

De Cuypere, Ludovic, Kristof Baten & Gudrun Rawoens. 2014.
A corpus-based analysis of the Swedish passive alternation. *Nordic Journal of Linguistics* 37(2), 199–223.

A corpus-based analysis of the Swedish passive alternation

Ludovic De Cuypere, Kristof Baten & Gudrun Rawoens

This article presents a corpus-based investigation of the motivations behind the use of the *s*-passive and the *bli*-passive in contemporary written Swedish. Following a probabilistic approach to language variation and building on observations in the literature, we examine the simultaneous effects of semantic and syntactic factors by means of a multivariate statistical analysis. Our corpus sample includes 1,197 passive sentences of three Swedish verbs, which alternate in their use of the passive (*acceptera* ‘accept’, *behandla* ‘treat’ and *välja* ‘choose’). The results suggest that the choice of passive form is significantly associated with five predictor variables: Subject Animacy, Subject Number, Modal Verb, *Av*-phrase (‘by’-phrase) and Aktionsart. Among these, Subject Animacy and Modal Verb appear to yield the strongest impact effect on the choice of passive form. The study adds to earlier research in that it allows for a more accurate analysis of the simultaneous effect and relative strength of each factor on the speaker’s choice of one of the passive forms.

Keywords morphological vs. periphrastic passive, multivariate statistical analysis, passive alternation, Swedish

Ludovic De Cuypere, Ghent University, Linguistics Department, Blandijnberg 2, B-9000 Ghent, Belgium. Ludovic.DeCuypere@UGent.be

Kristof Baten, Ghent University/Research Foundation Flanders (FWO-Vlaanderen), Linguistics Department, Blandijnberg 2, B-9000 Ghent, Belgium. Kristof.Baten@UGent.be

Gudrun Rawoens, Ghent University, Linguistics Department, Scandinavian Linguistics, Rozier 44, B-9000 Ghent, Belgium. Gudrun.Rawoens@UGent.be

1. INTRODUCTION

The three Mainland Scandinavian languages, Danish (DA), Norwegian (NW) and Swedish (SW),¹ have two basic ways of forming passive sentences: one morphological, by adding an *s*-suffix to the verb (as in (1a)), and one periphrastic, which involves an auxiliary, typically *bli*(*ve*), and the past participle of a main verb (as in (1b)) (see *Grammatik over det Danske Sprog* (Hansen & Heltoft 2011), *Norsk referansegrammatikk* (Faarlund, Lie & Vannebo 1997), *Svenska Akademiens grammatik* (Teleman, Hellberg & Andersson 1999, henceforth SAG).

(1) *s-passive and bli-passive of 'the egg is boiled'*

- a. DA: Ægg-et **koge-s**.
 NW: Egg-et **koke-s**.
 SW: Ägg-et **koka-s**.
egg-DEF boil-PASS
- b. DA: Ægg-et **bliver kogt**.
 NW: Egg-et **blir kokt**.
 SW: Ägg-et **blir kokat**.
egg-DEF becomes boiled

Both passive forms are productively used in everyday language, albeit to different extents and in a different distribution among the three languages (Laanemets 2012). The use of either one of the two structures appears to depend on a variety of lexical, semantic, syntactic and pragmatic factors (Engdahl 1999, 2001, 2006; Enger 2001; Laanemets 2004, 2009, 2012).

Traditional grammar descriptions (Mikkelsen 1911, Hansen & Heltoft 2011 for Danish; Western 1921, Hansen 1991, Faarlund et al. 1997 for Norwegian; Thorell 1973, Teleman et al. 1999 for Swedish) and previous single-study research (e.g. for Danish, Heltoft 1994, Heltoft & Falster Jakobsen 1996) have provided a set of distinctions to explain the differences in use between the two passives (see also Laanemets 2012).

The following distinctions have been proposed to account for the different uses in Danish and, to a lesser extent, Norwegian. The early grammarians (Mikkelsen 1911, Western 1921) linked the distinction of the two passives to Aktionsart. In their view, the morphological passive is preferred to describe ongoing events, whereas the periphrastic passive mostly describes the completion of events or the inception of events. Heltoft & Falster Jakobsen (1996) argued against this, claiming that Aktionsart only secondarily determines the use of either one of the passive forms (see also Engdahl 1999, Enger 2001). In these latter accounts the morphological passive is related to rules/commands, and to general events or norms which exist independently and objectively of the speaker. The periphrastic passive is associated with specific events or specific rules, which describe a subjective mood of the speaker (Heltoft & Falster Jakobsen, 1996). To illustrate these different uses, Heltoft & Falster Jakobsen (1996:203) give the following examples in (2) and (3).²

- (2) a. Inledningen **skrives** til sidst. (Danish)
Introduction.DEF write-PASS at last
 'The introduction is to be written last.'
- b. Der **tales** ikke mere dansk i Skåne.
there speak.PASS no more Danish in Scania
 'Danish is no longer spoken in Scania.'

- (3) a. Inledningen **bliver skrevet** til sidst. (Danish)
introduction.DEF becomes written at last
 'The introduction will be written last.'
- b. Der **bliver ofte talt** dansk i Skåne.
there becomes often spoken Danish in Scania
 'Danish is very often spoken in Scania.'

The use of the *s*-passive in (2) is said to indicate a rule or a general norm that exists independently of the speaker, whereas the use of the *bli*-passive in (3) appears to signal a specific event or rule that is subjectively anchored.

However, these straightforward distinctions cannot readily be extrapolated to Swedish (Enger 2001, Engdahl 2006). One major difference is that Danish and Norwegian only allow for morphological passivization with simplex tenses and that strong verbs can only take the morphological passive form in the present tense. Apart from these two grammatical restrictions, different semantic differences seem to apply as well (for Danish see Zola Christensen & Christensen 2005:122–124; Hansen & Heltoft 2011:Chapter V:651; and for Norwegian see Faarlund et al. 1997:513).

With respect to Swedish, several factors have been suggested in the literature to account for the choice of either one of the passive forms. However, it is generally acknowledged that no single factor fully determines the use of the two passive forms. Building on the probabilistic grammar approach to syntactic alternations (see Bod, Hay & Jannedy (2003) for a collection of studies working in this framework and Bresnan & Ford (2010) for an application of this approach to the English dative alternation), we hypothesize that the passive alternation is simultaneously motivated by multiple factors.

Using a corpus sample of $N = 1,197$ passive sentences, we examined the influence of the following six factors that have been suggested in the literature: Subject Animacy, Subject Person and Subject Number, the presence/absence of a Modal Verb and an *Av*-phrase ('by'-phrase), and Aktionsart. A logistic regression analysis was used to evaluate the effect of these factors on the use of these two passive forms.

The paper is structured as follows. Section 2 describes the Swedish passive forms and some possible factors that may influence which passive form is chosen. Section 3 outlines our corpus design, data annotation based on the factors and statistical method. The results are presented in Section 4. Section 5 discusses the results and concludes with some suggestions for future research.

2. THE SWEDISH PASSIVE ALTERNATION

This section focuses on the alternation between the morphological *s*-passive and the periphrastic *bli*-passive ('become'-passive) in Swedish. The *vara*-passive

(‘be’-passive) and *få*-passive (‘get’-passive) are not discussed because the present study is only concerned with the processual passive and not with the static passive as reflected by the two auxiliaries mentioned (see e.g. SAG IV:393, Larsson 2012). Processual passives emphasize the action/event, whereas static passives focus on the condition/state (see e.g. the examples mentioned in SAG IV:393: *Patienten blev hela tiden övervakad av vår personal* vs. *Patienten var hela tiden övervakad av vår personal* ‘The patient is continuously monitored by our staff’ with focus on the process and the situation, respectively). Moreover, most of the previous studies on the Swedish passive alternation which we relate to, have paid most attention to the choice between the *bli*-passive and the *s*-passive and little to no attention to the static passive forms. Furthermore, it should be noted that to the best of our knowledge only few studies are available on the Swedish passive alternation.

The *s*-passive is formed by adding *-s* to the verb form, which may additionally be inflected for tense. In contrast, the *bli*-passive is formed by adding the copula *bli* ‘become’ to the participle of the sentence verb, which agrees with the subject in number and gender. The sentences in (4), taken from SAG (IV:360), show an active sentence, the *s*-passive and *bli*-passive, respectively.

- (4) a. Min bror lagade stolen redan i förra veckan.
my brother repaired chair.DEF already in last week.DEF
 ‘My brother repaired the chair already last week.’
- b. Stolen lagades av min bror redan i förra veckan.
chair.DEF repaired.PASS by my brother already in last week.DEF
 ‘The chair was repaired by my brother already last week.’
- c. Stolen blev lagad av min bror redan i förra veckan.
chair.DEF became repaired by my brother already in last week.DEF
 ‘The chair was repaired by my brother already last week.’

According to SAG, almost every transitive verb can occur in both passive forms. However, a quantitative analysis by Laanemets (2004) shows that the *s*-passive is by far the most frequent alternate and that it has a wider distribution in terms of different tenses and different verb types (see also Engdahl 2001:87). In Laanemets’ (2004) corpus of 562 passive sentences, 91% occurs with an *s*-passive and only 1% with a *bli*-passive (the remaining 8% sentences were *vara*-passives). The frequencies reported on in the large-scale corpus study in Laanemets (2012) show a similar distribution: 97% *s*-passives versus 3% *bli*-passives in the Swedish newspaper corpus containing 19.7 million words and 90.4% and 9.6% respectively in the corpus containing fiction data with almost 2.5 million words (the *få*- and *vara*-passives were not included in this study) (Laanemets 2012:92).

With regard to the wider distribution of the Swedish *s*-passive, Laanemets’ corpus study further reveals that the morphological passive can occur in all tenses (i.e., the present, past and perfect tense). This finding is notably different from

Danish and Norwegian, where the use of the *s*-passive is restricted to the present tense and simple past (Laanemets 2004:97f.).³ According to Engdahl (2006:29–30), the differentiation of the two passive forms along the lines of tense in Danish and Norwegian (i.e. *s*-passive for simple tenses and *bli*(*ve*)-passive for perfect tenses) fits well with the claim that general events are preferably expressed by using the *s*-passive and particular events by using the *bli*(*ve*)-passive. Indeed, general events are usually described by means of simple tenses and particular events by means of perfect tenses. However, this claim does not hold for Swedish, because the *s*-passive is found in all tenses.

Regarding the *s*-passive in Swedish, SAG additionally mentions that certain intransitive verbs can occur in the passive voice, but nearly always in the form of an *s*-passive, as is illustrated in the passivization of the active sentence *Folk pratade mycket vid mötena* ‘People talked a lot at the meetings’ (SAG IV:360).

- (5) a. Det **pratades** mycket vid mötena.
it talked.PASS much at meetings.DEF
- b. *Det **blev** **pratad** myck et vid mötena.
it became talked much at meetings.DEF

In addition to the association between intransitive verbs and *s*-passive, SAG also lists other preferences. For example, verbs with experiencer or recipient roles, such as *trodde* ‘was thought’ and *hörde* ‘was heard’, mostly take an *s*-passive (SAG IV:397, Engdahl 2001:88). These restrictions are, however, not connected with the influencing factors regarding the passive alternation. As Engdahl (1999:8) justifiably points out, ‘[i]n order for there to be a choice of passive forms, both the *bli*-passive and the *s*-passive must be available for a given lexical item’. Engdahl (1999) found, for example, that frequent verbs, such as *tala* ‘speak’, *stänga* ‘close’ and *öppna* ‘open’, never occur with the *bli*-passive in the PAROLE corpus.⁴ Engdahl (2006:36) adds that *bli*-passives with those verbs are not grammatically impossible, but just ‘somewhat odd’.

In order to be able to assess the effect of a number of influencing factors, it is imperative to examine verbs that occur in both passive forms. Engdahl (1999:24) provides a non-exhaustive list of ten verbs that meet this requirement: *acceptera* ‘accept’, *avbryta* ‘interrupt’, *behandla* ‘treat’, *bjuda* ‘invite’, *erbjuda* ‘offer’, *framföra* ‘present’, *mörda* ‘murder’, *råna* ‘rob’, *rädda* ‘save’ and *välja* ‘choose’. In the present study, we selected three verbs from the list to test the effect of the influencing factors (see below).

Previous research, some of which is already mentioned, has come up with a number of possible influencing factors, including syntactic, lexical and semantics ones (see e.g. Engdahl 2001). The syntactic factor put forward in Engdahl (2001:88)

is associated with the fact that there are two syntactic constructions which can only contain morphological passives and not periphrastic ones, i.e. constructions with raising verbs (e.g. *Räntan antas vara 10 procent* ‘The interest is assumed to be 10%’) and impersonal passives (e.g. *Det bjöds på kaffe* ‘Coffee was served’). Lexically, it appears that for certain verbs slightly different meanings are instantiated depending on which passive forms they take. For instance, the verb *anta* means ‘accept’ or ‘pass (a law)’ with the *s*-passive (see (6)), but implies a more specific meaning of ‘being admitted to an education program’ in the *bli*-passive (see (7)) (Engdahl 2001:87).

- (6) När deklarationen om kvinnors rättigheter **antogs** av FN.
when declaration.DEF about women's rights was.adopted by UN
 ‘When the declaration of women’s rights was adopted by the UN.’
- (7) Så Karin Johnsson **blev antagen** som första utländska elev.
so Karin Johnsson became admitted as first foreign student
 ‘So Karin Johnsson was admitted as the first foreign student.’

As for semantics, one factor is related to Aktionsart. The *s*-passive is said to be used to express ongoing and/or repeated events, whereas the *bli*-passive would be reserved to completed and/or temporary events. As a consequence, *bli*-passives seem to focus on the inception of a new state (8) or the result of a changed state (9), whereas the *s*-passive seems to focus on the event itself (10) (Engdahl 2006:27).

- (8) Läkarna **blev** på kort tid intensivt **avskydda** av hela den övriga
doctors.DEF became in short time intensely despised by entire the other
personalen.
staff.DEF
 ‘The doctors were intensely despised by all the other staff in no time.’
 (SAG IV:399)
- (9) Här **blir** man fort **glömd** om man inte ser till att synas.
here becomes one soon forgotten if one not look after to be.seen
 ‘One is soon forgotten if one does not ensure to be seen.’
 (SAG IV:399)
- (10) Centrala provet i matematik **skrivs** i årskurs tre i
central test.DEF in mathematics write.PASS in grade three in
gymnasieskolan.
secondary.school.DEF
 ‘The national test in mathematics is written in the third grade in secondary school.’
 (SAG IV:398)

A second semantic factor, suggested by Sundman (1983) and elaborated by Engdahl (2006:27), is associated with the subject of the passive. According to this account, the *bli*-passive is preferred when the subject of a passive sentence can control

or influence the situation to a certain degree. Following SAG, Engdahl additionally links this control over the situation with animacy, as is illustrated in (11), where the animate participant, ‘the representative’ controls the efforts s/he makes in trying to get reelected (taken from Engdahl 1999:7).⁵

- (11) a. Representanten försökte **bli omvald**.
representative.DEF tried become re-elected
- b. ?Representanten försökte **omväljas**.
representative.DEF tried re-elected.PASS
 ‘The representative tried to get reelected.’

Engdahl (2006) further claims that control of the subject can be implied even if the subject is not expressed, as in imperative sentences, as in (12).

- (12) **Bli** inte **rånad** i Chicago!
become not robbed in Chicago
 ‘Don’t get robbed in Chicago!’

Example (12) is interpreted as ‘avoid putting yourself in a situation where you are likely to get robbed’. If *bli*-passives are assumed to be associated with animate subjects that are in control over a certain event (either explicit or implied), then logically *s*-passives will presumably be used with inanimate subjects, as well as with impersonal subjects as in example (13) taken from Engdahl (1999:11).

- (13) a. Det **dracks** mycket öl den kvällen.
it drunk.PASS much beer that evening.DEF
- b. ?Det **blev drucket** mycket öl den kvällen.
it became drunk much beer that evening.DEF
 ‘A lot of beer was drunk that evening.’

In (13) the impersonal subject can obviously not exert any influence on the situation. In impersonal passives, present-day Swedish only allows for the *s*-passive (SAG IV:359–404; Engdahl 2006:24; Lyngfelt 2010:178), even though the use of *bli*-passives is possible in a few restricted cases (see Engdahl 2006:39; Rawoens & Johansson 2014).⁶ Clearly, personal subjects and impersonal subjects reflect different levels of control. In the same vein, local (first and second person) and non-local (third person) subjects, as well as singular and plural subjects, might reflect different levels of control, which in turn might have a different effect on the choice of either one of the passive forms. The latter differentiation according to person (local vs. non-local) and number (singular vs. plural) has hitherto not been looked at, but is examined in the present study.

Another motivating factor that has been suggested in the literature is the role of the expressed agent by means of an *Av*-phrase ('by'-phrase). In passive sentences the agent in the *av*-phrase can be omitted, either because the reference to the agent is unclear, unknown, or simply because the specific agent is unimportant or irrelevant to mention. After all, agent demotion is what characterizes passivization. In her sample of the PAROLE corpus, Engdahl (2006:37) found that the agent in the *av*-phrase is only expressed in about 10–15% of the passive sentences, however, without any clear differentiation between *s*-passives and *bli*-passives. Laanemets (2012:124ff.), on the other hand, found a preference for the *bli*-passive.

Finally, it has been observed that the passive alternation can be associated with the use of modal/temporal auxiliaries. This has been especially pointed out for Danish where the presence of a modal verb appears to influence the choice of the passive form, in that a subjective interpretation or prediction motivates the use of the *s*-passive, and a non-subjective reading the use of the *bli*-passive (Heltoft & Falster Jakobsen 1996, Hansen & Heltoft 2011).

(14) Kaninerne kan **blive spist** av ræven.
rabbits can become eaten by fox.DEF
 'The rabbits can/may be eaten by the fox.'

(15) Spidspus kan ikke **spises**.
shrews can not eaten.PASS
 'Shrews are inedible.'

According to Heltoft & Falster Jakobsen (1996:209) the use of the *bli*-passive in (14) is associated with a subjective interpretation, while the *s*-passive in (15) signals a non-subjective interpretation. With respect to Swedish, Engdahl (1999) observed some similar tendencies, but at the same time she found that the auxiliaries *kan* 'can' and *ska* 'shall/will' are actually more commonly followed by an *s*-passive than by a *bli*-passive (with a ratio of >1000 *s*-passives: around 60 *bli*-passives); the auxiliary *vill* 'want to', however, seems to be more commonly connected with *bli*-passives. The choice of the passive forms does not seem to be so strictly associated with a (non-)subjective interpretation as in Danish. Nevertheless, it seems to be the mere presence or absence of a modal/temporal auxiliary already influences the choice of the passive form used.

To sum up, previous scholarship on the passive alternation in Danish, Norwegian and Swedish has revealed several factors that seem to motivate the use of one of the two passive forms. However, a multifactorial analysis of the different factors based on a representative corpus sample of actual language data remains wanting. As a result, it remains unclear which factors are actually involved and how strong their respective influence is on the speaker's choice of options. The next section elaborates on the methodology that we used to answer these questions.

	<i>Acceptera</i>		<i>Behandla</i>		<i>Välja</i>
accepteras	bli(r) accepterad/t/de		behandlas	bli(r) behandlad/t/de	väljs/väljas vald/t/da

Table 1. Search terms for the verbs *acceptera* ‘accept’, *behandla* ‘treat’ and *välja* ‘choose’.

Verb	Passive form	
	<i>s</i> -passive	<i>bli</i> -passive
<i>acceptera</i> ‘accept’	200	200
<i>behandla</i> ‘treat’	200	199
<i>välja</i> ‘choose’	198	200

Table 2. Number of observations by passive form and verb.

3. METHODOLOGY

3.1 Corpus

Starting from the list of verbs provided by Engdahl (1999), we selected three verbs taking both passive forms: *acceptera* ‘accept’ (N = 400), *behandla* ‘treat’ (N = 399), *välja* ‘choose’ (N = 398). We specifically chose these three verbs and not the others from Engdahl’s list because they had the highest number of *bli*-passives in the ‘The Swedish Language Bank,’ SPRÅKBANKEN.⁷ To illustrate this: the *bli*-passive forms of *välja* (i.e. *blir vald/t/a*, *blev vald/t/a* etc.) yielded over 4,000 attestations, while the *bli*-passive forms of *döda* ‘kill’ (not included in our study) yielded only 198 observations.

Using Korp, we searched the written subcorpora of Språkbanken for the following forms in present tense and thus retrieved more than 50,000 attestations (see Table 1).⁸

Since it was practically unfeasible to annotate all attestations manually, we selected the first 200 *s*-passives and the first 200 *bli*-passives of each verb from the random list provided by KORP. Our final corpus sample consists of N = 1,197 attestations – three sentences, which turned out to be false positives, were removed. Using a case-control research design, we selected an (almost) equal proportion of *bli*- and *s*-passives for each verb, as indicated in Table 2.

The examined proportion of passive forms implies that we overestimate the number of *bli*-passives (recall that, in reality, the *s*-passive is much more frequent than the *bli*-passive). As a downside to this sample design, however, we cannot evaluate the effect of the particular verbs under analysis (notice that, given the

number of observations in Table 2, the relationship between verb and passive form is fully independent). As a consequence, no conclusions about the effect of the verbs will be drawn in this study.

3.2 Predictor variables

The next paragraphs discuss the operationalization of each factor under analysis and state the associated hypotheses.

3.2.1 Subject Animacy: ‘animate’ vs. ‘inanimate’

Humans, animals and organizations/institutions were regarded as animate. In (16) *landet* ‘the country’ refers to a country’s citizens and was regarded as ‘animate’. Inanimate subjects are expected to occur with the *s*-passive, and logically, animate subjects are expected to prefer the *bli*-passive.

- (16) Sudans regering måste inse att landet måste samarbeta med FN för att **bli** helt **accepterat** i det internationella samfundet.

‘The Sudanese government needs to understand that the country has to collaborate with the UN to become completely accepted in the international society.’

(Europarl svenska)

3.2.2 Subject Person: ‘local’, ‘non-local’, ‘other’

Local was defined as the first- or second-person (singular or plural), while non-local as the third person (singular or plural). The label ‘other’ was used for cases with no clear contextual reference. The sentence in (17) illustrates a third person singular inanimate subject of the *s*-passive.

- (17) Amöbadysenteri **behandlas** med Metronidazol.

‘Amoebic dysentery is treated with Metronidazol.’

(Svenska Wikipedia)

We hypothesize that local subjects will be associated with the *bli*-passive, given that local subjects are always animate, whereas non-local subjects involve both animate and inanimate arguments. This is in line with the hypothesis that the difference between local and non-local subjects is associated with different levels of control (see Section 2). When no referent was explicitly mentioned, as in (18), we relied on the context in order to label the subject. Note that the agreement on the past participle in this example reveals a reference to an animate referent.

- (18) Det är svårt att **bli** fullt **accepterad** i det thailändska samhället, menar hon.
 ‘She thinks it is difficult to become fully accepted in the Thai society.’

(GP2012)

3.2.3 Subject Number: ‘singular’ vs. ‘plural’

The presence or absence of a plural marker was used to determine the category of this variable. When the subject was ‘other’, a value for singular or plural is of course missing. We expect that the effect of this variable will be in line with that of Aktionsart (see below), in the sense that ongoing general events are often expressed through the use of quantitatively undefined or plural arguments (SAG IV:328, 337). We hypothesize that ongoing events are associated with the *s*-passive. The same relation is expected to apply when plural subjects are involved.

3.2.4 Modal Verb: ‘present’ vs. ‘absent’

The presence or absence of a modal/temporal auxiliary (*kan, ska, måste, vill*) was used to determine the category of this variable.⁹ This means that the factor Modal Verb does not involve a semantic/functional interpretation (in terms of subjective and non-subjective interpretation, see above), but a formal differentiation between the presence or absence of a modal verb. The presence of modal/temporal auxiliaries is expected to be associated with the use of an *s*-passive, an example of which is provided in (19), where the auxiliary *kan* ‘can’ precedes the *s*-passive. This expectation follows from Engdahl’s (1999) observation that Swedish does not strictly agree with the Danish pattern and instead uses a modal verb more frequently with *s*-passive than with *bli*-passive.

- (19) Men, tillägger han, en direkt avsikt att döda kan inte **accepteras**.
 ‘But, he adds, a direct intention to kill cannot be accepted.’

3.2.5 Av-phrase: ‘present’ vs. ‘absent’

The presence or absence of an agent expressed by means of an *av*-phrase was used to determine the category of this variable. First, we expect that *av*-phrases will only be expressed to a minimal extent, i.e. in all probability to the extent of 10–15% observed in Engdahl (2006:37). An example with an *av*-phrase is given in (20).

- (20) Ledamotens värdighet härrör från att han **väljs** direkt av folket.
 ‘The members’ dignity originates from that he is elected directly by the people.’

Secondly, we expect that the presence of an *av*-phrase will more likely yield the use of the *s*-passive. This hypothesis runs counter to the finding by Laanemets (2012:124ff.) that the use of an *av*-phrase prefers a *bli*-passive. However, our reasoning is that the focus shifts from the subject to the ‘by’-agent, which implies a decreased control of

the subject over the situation, which would in turn motivate the use of an *s*-passive. This reasoning is thus in line with findings by Kirri (1975), Silén (1997) and Engdahl (2006).

3.2.6 Aktionsart: ‘completed vs. ‘ongoing’

To determine the Aktionsart of the passive sentence we relied on the distinction between ‘completed’ and ‘ongoing’ and its respective definitions in *SAG* (IV:324): ‘if an event implies an event with a natural endpoint, then the event is called completed. After the endpoint the event can no longer take place. An event without a natural endpoint is called ongoing’.¹⁰ The *SAG* (IV:398–399) further describes that *s*-passives are used with ongoing events, which do not have a clear beginning or ending (as in (21)), while *bli*-passives are used with completed events, especially when these have a clear outcome (as in (22)).¹¹

- (21) Detta dataregister **disponerades/?blev** **disponerat** bara av staben.
this data.register disposed.of.PASS/became disposed.of only by staff.DEF
 ‘This data register was disposed of by the staff.’
- (22) Vi **blev** **körda** till skolan av min syster.
we became driven to school.DEF by my sister
 ‘We were driven to school by my sister.’

It is important to note here that Aktionsart is associated with a combination of different factors, including the number and the (in)definiteness of the arguments present in the sentence, as well as accompanying adverbials. (See in this regard Hansen (2008), who discusses the role of phrasal adverbials and Husband (2012), who discusses the influence of quantization and definiteness.)

The *SAG* (IV:328) further relates the distinction between ongoing and completed events to unlimited vs. limited quantities and/or amounts of time. As illustrated in the examples under (23) (all taken from *SAG*), a limited quantity and/or a limited amount of time tends to refer to a completed event, while an unlimited quantity and/or amount of time generally indicates an ongoing event.

- | (23) <i>Ongoing event</i> | <i>Completed event</i> |
|---|---|
| a. Vi drack kaffe (i flera timmar).
<i>we drank coffee for several hours</i> | Vi drack kaffet (på fem minuter).
<i>we drank coffee.DEF in five minutes</i> |
| b. Maj skrev brev (i en timme).
<i>Maj wrote letters during one hour</i> | Maj skrev ett brev (på en timme).
<i>Maj wrote a letter in one hour</i> |

Apart from the time adverbials within the brackets, these examples also reveal an interplay with criteria such as number and definiteness/specificity. Plural arguments (*Maj skrev brev*) thus point to ongoing events (see predictor ‘Subject Number’ above), while definite or specific arguments (*Vi drack kaffet*; *Maj skrev ett brev*) point to completed events.

These criteria were used in the annotation process. In the case of specific and definite arguments, the Aktionsart was annotated as ‘completed’. The specific argument *ett manuskript* in (24) and the definite argument *klagomålet* in (25) signal completed events, which have a resultative end state: i.e. the acceptance of a certain manuscript and the treating of a certain complaint.

- (24) Frågan om när ett manuskript **blir accepterat** eller inte i en viss tidskrift är naturligtvis av stor vikt för författarna.

‘The question of when a manuscript is accepted or not in a particular journal is of course of great importance to the authors.’

(Medicinska texter)

- (25) De har som uppgift att hjälpa konsumenten så att klagomålet mot den utländska näringsidkaren **blir behandlat**.

‘It is their job to assist the consumer so that the complaint against the foreign trader is treated.’

(Finlandssvenska texter)

In the case of plural arguments, the Aktionsart was annotated as ‘ongoing’. Such plural arguments are unspecific and point towards general events or norms, which have neither a clear beginning nor a clear end state. The plural arguments in (26)–(28) do not involve specific ‘friends’, ‘violations’, or ‘girls’, and as a result the respective propositions indeed indicate general events or norms.

- (26) Vänner kan inte **väljas** enbart rationellt, då är det fråga om någonting annat än vänskap.

‘Friends cannot be selected purely rationally, then it is a question about something else than friendship.’

(Finlandssvenska texter)

- (27) Kränkningar av svenskt territorium kan aldrig **accepteras**.

‘Violations of the Swedish territory can never be accepted.’

(Tidningstexter)

- (28) Oftast tjejer **blir accepterade** för dom är alltid snälla med mig, kram på er!

‘Girls are usually accepted because they are always nice to me, hug on you!’

(Bloggmix)

In case both criteria were combined, i.e., a combination of plural and definiteness, the Aktionsart was annotated as ‘completed’. In (29), for example, ‘the girls’ refer to a specific group of girls, in contrast to (28), where ‘girls’ are referred to in a general sense. The proposition in (29) is completed, as that specific group of girls will get accepted.

(29) Men tjejerna som börjar **blir** automatiskt **accepterade**.

‘But the girls who start become automatically accepted.’

(Tidningstexter)

Following the *SAG*, we want to emphasize that the association between bare plurals and ongoing events on the one hand and definiteness/specificity and completed events on the other hand, is to be understood as a tendency and not a clear-cut rule, which fully squares with the probabilistic approach that we subscribe to. The use of certain adverbials can change the Aktionsart.

(30) Jag hoppas verkligen att denna strävan **blir** allmänt **accepterad** inom alla EU-institutioner, även här i parlamentet.

‘I truly hope that this pursuit is generally accepted in all EU-institutions, even here in parliament.’

(Europarl svenska)

In (30) the use of *allmänt* ‘generally’ induces the interpretation of an ongoing event, even though the argument *denna strävan* ‘this pursuit’ is specific and comes with a definite article.¹²

As mentioned above, the classical expectation with regard to Aktionsart is that *s*-passives are used with ongoing events, and *bli*-passives with completed events (*SAG* IV:389–399; see Mikkelsen 1911 for Danish and Western 1921 for Norwegian). Engdahl (2006) casts doubts, however, about the weight of Aktionsart. According to Engdahl, other factors, such as animacy, have a stronger effect. The same argument was made for Danish by Heltoft & Falster Jakobsen (1996) who ascribe more weight to the subjective/objective mood of the clause (see also Enger 2001). Our multivariate analysis enables us to weigh the effect of this variable in relationship to the other influencing factors under analysis.

3.3 Statistical analysis

Logistic regression modelling is a widely used statistical method to analyze dichotomous outcomes (in our case, *s*-passive vs. *bli*-passive). The statistical model is now also commonly applied in corpus-linguistic research (e.g. Bresnan et al. 2007 which has become a textbook example in quantitative linguistics).¹³ The logistic regression model is flexible, in that it may include both continuous and categorical variables and the results are relatively easy to interpret.

With six predictor variables (excluding interactions), the logistic regression model can be formulated as in (i):

$$\log_e \left(\frac{\pi_i}{1 - \pi_i} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_6 X_6 \quad (i)$$

whereby π_i equals the probability of an *s*-passive and $1 - \pi_i$ equals the probability of a *bli*-passive. Then, the ratio $\pi_i / (1 - \pi_i) =$ the odds of an *s*-passive. X_{1-6} indicates the six predictor variables and β_{1-6} their six coefficients, which basically express the mathematical relationship between the outcome variable and the predictor variables. The (linear) logistic regression model in (i) results from the logit transformation of the (non-linear) logistic response function in (ii):

$$\pi_i = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_6 X_6)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_6 X_6)} \quad (\text{ii})$$

In the next section, which presents the results of our statistical data analysis, we will show how this logistic response function in (ii) can be used to estimate the probability of an *s*-passive based on a particular combination of factors, and consequently also for each specific observation.

4. RESULTS

We present the results of our statistical data analysis in two consecutive parts. First, by way of introducing the data, we present the bivariate results, which evaluate the association between the outcome variable and the predictor variables separately. Then we present the results of the logistic regression analysis, which evaluates the simultaneous effects of the predictor variables.

The bivariate distributions of all the predictor variables under analysis are given in Table 3. The statistical significance of these distributions was tested by means of chi-square tests. Cramér's V coefficient was used to estimate the effect sizes. Row percentages and standardized residuals are also reported. The latter allow for a more accurate interpretation of the effects than the row percentages; standardized residuals larger than |2| indicate a significant effect.

Following this analysis, we have strong evidence that the choice of passive form is significantly associated with four predictor variables: Subject Animacy and Subject Person, Modal Verb and *Av*-phrase. The observed tendencies are that the *s*-passive is positively associated with an inanimate and non-local subject, the presence of a modal verb, and by the use of an *av*-phrase. Based on these bivariate statistics, no evidence is found for Aktionsart and Subject number. Note that the distribution is strongly skewed in the case of Subject Person and *Av*-phrase: the large majority of observations includes a non-local subject and lacks an *av*-phrase. Finally, the bivariate results further suggest that Subject Animacy and Modal Verb yield the strongest effect.

Table 4 gives the model estimates of our logistic regression analysis. Model estimates, except for Odds Ratios and their 95% Confidence Intervals, were computed with functions from Harrell's (2013) rms package.¹⁴

	<i>s</i> -passive 599 (50%)	<i>bli</i> -passive 598 (50%)	Total	<i>p</i> -value (effect size)
Subject Animacy				< .0001 (.43)
inanimate	424 (72%, 14.8)	167 (28%, -14.8)	591	
animate	175 (29%, -14.8)	431 (71%, 14.8)	606	
Modal Verb				< .0001 (.39)
present	236 (86%, 13.6)	38 (14%, -13.6)	274	
absent	363 (39%, -13.6)	560 (61%, 13.6)	923	
Subject Person				< .0001 (.15)
non-local	588 (53%, 7.17)	523 (47%, -7.17)	1111	
other	6 (20%, -3.33)	24 (80%, 3.33)	30	
local	5 (9%, -6.30)	51 (91%, 6.30)	56	
Av-phrase				< .0001 (.14)
present	79 (73%, 5.04)	29 (27%, -5.04)	108	
absent	520 (48%, -5.04)	569 (52%, 5.04)	1089	
Aktionsart				= .059 (.05)
completed	292 (53%, 1.95)	258 (47%, -1.95)	550	
ongoing	307 (47%, -1.95)	340 (52%, 1.95)	647	
Subject Number*				= .35 (.03)
singular	350 (50%, -0.98)	355 (50%, 0.98)	705	
plural	244 (53%, 0.98)	220 (47%, -0.98)	464	

*28 missing values

Table 3. Bivariate statistics.

Based on our logistic regression model we have evidence for the effect of five predictor variables: Subject Animacy and Subject Number, Av-phrase, Aktionsart, and Modal Verb; Subject Person is not retained as significant in this multivariate model.

The direction of the effects corroborates what we hypothesized, except for Aktionsart: all else being equal, the *s*-passive is more probable than the *bli*-passive when the Subject is inanimate and plural, when there is an *av*-phrase and modal verb present, and when a completed Aktionsart is expressed. According to this model, the presence of a modal verb yields the strongest impact effect on the choice of passive construction. The odds ratio estimate indicates that the odds of a sentence with a modal verb being used with an *s*-passive are about nine times the odds of a sentence without a modal, with a 95% confidence interval of 6.07–13.77. The second strongest effect is associated with the animacy of the Subject. Here the odds of an animate Subject to be used with an *s*-passive is nearly seven times that of an inanimate Subject. Notice further that although Aktionsart and Subject Number are significant, their estimated impact effects are very low; an odds ratio of slightly over 1 basically means that there is only a small preference for the *s*-passive when a plural Subject is used or when a completed action is expressed.

	Est. Coef. (s.e.)	Wald Z	<i>p</i> -value	Odds Ratio	95% CI OR
Intercept	-1.789 (0.15)	-11.32	< .0001		
Subject Animacy (inanimate)	1.9147 (0.15)	12.94	< .0001	6.785	4.98–9.25
Av-phrase (present)	1.393 (0.26)	5.28	< .0001	4.025	2.40–6.76
Aktionsart (completed)	0.300 (0.14)	2.07	.0385	1.350	1.02–1.79
Modal Verb (present)	2.213 (0.21)	10.57	< .0001	9.142	6.07–13.77
Subject Number (plural)	0.579 (0.15)	3.81	.0001	1.784	1.32–2.40

s.e. = standard error of the estimated coefficient

Table 4. Logistic regression model estimates (success = *s*-passive).

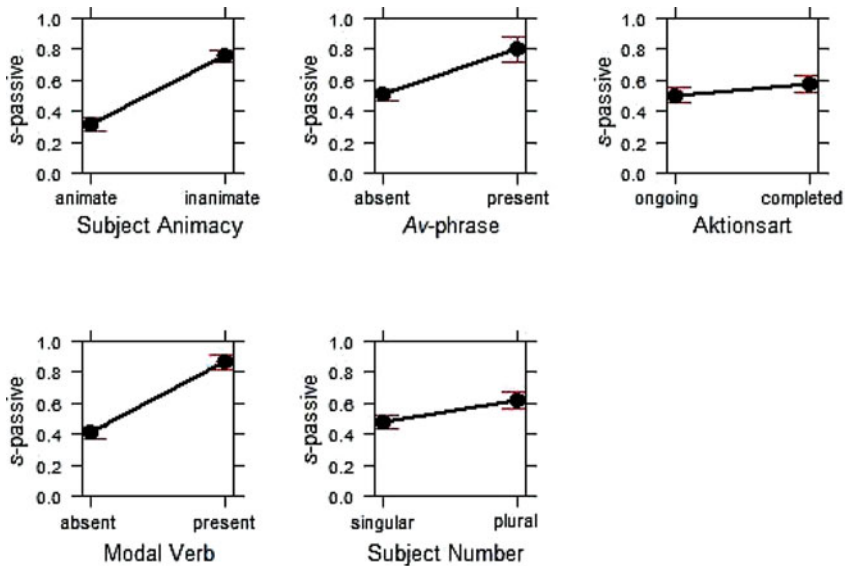


Figure 1. (Colour online) Predicted probabilities and 95% Confidence Intervals for the significant variables of the logistic regression model in Table 4.

The impact effects associated with each factor may also be visualized in terms of probabilities (i.e., on a scale from 0 to 1), as in Figure 1.¹⁵

The plots nicely illustrate the opposite effects that the different levels of the variables have on the choice of *s*-passive. Thus, with an *av*-phrase present, the probability of an *s*-passive is on average approximately 80%.¹⁶ Notice again the small probability differences for Aktionsart and Subject Number.

Finally, we can estimate the predicted probability of each observation or combination of factors by means of the logistic response function given in (ii). By way of illustration, we compute the probability of an *s*-passive based on three different combination of factors.

The predicted probability of an *s*-passive form of a sentence with a singular, animate Subject, without an *Av*-phrase, with a completed Aktionsart but without a Modal Verb, is estimated to be 18%. A probability lower than 50% basically means that the *bli*-passive is actually predicted for this particular combination of factors. The predicted probability is obtained by solving the equation in (iii). The estimated coefficients associated with each variable (see Est. Coeff. in Table 3) are multiplied either by 1 or by 0 depending on which level we are interested in. Thus, ‘inanimate’ = 1 vs. ‘animate’ = 0, ‘present’ = 1 vs. ‘absent’ = 0, etc.

Probability of s-passive

$$= \frac{\exp(-1.789 * 1 + 1.9147 * 0 + 1.393 * 0 + 0.300 * 1 + 2.213 * 0 + 0.579 * 0)}{1 + \exp(-1.789 * 1 + 1.9147 * 0 + 1.393 * 0 + 0.300 * 1 + 2.213 * 0 + 0.579 * 0)}$$

$$= 0.18$$

(iii)

Now, changing the status of Subject Animacy to ‘inanimate’ and that of Modal Verb to ‘present’ (two variables for which we found a large impact effect), we find that the probability shifts to 93% (see equation (iv)). Thus, according to our model the use of the *s*-passive is very likely, given the specific combination of factors.

Probability of s-passive

$$= \frac{\exp(-1.789 * 1 + 1.9147 * 1 + 1.393 * 0 + 0.300 * 1 + 2.213 * 1 + 0.579 * 1)}{1 + \exp(-1.789 * 1 + 1.9147 * 1 + 1.393 * 0 + 0.300 * 1 + 2.213 * 1 + 0.579 * 1)}$$

$$= 0.93$$

(iv)

If we also change Aktionsart to ‘ongoing’, then the probability slightly lowers to 91% (see equation (v)), which reflects the low impact effect associated with Aktionsart.

Probability of s-passive

$$= \frac{\exp(-1.789 * 1 + 1.9147 * 1 + 1.393 * 0 + 0.300 * 0 + 2.213 * 1 + 0.579 * 1)}{1 + \exp(-1.789 * 1 + 1.9147 * 1 + 1.393 * 0 + 0.300 * 0 + 2.213 * 1 + 0.579 * 1)}$$

$$= 0.91$$

(v)

The next section discusses our findings in comparison with the previous research on the passive alternation.

5. DISCUSSION

The present study of three verbs *acceptera* ‘accept’, *behandla* ‘treat’ and *välja* ‘choose’ that alternate between the *s*-passive and the *bli*-passive investigates the simultaneous effects of six factors as well as the relative strength of each factor separately.

The results show that the strongest effect is obtained for the factor Modal Verb. This finding corroborates earlier studies on the passive alternation. Engdahl (1999), for example, assumed that particularly the modal/temporal auxiliaries *kan*, ‘can’ and *ska*, ‘will’ are preferably associated with the *s*-passive. Since our corpus not only contains the modals mentioned, but also includes other modal/temporal auxiliaries as *måste* ‘must, have to’ in (31), we can conclude that the presence of any modal/temporal auxiliary motivates the use of an *s*-passive.

(31) Det är en funktionell del av gemenskapen och den måste **behandlas** som en sådan.

‘It is a functional part of the community and it must be treated as such.’

(Europarl svenska)

Engdahl (1999:19) and Laanemets (2012:170) note, however, that the modal *vill* ‘want to’ may be an exception, as it seems to be more common to use *vill* with the *bli*-passive. Our data confirms this: of the 38 *bli*-passive sentences with modal verbs, 13 include *vill*, whereas of the 236 *s*-passive sentences with modal verbs, only one involves *vill* (32).

(32) Om EU vill **accepteras** av sina invånare måste de uppleva att vi arbetar för dem.

‘If the EU wants to be accepted by its citizens they need to experience that we work for them.’

(Europarl svenska)

It would of course be worthwhile to investigate the influence of the different modal verbs in closer detail, but since this was the first investigation using a probabilistic approach, we limited the factor Modal to the distinction present vs. absent, leaving more fine-grained distinctions for further research.

Other strong effects were obtained for the factors Animacy and *Av*-phrase. The finding for Animacy is in line with the literature, which has repeatedly stated that inanimate subjects tend to occur with the *s*-passive whereas animate subjects appear to prefer the *bli*-passive. With respect to the factor *Av*-phrase, our results first of all corroborate the observation that the use of an *av*-phrase is rare; only 9% of our observations include an *av*-phrase. We further found that the use of an *av*-phrase is associated with a preference for the *s*-passive. This contradicts Laanemets’ (2012) observation that the *av*-phrase is associated with a preference for the *bli*-passive.

As explained in our discussion of the variables in Section 2, we believe that the preference for the *s*-passive may be explained by the fact that the focus shifts from the subject to the ‘by’-agent, which is thus attributed more control of the situation compared to the subject, as is illustrated in (33).

- (33) Presidenten väljs av parlamentet och utser sedan en statsminister vars regering måste **godkännas** genom en omröstning i Saeiman.
 ‘The president is elected by the parliament and then appoints a prime minister whose government must be approved through a re-election in Saeiman.’
 (Bloggmix)

Regarding agents, it can furthermore be noted that the ‘by’-agent may be expressed in alternative ways, for instance through using adverbials, as *inom alla EU-institutioner* and *i parlamentet* in (34). However, we did not label such instances as *av*-phrase. This use of alternatives with regard to expressing the ‘by’-agent is something future studies could examine in more detail.

- (34) Jag hoppas verkligen att denna strävan **blir** allmänt **accepterad** inom alla EU-institutioner, även här i parlamentet.
 ‘I truly hope that this pursuit is generally accepted in all EU-institutions, even here in parliament.’
 (Europarl svenska)

The remaining factors that yielded an effect were Subject Number and Aktionsart. Our analysis suggests that their effect is relatively low. With regard to Subject Number the use of the *s*-passive is found to be associated with plural subjects. Possibly, the use of plurals coincides with the expression of general events or general norms (see SAG IV:328–329), as in (35), which squares with the finding that the *s*-passive is associated with a general event or norm (Engdahl 1999, 2006).

- (35) I övriga trafikslag har dessa risker kunnat minimeras genom ett systematiskt säkerhetsarbete och tydliga regler där inga överträdelser **accepteras**.
 ‘In other types of transport these risks have been minimized through systematic security work and clear rules where no violations are accepted.’
 (Finlandssvenska texter)

Our results for the factor Aktionsart corroborate Engdahl’s (2006) observation that the influence of Aktionsart is very weak. This finding is in line with the observations of the Danish passive use discussed by Heltoft & Falster Jakobsen (1996). However, in contrast to the expected pattern, our results suggest that the *s*-passive is associated with a completed event, rather than with an ongoing event. In (36), for example, we had expected the use of an *s*-passive, because the sentence expresses an ongoing event. The same applies to (34) above, where an ongoing event, which is particularly

reinforced by the use of *allmänt* ‘generally’, is used together with a *bli*-passive. At this stage, we have no explanation for this finding.

(36) Men – är det inte BRA om det **blir accepterat** även för killar att vara lucia, och att ha vilka kläder man vill?

‘But – is it not GOOD if it is accepted that even boys are Lucia, and that they dress the way they want?’

(Bloggmix)

Finally, Subject Person is the only factor that did not appear to be significant. We related this factor to the notion of control – a local subject would reflect a higher level of control and accordingly prefer a *bli*-passive. No evidence was found for this influence.

6. CONCLUSIONS AND FUTURE PROSPECTS

In the context of the special issue, we have taken a corpus-linguistic and probabilistic approach to examine a contemporary research question in Nordic linguistics. Previous research on the Swedish passive alternation provided us with several factors that potentially motivate the choice of either one of two passive forms. A statistical evaluation of the effects of the proposed factors had hitherto not been conducted. Aiming at filling that gap, the present study performed a statistical analysis on the various factors that can influence the choice of the morphological or periphrastic passive starting from a data set of three verbs that can alternate. Our study thus adds to earlier research in that it allows for a more accurate analysis of the simultaneous effect and relative strength of each factor on the speaker’s choice of one of the passive forms.

Our study only included three verbs, which may raise questions about the generalizability of our findings. Taking into account the semantic diversity of the three verbs in the study, we could conjecture that the findings are applicable more generally and more widely than the study itself. This is merely speculative, however, and more research is required.

Two directions of future research include examining a larger corpus sample and testing the statistical model(s) with a psycholinguistic experiment. A larger corpus sample should include a larger number of verbs, which will shed more light on possible preferential differences between the verbs. With a larger corpus sample one could also retain the proportional difference between the two passive forms. As mentioned, the fact that we used a case-controlled design overestimates the probability of an *s*-passive.

In addition to a larger corpus study, one could apply the ‘100-split task’, a psycholinguistic test originally developed by Bresnan (2007) in the context of

the English dative alternation. The 100-split task is a grammaticality judgment experiment that aims to capture the probabilistic preferences of speakers. In this test, participants are asked to ‘rate the naturalness of alternative forms as continuations of a context by distributing 100 points between the alternatives’ (Ford & Bresnan 2013:300). The average score for each sentence is then compared to the predicted probabilities estimated by a statistical model. Earlier tests on the dative alternation show a remarkable correlation between the scores of the participants and the statistical models, which suggests that probabilistic models of language variation accurately capture the intuitions of the speakers. Together with the results of an observational corpus study, experimental research provides strong converging evidence for the probabilistic approach to syntactic variability.

ACKNOWLEDGEMENTS

The authors wish to thank two anonymous *Nordic Journal of Linguistics* reviewers for their valuable comments and suggestions and Sara Van Herreweghe for her help with the data retrieval. Gudrun Rawoens also expresses her gratitude to the foundation *Petra och Karl Erik Hedborgs stiftelse* for a research stay at Stockholm University in 2014.

NOTES

1. For passivization in the Insular Scandinavian languages Icelandic and Faroese see e.g. Barðdal & Molnár (2003) and Askedal (2010).
2. The translations are taken from Heltof & Falster Jakobsen (1996).
3. It can further be noted, however, that the use of the *s*-passive in the simple past in Danish and Norwegian, even though possible, is less frequent than the *s*-passive in the present tense (see the numbers in Figures 5.5 and 5.6 in Laanemets 2012:97).
4. Engdahl (1999) does not provide information about the frequency of these verbs. A search in the PAROLE corpus (conducted on 10 June 2014) yields the following numbers: *tala* ‘speak’ (13,309), *stänga* ‘close’ (3,256) and *öppna* ‘open’ (4,611).
5. The grammaticality judgment indicated by the question marks in examples (11) and (13) is taken from the original source.
6. In this respect, Swedish differs from Norwegian and Danish that do allow the periphrastic passive in impersonal passives (Áfarli 2006, Engdahl 2006:40). For comparative figures on the distribution of the *s*-passive and *bli*-passive in impersonal passives, see also Laanemets (2012:120, Table 7.5).
7. The corpora in Språkbanken can be searched freely through the webconcordancer Korp: <http://spraakbanken.gu.se/korp/>
8. Akademiska texter, August Strindberg, Finlandssvenska texter, Medicinska texter, Parallella material, Skönlitteratur, Tidningstextere, Tidskrifter, Dramawebben, LSBarT, Parole, Psalmboken, SNP 78–79, SUC 2.0, SUC 3.0, Svenska partiprogram och

valmanifest 1887–2010, Bloggmix, Europarl svenska, Svensk författningssamling, Svenska Wikipedia, SweWac.

9. All modal verbs (see *SAG IV*:282ff.) were taken into account, but only the verbs mentioned occurred in our dataset.
10. Our translation of: ‘om en aktion innebär att ett skeende har en naturlig slutpunkt kallas aktionen avgränsad (. . .) Efter slutpunkten kan skeendet inte längre äga rum. En aktion utan naturlig slutpunkt kallas oavgränsad’.
11. The grammaticality judgment indicated by the question mark is taken from the original source.
12. The annotations were performed by two annotators (authors two and three), who each annotated half of the observations. Random checks were performed to compare each other’s annotations and the few cases of doubt that each individual annotator had, were discussed in group. No formal test was performed to evaluate the inter-annotator agreement.
13. Baayen (2008:195–208) offers a more detailed (but very practical) introduction to logistic regression modeling in the context of linguistic research.
14. All statistical analysis was done in R version 3.0.1 (16 May 2013). Model diagnostics indicate a very good fit of our model. The le Cessie-van Houwelingen-Copas-Hosmer unweighted sum of squares test (see Hosmer et al. 1997) does not give evidence for a lack of fit of a logistic regression model ($p = .067$). The bootstrap validated C index of concordance = .82, which indicates that the model has a very good general prediction rate. Multicollinearity does not seem to be a problem, as all Variable Inflation Factor (VIF) values are below 1.10 (Subject Number). Model diagnostics (except for VIF) were computed with functions from Harrell’s (2013) rms package. VIF was computed with the VIF function from Fox & Weinberg’s (2011) car package.
15. Figure 1 was created with the all Effects function from Fox (2003) effects package.
16. ‘On average’ means that the other variables are set to their mean, which is equivalent to setting the distribution of the variables to their distribution in the data.

REFERENCES

- Áfarli, Tor. 2006. Passive and argument structure. In Werner Abraham & Larisa Leisiö (eds.), *Passivization and Typology: Form and Function* (Typological Studies in Language 68), 373–382. Amsterdam: John Benjamins.
- Askedal, John Ole. 2010. Germanic passive constructions. In John Ole Askedal, Ian Roberts & Tomonori Matsushita (eds.), *Noam Chomsky and Language Descriptions*, 75–110. Amsterdam: John Benjamins.
- Baayen, Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Barðdal, Jóhanna & Valéria Molnár. 2003. The passive in Icelandic – compared to Mainland Scandinavian. In Jorunn Hetland & Valéria Molnár (eds.), *Structures of Focus and Grammatical Relations* (Linguistische Arbeiten 477), 231–260. Tübingen: Max Niemeyer.
- Bod, Rens, Jennifer Hay & Stefanie Jannedy (eds.). 2003. *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
- Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in Search of Its Evidential Base*, 77–96. Berlin: Mouton de Gruyter.

- Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen. 2007. Predicting the Dative Alternation. In Gerlof Boume, Irene Kraemer & Jan Zwarts (eds.), *Cognitive Foundations of Interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.
- Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1), 186–213.
- Engdahl, Elisabet. 1999. The choice between *bli*-passive and *s*-passive in Danish, Norwegian and Swedish. Ms., Göteborg University. http://www.svenska.gu.se/digitalAssets/1336/1336829_engdahl-nordsem-passivechoice-1999.pdf (accessed 15 December 2013).
- Engdahl, Elisabet. 2001. Valet av passivform i modern svenska. *Svenskans beskrivning* 24, 81–90. [Linköping Electronic Conference Proceedings, <http://www.ep.liu.se/ecp/006/007/ecp006007b.pdf>, accessed 15 December 2013]
- Engdahl, Elisabet. 2006. Semantic and syntactic patterns in Swedish passives. In Benjamin Lyngfelt & Torggrim Solstad (eds.), *Demoting the Agent: Passive, Middle and Other Voice Phenomena*, 21–45. Amsterdam: John Benjamins.
- Enger, Hans-Olav. 2001. Om *s*-passivens relasjoner til modalitet, aspekt og kasus. *Tijdschrift voor Skandinavistiek* 22(2), 411–133.
- Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Ford, Marilyn & Joan Bresnan. 2013. Using convergent evidence from psycholinguistics and usage. In Manfred Krug & Julia Schlüter (eds.), *Research Methods in Language Variation and Change*, 295–312. Cambridge: Cambridge University Press.
- Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8(15), 1–27.
- Fox, John & Sanford Weinberg. 2011. *An R Companion to Applied Regression*, 2nd edn. Thousand Oaks, CA: Sage.
- Hansen, Eskil. 1991. Passiv i moderne norsk. I: *Språknytt. Språkrådet*. [http://www.sprakradet.no/Toppmeny/Publikasjoner/Spraaknytt/Arkivet/Eldre/Passiv_i_moderne_norsk, accessed 14 May 2014]
- Hansen, Erik & Lars Heltoft. 2011. *Grammatik over det Danske Sprog*, vols. I–III. Odense: Danske Sprog- og Litteraturselskab (DSL) og Syddansk Universitetsforlag.
- Hansen, Maj-Britt Mosegaard. 2008. *Particles at the Semantics/Pragmatics Interface: Synchronic and Diachronic Issues, a Study with Special Reference to the French Phasal Adverbs*. Oxford: Elsevier.
- Harrell, Frank E. Jr. 2013. rms: Regression Modeling Strategies. R package version 4.0–0. <http://CRAN.R-project.org/package=rms>.
- Heltoft, Lars 1994. *S*-modus og perifrastisk modus. In Nils Jörgensen, Christer Platzack & Jan Svensson (eds.), *Språkbruk, grammatik och språkförändring. En festskrift till Ulf Teleman*, 155–165. Lund: Lund University.
- Heltoft, Lars & Lisbeth Falster Jakobsen. 1996. Danish passives and subject positions as a mood system: A content analysis. In Elisabeth Engberg-Pedersen, Michael Fortescue, Peter Harder, Lars Heltoft & Lisbeth Falster Jakobsen (eds.), *Content, Expression and Structure: Studies in Danish Functional Grammar*, 199–234. Amsterdam: John Benjamins.
- Hosmer, David W., Trina Hosmer, Saskia le Cessie & Stanley Lemeshow. 1997. A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine* 16, 965–980.

- Husband, Edward Matthew. 2012. Some structural analogies between existential interpretation and telicity. In María Cristina Cuervo & Yves Roberge (eds.), *The End of Argument Structure?* (Syntax and Semantics 38), 105–129. Bingley: Emerald.
- Kirri, Arto. 1975. *Studier över passivkonstruktioner i nysvenskt skriftspråk* (Meddelanden från Institutionen för nordisk filologi vid Helsingfors universitet. Serie A, nr. 2). Helsingfors: Helsingfors universitet.
- Laanemets, Anu. 2004. *Dannelse og anvendelse af passiv i dansk, norsk og svensk* (Nordistica Tartuensia 11). Tartu: Tartu University.
- Laanemets, Anu. 2009. The passive voice in written and spoken Scandinavian. In Marc Fryd (ed.), *The Passive in Germanic Languages, GAGL 49*, 144–166. Groningen: University of Groningen, Center for Language and Cognition Groningen.
- Laanemets, Anu. 2012. *Passiv i moderne dansk, norsk og svensk. Et korpusbaseret studie af tale- og skriftsprog*. Tartu: University of Tartu Press.
- Larsson, Ida. 2012. Inte helt passiv: konstruktion med *få* + particip i tal och skrift. *Språk och stil* NF 22(2), 26–56.
- Lyngfelt, Benjamin. 2010. En akademiledamot pensioneras inte – om konstruktioner med *s*-verb i svenskan. In Kristinn Jóhannesson, Ida Larsson, Erik Magnusson Petzell, Sven-Göran Malmgren, Lena Rogström & Emma Sköldberg (eds.), *Bo65. Festskrift till Bo Ralph*, 178–192. Göteborg: Meijerbergs arkiv för svensk ordforskning.
- Mikkelsen, Kristian. 1911. *Dansk Ordføjningslære*. København: Hans Reitzels Forlag.
- R Core Team. 2013. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/> (accessed 5 November 2013).
- Rawoens, Gudrun & Annika Johansson. 2014. A corpus-based contrastive study of impersonal passives in Swedish and Dutch. Ms., Ghent University and Stockholm University.
- Silén, Beatrice. 1997. *Agentadverbialet i modern svenska* (Studier i nordisk filologi 76). Helsingfors: Svenska litteratursällskapet i Finland. [Ph.D. dissertation]
- Sundman, Marketta. 1983. Control, subject and voice in Swedish. In Fred Karlsson (ed.), *Papers from the Seventh Scandinavian Conference of Linguistics*, 101–112. Helsinki: University of Helsinki, Department of General Linguistics.
- Teleman, Ulf, Staffan Hellberg & Erik Andersson. 1999. *Svenska Akademiens Grammatik*, vol. IV. Stockholm: Svenska Akademien.
- Thorell, Olof. 1973. *Svensk grammatik*. Stockholm: Esselte Studium.
- Western, August. 1921. *Norsk riksmåls-grammatikk for studerende og lærere*. Kristiania: Aschehoug.
- Zola Christensen, Robert & Lisa Christensen. 2005. *Dansk Grammatik*. Odense: Syddansk universitetsforlag.