistic variation is constrained by universal ordering principles, for instance the cartographic framework ([Cinque 1999 et seq.](#)) and generative approaches to the syntax–semantics interface (i.a., [Travis 2010](#); [Ramchand 2018](#)). The connection between parallel corpus methods and this type of theoretical work is underdeveloped, and HTM can be seen as a subtype of CCL particularly suited to bridging this gap. This paper is part of a larger research project aiming to make this connection.

The aim of this paper is thus to expand the scope of HTM. It does so by applying it to viewpoint aspect in Mandarin Chinese and Dutch. As an aspect-oriented language, Mandarin has a set of verbal particles encoding particular types of viewpoint aspect (i.a., [Chao 1968](#); [Li and Thompson 1981](#); [Yang 1995](#); [Dai 1997](#); [Xiao and McEnery 2004](#); [Wiedenhof 2016](#); [Zhao 2022](#)). Dutch is on the other end of the spectrum, with periphrastic aspectual expressions—like those in (2b,c)—characterized by optionality and limited systematicity (cf. [Boogaart and Bogaards Forthcoming](#)).

The present study takes two Mandarin viewpoint aspect markers and examines how they are translated to Dutch in a specially assembled parallel corpus drawing on five Mandarin novels and their Dutch translations. The two selected aspect markers are 起来 *-qilai* and 下去 *-xiaqu*, which encode “ingressive” (also called “inchoative” and “inceptive”) and “continuative” (also called “successive” and “continuous”) aspect, respectively.² These particular types of viewpoint aspect were chosen because they vary in conceptual complexity yet are similar enough to be compared and operationalized simultaneously—that is, using the same conceptual building blocks.

This study innovates in three ways vis-à-vis previous studies applying (H)TM or MParT. First, two distinct conceptual categories are involved at once rather than just one; examples of the latter include parallel studies into motion events ([Verkerk 2014](#)), speech representation ([Verhagen 2012](#)), indefinite pronouns ([Beekhuizen et al. 2017](#)) and definiteness ([Bremmers et al. 2021](#)). Working with two types of aspect makes it possible to compare the (non-)marking of each type as a proxy for obligatoriness in the target language.

Second, the types of viewpoint aspect under study differ conceptually from previous (H)TM studies into viewpoint aspect, which operate at the level of the perfective–perfect and perfective–imperfective distinctions ([Dahl 2014](#); [van derKlis et al. 2017, 2020, 2021](#); [Bogaards 2019a](#)). I will try to show that ingressivity and continuativity can be incorporated in a parallel corpus study by defining internal-temporal templates based on selection theory ([Bickel 1997](#)).

Third, the aspect markers under study are low-frequency expressions, so rather than extracting a set number of sentences featuring a given aspect marker from one novel (van der Klis et al. 2020, 2021) or a number of novels (Bogaards 2019a), this study is based on a corpus of five full novels, amounting to almost 30,000 translation pairs. Working with a corpus of this size is an attempt to work out the scale necessary to apply HTM to forms with lower frequency.

The paper is structured as follows. Section 2 goes over the basic assumptions of HTM, its relation to other types of parallel corpus research, and two issues raised by the design of this study. Then, Section 3 reviews previous research on viewpoint aspect in Dutch, aiming to give an overview of the expressions already known to be tied to ingressive and continuativity. Next up are the Mandarin aspect markers: in Section 4, the aspectual values encoded by 起来 *-qilai* and 下去 *-xiaqu* are described on the basis of Mandarin reference grammars and previous studies, and operationalized such that (i) it is clear how they carve up the aspectual domain within their ‘native’ system and (ii) they can be uniformly applied to the target language. Section 5 reports on corpus assembly and annotation and Section 6 presents the results: a discussion of the types of ingressive and continuative aspect marking found in a non-aspect-oriented language like Dutch, and a comparison of (non-)marking frequency of these types of viewpoint aspect. Section 7 concludes by giving an overview of the ‘discovered’ means of expression and reflecting on methodological and theoretical points raised by the study.

2. Heuristic Translation Mining

At its core, heuristic translation mining (HTM) involves systematically studying parallel corpus items—i.e., original and translated language material aligned sentence-by-sentence—with the aim of identifying means of expression in the language of translation (Lu and Verhagen 2016, p. 12; Bogaards 2019a, p. 161). This section starts out by reviewing HTM’s basic assumptions and relating the method to similar parallel corpus approaches. Then, I raise two additional issues: selecting conceptual content and discourse type (discourse vs. narrative), and the relation between parallel and non-parallel corpora.

2.1. Basic Assumptions

HTM takes several basic assumptions from multiple parallel text analysis (MPaT), specifically regarding *sensibility*, *heterogeneity* and *asymmetry*. These principles both overlap with and deviate from corpus-based contrastive linguistics (CCL) and translation mining (TM). The sensibility assumption bears on the text producers of the parallel corpus material: the translator, faced with the task of verbalizing the conceptual content from the original text in the language of translation, is assumed to come up with sensible strategies to complete this task (Lu and Verhagen 2016, p. 170). In this sense, HTM treats translations as large surveys with a limited number of respondents.³

An intervening factor is crosslinguistic heterogeneity: languages vary in the means of expression they offer their speakers; they may “provide different conventional construals” (Verhagen 2012, p. 1). Especially at the level of grammatical notions such as viewpoint aspect, the consequence is that (aspect) marking in the original need not entail marking or even equivalent construal in the translation. Corpus annotation for HTM, therefore, requires labels for ‘unmarked’ and ‘untranslated’ items (Bogaards 2019a, pp. 170–73).⁴ These labels have direct counterparts in corpus-based contrastive linguistics (CCL), namely, ‘zero correspondence’ and ‘divergent correspondence’ (Johansson 2007, pp. 25–26).

The third basic principle drawn from MPaT pertains to the methodological status of the languages under study, which is asymmetrical: the source language (here: Mandarin) is used to gather new insights about the target language (here: Dutch), not the other way around. This is because conceptual content (here: ingressive and continuative aspect) is defined in terms of particular formal manifestations in one language. Naturally, the analysis can be reversed afterwards, by assembling a new corpus drawing on, for instance, Dutch original texts and Mandarin translations. But this requires a separate, extra step; in its

basic application, HTM is a “one-way street” (Bogaards 2019a, p. 162). In this sense, HTM has a narrower focus than CCL, which treats translations with caution but does generally pursue bidirectional comparison (e.g., Xiao and McEneaney 2004; Johansson 2007). HTM’s monodirectionality is in line with its focus on maximally divergent expression—in the present case, aspect-oriented versus non-aspect-oriented languages.

HTM shares with MParT its emphasis on the detailed qualitative analysis of parallel sentence pairs as a way of getting a grip on crosslinguistic variation. However, it aligns with TM in that it also seeks to quantify the corpus data, contrary to MParT. The main difference between HTM and TM is the monodirectionality discussed previously, which TM does not apply as strictly. TM starts out by identifying a *crosslinguistic* formal category (or a set of them), for instance, the *have-Perfect* in Germanic and Romance languages (Dahl 2014; van der Klis et al. 2017, 2020, 2021). The goal is then to study the crosslinguistic variation of this category based on differences in distribution, visualized by means of multidimensional scaling (van der Klis and Tellings 2022). Conceptual content is thus not defined in terms of one or more forms in one language, but by forms shared by all languages under study—based on the initial observation that the same form exists across these languages. By extension, TM does not pursue HTM’s heuristic goals: rather than ‘discovering’ new means of expression in one language, TM aims to map the distribution of known ones in several languages. Hence, TM and HTM are complementary approaches: they map different things.

2.2. Conceptual Content and Discourse Type

It is well known that aspect may interact with discourse type (de Swart 2007). In the case of Dutch, use of the Perfect (vis-à-vis the Simple Past) depends on whether one is dealing with *narrative* or *non-narrative* discourse (Boogaart 1999; Le Bruyn et al. 2019; van der Klis et al. 2020, 2021).⁵ In the English and Dutch data collected by Le Bruyn et al. (2019, p. 168), the Perfect only shows up in non-narrative discourse, never in narrative discourse. The question is whether this restrictive interaction is a property of viewpoint aspect in general, or of subtypes like (im)perfectivity in particular. In the present case: do ingressivity (i.e., construing the start of a situation) and continuativity (i.e., construing the continuation of a situation) depend on this discourse-level distinction for their realization?

One known way of realizing ingressivity and continuativity in Dutch is with the (semi-)auxiliaries *beginnen* ‘start’ and *blijven* ‘stay’, respectively (cf. Section 3). They can be used in both narrative sequencing, as in (3), and non-narrative dialogue, as in (4).

- (3) a. Toen *begon* Kailyn te zingen.
‘Then Kailyn started to sing.’
b. Vervolgens *bleef* ze maar zingen.
‘After that she just kept on singing.’
- (4) a. “Wat doe je in de keuken?” “Ik *ben begonnen* met lunch maken.”
“What are you doing in the kitchen?” “I’ve started making lunch.”
b. “Schilder je nog?” “Ja, ik *ben blijven* schilderen tijdens de pandemie.”
“Are you still painting?” “Yes, I kept on painting during the pandemic.”

The difference between (3) and (4) lies not in the expression of ingressivity/continuativity (by *beginnen* and *blijven* in both cases), but in the tense marking on these verbs: simple past in (3), present perfect in (4). If one abstracts away from the viewpoint contributed by the tense markings in these clauses, then the construal and formal realization of ingressivity and continuativity in (3) and (4) are stable: *beginnen* and *blijven* mark initial and continuative points regardless and are therefore not sensitive to discourse type.

The more general point is: interaction with discourse-level distinctions such as ‘narrative/non-narrative’ seems to depend on the specific aspectual categories under study. While the Dutch Perfect is restricted to non-narrative discourse, there is no reason to suspect that ingressivity and continuativity are as well. For this reason, this study will not make

this distinction in the corpus annotation, contrary to previous (H)TM studies, which deal with (im)perfectivity (e.g., [Le Bruyn et al. 2019](#); [Bogaards 2019a](#)).

2.3. Parallel and Non-Parallel Corpora

Parallel corpora have their limitations: they are largely comprised of translated discourse, which is different from non-translated discourse (e.g., [Vandevoorde et al. 2016](#)), and they are restricted in size by the limited availability of high-quality translations. To make this more tangible: the (large) corpus assembled for the present study, consisting of five full novels, contains 413,094 Dutch tokens. By comparison, the SoNaR corpus of written Dutch ([Oostdijk et al. 2013](#)) contains more than 500 million—a ratio of roughly 1 to 1200. SoNaR also covers a variety of genres, whereas the corpus assembled here draws only on novels. For these reasons, HTM is said to produce results with strictly *hypothetical* status ([Bogaards 2019a](#), p. 188), that is, hypotheses of the type “*beginnen* is a frequent ingressive marker in Dutch”, which are to be tested on a larger amount of non-translated linguistic data. A related point is that these hypotheses are likely not exhaustive; due to the limited size and composition of parallel corpora, there is always the possibility that certain means of expression were missed by the heuristic.

In HTM, parallel and non-parallel corpora should thus be used side-by-side. Two interrelated studies illustrating this point are [Bogaards \(2019a, 2019b\)](#): the former investigated the imperfective domain in Dutch by using the Mandarin imperfective marker 着 *-zhe* in a parallel corpus, which among other things yielded the construction ‘posture verb + past participle’ (e.g., *begraven liggen* ‘lie buried’); the latter then examined the construction’s meaning and restrictions on its productivity by examining about 8000 attestations from the SoNaR corpus. In the same vein, this study aims to produce testable hypotheses concerning ingressivity and continuativity in Dutch.

3. Ingressive and Continuative Aspect in Dutch

What is already known about ingressive and continuative aspect in Dutch? After presenting a very general definition of these types of viewpoint aspect, this section will survey the means of expression associated with them in the literature.

According to [Dik and Hengeveld \(1997, pp. 221–25\)](#), ingressivity and continuativity belong to a particular subtype of viewpoint aspect, namely, *phasal* aspect distinctions. Phasal aspect operates on situation aspect (e.g., States, Activities in the sense of [Vendler 1967](#)) in that phasal distinctions divide situations up into “phase[s] of development [. . .] in terms of beginning-continuation-end” ([Dik and Hengeveld 1997, p. 221](#)).⁶ The ingressive viewpoint focuses on the first of these, i.e., the initial temporal boundary; and the continuative viewpoint focuses on a temporal point somewhere in between the initial and terminal boundaries.⁷ [Dik and Hengeveld \(1997, p. 225\)](#) provide the simple English examples in (5) to illustrate these general definitions.⁸

- | | | | |
|-----|----|--------------|--------------------------|
| (5) | a. | Ingressive | ‘John started crying.’ |
| | b. | Continuative | ‘John continued crying.’ |

Having formulated a working definition of these types of viewpoint aspect, we will turn to previous treatments of these notions in Dutch grammar.

3.1. Ingressivity in Dutch

The *Syntax of Dutch* ([Broekhuis et al. 2015](#); [Broekhuis and Corver 2015](#)) discusses the verb *beginnen* ‘start’ from (3) and (4) as an “inchoative”—which is interchangeable in this context with “ingressive”—verb, as well as the “aspectual verbs” *gaan* ‘go’ and *komen* ‘come’ combined with a bare infinitive ([Broekhuis et al. 2015, pp. 150–56](#)). These verbs take either a VP-complement (e.g., *gaan lopen* ‘begin walking’) or TP-complement (e.g., *beginnen te lopen* ‘begin walking’). They also mention verbs that do not take a VP- or TP-complement, namely, *aanvangen met* ‘commence with’ and *overgaan tot* ‘take to’. These other verbal expressions

(except *overgaan tot*) are mentioned by Vandevoorde et al. (2016, p. 136) as well, who also add *een aanvang nemen* ‘commence’, *starten* ‘start’ and *aanvatten* ‘commence’.

A more elaborate category of ingressive markers in Dutch is described in the general reference grammar *Algemene Nederlandse Spraakkunst* (ANS: Haeseryn et al. 1997): the prepositional pattern shown previously in (2b), which combines with the main verb *zijn* ‘be’ to encode progressive aspect, may also be licensed by other verbs to encode ingressivity—simply called “starting” in the ANS (Haeseryn et al. 1997, pp. 1048–54). This pattern is called the ‘*aan het*-construction’. The ingressive verbs licensed by the construction subdivide into causative and non-causative forms; (6) illustrates both: *krijgen* ‘obtain’ in (6a) is causative, *slaan* ‘hit’ in (6b) non-causative.

- (6) a. De docent *krijgt* zijn leerlingen *aan het lezen*.
‘The teacher gets his students to start reading.’
- b. De leerlingen *slaan* enthousiast *aan het lezen*.
‘The students enthusiastically start reading.’

Van Pottelberge (2004, pp. 27–51), Booij (2010, pp. 146–68) and Bogaards (2020a, pp. 59–92) further developed the ANS classification. Van Pottelberge noted that certain modal verbs get an ingressive reading when used in this construction.⁹ Table 1—which covers part of the full classification in Bogaards (2020a, p. 88)—lists the 14 verbs with ingressive meaning in this construction, subcategorized by causativity and modality.

Table 1. Ingressive verbs licensed by the ‘*aan het*-construction’.

Ingressive Verb	[±CAUSATIVE]	[±MODAL]
<i>gaan</i> ‘go’, <i>komen</i> ‘come’, <i>(ge)raken</i> ‘get’, <i>slaan</i> ‘hit’, <i>zich zetten</i> ‘put oneself’	–	–
<i>krijgen</i> ‘obtain’, <i>maken</i> ‘make’, <i>zetten</i> ‘put’, <i>brengen</i> ‘bring’	+	–
<i>kunnen</i> ‘can’, <i>moeten</i> ‘must’, <i>mogen</i> ‘may’, <i>willen</i> ‘want’	–	+

A final construction linked to ingressivity is the verb *worden* ‘become’ in its function as a “change-of-state copula” (Hanegreefs 2004). *Worden* selects an adjectival or nominal complement (e.g., *boos worden* ‘become angry’, *vader worden* ‘become a dad’) and construes the beginning of the state denoted by the adjective or noun.¹⁰

Summing up, the literature mentions five strategies for ingressivity in Dutch:

- *Beginnen* ‘start’ + *te* + V_{INF};
- *Gaan* ‘go’ and *komen* ‘come’ + V_{INF};
- Ingressive verb set (Table 1) + *aan*-PP;
- Change-of-state copula *worden* ‘become’ + A/N;
- Other verbal strategies: *aanvangen met* ‘commence with’, *overgaan tot* ‘take to’, *een aanvang nemen* ‘commence’, *starten* ‘start’, *aanvatten* ‘commence’ ...

3.2. Continuity in Dutch

The verb set of which the ingressive verbs in Table 1 are a subset also includes two continuative verbs: *blijven* ‘stay’ and its causative counterpart *houden* ‘hold’ (Haeseryn et al. 1997, p. 1050). They are illustrated in (7).

- (7) a. De docent *houdt* zijn leerlingen *aan het lezen*.
‘The teacher keeps his students reading.’
- b. De leerlingen *blijven* enthousiast *aan het lezen*.
‘The students keep on reading enthusiastically.’

The verb *blijven* ‘stay’ also combines with a VP (rather than a PP), e.g., *blijven lezen* ‘keep on reading’ (Broekhuis and Corver 2015, pp. 1019–20), and as a copula similar to *worden*, e.g., *boos blijven* ‘stay angry’ (Broekhuis and Corver 2015, p. 796). Another verb that Broekhuis and

Corver mention, again in passing, is *vervolgen* ‘continue’ (Broekhuis and Corver 2015, p. 695). Finally, there is the ‘expressive *en maar*-construction’, which encodes continuativity while evaluating the continuation negatively (Broekhuis and Corver 2017; Anthonissen et al. 2019, pp. 1148–50). An example from Broekhuis and Corver (2017, p. 311) is given in (8).

- (8) *En maar leugens verspreiden!* Houd daar toch mee op!
 ‘You just keep on spreading lies! Stop doing that!’

In sum, we can find five types of strategies in the literature for encoding continuativity in Dutch, three of which involve *blijven* ‘stay’:

- *Blijven* + VINF;
- Continuative copula *blijven* + A/N;
- Continuative verb (*blijven* or *houden* ‘hold’) + aan-PP;
- *En maar* + V_{INF};
- Other verbal strategies: *vervolgen* ‘continue’ . . .

4. Operationalizing Viewpoint Aspect

This section aims to operationalize the viewpoint-aspectual values encoded by Mandarin 起来 *-qilai* and 下去 *-xiaqu* for HTM corpus annotation. This is no trivial task for two reasons. First, both heuristic elements are polysemous and have several uses besides their aspectual ones. These meanings need to be teased apart and accounted for during annotation. Second, corpus annotation based on the markers’ aspectual meaning must be transparent and consistent. This is only possible if the conceptual content under study is made explicit and applicable. At the end of this section, I will propose that one way of doing so is by working with viewpoint-aspectual templates defined by temporal boundaries and phases. Before that, I will go into the non-aspectual and aspectual semantics of *-qilai* and *-xiaqu*. Items from the corpus will be used as examples. (See Supplementary Materials).

4.1. Non-Aspectual Uses

As aspect markers, both *-qilai* and *-xiaqu* derive from lexical verbs denoting movement (Wu 2000). Synchronically, they still exist as lexical verbs with opposite meanings in terms of spatial direction: *qilai* ‘stand up’ and *xiaqu* ‘go down’, see (9) and (10) below.¹¹

- (9) 哦, 小夕, 你刚起来吗?
 O, Xiaoxi, ni gang qilai-ma?
 oh Xiaoxi you Just now QILAI-QST
 ‘Oh Xiaoxi did you only get up just now?’

- (10) 可是我哭着下去找他。
 Keshi wo ku-zhe xiaqu zhao ta.
 but I cry-IMP XIAQU search him
 ‘But I went downstairs crying to look for him.’

A closely related use is as directional “Resultative Verb Compounds” (Li and Thompson 1981, pp. 58–65; Cherici 2019), i.e., directional RVCs, where *-qilai* and *-xiaqu* attach to a main verb and specify the situation denoted by that verb to involve completed upward or downward movement (see (11) and (12)). It is easy to see how these types of spatial movement relate to ‘standing up’ and ‘going down’.

- (11) 小染想跳起来。
 Xiaoran xiang tiao-qilai
 Xiaoran want jump-QILAI
 ‘Xiaoran wanted to jump upright.’

- (12) 吴敏把头低了下去, 没有答腔。
 Wu Min ba tou di-le-xiaqu, meiyou daqiang
 Wu Min take head bow-PRF-XIAQU not.have answer
 ‘Wu Min bowed his head down, and didn’t answer.’

Abstracting away from ‘up’ and ‘down’, *-qilai* and *-xiaqu* may also function as RVCs without directional meaning (Xiao and McEnergy 2004, pp. 227–28; Xiong 2017, p. 51). In these cases, they indicate not completed movement but completed event (see (13) and (14)).

- (13) 我的两只手被捆绑起来吊了两天。
 Wo de liang zhi shou bei kungbang-qilai diao-le liang tian.
 I SUB two CLF hand by bind-QILAI hang-PRF two days
 ‘They had bound my hands together and left me hanging like that for two days.’
- (14) 只是看也不看地咽下去。
 Zhishi kan ye bu kan de yan-xiaqu.
 only look also not look SUB swallow-XIAQU
 ‘She swallowed everything without looking.’

For *-xiaqu*, these are all non-aspectual types. But for *-qilai* there is one more category, which Xiong (2017) terms “semantically opaque *qilai*”, covering cases where “it is hard to flesh out the semantic contribution of *qilai*” (Xiong 2017, p. 51). This category covers multiple types, including impersonal and evaluative uses (Wang and Wu 2018), as well as the *qilai*-middle, which is illustrated in (15). Since the present study is concerned with ingressive *-qilai*, I will adopt Xiong’s ‘opaque’ category as a cover term here for all non-aspectual uses of *qilai* other than lexical, directional and resultative ones.

- (15) 行走起来，非常不方便，鞋子里渗进雪水，冻得两只脚又僵又痛。
 Xingzou-qilai feichang bu fangbian xiezi-li shen-jin xue shui
 walk-QILAI very not convenient shoe-inside seep-in snow water
 dong de liang-zhi jiao you jiang you tong.
 freeze RES two-CLF foot also stiff also ache
 ‘It walks very unpleasantly, the snow water gets into your shoes and your feet get so cold that they become both stiff and painful.’

There are thus three non-aspectual uses of *-xiaqu* (lexical, directional, resultative) and four non-aspectual uses of *-qilai* (lexical, directional, resultative, opaque) to be accounted for in the corpus annotation. Next up is the aspectual semantics of both markers.

4.2. Ingressive *-qilai*

While *-qilai* synchronically retains its lexical meaning of ‘standing up’ as well as directly extended uses such as directionality, it is also widely analyzed as a viewpoint aspect marker encoding ingressivity (i.a., Chao 1968; Dai 1997; Xiao and McEnergy 2004; Shyu et al. 2013; Xiong 2017; Wang and Wu 2018). As an ingressive marker, *-qilai* attaches to a main verb and contributes an aspectual viewpoint to the situation denoted by that verb. Xiao and McEnergy define its aspectual contribution as “a situation has started and will continue for some time” (Xiao and McEnergy 2004, p. 219), which aligns with the general definition given earlier in terms of phases of development: *-qilai* selects the beginning phase. In terms of situation aspect, ingressive *-qilai* is sensitive to durativity, being felicitous with all situation types except Achievements.

As for its syntax, *-qilai* may be split if it attaches to a verb that consists of two syllables or licenses some sort of object. Several elements may come between *qi* and *lai*. For example, in (16), *-qilai* attaches to the bisyllabic verb 跳舞 *tiaowu* ‘dance’, producing the structure 跳起舞来 *tiao-qi wu-lai*. In (17), *-qilai* is selected by the verb–object constituent 开玩笑 *kai wanxiao* ‘crack jokes’, in which case *qi* comes in between the verb and its object: 开起玩笑来 *kai-qi wanxiao-lai*. These split patterns have to be accounted for in the corpus.

- (16) 几个年轻人跟着彼头的“黑鸟”歌跳起舞来了。
 Ji-ge nianqingren genzhe Bitou de “Heiniao” ge
 Several-CLF young.people follow Beatles SUB Blackbird song
 tiao-qi wu-lai-le
 jump-QI dance-LAI-PRF
 ‘A couple of young people started dancing to *Blackbird* by the Beatles.’

- (17) 她竟给男人开起玩笑来。
 Ta jing gei nanren kai-qi wanxiao-lai,
 she unexpectedly give man open-QI joke-LAI
 ‘She even started cracking jokes to the man.’

Finally, there is some discussion in the literature on whether the ingressive viewpoint manifested by *-qilai* is a subtype of perfective or imperfective aspect. [Xiao and McEnergy \(2004\)](#) consider it an imperfective viewpoint because it leaves a situation’s terminal boundary unfocused. [Smith \(1997\)](#) argues that it is neither, and introduces an additional category at the level of (im)perfectivity—‘neutral viewpoint aspect’—to account for their relationship. A third way out is to extend one’s definition of (im)perfectivity to include not just totality but also transition sensitivity ([Koss et al. 2022](#)), which results in $2 \times 2 = 4$ categories (in which case ingressivity would be non-total but transition-sensitive). For a heuristic study like the present one, however, the question whether *-qilai*’s ingressivity abstracts to perfectivity or imperfectivity is not relevant. We are interested in the specific means of expression translators come up with for this conceptual content, not its theoretical status in relation to similar concepts.

For *-qilai*, we now have input for both the operationalization (initial boundary and subsequent phase) and the corpus annotation (split structures). The aspectual use of *-xiaqu* is up next.

4.3. Continuative *-xiaqu*

-Xiaqu runs parallel to *-qilai* in its development as an aspect marker with its lexical source still intact. Besides its lexical meaning of ‘going down’ and immediate directional and resultative extensions, it has also been analyzed as a “successive” or “continuative” aspect marker ([Chao 1968](#); [Yang 1995](#); [Dai 1997](#); [Kang 1999](#); [Xiao and McEnergy 2004](#)). In terms of situation type, *-xiaqu* is sensitive to the same parameter as *-qilai*, namely, durativity ([Xiao and McEnergy 2004](#)). It is thus likewise incompatible with Achievements.

-xiaqu is both more complex and less complex than *-qilai*. Syntactically, it is less complex, as it cannot be split.¹²

Semantically, however, *-xiaqu*’s “continuativity” covers two types of viewpoint aspect. The general characterization given by Xiao and McEnergy for the viewpoint encoded by *-xiaqu* is twofold: it “presupposes the prior initiation of a situation” and “focuses on the resumptive point and its continuative stage” ([Xiao and McEnergy 2004](#), p. 227). The conceptual complexity is manifested in the space between the prior initiation and resumptive point: these may be contiguous, as in (18), where the ‘staring’ is followed directly by ‘keeping on staring’. Crucially, though, *-xiaqu* also allows the prior initiation and resumptive point to be separated by a period in which the ‘continued’ situation does not hold. In this case, *-xiaqu* construes a special type of ingressive point: one preceded by prior initiation. This is illustrated by (19), a request to resume talking, which implies both that the other party talked before and that they stopped talking.

- (18) 姜子牙再那么盯下去我就变成尼姑了。
 Jiang Ziya zai name ding-xiaqu wo jiu bianchengnigu le.
 Jiang Ziya again like.that gaze-XIAQU I just become nun PRF
 ‘Jiang Ziya, if you keep on staring at me like that, I will turn into a nun.’
- (19) “你说下去吧！妈，不要停。”
 “Ni shuo-xiaqu-ba! Ma, bu yao ting.”
 you speak-XIAQU-SUG mother not will stop
 ‘Resume talking, mother, don’t stop.’

Xiao and McEnergy term cases like (18) and (19) “uninterrupted” and “interrupted” continuativity, respectively ([Xiao and McEnergy 2004](#), p. 233). Indeed, the (lack of) interruption produces the two kinds of continuative points in (18) and (19): due to interruption, *-xiaqu* in (19) construes a special kind of ingressive point, whereas there is no interruption and hence no ingression in (18); the situation is already ongoing. Instead, cases like (18) are

more closely related to ‘stopping’ than ‘starting’—that is, they boil down to the negation of ‘stopping’ (Jackendoff 1990, p. 44; Lieber and Baayen 1997, p. 804). This is reflected by the fact that the semantic contribution of *-xiaqu* in cases like (18) can be paraphrased as ‘not stop’ (‘if you don’t stop staring at me like that . . . ’). Moreover, when non-interrupted continuative *-xiaqu* is negated, it gets the meaning ‘stop’, i.e., ‘not not stop’, cf. (20). Conversely, if a sentence like (19) is negated, it would mean “not resume” (the same goes for *-qilai*, which would mean “not begin”).

(20)	小纲, 别说下去了!		
	Xiaogang,	<i>bie</i>	<i>shuo-xiaqu-le!</i>
	Xiaogang	not	speak-XIAQU-PRF
	‘Xiaogang, stop talking now!’		

This means that Xiao and McEnergy’s (un)interrupted distinction should play a central role in the present corpus study. To highlight the distinction, I will use two more distinct labels, reserving “continuative” for cases like (18) and (20), where the situation is ongoing and was not terminated; and using “resumptive” for cases like (19), where the situation *was* terminated and *-xiaqu* profiles a special type of initial (i.e., resumptive) point. A related point for annotation is that negated instances of continuative *-xiaqu* (cf. (20)) should not be part of the heuristic, as they do not translate to a negated version of a continuative marker, but to a notion that is not itself continuative (i.e., ‘stopping’).

We now have all the input we need for *-xiaqu* in order to operationalize (prior initiation and *either* not stopping *or* starting again after stopping) and annotate (no split structures and exclusion of negated instances of the continuative marker).

4.4. Operationalization

As the previous two sections made clear, the aspect markers *-qilai* and *-xiaqu* both deal with boundaries and phases in time, albeit different kinds of temporal points in varying constellations. In these terms, they interact with the situations (lexicalized by verbs or adjectives) that they project their aspectual viewpoint onto. By that I mean: situations are likewise made up of temporal phases (e.g., ‘eating a sandwich’) and boundaries (e.g., the first and last bites of that sandwich), and what aspectual viewpoint does is to impose or focus those boundaries and/or phases. This view corresponds to what are called “(Radical) Selection Theories of Aspect” (i.a., Breu 1994; Bickel 1997; Sasse 2002, pp. 222–25; Croft 2012, pp. 49–52; Koss et al. 2022; Bogaards Forthcoming), which assume aspectual bidimensionality (i.e., situation and viewpoint) and claim that viewpoint markers (e.g., ingressive and continuative) map phasal structure already (partly) present in the aspectual structure of situations. They thus have *selectional* preferences.¹³ The aspectual building blocks shared by both situation and viewpoint aspect are temporal boundaries (notated as τ) and middle phases (notated as ϕ) (Bickel 1997). Boundaries τ are punctual and involve change (e.g., telicity); phases ϕ are durative and may be either dynamic or stative.

My claim is that these building blocks can be used to construct templates that make the viewpoint-aspectual value of the heuristic elements explicit and, at a more practical level, lay out a way to annotate the corpus data consistently and transparently.

So, what do ingressive *-qilai* and continuative/resumptive *-xiaqu* look like in terms of these building blocks? Although Xiao and McEnergy (2004) do not explicitly subscribe to the type of aspectual theory discussed above, they do work with temporal boundaries to visualize *-qilai*’s and *-xiaqu*’s aspectual values. I will therefore base my templates on their work (cf. Xiao and McEnergy 2004, pp. 223–33), but attach more weight to (i) the phases construed by boundaries, (ii) the dissimilarities between *-xiaqu*’s continuative and resumptive meanings and (iii) the similarities between resumptivity and ingressivity.

Starting with *-qilai*, the ingressive viewpoint encoded by this marker is quite simple, consisting of just one initial boundary τ and one middle phase ϕ in which the situation denoted by the main verb holds. This is visualized in Figure 1 across a temporal axis t ; the ‘initial’ status of τ is captured by including an additional phase preceding it during which the same situation does not hold ($\neg\phi$). Put more generally, initial boundaries τ are those in

between two phases $\neg\phi$ and ϕ . The braces map these phases onto the temporal axis, and their (half-)dottedness is meant to indicate that they are not (fully) in focus; the emphasis lies on the initial point.

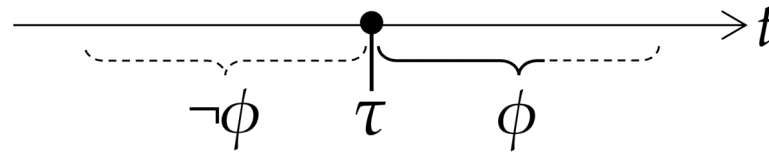


Figure 1. Schematic representation of ingressive aspect built out of boundaries τ and phases ϕ .

Schematizing continuative *-xiaoqu* using the same building blocks and conventions, the viewpoint in Figure 2 is somewhat more complex; in addition to the focused boundary τ_2 there is an unfocused preceding boundary τ_1 , which is initial (between $\neg\phi$ and ϕ), contrary to τ_2 (between ϕ_1 and ϕ_2). Normally, a boundary following a phase would itself constitute a terminal boundary (in between ϕ and $\neg\phi$), but here the second phase is not negated. The analysis in Section 4.3 would account for this by claiming that ϕ_2 in Figure 2 corresponds to $\neg(\neg\phi)$.

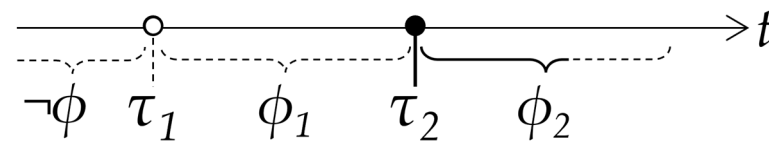


Figure 2. Schematic representation of continuative aspect built out of boundaries τ and phases ϕ .

Resumptive *-xiaoqu* is the most complex (see Figure 3), featuring three boundaries and four phases: two where the situation holds, two where it does not. The focused boundary τ_3 is sandwiched between $\neg\phi$ and ϕ , making it initial. However, there is an additional unfocused complete phase preceding τ_3 defined by an initial boundary τ_1 and a terminal boundary τ_2 . In this sense, resumptivity constitutes a focused initial boundary embedded in a more complex viewpoint. I would thus argue that in its focused region, resumptivity is in fact more similar to ingressivity than to continuativity.¹⁴

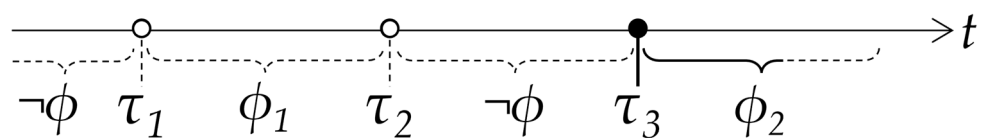


Figure 3. Schematic representation of resumptive aspect built out of boundaries τ and phases ϕ .

How do the templates in Figures 1–3 help the annotator? In two ways: first, the annotator can test whether the viewpoint is present in the translation by checking whether the situation in question also holds without a given marker. If so, they are dealing with continuativity; if not, with either ingressivity or resumptivity. Second, they can draw on the varying complexity of the viewpoints to motivate their annotation. The ingressive viewpoint construes only one temporal boundary, the continuative viewpoint two, the resumptive viewpoint three. Only if these complex aspectual structures are encoded by a given means of expression, can the translation be said to be marked for that type of viewpoint aspect. Taken together, working with these templates gives the annotator explicit and transparent schemas to fall back on when determining whether a specific type of viewpoint aspect is explicitly marked in a translation item in the corpus.

5. Corpus Assembly and Annotation

This study is part of a larger project aiming to develop and apply HTM in the domain of viewpoint aspect, with special attention for Mandarin and Dutch as aspect-oriented and

non-aspect-oriented languages. The parallel corpus that the present study is based on is being assembled as part of this project. Corpus assembly is currently in an intermediate stage: five full novels and their translations have been digitized (via scanning and OCR) and aligned sentence-by-sentence, out of a planned twelve novels total. This study is the first application of HTM to a corpus of this size. Table 2 lists the Mandarin novels and Dutch translations with their authors and translators.

Table 2. Novels and translations in the parallel corpus.

Original	Author	Translation	Translator(s)
桑青与桃红 <i>Sangqing yu Taohong</i> 'Mulberry and peach'	聂华 Nie Hualing	<i>Twee Chinese vrouwen</i> 'Two Chinese women'	Anne Sytske Keijser
孽子 <i>Niezi</i> 'Unworthy sons'	白先勇 Pai Hsien-Yung	<i>Jongens van glas</i> 'Boys made of glass'	Mark Leenhouts
只有一个太阳 <i>Zhi you yige taiyang</i> 'There's only one sun'	张洁 Zhang Jie	<i>Er is maar één zon</i> 'There's only one sun'	Koos Kuiper
十爱 <i>Shi ai</i> 'Ten loves'	张悦然 Zhang Yueran	<i>Tien liefdes</i> 'Ten loves'	Various, edited by Annelous Stiggelbout
第七天 <i>Di qi tian</i> 'The seventh day'	余华 Yu Hua	<i>De zevende dag</i> 'The seventh day'	Jan De Meyer

In its current state, the parallel corpus comprises 29,668 aligned sentence pairs: 611,232 characters on the Mandarin side, 413,094 words on the Dutch one. The basis for determining sentence pairs was orthographic: the full stop (“ . ” in Mandarin orthography) was taken as a boundary to distinguish corpus items. But as became clear during the alignment process, sentence boundaries often shift in translation; shorter Mandarin sentences are sometimes sewn together by Dutch translators, and long-winded Mandarin sentences are generally broken up into two or more separate clauses in Dutch. To ensure the quality of alignment, all sentence pairs in the corpus were therefore aligned manually.

With the parallel corpus for this study in place, the next step was to extract all instances of 起来 *-qilai* and 下去 *-xiaqu*. For *-xiaqu* this was simple enough, as it cannot be split. However, for split cases of *-qilai* a number of tokens had to be allowed between *qi* and *lai*. To extract as many instances as possible while also limiting noise, a maximum of three tokens was allowed. This yielded a total of 1074 items featuring *qi*{0,3}*lai* and 230 items featuring *xiaqu*. Noise—that is, items that did not correspond to either the non-aspectual or aspectual uses of *xiaqu* or *qilai* discussed in Section 4—was very rare: none for *-xiaqu*, only seven for *-qilai*. Sentence (21) is an example of noise: in this sentence, *qi* is part of the word 一起 *yiqi* ‘together’, and *lai* is part of the verb 过来 *guolai* ‘come over’.

- (21) “我们是一起过来的。”男声回答。
 “Women shi yiqi guolai de.” Nan sheng huida.
 we be together come.over SUB man voice answer
 ‘“We came here together,” the male voice answered.’

The first annotation round concerned coding all 1304 items for the usage types laid out in Section 4: lexical (*qilai* ‘stand up’, *xiaqu* ‘go down’), directional (‘up’ and ‘down’) and aspectual. For *-qilai* there was also the leftover category ‘opaque’ for items corresponding to none of these uses (e.g., middles). Finally, for *-xiaqu* the specific type of aspect had to be determined: continuative or resumptive. Sometimes this became clear from the immediate co-text (e.g., (18)–(20)), sometimes the wider co-text had to be consulted. Sentence (22a) is an example of a co-textually determined ‘continuative’ annotation: the direct co-text was not entirely conclusive, so the sentence directly before it—(22b)—had to be checked in order to pin down the aspectual type. In this way, all items could be annotated with certainty.

- (22) a. 她仍然用中文不停地说下去。
Ta rengran yong Zhongwen bu ting de shuo-xiaqu.
 she still use Chinese not stop SUB speak-XIAQU
 ‘She continues/resumes speaking in Chinese without stopping.’
- b. 他打了几次手势叫桃红住嘴。
Ta da-le ji-ci shoushi jiao Tao Hong zhuzui.
 he hit-PRF several-CLF gesture call Tao Hong shut.up
 ‘He gestures to Tao Hong to stop talking.’

Tables 3 and 4 show the annotation results from the first round for both markers. Everything except the ingressive, continuative and resumptive categories will be left out of consideration from this point on.

Table 3. Types of *-qilai* in the corpus, absolute and relative frequencies.

Type	n	%
Ingressive	452	42.1%
Directional	400	37.2%
Opaque	115	10.7%
Resultative	85	7.9%
Lexical	15	1.4%
Noise	7	0.7%
Total	1074	100%

Table 4. Types of *-xiaqu* in the corpus, absolute and relative frequencies.

Type	n	%
Directional	104	45.2%
Continuative	56	24.3%
Resultative	40	17.4%
Resumptive	26	11.3%
Lexical	4	1.7%
Noise	0	0%
Total	230	100%

An additional concern for the annotation of continuative *-xiaqu* raised in Section 4 involves negation, which—because it results in the interpretation ‘stopping’—is not itself continuative. Within the continuative subtype, negation at the clause-level was therefore also annotated. There were 10 negated continuative items total, which were not included in the second round of annotation. The total number of tokens to be involved in the analysis is listed in Table 5.

Table 5. Total number of tokens of aspectual *-qilai* and *-xiaqu* to be involved in HTM analysis.

Aspect Marker	n
<i>-qilai</i>	452
<i>-xiaqu</i>	72
Total	524

Then came the second annotation round—the core business of HTM: determining the expression(s) that the aspectual content (viz., ingressive, continuative, resumptive) was translated with. Following from the basic assumptions laid out in Section 2, not all content is necessarily *translated* or *marked*. The difference between the two is the following. *Untranslated* items either leave out the part of the clause with the aspect marker in the translation or deviate so strongly from the original structure that the locus of viewpoint aspect cannot be pinpointed. *Unmarked* items follow the original structure but do not mark the viewpoint aspect explicitly.

Sentence (23) exemplifies an untranslated item, (24) an unmarked one. From here on out, the Dutch translation from the parallel corpus will be given under (a), and my own English translation of the Dutch translated corpus item under (b). In (23), the reporting clause containing *-qilai* was simply left out in the translation; in (24), the verb marked by *-qilai* (尖叫 *jianjiao* ‘shriek’) was translated to *riep uit* ‘exclaimed’, without any marking.

- (23) “是匈奴吧？”吴每笑了起来。
 “*Shi Xiongnu ba?*” Wu Min *xiao-le-qilai.*
 be Huns SUG Wu Min laugh-PRF-QILAI
 a. ‘Ah, Hunnen dan.’
 b. ‘Ah, the Huns then.’
- (24) 他看着就尖叫起来: [...]
Ta zhe jiu jianjiao-qilai:
 he look-IMP just shriek-QILAI
 a. ‘Al lezende riep hij uit: [...].’
 b. ‘While reading he exclaimed: [...].’

All remaining corpus items (i.e., neither untranslated nor unmarked) were marked for ingressivity, continuativity or resumptivity in some way. The labels for their aspect marking strategies were designed and assigned in two ways. First, if the marking was the same as one of the strategies mentioned in previous work (Sections 3.1 and 3.2), a label was assigned immediately. If not, the second way to work cyclically: first a preliminary assessment was made and later on—if the same strategy showed up more frequently—a definitive label was assigned to all of the corpus items featuring that strategy. In this way, all 524 relevant corpus items were ‘mined’ heuristically for viewpoint-aspectual means of expression. The results of this exercise will be presented in the next section.

6. Results

This section reports on the results from the HTM annotation: 452 tokens of aspectual *-qilai* and 72 tokens of aspectual *-xiaqu* were coded for the locus of viewpoint aspect in the Dutch translations. For each marker, I will present a quantitative overview of the ‘discovered’ categories, and then go over each of the categories, providing a more qualitative analysis based on specific examples. The idea is that these characterizations serve as hypotheses of the type “X encodes ingressive/continuative/resumptive aspect”, which may then be tested in future research on non-parallel corpora with a view to establishing the precise (non-)aspectual semantics and restrictions on productivity. At the end of this section, I will compare (non-)marking frequency of these types of viewpoint aspect and offer a tentative explanation for marking differences in terms of selection theory.

6.1. Ingressive *-qilai*

Table 6 gives a quantitative overview of all translation strategies (including unmarked and untranslated items) for ingressive *-qilai*. The labels are presented at two levels of abstraction: supercategories, which constitute either the head of a syntagmatically variable construction (e.g., *gaan* ‘go’) or a categorial abstraction over individual translations (i.e., adverbial and other verbal strategies); and subcategories, which specify either the different types of complement selected by a head (e.g., [*gaan* PP] and [*gaan* V_{INF}] under *gaan*) or the specific strategies corresponding to an abstracted category (e.g., *nu* ‘now’ and *voor het eerst* ‘for the first time’ under adverbial strategies). Adverbial strategies are cases where the ingressive value of *-qilai* is encoded by means of an extra AdvP, that is, one that was not present in the original; other verbal strategies are verbs that do not select a VP- or TP-complement.

Table 6. Strategies for translating ingressive *-qilai* to Dutch.

Supercategories			Subcategories		
Translation	n	%	Translation	n	%
UNMARKED	172	38.1%	UNMARKED	172	38.1%
<i>beginnen</i> 'begin'	146	32.3%	<i>beginnen te</i> V _{INF}	141	31.2%
			<i>beginnen</i>	5	1.1%
<i>worden</i> 'become'	41	9.1%	<i>worden</i> A	37	8.2%
			<i>worden</i> N	4	0.9%
<i>barsten</i> 'burst'	20	4.4%	<i>uit-barsten</i> PP	19	4.2%
			<i>los-barsten</i>	1	0.2%
<i>gaan</i> 'go'	11	2.4%	<i>gaan</i> V _{INF}	7	1.6%
			<i>gaan</i> PP	4	0.9%
<i>komen</i> 'come'	10	2.2%	<i>komen</i> PP	4	0.9%
			<i>komen</i> V _{INF}	2	0.4%
			<i>komen</i> A	2	0.4%
			<i>komen</i>	2	0.4%
<i>krijgen</i> 'obtain'	10	2.2%	<i>krijgen</i> N	7	1.6%
			<i>krijgen</i> A	2	0.4%
			<i>krijgen</i> PP	1	0.2%
ADVERBIAL STRATEGY	9	2%	adverbial past participle (e.g., <i>beschaamd</i> 'embarrassed')	4	0.9%
			other adverb (e.g., <i>schor</i> 'hoarse')	2	0.4%
			<i>nu</i> 'now'	1	0.2%
			<i>voor het eerst</i> 'for the first time'	1	0.2%
			<i>nog nooit zo</i> 'never like that'	1	0.2%
OTHER VERBAL STRATEGY	9	2%	<i>een gesprek aanknopen</i> 'strike up a conversation'	1	0.2%
			<i>de oren spitsen</i> 'prick up one's ears'	1	0.2%
			<i>verdriet doen</i> 'give (sb.) pain'	1	0.2%
			<i>ontvouwen</i> 'unfold'	1	0.2%
			<i>oplaaien</i> 'flare up'	1	0.2%
			<i>overgaan tot</i> 'take to'	1	0.2%
			<i>plaatsmaken voor</i> 'make way for'	1	0.2%
<i>volgen</i> 'follow'	1	0.2%			
<i>zich mengen in</i> 'join in'	1	0.2%			

Table 6. Cont.

Supercategories			Subcategories		
Translation	n	%	Translation	n	%
UNTRANSLATED	6	1.3%	UNTRANSLATED	6	1.3%
<i>moeten</i> 'must'	5	1.1%	<i>moeten</i> V _{INF}	5	1.1%
<i>raken</i> 'get'	5	1.1%	<i>raken</i> PP	5	1.1%
<i>maken</i> 'maken'	4	0.9%	<i>maken</i> A	3	0.7%
			<i>maken</i> PP	1	0.2%
<i>zetten</i> 'put'	2	0.4%	<i>op-zetten</i>	1	0.2%
			<i>in-zetten</i>	1	0.2%
<i>schieten</i> 'shoot'	2	0.4%	<i>schieten</i> PP	2	0.4%
Total	452	100%	total	452	100%

What is immediately apparent from the overview in Table 6 is that 'unmarked' is the largest supercategory (38.1%), although all marked cases taken together do constitute the majority (60.6%). Untranslated cases are very infrequent (1.3%). Among marked cases,

strategies from previous work (Section 3.1) are well represented: *beginnen* (32.3%), copular *worden* (9.1%) and verbs from the ingressive set (*gaan, komen, krijgen, moeten, raken, maken, zetten*: cumulatively 7.5%) together make up nearly half of all items (48.9%). Other specific verbs taking a complement are less frequent (*barsten, schieten, laten*: cumulatively 5%), and least frequent are adverbial and other verbal strategies (both 2%). Overall, it seems that the Dutch linguistics literature has a decent grip on ingressivity: all mentioned strategies occur in the corpus.

I will first discuss the marked categories in order of frequency: *beginnen* and *worden* (Section 6.1.1), the ingressive set (Section 6.1.2), specific verbs selecting a complement (Section 6.1.3), and other verbal and adverbial strategies (Section 6.1.4). Last up are unmarked cases (Section 6.1.5).

6.1.1. Known Strategies from Previous Work: *Beginnen* and *Worden*

The prototypical Dutch ingressive verb, *beginnen* ‘begin’, is the most frequent marking strategy. It occurs both with a clausal complement headed by *te* ‘to’ and without (see (25) and (26)).

- (25) 杨教头喇地一下，打开摺扇，扇了起来。
 Yang jiaotou shua de yixia dakai zheshan, shan-le-qilai.
 Yang master rustle SUB once open fan fan-PRF-QILAI
 a. ‘Meester Yang sloeg zijn waaier open en begon ermee te wapperen.’
 b. ‘Master Yang opened his fan and started fanning with it.’
- (26) 阁楼屋顶的声音又响起来了。
 Gelou wuding de shengyin you xiang-qilai-le.
 attic roof SUB sound again make.sound-QILAI-PRF
 a. ‘De geluiden op zolder beginnen weer.’
 b. ‘The sounds in the attic start again.’

Copular *worden* ‘become’ is also present, with both adjectival and nominal complements (see (27) and (28)). As a copula, *worden* attributes the property denoted by the adjective or noun to its subject (in (27) and (28): *boos* ‘angry’ and *een cultuurland* ‘a cultured country’); specifically, it signals the start of that attribution.

- (27) 我气起来就打了他一记耳光。
 Wo qi-qilai jiu da-le ta yi-ji erguang,
 I angry-QILAI just hit-PRF him a-CLF slap.on.the.face
 a. ‘Ik werd kwaad en gaf hem een draai om zijn oren...’
 b. ‘I got upset and gave him a whack around the ears...’
- (28) 幸亏文化故国至少在一百年内还不会很快地文化起来。
 Xinkui wenhua guguo zhishao zai yibai-nian-nei
 fortunately/culture ancient country at least at one.hundred-years-within
 hai bu hui hen kuai de wenhua-qilai,
 still not will very quick SUB culture-QILAI
 a. ‘Gelukkig zou dit land met zijn oude cultuur tenminste de eerstkomende honderd jaar niet gauw weer een cultuurland worden.’
 b. ‘Fortunately, this country with its ancient culture wouldn’t become a cultured country for at least the next one hundred years.’

These findings underscore previous work on ingressivity in Dutch, but do not add anything new. The next section will offer findings that are (partially) novel.

6.1.2. Verbs from the Ingressive Set: *Gaan, Komen, Krijgen, Moeten, Raken, Maken, Zetten*

Seven out of the fourteen verbs in the Dutch ingressive verb set (Section 3.1) are represented in the corpus. Crucial about previous work on the ingressive verb set is that these verbs are said to be in a syntagm with a situational PP headed by the preposition *aan* (i.a., Van Pottelberge 2004; Bogaards et al. Forthcoming; Bogaards). However, these verbs also take other types of complements in the corpus: PPs headed by other prepositions, verbs, nouns and adjectives. Table 7 gives an overview of the specific combinations, providing one or two examples for each one. I will not go into each individual combination here, but the upshot is that the verbs from the ingressive set license a wider set of complements than just *aan*-PPs. Some of these complements may in fact be more frequent than PPs headed by *aan*, given that they are attested here contrary to many combinations with *aan*. As for the

verbs themselves, it is striking that only one causative and one modal were found in the corpus (*moeten* and *maken*).¹⁵

Table 7. Types of complements selected by verbs from the ‘ingressive set’.

Verb	Type of Complement			
	PP	V _{INF}	N	A
<i>gaan</i> ‘go’	<i>aan het werk gaan</i> ‘start working’ <i>op de vuist gaan</i> ‘start a fight’	<i>gaan vechten</i> ‘start fighting’ <i>gaan zitten</i> ‘sit down’	-	-
<i>komen</i> ‘come’	<i>op gang komen</i> ‘get going’ <i>in beweging komen</i> ‘start moving’	-	-	<i>los komen</i> ‘get loose’
<i>krijgen</i> ‘obtain’	<i>het aan zijn hart krijgen</i> ‘start having problems with one’s heart’	-	<i>ruzie krijgen</i> ‘get into a fight’ <i>moed krijgen</i> ‘become brave’	<i>het druk krijgen</i> ‘become busy’
<i>moeten</i> ‘must’	-	<i>moeten lachen</i> ‘start laughing’	-	-
<i>raken</i> ‘get’	<i>aan de praat raken</i> ‘end up chatting’ <i>in de war raken</i> ‘become confused’	-	-	-
<i>maken</i> ‘make’	<i>aan het hoesten maken</i> ‘make (sb.) cough’	-	-	<i>lui maken</i> ‘make (sb.) lazy’ <i>bedroefd maken</i> ‘make (sb.) sad’

Finally, *zetten* also occurs in the corpus, but prefixed with *in-* or *op-* (i.e., *inzetten* ‘set in’ and *opzetten* ‘put on’). These have ingressive meaning by themselves. One of them is illustrated in (29).

- (29) 他的声音也变得幼稚温柔起来。
Ta de shengyin ye biande youzhi wenrou-qilai.
 he SUB voice also change childish tender-QILAI
 a. ‘Hij zette er ook een lief kinderstemmetje bij op.’
 b. ‘He also put on a sweet childish voice.’

In sum, the parallel corpus reveals that the ‘ingressive verb set’ in Dutch is more heterogeneous than previously thought.

6.1.3. Other Verbs Selecting a Complement: *Barsten, Schieten*

The corpus also features three complement-selecting verbs not in the ingressive verb set: *barsten* ‘burst’, *schieten* ‘shoot’ and *laten* ‘let’. The first two of these combine with PPs denoting laughter, crying or related actions, such as *in de lach* ‘lit. in the laugh’, *in huilen* ‘lit. in crying’, and *in gillen* ‘lit. in shouting’. *Barsten* is prefixed with *uit-* ‘out’ in these cases (i.e., *uitbarsten* ‘burst out’). The punctual semantics of ‘bursting’ and ‘shooting’ appears to map to the initial boundary of these situations. This might be facilitated by the fact that the initiation of laughter and crying is usually accompanied by some vehemence. This is likewise lexicalized in English by ‘bursting into laughter/tears’. Corpus items (30) and (31) illustrate *schieten* and *uitbarsten*.

- (30) 他们两个互相看看后又嘻嘻笑起来。
Tamen liang-ge huxiang kankan-hou you xixi xiao-qilai.
 they two-CLF each.other look-after again happy laugh-QILAI
 a. 'De twee keken elkaar aan en *schoten* opnieuw in de lach.'
 b. 'The two of them looked at each other and burst into laughter again.'
- (31) 听到这里, 司马南江忍不住大笑起来。
Tingdao zheli, Sima Nanjiang renbuzhu daxiao-qilai.
 hear here Sima Nanjiang cannot.help big.laugh-QILAI
 a. 'Toen Sima Nanjiang dat hoorde, kon hij zich niet inhouden en *barstte* in hard lachen uit.'
 b. 'When Sima Nanjiang heard that, he couldn't help himself and burst out in loud laughter.'

Barsten and *schieten* appear to impose selectional restrictions with regard to their PP-complement (e.g., one can *in lachen uitbarsten* 'burst out in laughter' but not *in eten uitbarsten* 'burst out in eating').

6.1.4. Adverbial and Other Verbal Strategies

Certain adverbs and verbs in the corpus appear to lexicalize focus on a situation's initial boundary. Examples of such adverbs are *nu* 'now' and *voor het eerst* 'for the first time'. A different adverbial strategy is to relegate the Mandarin main verb marked by *-qilai* to the status of adverbial modifier—for instance, as a past participle—and attribute a property to the former verb's subject that way. An ingressive interpretation is still possible in these cases, though, as (32) illustrates: by functioning as an adverb to the main verb *liet zakken* 'lowered', *beschaamd* 'embarrassed' allows for the interpretation that the state of 'being embarrassed' started at this point, rather than holding already.

- (32) 吴敏不好意思起来, 头一直俯着。
Wu Min buhaoyisi-qilai, tou yizhi fu-zhe.
 Wu Min feel embarrassed-QILAI head straight look.down-IMP
 a. 'Wu Min liet *beschaamd* zijn hoofd zakken.'
 b. 'Wu min lowered his head in shame.'

Verbs or verbal expressions lexicalizing ingressivity in the corpus include *de oren spitsen* 'prick up one's ears', *ontvouwen* 'unfold', *plaatsmaken voor* 'make way for' and *een gesprek aanknopen* 'strike up a conversation', which is illustrated in (33).

- (33) 并且不经他人介绍就和司马南江交谈起来。
Bingqie bu jing taren jieshao jiu he Sima Nanjiang jiaotan-qilai.
 and not through another introduce just with Sima Nanjiang chat-QILAI
 a. 'En zelfs zonder aan hem voorgesteld te zijn *een gesprek* met hem *aanknoopten*.'
 b. 'And even without having been introduced to him, struck up a conversation with him.'

Means of expression of this type, where ingressivity is subtly lexicalized, are the most difficult to identify using introspection because they are not very frequent. Indeed, all of them are hapaxes in the parallel corpus. This is where HTM is of clear added value. With a larger parallel corpus, there would likely be more lexicalized expressions 'discovered'.

6.1.5. Unmarked Cases

The largest category of translations for ingressive *-qilai* is unmarked. Sentences (34)–(36) are examples of these unmarked corpus items.

- (34) 屋顶的声音又响起来了。
Wuding de shengyin you xiang-qilai-le.
 roof SUB sound again make.sound-QILAI-PRF
 a. 'Het geluid op het dak is weer terug.'
 b. 'The sound on the roof is back again.'

- (35) 郭老摇头笑了起来。
 Guo lao yaotou xiao-le-qilai.
 Guo old shake head laugh-PRF-QILAI
 a. 'Hij *lachte* hoofdschuddend.'
 b. 'He laughed, shaking his head.'
- (36) 他突然沉默起来,
 Ta turan chenmo-qilai.
 he suddenly silent-QILAI
 a. 'Opeens *zweeg* hij.'
 b. 'Suddenly he became silent.'

In (34), the locus of ingressivity is the stative predicate *is terug* 'is back', in (35) the durative and atelic predicate *lachte* 'laughed', and in (36), the punctual and telic predicate *zweeg* 'became silent'. Their unmarked ingressive interpretations in Dutch appear to be licensed given the right context, and provided that it is clear what the initial boundary should roughly look like. However, with the exception of (36), which independently features an adverb suggesting ingressivity (i.e., 突然 *turan* 'suddenly'), these items could also receive non-ingressive readings where 'being back' and 'laughing' are simply ongoing. It thus seems that, at least in the parallel corpus, Dutch is more often underspecified for ingressivity than Mandarin.

6.2. Continuative/Resumptive -*xiaqu*

The two types of viewpoint aspect encoded by *-xiaqu*—continuative and resumptive—were treated separately for the purposes of annotation (cf. Sections 4 and 5). Table 8 lists the translation strategies for its continuative sense, according to the same format as *-qilai*: abstracted supercategories subdividing into subcategories. Contrary to the ingressive overview, the most frequent strategies are in fact not known from previous work (Section 3.2): the verbal prefixes *door-* 'through' and *verder-* 'further', which make up more than half of all items (54.3%), are thus an aspectual 'discovery' in this sense. The verbs from the continuative verb set (*blijven* and *houden*: cumulatively, 23.9%) are represented, but not with the types of PP-complement discussed in the literature. There are a few adverbial strategies (10.9%), and just one unmarked case (2.2%). Another prefix not known from the literature—*voort-* 'forward'—also occurs, but only once (2.2%). Conspicuously absent from the corpus with previous work in mind is the *en maar*-construction (cf. Section 3.2); this is likely because it is a colloquial construction not or rarely found in the literary genre (although it could have popped up in embedded dialogue).

For resumptive *-xiaqu*, the quantitative picture in Table 9 is largely but not entirely different. The prefixes *verder-* and *door-* are both represented, but in a rather different distribution: *verder-* (65.4%) is almost three times as frequent ($65.4/23.9 = 2.7$), and there is only one instance of resumptive *door-* (3.8%). A chi-squared test (with Yates correction for small sample sizes) shows that the difference in the distribution of continuative and resumptive interpretations over *door-* and *verder-* is significant ($\chi^2 = 9.61$, $df = 1$, $p = 0.002$).¹⁶ I will return to this finding in Section 6.2.5. *Voort-* 'further' occurs once here too, and there is one instance of a verbal strategy: *vervolgen* 'proceed' (both 3.8%). A notable strategy for resumptivity is the use of an ingressive verb (*gaan*, *beginnen*) combined with the adverb *weer* 'again' (19.2%). This aligns with the idea voiced in Section 4.4 that resumptivity basically constitutes a special type of ingressivity embedded in a more complex viewpoint. A final observation is that there are zero cases of untranslated or unmarked resumptivity.

Table 8. Strategies for translating continuative *-xiaqu* to Dutch.

Supercategories			Subcategories		
Translation	n	%	Translation	n	%
<i>door-</i> ‘through’	14	30.4%	doorgaan (te/met V _{INF})	9	19.6%
			door-V	5	10.9%
<i>verder-</i> ‘further’	11	23.9%	verder V	10	21.7%
			verdergaan met V _{INF}	1	2.2%
<i>blijven</i> ‘stay’	10	21.7%	blijven V _{INF}	8	17.4%
			blijven door-V _{INF}	1	2.2%
			aan-blijven	1	2.2%
ADVERBIAL STRATEGY	5	10.9%	<i>aan een stuk door</i> ‘continuously’	2	4.3%
			<i>steeds</i> ‘continually’	1	2.2%
			<i>nog langer</i> ‘even longer’	1	2.2%
			<i>in één adem</i> ‘in one breath’	1	2.2%
UNTRANSLATED	3	6.5%	UNTRANSLATED	3	6.5%
UNMARKED	1	2.2%	UNMARKED	1	2.2%
<i>houden</i> ‘keep’	1	2.2%	uit-houden	1	2.2%
<i>voort-</i> ‘forward’	1	2.2%	voort-V	1	2.2%
Total	46	100%	Total	46	100%

Table 9. Strategies for translating resumptive *-xiaqu* to Dutch.

Supercategories			Subcategories		
Translation	n	%	Translation	n	%
<i>verder-</i> ‘further’	17	65.4%	verder V	14	53.9%
			verdergaan met V _{INF}	3	11.5%
INGRESSIVE VERB + ADVERBIAL STRATEGY	5	19.2%	<i>weer gaan</i> ‘go again’	3	11.5%
			<i>weer beginnen</i> ‘begin again’	2	7.7%
OTHER VERBAL STRATEGY	2	7.7%	<i>vervolgen</i> ‘proceed’	2	7.7%
<i>door-</i> ‘through’	1	3.8%	doorgaan met V _{INF}	1	3.8%
<i>voort-</i> ‘forward’	1	3.8%	voort-V	1	3.8%
N	26	100%	N	26	100%

In what follows, the markers from Tables 8 and 9 will be discussed in order of frequency: prefixes (Section 6.2.1), verbs from the continuative set (Section 6.2.2) and adverbial and other verbal strategies (Section 6.2.3). The section concludes with a discussion of the most frequent markers from Tables 8 and 9 in terms of their overlap and specialization with regard to continuativity and resumptivity.

6.2.1. Prefixes: *Verder-*, *Door-*, *Voort-*

The most striking ‘discovery’ of the present HTM exercise is the prevalence of prefixes not mentioned in the literature on Dutch viewpoint aspect: *verder-* ‘further’, *door-* ‘through’ and *voort-* ‘forward’. *Verder-* and *voort-* occur in both continuative and resumptive interpretations. Sentence (37) exemplifies a continuative interpretation of *verder-*, (38) a resumptive one. Formally, *verder-* may either select a main verb (e.g., *verder leven* ‘keep living’ in (37)) or occur in the combination *verdergaan met* (e.g., *verdergaan met praten* ‘resume talking’ in (38)). *Voort-* seems to be more limited in what it attaches to, but the precise restrictions are not clear from the parallel corpus.

- (37) 有一次我听到女的一边哭泣一边说，再也不愿意和他这个穷鬼过下去了，
 Youyici wo tingdao niude yibian kuqi yibian shuo, zaiyebu
 once I hear woman one.side weep one.side speak not.any.more
 yuanyi he ta zhe-ge qiong gui guo-xiaqu-le.
 wish with him this-CLF poor ghost live-XIAQU-PRF
 a. 'Op een keer hoorde ik haar huilend zeggen dat ze niet meer verder wilde leven met een armoedzaaier zoals hij.
 b. 'One time I heard her cry and say that she didn't want to keep on living with a bum like him.'

- (38) 我上了炕还得说下去。
 Wo shang-le kang hai dei shuo-xiaqu.
 I upon-PRF kang still must speak-XIAQU
 a. 'Als ik op de kang lig wil ik verdergaan met praten.'
 b. 'When I lay down on the kang I want to resume talking.'

Formally, the prefix *door-* mirrors *verder-*, occurring attached to a main verb or in the combination *doorgaan met*. *Door-* has a clear preference for continuativity, with only one attestation being resumptive. Both interpretations are illustrated below: (39) continuative, (40) resumptive.¹⁷

- (39) 他们再那样子谈下去，我就要下车了。
 Tamen zai na yangzi tan-xiaqu, wo jiu yao xiache-le.
 they again that way chat-XIAQU I just want get off-PRF
 a. 'Als zij in deze trant *door*praten, wil ik eruit.'
 b. 'If they keep on talking in this way, I want to get off.'
- (40) 她继续唱下去。
 Ta she jixu chang-xiaqu.
 she proceed sing-XIAQU
 a. 'Ze gaat *door* met zingen.'
 b. 'She resumes singing.'

Further research into these prefixes is desirable, since they have received little treatment in previous work—much in favor of the verb *blijven*, which is up next.

6.2.2. Verbs from the Continuative Set: *Blijven*, *Houden*

In previous work (Section 3.2), *blijven* 'stay' was associated with continuativity, and in various syntagms: with an infinitive, a PP headed by *aan* 'on', and as a copula with a nominal or adjectival complement. It is, however, only attested in the corpus with an infinitive, and it is overall less frequent than two of the prefixes discussed before (*door-* and *verder-*). More strikingly, *blijven* is restricted to continuative (i.e., non-resumptive) interpretations, which is not made explicit in previous work. All of this suggests that *blijven* occupies a more central position in the Dutch viewpoint aspect literature than is empirically justified.

Blijven's causative counterpart *houden* 'keep' is likewise restricted to non-resumptive continuativity. *Houden* is attested once, also not with the type of complement mentioned in the literature (*aan*-PP), but with a prefix that allows it to be used non-causatively (i.e., *uithouden* 'endure'). This attestation is shown in (41).

- (41) “我们撑船人可以这样子活下去，你们舱客就不能吗？”
 “Women chengchuan ren keyi zheyangzi huo-xiaqu, nimen
 we punt.boat people can like.this live-XIAQU you.PL
 cang ke jiu bu neng ma?”
 cabin customer just not can QST
 a. 'Als wij vaargasten het zo kunnen *uithouden*, kunnen jullie passagiers dat toch ook?'
 b. 'If us sailors can endure it like this, you passengers should be able to as well, right?'

6.2.3. Adverbial and Other Verbal Strategies

Continuative and resumptive viewpoint aspect are realized in the translations by distinct adverbial and/or other verbal strategies. For continuativity, there are four adverbs in the corpus, all of which boil down to 'not stopping during the situation': *aan een stuk door* 'continuously', *steeds* 'continually', *in één adem* 'in one breath', and *nog langer* 'even longer', which is illustrated in (42).

- (42) 再盯下去我就变成姨太太了。
 Zai ding-xiaqu wo jiu biancheng yitaitai-le.
 again stare-XIAQU I just change concubine-PRF
 a. 'Als je me *nog langer* aanstaart zal ik in de concubine veranderen.'
 b. 'If you stare at me me any longer I'll transform into the concubine.'

For resumptivity there is one case of this type of strategy, but a notable one: the use of an ingressive verb (*beginnen* or *gaan*—cf. Section 6.1.1) with an adverb indicating the return of a former situation (i.e., *weer* ‘again’). This speaks for the analysis of resumptivity as a special type of ingressivity with additional complexity; precisely this complexity (i.e., the situation held previously) is contributed by *weer*. Corpus item (43) illustrates this.

- (43) 她把钱收到裙子口袋里，推起她的车子，顶着风，吃力地行走下去，
Ta *jiang* *qian* *shoudao* *qunzi* *koudai-li*, *tui* *qi* *ta*
 she take money receive skirt pocket-in push rise she
de *Chezi*, *ding-zhe* *Feng*, *chili* *de* *xingzou-xiaqu*.
 sub cart go.against-IMP wind strain SUB walk-XIAQU
 a. 'Ze stak het geld weg in haar rok en begon weer moeizaam haar kar tegen de wind in te duwen.'
 b. 'She put the money away in her skirt and started to push her cart again, with difficulty, against the wind.'

6.2.4. Unmarked Case

There is one unmarked continuative case in the corpus, and zero unmarked resumptive ones. How can continuativity go unmarked? Given the conceptual structure of continuativity schematized in Section 4.4, there are at least two co(n)textual prerequisites: (i) the situation may not have been interrupted and (ii) the continuative (i.e., ‘not stopping’) boundary must somehow be salient, otherwise it would be necessary to mark it. The one unmarked case, shown in (44), indeed meets these criteria: the situation—i.e., *活huo* ‘live’—clearly cannot be interrupted, and the sentences preceding it specify that anaphoric *er* ‘it’ refers to afflictions of old age.¹⁸ Hence, a kind of abstract ‘not stopping’ boundary can be understood from the transition from not suffering to suffering from these afflictions.

- (44) 要不他们怎么活下去。
Yaobu *tamen* *zenme* *huo-xiaqu*.
 otherwise they how live-XIAQU
 a. 'Hoe viel er anders mee te leven?'
 b. 'How else would they be able to live with it?'

Generally, though, continuativity is basically always marked in the Dutch translations—contrary to ingressivity. Section 6.3 will speculate on why this might be.

6.2.5. Specialization and Overlap in Continuativity and Resumptivity

Aspectual *-xiaqu* covers both continuativity and resumptivity. The most frequent Dutch translation strategies (*verder-*, *door-*, *blijven*) differ in this respect. To get a better grip on this variation, Table 10 shows the frequency of the Dutch strategie encoding continuativity or resumptivity.

Table 10. Frequency of *verder-*, *door-* and *blijven* encoding continuative/resumptive aspect in the corpus.

Marker	Continuative		Resumptive		Total	
	n	%	n	%	n	%
<i>verder-</i> ‘further’	11	39.3%	17	60.7%	28	100%
<i>door-</i> ‘through’	14	93.3%	1	6.7%	15	100%
<i>blijven</i> ‘stay’	10	100%	0	0%	10	100%

Table 10 makes it clear that Dutch *verder-* covers both types, although it prefers resumptivity. As shown at the start of Section 6.2, this association is statistically significant. The other two markers, on the other hand, have a clear preference for continuativity: *door-* gets a resumptive interpretation only once, as opposed to 14 continuative interpretations; *blijven* is restricted to continuativity.

So contrary to Mandarin *-qilai* and the Dutch ingressive strategies, which differed chiefly in (non-)marking (the Dutch sentences often being underspecified for ingressivity), the Mandarin and Dutch continuative strategies vary in a different way: where *-xiaqu* takes care of both continuativity and resumptivity, Dutch has a specialized marker for continuativity and two markers with a strong preference for just one of the two (*door-* for continuativity, *verder-* for resumptivity).

6.3. Marking Frequency and Aspectual Complexity

HTM with at least two heuristic elements makes it possible to compare marking frequency. As the previous sections showed, there are considerable differences in the distribution of (un)marked translations over the viewpoint aspect markers: *-qilai* was marked less frequently than *-xiaqu*, and there was also a slight difference between *-xiaqu*'s continuative and resumptive readings. This comparison can be enhanced with marking frequencies from previous HTM research. Bogaards (2019a) applied HTM using the Mandarin “imperfective” verbal marker 着 *-zhe*, using a similar novel/translation-based corpus and equivalent criteria for ‘unmarkedness’ and ‘untranslatedness’. Table 11 shows marking frequency for not only the markers used in the present study, but also imperfective *-zhe*, taken from Bogaards (2019a, p. 174).

Table 11. Distribution of marked/unmarked translations over translated aspect markers¹⁹.

Type of Aspect	Marked		Unmarked		Total	
	n	%	n	%	n	%
Ingressive	274	61.4%	172	38.6%	446	100%
Continuative	42	97.7%	1	2.3%	43	100%
Resumptive	26	100%	0	0%	26	100%
Imperfective	127	26.4%	354	73.6%	481	100%

Now there is a threefold difference: in the Dutch translations, *-zhe*'s imperfectivity goes unmarked most of the time (73.6%), *-qilai*'s ingressivity is marked in the majority of items (61.4%) while *-xiaqu*'s continuativity is marked basically always (97.7%) and its resumptivity in all cases. Three separate chi-squared tests (with Yates correction) show that (i) the difference in marking between continuative and resumptive translations is not significant ($\chi^2 = 0.07$, $df = 1$, $p = 0.798$), (ii) the difference between continuative/resumptive and ingressive and ingressive translations is significant ($\chi^2 = 35.26$, $df = 1$, $p < 0.001$), (iii) as is the difference between ingressive and imperfective translations ($\chi^2 = 115.7$, $df = 1$, $p < 0.001$). Thus, the marking hierarchy is imperfective > ingressive > continuative/resumptive. How can these differences in marking be accounted for? A straightforward explanation would be that Dutch grammaticalizes these notions to different degrees. But there may be a more general force at play here. I will put forward a tentative explanation for the marking differences in Table 11.

As was claimed in Section 4.4, the ingressive, continuative and resumptive viewpoints vary in aspectual complexity. That is, the basic building blocks (boundaries τ and phases ϕ) required to model each viewpoint differ in both number and organization. As the conceptual structures in Figures 1–3 make clear, ingressivity is made up out of three building blocks ($\neg\phi$, τ , ϕ), continuativity out of five ($\neg\phi$, τ_1 , ϕ_1 , τ_2 , ϕ_2) and resumptivity out of seven ($\neg\phi$, τ_1 , ϕ_1 , τ_2 , $\neg\phi$, τ_3 , ϕ_2). The imperfective viewpoint, on the other hand, is simpler; its representation in these terms is ' ϕ ' along a temporal axis t . In this sense, the number of building blocks required for each viewpoint mirrors the degree to which they are marked in the translations: more complexity means more marking. I would therefore like to suggest that, in the domain of viewpoint aspect, complexity correlates with obligatory marking. Clearly, this claim requires more empirical evidence to be at all convincing; I am positing it here merely as a general hypothesis. HTM offers a way of testing this type of claim, as working with translations in this manner facilitates quantifying (non-)marking of conceptual content.

7. Conclusions

This study conducted a heuristic translation mining (HTM) analysis in the viewpoint-aspectual domain—specifically: ingressive, continuative and resumptive aspect—in Mandarin Chinese and Dutch. The aims were threefold: (i) methodological, to expand HTM by working with several and more specific heuristic elements; (ii) descriptive, to heuristically map out ingressivity, continuativity and resumptivity in a non-aspect-oriented language

like Dutch; and (iii) theoretical, to contribute to general research into crosslinguistic variation in encoding specific conceptual (aspectual) content.

Methodologically, the study drew on selection theories of aspect (e.g., Bickel 1997) to construct viewpoint-aspectual templates in terms of phases ϕ and boundaries τ , which were used to annotate the corpus data consistently and analyze them in a transparent way. A second finding having to do with method is that aspect markers at this level of specificity are in fact quite rare. In its current state of 29,668 parallel pairs and 611,232 characters, the corpus yielded 452 ingressive tokens, 56 continuative tokens and only 26 resumptive ones. The larger parallel corpus that is being assembled for the research project that this study is a part of is therefore by no means an unnecessary luxury. Third and last, the use of several distinct markers made it possible to compare marking frequency, which is not possible without some kind of quantifiable comparative method such as HTM.

Descriptively, the HTM analysis showed that the literature on Dutch viewpoint aspect has quite a solid grip on ingressivity, but not so much on continuativity (and not at all on resumptivity). Newly ‘discovered’ means of expression for ingressivity are:

- Other complements for the verbs from the ingressive set: various prepositional, nominal and adjectival complements;
- Other verbs taking a complement: *barsten* ‘burst’ and *schieten* ‘shoot’ selecting a PP with specific restrictions (e.g., *in lachen* or *in de lach* ‘in laughter’);
- Adverbial and other verbal strategies;
- Unmarked verbs.

For continuativity, the focus in previous work was on the verb *blijven* ‘stay’ and its causative counterpart *houden* ‘hold’. Because of that, the more frequent markers (in the present corpus) *door-* and *verder-* were overlooked, and resumptivity largely ignored. The new findings for these types of aspect in Dutch are:

- Prefixes *door-* ‘through’, *verder-* ‘further’ and *voort-* ‘forward’, with varying preferences for continuativity and resumptivity;
- Adverbial strategies for continuativity;
- Ingressive verb + ‘return to former situation’-adverb (e.g., *weer* ‘again’) for resumptivity.

These lists constitute testable hypotheses for the Dutch viewpoint-aspectual domain and as such provide direct input for further research in this field. That is to say, the hypotheses that these constructions encode ingressive/continuative/resumptive aspect may be specified by examining their (non-)aspectual semantics and restrictions on productivity. At a more general level, these lists also indicate where specific types of viewpoint aspect may be encoded in a non-aspect-oriented language: rather than in grammatical particles, they may be articulated by specific verbs, lexical and adverbial strategies, prefixes, and verb–adverb combinations. In the long run, this may also feed into linguistic theories on the crosslinguistic ordering of functional projections (e.g., cartographic approaches in the spirit of Cinque 1999); if one postulates that continuative viewpoint aspect occupies a fixed syntactic position crosslinguistically, for example, then Mandarin *-xiaqu* and Dutch *door-*, *verder-* and *voort-* present a specific case to support or challenge such an assumption. In this sense, these hypotheses are theoretically relevant for future research.

Theoretically, this study found a (non-)marking difference between Mandarin and Dutch regarding ingressivity (but not for continuativity/resumptivity). In the parallel corpus, Dutch left its verbs underspecified for ingressive aspect in 38.6% of cases. Of course, this finding is not very telling if the analysis is not reversed: the question is how frequently Mandarin translations of Dutch originals with, say, the marker *beginnen* are marked for ingressive aspect (and whether there are other ways of doing so than with *-qilai*). A more tangible finding is the difference in marking *across* viewpoint aspects within Dutch, especially in comparison to earlier (H)TM research into imperfectivity in Mandarin and Dutch: whereas imperfectivity was marked in only 26.4% of cases, ingressivity was marked 61.4% of the time and continuativity/resumptivity basically always. Based on this marking hierarchy, the tentative hypothesis was formulated that aspectual complexity (in terms of

selection theory) correlates with obligatory marking. HTM provides the opportunity to test this type of hypothesis quantitatively and without needing to rely on intuitions.

New (aspectual) discoveries can be made by extending HTM; doing so requires assembling larger parallel corpora and being specific about the conceptual content under study. Ironing out these methodological issues makes HTM a useful empirical tool for investigating the expression of viewpoint aspect across languages.

Supplementary Materials: The following are available online at www.mdpi.com/article/10.3390/languages7030158/s1.

Funding: This research was conducted as part of a PhD project funded by the Dutch Research Council (NWO), grant number PGW20.013.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original sentences featuring *-qilai* and *-xiaqu* and their translations from the parallel corpus, as well as the annotations, are available online at the following link: https://www.maartenbogaards.nl/resources/bogaards_2022_languages-corpus.xlsx.

Acknowledgments: I would like to thank three anonymous reviewers and the special issue editor for their challenging yet constructive criticism. I particularly appreciate Bert Le Bruyn's efforts in synthesizing the reviews and actively thinking along with me. Thanks to Ronny Boogaart for valuable comments on earlier versions of this work. Needless to say, all remaining errors are my own.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Notes

¹ Cf. Le Bruyn et al. (2022) for a meta-overview of these and more research traditions.

² In the Chinese linguistics literature, “inchoative” and “inceptive” refer to the beginnings of States and Events, respectively, i.e., a distinction motivated by the interaction with situation type. “Ingressive” is then used as a cover term for both, which is why I use it here. Cf. Bogaards (2020b, pp. 9–13) and Wierenga (2022, pp. 3–27) for discussion.

³ The analogy with linguistic surveys is not just metaphorical: a possible research method is to hire professional translators as actual ‘survey respondents’ to translate a specially selected set of items, as in Johansson (2007, chp.11).

⁴ These labels will be discussed in more detail in Section 5.

⁵ Non-narrative discourse corresponds largely to dialogue, except that narrative discourse may sometimes also show up in dialogue (Le Bruyn et al. 2019, p. 167 call this “mini-narrative discourse”).

⁶ This approach will be fleshed out in Section 4.4.

⁷ Dik and Hengeveld do not distinguish between two types of ‘continuing’: interrupted (“After stopping for a bit, John continued crying again”) and non-interrupted (“John continued crying for hours on end”). Their definition above only covers non-interrupted continuation. Section 4 will argue that these two types should be distinguished (following Xiao and McEnery 2004).

⁸ Dik and Hengeveld use the term “continuous” for (5b), but this term is usually reserved for an intermediate aspectual distinction below imperfectivity and above (non-)progressivity (cf. Comrie 1976, p. 25); I therefore prefer “continuative”.

⁹ Van Pottelberge (2004, p. 36) gives the example in (i), in which the modal verb *willen* ‘want’ yields an ingressive interpretation when combined with *aan het bouwen* ‘building’ (i.e., ‘want to start building’).

(i) Nu de Raad van State geen bezwaar heeft, wil het Zeeuwse nutsbedrijf zo snel mogelijk *aan het bouwen*. ‘Now that the Council of State has no objection, the utility company from Zeeland wants to start building as soon as possible.’

¹⁰ As one reviewer points out, *worden* has a broader semantics in that it can also construe a pre-phase to a change-of-state. For example, in springtime one can say *de bladeren worden groen* ‘the leaves are turning green’, which denotes a runup process to a final state of being entirely green.

¹¹ This paper uses the following abbreviations for glossing: CLF = classifier; IMP = imperfective aspect marker; PRF = perfective aspect marker; QST = question marker; RES = resultative marker; SUB = subordinating marker; SUG = suggestion marker.

¹² This seems to be related to aspectual *-xiaqu* being unable to license verbs with objects. To illustrate, (ii) and (iii)—which add the standard object 歌 *ge* ‘song’ to the predicate 唱 *chang* ‘sing’ in corpus item (i)—were both judged unacceptable by a native speaker informant.

- | | | | |
|-------|--|-----------------|----------------------------|
| (i) | 她继续唱下去。
Ta
she
'She continues singing.' | jixu
proceed | chang-xiaqu.
sing-xiaqu |
| (ii) | *她继续唱歌下去。
*Ta jixu changge-xiaqu. | | |
| (iii) | *她继续唱下歌去。
*Ta jixu chang-xia ge-qu. | | |

- 13 This is not to say that viewpoint-aspectual markers *only* select the types of predicates they prefer. 'Mismatches' between the two dimensions may be exploited for aspectual coercion, resulting, for instance, in a shift in situation type (cf. de Swart 1998).
- 14 One reviewer raises the question how continuative and resumptive aspect are similar rather than different, given that these two viewpoints are mapped to the same marker in Mandarin. One possibility is that the focused phase ϕ_2 in both cases corresponds to $\neg(\neg\phi)$, albeit in different ways: for continuativity this would be the negation of a 'stopping' presupposition; for resumptivity it would be the negation of actual stopping (i.e., the intermediate $\neg\phi$).
- 15 This raises the question (as noted by a reviewer) whether these types of verbs are less well-suited to encoding ingressivity. Alternatively, it might be a matter of frequency. In any case, this question could be addressed by means of a non-parallel corpus.
- 16 All statistical tests in this paper were performed at <https://www.reuneker.nl/files/chi> (accessed on 16 June 2022).
- 17 The resumptivity of (40) was determined by looking at the preceding sentences, which translate to "I tell her to stop singing. [...] She tells me that this is the first song I taught her."
- 18 Translated to English, the three preceding sentences read: "People of that age usually respond rather slowly. If someone is so far gone that they don't realize they're annoying others, they're all the more annoying. But maybe that was their good fortune? How else would they be able to live with it?"
- 19 The total number of markers differs slightly from those in Tables 6, 8 and 9 because *untranslated* items were excluded here.

References

- Aijmer, Karin, Bengt Altenberg, and Mats Johansson, eds. 1996. *Languages in Contrast: Papers from a Symposium on Text-Based Cross-Linguistic Studies, Lund 4–5 March 1994*. Lund: Lund University Press.
- Anthonissen, Lynn, Astrid De Wit, and Tanja Mortelmans. 2019. (Inter)subjective uses of the Dutch progressive constructions. *Linguistics* 57: 1111–59. [CrossRef]
- Barlow, Michael. 2008. Parallel texts and corpus-based contrastive analysis. In *Current Trends in Contrastive Linguistics*. Edited by María de los Ángeles Gómez González, Lachlan Mackenzie and Elsa González Álvarez. Amsterdam: John Benjamins, pp. 101–21.
- Beekhuizen, Barend, Julia Watson, and Suzanne Stevenson. 2017. Semantic typology and parallel corpora: Something about indefinite pronouns. Paper presented at the 39th Annual Conference of the Cognitive Science Society, London, UK, July 26–29; Edited by Glenn Gunzelmann, Andrew Howes, Thora Tenbrink and Eddy Davelaar. Austin: Cognitive Science Society, pp. 112–17.
- Bertrand, Anne, Yurika Aonuki, Sihwei Chen, Joash Gambarage, Laura Griffin, Marianne Huijsmans, Lisa Matthewson, Daniel Reisinger, Hotze Rullmann, Raiane Salles, and et al. 2022. Nobody's Perfect. *Languages* 7: 148. [CrossRef]
- Bhat, Shankara. 1999. *The Prominence of Tense, Aspect and Mood*. Amsterdam: John Benjamins.
- Bickel, Balthasar. 1997. Aspectual scope and the difference between logical and semantic representation. *Lingua* 102: 115–31. [CrossRef]
- Binnick, Robert, ed. 2012. *The Oxford Handbook of Tense and Aspect*. Oxford: Oxford University Press.
- Bogaards, Maarten. 2019a. A Mandarin Map for Dutch Durativity: Parallel text analysis as a heuristic for investigating aspectuality. *Nederlandse Taalkunde* 24: 157–93. [CrossRef]
- Bogaards, Maarten. 2019b. Posture verbs combined with past participles in Dutch: Fixed or productive patterns? In *Linguistics in the Netherlands 2019*. Edited by Janine Berns and Elina Tribushinina. Amsterdam: John Benjamins, pp. 67–82. [CrossRef]
- Bogaards, Maarten. 2020a. Beyond Progressive Aspect: On Situational PPs, Boundary Effectuation and the Dutch Preposition *aan*. Master's thesis, Leiden University, Leiden, The Netherlands. Available online: <https://hdl.handle.net/1887/137598> (accessed on 16 June 2022).
- Bogaards, Maarten. 2020b. Manifestaties van Ingressiviteit: Ingressief, Inceptief en Inchoatief Aspect in het Mandarijn en pre-Tang-Chinees. Bachelor's thesis, Leiden University, Leiden, The Netherlands. Available online: <https://hdl.handle.net/1887/136109> (accessed on 16 June 2022).
- Bogaards, Maarten. Forthcoming. *Prospectief Aspect in het Nederlands. Over op het Punt Staan en Gerelateerde Constructies*. Leiden: Nederlandse Taalkunde.
- Bogaards, Maarten, Ronny Boogaart, and Sjeff Barbiers. Forthcoming. Bogaards, Maarten, Ronny Boogaart, and Sjeff Barbiers. Forthcoming. The syntax of progressive and ingressive *aan*-constructions in Dutch. In *Linguistics in the Netherlands 2022*. Edited by Jorrig Vogels and Sterre Leufkens. Amsterdam: John Benjamins.
- Bogaards, Maarten. Under review. Top-down vs. bottom-up approaches to aspect: The case of the Dutch prepositional progressive. Under review with *Journal of Germanic Linguistics*.
- Boogaart, Ronny. 1999. Aspect and Temporal Ordering: A Contrastive Analysis of Dutch and English. Ph.D. thesis, Vrije Universiteit, Amsterdam, The Netherlands. Available online: https://www.lotpublications.nl/Documents/027_fulltext.pdf (accessed on 16 June 2022).

- Boogaart, Ronny, and Maarten Bogaards. Forthcoming. Aspectualiteit. In *Elektronische Algemene Nederlandse Spraakkunst (e-ANS)*, 3rd ed. Edited by Maaïke Beliën. Amsterdam: Institute for the Dutch Language.
- Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Bremmers, David, Jianan Liu, Martijn van der Klis, and Bert Le Bruyn. 2021. Translation mining: Definiteness across languages. *Linguistic Inquiry*. Available online: https://doi.org/10.1162/ling_a_00423 (accessed on 16 June 2022). [CrossRef]
- Breu, Walter. 1994. Interactions between lexical, temporal and aspectual meanings. *Studies in Language* 18: 23–44. [CrossRef]
- Broekhuis, Hans, and Norbert Corver. 2015. *Syntax of Dutch: Verbs and Verb Phrases*. Amsterdam: Amsterdam University Press, vol. 2.
- Broekhuis, Hans, and Norbert Corver. 2017. The expressive *en maar*-construction. In *Crossroad Semantics: Computation, Experiment and Grammar*. Edited by Hilke Reckman, Lisa Cheng, Maarten Hijzelendoorn and Rint Sybesma. Amsterdam: John Benjamins, pp. 305–25.
- Broekhuis, Hans, Norbert Corver, and Riet Vos. 2015. *Syntax of Dutch: Verbs and Verb Phrases*. Amsterdam: Amsterdam University Press, vol. 1.
- Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Cherici, Alessia. 2019. A corpus-based study of Chinese RVC *qilai* and its interaction with state predicates. *Buckeye East Asian Linguistics* 4: 35–45.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads*. Oxford: Oxford University Press.
- Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Croft, William. 2012. *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.
- Cysouw, Michael, and Bernhard Wälchli. 2007. Parallel texts: Using translational equivalents in linguistic typology. *STUF Language Typology and Universals* 60: 95–99. [CrossRef]
- Dahl, Östen. 2007. From questionnaires to parallel corpora in typology. *STUF Language Typology and Universals* 60: 172–81. [CrossRef]
- Dahl, Östen. 2014. The perfect map: Investigating the cross-linguistic distribution of TAME categories in a parallel corpus. In *Aggregating Dialectology, Typology and Register Analysis*. Edited by Benedikt Szmrecsanyi and Bernhard Wälchli. Berlin: De Gruyter Mouton, pp. 268–89.
- Dai, Yaojing. 1997. *Xiandai Hanyu Shiti Xitong Yanjiu [A Study of Aspect in Modern Chinese]*. Hangzhou: Zhejiang Educational Press.
- de Swart, Henriëtte. 1998. Aspect shift and coercion. *Natural Language & Linguistic Theory* 16: 347–85.
- de Swart, Henriëtte. 2007. A cross-linguistic discourse analysis of the Perfect. *Journal of Pragmatics* 39: 2273–307. [CrossRef]
- de Swart, Henriëtte, Jos Tellings, and Bernhard Wälchli. 2022. Not ... Until Across European Languages: A Parallel Corpus Study. *Languages* 7: 56. [CrossRef]
- Dik, Simon, and Kees Hengeveld. 1997. *The Theory of Functional Grammar, Part 1: The Structure of the Clause*, 2nd ed. Berlin: De Gruyter Mouton.
- Granger, Sylviane, and Marie-Aude Lefer. 2020. A two-pronged approach to corpus-based crosslinguistic studies. *Languages in Contrast* 20: 167–83. [CrossRef]
- Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij, and Maarten van den Toorn, eds. 1997. *Algemene Nederlandse Spraakkunst*, 2nd ed. Groningen: Martinus Nijhoff.
- Hanegreefs, Hilde. 2004. The Dutch change-of-state copula *worden* and its Spanish counterparts: A matter of aspect and voice. *Annual Review of Cognitive Linguistics* 2: 1–29. [CrossRef]
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge: MIT Press.
- Johansson, Stig. 2007. *Seeing Through Multilingual Corpora*. Amsterdam: John Benjamins.
- Kang, Jian. 1999. The Composition of the Perfective Aspect in Mandarin Chinese. Ph.D. thesis, Boston University, Boston, MA, USA. Available online: <https://www.proquest.com/docview/304493825> (accessed on 16 June 2022).
- Knotková, Magdaléna, and Wei-Lun Lu. 2020. Rendering, generalization and variation. On the use of multiple parallel texts as a comparative method in cognitive poetics. *Cognitive Linguistic Studies* 7: 201–21. [CrossRef]
- Koss, Tom, Astrid De Wit, and Johan van der Auwera. 2022. The aspectual meaning of non-aspectual constructions. *Languages* 7: 143. [CrossRef]
- Le Bruyn, Bert, Martijn van der Klis, and Henriëtte de Swart. 2019. The Perfect in dialogue: Evidence from Dutch. In *Linguistics in the Netherlands 2019*. Edited by Janine Berns and Elina Tribushinina. Amsterdam: John Benjamins, pp. 162–75. [CrossRef]
- Le Bruyn, Bert, Martin Fuchs, and Martijn van der Klis. 2022. Parallel corpus research and target language representativeness: The contrastive, typological and translation mining traditions. *Languages* 7. Proof.
- Li, Charles, and Sandra Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Lieber, Rochelle, and Harald Baayen. 1997. A semantic principle of auxiliary selection in Dutch. *Natural Language and Linguistic Theory* 15: 789–845. [CrossRef]
- Lu, Wei-lun, and Arie Verhagen. 2016. Shifting viewpoints: How does that actually work across languages? An exercise in Parallel Text Analysis. In *Viewpoint and the Fabric of Meaning*. Edited by Barbara Dancygier, Wei-lun Lu and Arie Verhagen. Berlin: De Gruyter Mouton, pp. 169–90.

- Lu, Wei-lun, Arie Verhagen, and I-wen Su. 2018. A multiple parallel-tekst approach for viewpoint research across languages: The case of demonstratives in English and Chinese. In *Expressive Minds and Artistic Creations: Studies in Cognitive Poetics*. Edited by Szilvia Csábi. Oxford: Oxford University Press, pp. 131–57.
- Oostdijk, Nelleke, Martin Reynaert, Véronique Hoste, and Ineke Schuurman. 2013. The construction of a 500-million word reference corpus of contemporary written Dutch. In *Essential Speech and Language Technology for Dutch: Results by the STEVEN-Programma*. Edited by Peter Spyns and Jan Odijk. Dordrecht: Springer, pp. 219–47.
- Ramchand, Gillian. 2018. *Situations and Syntactic Structures*. Cambridge: MIT Press.
- Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6: 199–271. [CrossRef]
- Shyu, Shu-ing, Yu-Fan Wang, and Zhi-jie Lin. 2013. An approximation to secondary predication structure: A case of V-qilai in Mandarin Chinese. *Language and Linguistics* 14: 701–36.
- Smith, Carlota. 1997. *The Parameter of Aspect*, 2nd ed. Dordrecht: Springer.
- Travis, Lisa. 2010. *Inner Aspect*. Dordrecht: Springer.
- van der Klis, Martijn, and Jos Tellings. 2022. Multidimensional scaling and linguistic theory. *Corpus Linguistics & Linguistic Theory*. Advance online publication. [CrossRef]
- van derKlis, Martijn, Bert Le Bruyn, and Henriëtte de Swart. 2017. Mapping the PERFECT via Translation Mining. Paper presented at the 15th Conference of the European Chapter of the Association for Computational Linguistics, Valencia, Spain, April 3–7; Edited by Mirella Lapata, Phil Blunsom and Alexander Koller. Valencia: Association for Computational Linguistics, vol. 2, pp. 497–502.
- van der Klis, Martijn, Bert Le Bruyn, and Henriëtte de Swart. 2020. *Temporal Reference in Discourse and Dialogue*. Manuscript submitted. Available online: <https://time-in-translation.hum.uu.nl//media/news/2020/01/21/draft-hp.pdf> (accessed on 16 June 2022).
- van der Klis, Martijn, Bert Le Bruyn, and Henriëtte de Swart. 2021. A multilingual corpus study of the competition between PAST and PERFECT in narrative discourse. *Journal of Linguistics* 58: 423–457. [CrossRef]
- Van Pottelberge, Jeroen. 2004. *Der am-Progressiv: Struktur und Parallele Entwicklung in den Kontinentalwestgermanischen Sprachen*. Tübingen: Narr.
- Vandevoorde, Lore, Gert de Sutter, and Koen Plevoets. 2016. On semantic differences between translated and non-translated Dutch. In *Empirical Translation Studies: Interdisciplinary Methodologies Explored*. Edited by Ji Meng. Sheffield: Equinox, pp. 128–46.
- Vendler, Zenon. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.
- Verhagen, Arie. 2012. Construal and Stylistics—Within a language, across contexts, across languages. Paper presented at Stylistics Across Disciplines Conference, Leiden, The Netherlands, June 18; Available online: <https://openaccess.leidenuniv.nl/handle/1887/20133> (accessed on 16 June 2022).
- Verkerk, Annemarie. 2014. Where Alice fell into: Motion events from a parallel corpus. In *Aggregating Dialectology, Typology and Register Analysis*. Edited by Benedikt Szmrecsanyi and Bernhard Wälchli. Berlin: De Gruyter Mouton, pp. 324–54.
- Wang, Arthur, and Hung-Hsi Wu. 2018. The syntax and semantics of Chinese impersonal V-qilai construction. *Taiwan Journal of Linguistics* 16: 1–34.
- Wiedenhof, Jeroen. 2016. *A Grammar of Mandarin*. Amsterdam: John Benjamins.
- Wierenga, Roné. 2022. Inchoatiewe niehoofwerkwoordkonstruksies in Afrikaans: 'n Korpusondersoek. Master's thesis, Noordwes-Universiteit, Potchefstroom, South Africa.
- Wu, Xiu-Zhi Zoe. 2000. Grammaticalization and the Development of Functional Categories in Chinese. Ph.D. thesis, University of Southern California, Los Angeles, CA, USA. Available online: <https://www.proquest.com/docview/304627096> (accessed on 16 June 2022).
- Xiao, Richard, and Tony McEnery. 2004. *Aspect in Mandarin Chinese: A Corpus-Based Study*. Amsterdam: John Benjamins.
- Xiong, Jiajuan. 2017. *Chinese Middle Constructions: Lexical Middle Formation*. Singapore: Springer Nature.
- Yang, Suying. 1995. The Aspectual System of Chinese. Ph.D. thesis, University of Victoria, Victoria, BC, Canada. Available online: https://dspace.library.uvic.ca/bitstream/handle/1828/9443/Yang_Suying_PhD_1995.pdf (accessed on 16 June 2022).
- Zhao, Ruoying. 2022. Decomposing Perfect Readings. *Languages* 7. Revision.