



10-17-2013

# The Evolutionary Trajectory of the Icelandic New Passive

Anton Karl Ingason  
*University of Pennsylvania*, [ingason@ling.upenn.edu](mailto:ingason@ling.upenn.edu)

Julie Anne Legate  
*University of Pennsylvania*, [jlegate@ling.upenn.edu](mailto:jlegate@ling.upenn.edu)

Charles Yang  
*University of Pennsylvania*, [charles.yang@ling.upenn.edu](mailto:charles.yang@ling.upenn.edu)

---

# The Evolutionary Trajectory of the Icelandic New Passive

## **Abstract**

We examine the diachronics of a New Passive construction in Icelandic and use Yang's model of language learning and change to explain its rapid rise. The New Passive has been spreading at the expense of a Canonical Passive in the recent past 50 years. Applying empirical measurements from the IcePaHC corpus, we show that our model can be used to account for the spread of the New Passive and the rate of change. The model also has implications for the actuation of the change.

# The Evolutionary Trajectory of the Icelandic New Passive\*

Anton Karl Ingason, Julie Anne Legate and Charles Yang\*

## 1 Introduction

The Icelandic New Passive<sup>1</sup> is a syntactic variant which first appeared around 1950 (Maling and Sigurjónsdóttir 2002). Since its emergence, The New Passive (NP) has been gaining ground at the expense of a functionally equivalent Canonical Passive (CP). The rapid spread of the NP is exemplified by a striking generational difference. Young speakers are much more likely to use the NP and to rate it as acceptable when asked to give a syntactic judgment. This state of affairs offers an excellent opportunity to study the evolution of syntactic change in progress, which is the topic of this paper. We use the Variational Model (VM) of Yang (2002) to analyze the NP change as a mechanical process of learning from input data. The model predicts the direction of the change in the population and the rate of change across generations. Furthermore, our analysis has implications for the actuation of the change.

The core data are shown in (1). In a Canonical Passive, the theme/patient of the active (1a), ‘the boy’, surfaces as the subject of the passive (1b). The main verb in the passive shows passive morphology. The CP is parallel to the English passive in these respects. The New Passive is shown in (1c). In the NP, the underlying theme/object stays an object, passive morphology appears, and an expletive *það* ‘there’ is inserted in first position.

- (1) a. Álfurinn lamdi strákin. (Active)  
the.elf.SBJ beat the.boy.OBJ  
‘The elf beat the boy’  
b. Strákurinn var laminn. (Canonical Passive; CP)  
the.boy.SBJ was beaten.PASS  
‘The boy was beaten’  
c. Það var lamið strákin. (New Passive; NP)  
there.EXPL was beaten.PASS the.boy.OBJ  
‘The boy was beaten’

The nature of trajectory of change has received less attention than the formal properties of the CP and the NP. This paper presents a mathematical model of the evolutionary trajectory of the Icelandic New Passive, a step towards predictive modelling in the study of syntactic change. Mathematical models are increasingly opening up fruitful avenues in historical syntax. For example, the Constant Rate Effect (Kroch 1989) showed that quantitative reasoning can draw on usage frequencies and inform the analysis of linguistic competence during a change in progress. The Variational Model (VM) of Yang (2002) derives the S-shaped curve of change from a hypothesized language acquisition mechanism, and the Death Rattle Effect (Ingason 2010) in the erosion of non-default grammar rules is another derived consequence of the VM, and gives a fully mechanical account of what superficially appears to be a sociolinguistic effect of hypercorrection.

By developing mathematical models of mechanical change in historical syntax, we can entertain the possibility that diachronically predictive analysis is possible. If we do not have to resort to social pressures as explanations for change, we may be able to move beyond fitting curves to developments that already happened. Importantly, this line of research has not fully engaged with major syntactic change in progress. Previous studies on major reorganization of clausal syntax, while important, generally discuss changes that went to completion a long time ago. Therefore, the opportunity to

---

\*Parts of this and related work were presented at NVA 41 in Bloomington, LSA 87 in Boston and Non-Canonically Case-Marked Subjects within and across Languages and Language Families in Iceland. We would like to thank the audiences at these events for valuable comments.

<sup>1</sup>This construction is also referred to as ‘The New Construction’ or ‘The New Impersonal Construction’. The different names reflect a debate on the status of the construction as an active or a passive. These issues are orthogonal to deriving the evolutionary trajectory. We discuss the syntactic analysis in an appendix.

study the rapidly spreading Icelandic New Passive is welcome. It is also a major change in the sense that it is not limited to a few lexical items; it affects all passives in the language.

The Icelandic New Passive will be a test case for a general theory of change. We will explain what is driving the spread of the change. We will take steps towards predicting the rate of change. Finally, we will discuss why the change started when it did and not at some other time in history. Explaining linguistic change is a difficult task and many questions will remain unanswered. However, we hope that our attempt to discuss the evolutionary trajectory of the New Passive within a precise formal framework will encourage further progress. The paper is organized as follows. Section 2 provides linguistic background, explaining the syntactic structures involved. Section 3 discusses the history of the New Passive and its social environment. In Section 4, we provide a model of the NP's spread in terms of linguistic input to children. We extend the model to predicting the rate of change in Section 5, deriving a specific evolutionary trajectory throughout the period 1950-2050. In Section 6, we show how favorable conditions emerged for the actuation of the change in the early 1900's. Section 7 concludes. Section 8 is a brief appendix outlining our syntactic analysis.

## 2 Linguistic Background

The syntactic difference between the Canonical Passive and the New Passive in (1) is that in contrast to the CP, the underlying object stays an object in the NP. There is no difference in meaning between (1b) and (1c) and for a speaker who has the NP, either variant can be used in the same situation. The syntax of the two variants has been studied extensively in a number of publications and the main facts are now well documented (Maling and Sigurjónsdóttir 2002, Barðdal and Molnár 2003, Eythórsson 2008, Jónsson 2009, Sigurðsson 2011, Árnadóttir et al. 2011, Ingason et al. 2012a, Sigurðsson 2012, Sigurðsson and Stefánsdóttir 2013). The NP shares formal properties of both actives and passives; this is a widely discussed topic. Consider the NP example in (2) which has a passive participle and *by*-phrase.

- (2) Það var skoðað bílinn af bifvélavirkjanum. (NP)  
 there was inspected.PASS the.car **by** the.car mecahnic  
 'The car was inspected by the car mechanic'  
 Example from Jónsson (2009:294)

The passive morphology in (2) is a typical property of passives. The possibility to express the agent of the active in a *by*-phrase is also a feature which is strongly associated with passives. For speakers who have the NP, the possibility of including a *by*-phrase is well attested in experiments (Sigurðsson 2012, Sigurðsson and Stefánsdóttir 2013). The construction has indeed been described as a passive by some linguists (Eythórsson 2008, Jónsson 2009). However, the NP also lines up with actives in some regards. In (3), 'the boy', which is the patient/theme, shows nominative morphology in the CP and accusative morphology in the NP.

- (3) a. Strákurinn var laminn. (Canonical Passive; CP)  
 the.boy.SBJ.NOM was beaten.PASS  
 'The boy was beaten'  
 b. Það var lamið strákin. (New Passive; NP)  
 there.EXPL was beaten.PASS the.boy.OBJ.ACC  
 'The boy was beaten'

In Canonical Passives, the underlying object becomes the subject of the passive and appears with nominative case. This is a well known property and it is familiar from English and other languages. The fact that 'the boy' has accusative case in (3b) is therefore unusual if the sentence is a passive. For arguments that the NP is an impersonal active, see Maling and Sigurjónsdóttir (2002).

A third view is that the labels 'active' and 'passive' are not a central problem for the syntactic analysis. The NP shares properties associated with both labels and any formal analysis will have to account for that reality. At that point, the use of the voice labels becomes a matter of definitions

where the NP falls into empirical middle ground. Issues surrounding this labeling problem are discussed by Sigurðsson (2012), and Wood (2013) develops an argument for the conclusion that a ‘passive’ is not a formal primitive but a fairly loose cover term for a set of properties that are all attested independently of passives. Our syntactic analysis is in the spirit of this third view and we position the NP within a typology of passive-like constructions. The analysis is outlined in an appendix.

### 3 History and Social Environment

The earliest reported use of the NP is around 1950, according to Maling and Sigurjónsdóttir (2002). In their syntactic judgment experiment, a sharp apparent time change was revealed in the results. Younger speakers were much more likely than older speakers to rate the NP as acceptable. A sharp generational difference was confirmed in a later study (Thráinsson et al. forthcoming). Geographically, the change is on the rise everywhere in Iceland, but parts of the capital are lagging behind to a measurable degree. A negative correlation was found between education and acceptability of the NP. We point to the references above for further details. The social facts are compatible with various interpretations, including one under which the usage of a specific speaker reflects a mechanical language acquisition process. Under this view, sociolinguistic differences in usage frequencies are expected among individuals who are exposed to different input patterns, even if the entire speech community is undergoing the same change.

In contrast to our proposal, it is possible that social forces are relevant to the purpose of driving the change. For example, some aspect of social identity might be relevant for the mechanism of spread. We will not pursue this possibility here; the question remains open. However, there is a reason to believe that the NP change is a particularly good candidate for a type of variation that is not sensitive to social evaluation. Labov and Harris (1986:21) observe that not all types of linguistic structure are equal in terms of social impact:

Abstract linguistic structure has little or no social impact on members of the community. The interface of language and society is narrow, and primarily on the surface: the words and sounds of the language.

In an attempt to formalize the intuition in the above quote, Ingason et al. (2012b) tested a strong hypothesis of “Antisocial Syntax”, that “word order is not socially evaluated, unless it can be identified with specific lexical or phonological material.” The study looked at potential style-shifting in the word order of English and Icelandic during periods when changes were taking place. Shifted usage frequencies in different social contexts are a central piece of evidence for social impact, regardless of the nature of these effects (see Labov 1972, Bell 1984). The findings suggested that apparent style-shifting in abstract syntactic structures disappears when shallow surface exceptions are excluded. We leave future investigation of social forces in the evolution of the New Passive for future work. The following section introduces a model that accounts for the central facts without taking such forces into account, a line of research worth pursuing in the interest of parsimony.

### 4 Spread of the New Passive

The Variational Model (Yang 2002) models how a child gradually makes a binary choice (e.g., *A* vs. *a*) based on evidence from linguistic input. When the child hears an input token/sentence, the *A* grammar is selected with probability *p* (and *a* with prob. *q*). Grammar *A* gets a reward/penalty based on whether *A* is compatible with the input; *p* and *q* are updated for future interactions (cf. Bush and Mosteller 1951). When the child hears a sentence and selects *A*, the new values for *p* and *q* will be as follows.

$$(4) \quad \text{a. If } A \text{ is compatible } (\checkmark A) \text{ with input, } \begin{cases} p' = p + \gamma q \\ q' = (1 - \gamma)q \end{cases}$$

- b. If  $A$  is **not compatible** ( $*A$ ) with input,  $\begin{cases} p' = (1 - \gamma)p \\ q' = q + \gamma p \end{cases}$

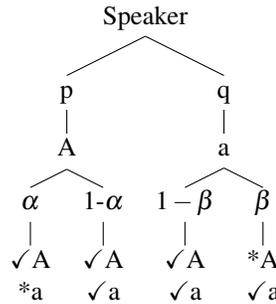
If all the input comes from the same grammar, this grammar will end up being selected by the child since it will be the only hypothesis that is never penalized. Language change introduces conflicting evidence in the input to children. Assuming a learning mechanism of type (4), each of the competing grammars will be penalized by some types of input, because each grammar produces some output that is incompatible with the other grammar. Given our assumptions, it is possible to derive the expected usage probabilities in the next generation of speakers. We do so in terms of a penalty probability, the probability with which a grammar is not compatible with input. Let  $C_A$  and  $C_a$  be the penalty probabilities for  $A$  and  $a$  respectively. If a child learns for a while according to (4), her resulting usage probabilities will approach the following values (see Yang 2002 for details).

(5)

$$\lim_{t \rightarrow \infty} p_t = \frac{C_a}{C_A + C_a} \quad \lim_{t \rightarrow \infty} q_t = \frac{C_A}{C_A + C_a}$$

The penalty probability can be measured in a corpus. Consider the illustration below which maps out the types of events that influence the resulting acquired usage probabilities.

(6)



In a given generation, a proportion  $p$  of the input is generated by grammar  $A$ . A proportion  $\alpha$  of these tokens are only compatible with grammar  $A$  and penalize  $a$ . This is illustrated in (6). We can think of  $\alpha$  as the evolutionary advantage of  $A$  because this value represents the “ability” of  $A$  to generate output that is incompatible with  $a$ . The overall probability of  $a$  being penalized, its penalty probability, is the product of the two values,  $C_a = \alpha p$ . The penalty probability of  $A$  is calculated in a parallel fashion,  $C_A = \beta q$ . This result is shown in (7).

(7)

$$C_a = \alpha p \quad C_A = \beta q$$

We can now combine (5) and (7) and write out the usage probability of  $A$  in the next generation,  $p'$  as a function of the penalty probabilities of the two grammars as shown in (8).

(8)

$$p' = \frac{\alpha p}{\alpha p + \beta q} = \frac{p}{p + (1-s)q} \quad \text{where } s = \frac{\alpha - \beta}{\alpha}$$

The prediction that follows is that if  $\alpha > \beta$ , then grammar  $A$  will eventually beat grammar  $a$ . In other words, when two incompatible grammatical analyses are used in a linguistic community, the grammar that has a greater evolutionary advantage will eventually take over. The values of  $\alpha$  and  $\beta$  can be estimated using a corpus and we now proceed to show that the New Passive is predicted to spread at the expense of the Canonical Passive. We used the IcePaHC corpus (Wallenberg et al. 2011)<sup>2</sup> to estimate  $\alpha$  and  $\beta$  in the case of the New Passive and the Canonical Passive. IcePaHC is a monolingual CP corpus and in order to estimate the advantage of the NP, we converted the corpus

<sup>2</sup>The corpus is described in Rögnavaldsson et al. (2011a,b, 2012)

into a monolingual NP corpus. The advantage of the NP comes from sentences which surface as type (9a) when generated by the CP, namely passives that are **definite-theme-initial**, and where the theme is morphologically ambiguous between the NP and the CP analysis (i.e., between nominative and accusative case).

- (9) a. Barnið var lamið.  
**the.child.NOM/ACC** was beaten  
 ‘The child was beaten’ (✓CP, ✓NP)  
 b. Það var lamið barnið.  
 there was beaten **the.child.NOM/ACC**  
 ‘The child was beaten’ (\*CP, ✓NP)

The New Passive system produces sentences of the above type as (9b), which is not compatible with the Canonical Passive. In the CP grammar, the definite theme is a subject and cannot appear in this low object position. The ungrammaticality of (9b) under a CP analysis follows from a definiteness effect in Icelandic which is parallel to the English contrast in (10).

- (10) a. There was **a child** beaten.  
 b. \*There was **the child** beaten.

Examples of type (9b) are only advantageous to the NP grammar and not the CP, because they can only be analyzed by the NP grammar. Under a NP analysis the sentence-initial theme can be analyzed as a fronted, topicalized constituent. We disregard examples of type (1) where there is a morphological contrast between the NP and the CP. The case morphology in such examples means that the grammars are mutually incompatible and therefore the examples would not affect the dynamics of change. Case syncretism as seen in (9a) is widespread in the language and these tokens provide the relevant data for the child. The indefinite theme variant of CP (9a) is produced like NP (9b) but with an indefinite theme. This does not violate the definiteness effect and therefore such examples are compatible with both grammars. About 50% of the relevant passives are advantageous to the NP, so we estimate the NP’s evolutionary advantage as  $\alpha \approx 0.5$ .

There are also cases where the Canonical Passive enjoys an advantage over the New Passive. The advantage of the CP,  $\beta$ , comes from examples of type (11a) where a constituent has been fronted and an indefinite theme appears in the V2 subject position. Because Icelandic is a verb second language, the subject appears after the auxiliary when any other element appears in first position.

- (11) a. Stundum var barn lamið.  
 sometimes was **child.NOM.ACC** beaten  
 Sometimes, a child was beaten. (✓CP, \*NP)  
 b. Stundum var lamið barn.  
 sometimes was beaten **child.NOM.ACC**  
 Sometimes, a child was beaten. (✓CP, ✓NP)

This **XP>auxiliary>indefinite Theme** word order is advantageous to the CP because the theme is an object in the NP and must remain low as in (11b). This high subject position is not compatible with the NP. Again, examples which lack case syncretism do not affect the dynamics of change. The definite theme parallel of (11a) does not affect the change since the two output variants of the respective grammars are mutually incompatible. Sentence type (11a) which is advantageous to the CP occurs in about 5% of relevant passives so we estimate that  $\beta \approx 0.05$ .

The advantage values are approximations and they can be refined. A more representative child-input corpus could be used and certain nuances could be worked out for some of the examples. However, the crucial fact is that by an overwhelming difference,  $\alpha > \beta$ , here  $0.5 > 0.05$ . In any Icelandic corpus,  $\alpha > \beta$  will hold, and therefore the New Passive is predicted to beat the Canonical Passive and rise to complete acceptance, which is consistent with reports of its rapid spread. We should note that Yang (2002) is not a model of actuation so the spread prediction does assume that the new variant has been introduced. We will turn back to the actuation problem below, but first we discuss implications for the rate of change.

## 5 Rate of Change

In this section we show how our approach can be used to develop analyses about the rate of change. This type of analysis is more sensitive to having exact measures of  $\alpha$  and  $\beta$  from language acquisition input, but we can proceed on the basis of our current values for now. Empirical refinements of the two values can be inserted in the equations below to obtain a more precise prediction. Importantly, the shape of the curve is derived from monolingual corpus data which means that we are not fitting a curve to an already observed change.

We use  $p$  and  $q$  to indicate the usage probabilities of the two grammars in a given generation. According to result (8), the next generation of speakers will have usage probabilities  $p'$  and  $q'$ .

(12)

$$p' = \frac{p}{p + (1-s)q} \quad q' = \frac{(1-s)q}{p + (1-s)q}$$

The equations in (12) give us the predicted usage probabilities in the following generation in terms of the current generation probabilities and a value  $s = (\alpha - \beta/\alpha)$  (see previous section). The value  $s$  reflects the relative difference in evolutionary advantage between the two grammars. Using this result, we divide  $q'/p'$  and obtain the following.

(13)

$$\frac{q'}{p'} = (1-s)\frac{q}{p}$$

From (13) we can see that the predicted value of  $q/p$  in the subsequent generation can be calculated by multiplying the value of  $q/p$  in the current generation with  $(1-s)$ . This means that for any generation  $t$ , the following holds:

(14)

$$\frac{q_t}{p_t} = (1-s)^t \frac{q_0}{p_0}$$

Solving for  $t$  gives an expression for the number of generations it takes for given initial values of  $p_0$  and  $q_0$  to change to some other values  $p_t$  and  $q_t$  at time  $t$ , for the value of  $s$  that comes from the penalty probabilities of the two grammars.

(15)

$$t = \frac{\log \frac{q_t}{p_t} - \log \frac{q_0}{p_0}}{\log(1-s)}$$

We can insert our estimates for  $\alpha$  and  $\beta$  to compute the value of  $s$ . If the advantage of the NP is  $\alpha = 0.5$  and the advantage of the CP is  $\beta = 0.05$ , then  $s$  can be calculated as below.

(16)

$$s = \frac{\alpha - \beta}{\alpha} \approx \frac{0.5 - 0.05}{0.5} \approx 0.9$$

The Variational Model is not a model of actuation so it does not explain how 0% usage of the new grammar becomes 1%. Due to the mathematical properties of the S-shaped curve, the last stages of the change going to completion are not accounted for. The curve only approaches 100% usage of the winning variant, and does so into infinity. Nevertheless, changes do go to completion, so a grammar which is being replaced can at some point go out of use entirely. We can simplify the picture by paying attention to the transition from 1% usage to 99% usage. Let  $p$  rise from 1% to 99%. In the very beginning,  $p_0 = 0.01$  and  $q_0 = 0.99$ , but at the point when the NP grammar has taken over, at time  $t$ , the probabilities are reversed,  $p_t = 0.99$  and  $q_t = 0.01$ . Inserting these values and  $s$  from (16) into (15), we obtain the following.

$$t \approx -4.0/\log(1-s)$$

$$t \approx -4.0 / \log(0.1) = 4$$

The conclusion is that given the current assumptions, we predict the change to go to completion in 4 generations, which amounts to about 100 years. If the change started in 1950, language acquisition of a child now should result in around 60% usage of the NP. This seems compatible with reports in the literature (Thráinsson et al. forthcoming). Moreover, in the absence of independent developments, children should no longer acquire the CP around 2050. More work remains to be done in terms of refining the envelope of relevant examples and extending the comparative dimension of this work to a wider range of empirical cases. However, the pleasing aspect of this result is that the shape of the curve is derived from a monolingual corpus. The ultimate importance of such a prediction is not whether our current estimates are sufficiently accurate, but rather that we have a line of inquiry that derives falsifiable S-curves instead of fitting them to changes that already took place. Our work addresses the lack of proposals for a change mechanism in linguistics noted by Kroch (1989).

[...] given the mathematical simplicity and widespread use of the logistic, its use in the study of language change seems justified, even though, unlike in the population genetic case, no mechanism of change has yet been proposed from which the logistic form can be deduced.

The spread of the NP is predicted to follow a specific S-shaped curve which is derived from the hypothesized causes of particular acquired probabilities across generations. Our causal account is therefore not a mere statistical description of what happened. Before concluding, we discuss the implications of our work for the actuation of the change.

## 6 Actuation

It is very challenging to explain why linguistic change starts when it does rather than at some other time in history. The “actuation problem” (Weinreich et al. 1968) cannot be fully solved by a model that focuses on the trajectory of a change that has already started. However, since the New Passive has an enormous advantage over the Canonical Passive, it is interesting to consider whether  $\alpha > \beta$  was always true. We can entertain the possibility that a change which enjoys such a large evolutionary advantage is likely to take off in a population, whichever the reasons for the initial innovation.

Curiously, and apparently casting doubt on our model, the advantage of the NP remains stable over the course of history if the grammar of Icelandic is otherwise stable. The IcePaHC corpus (Wallenberg et al. 2011) allows us to break the distribution of passives down by period, and no interesting changes in the frequencies behind  $\alpha$  and  $\beta$  take place. However, the grammar of Icelandic was not stable, so the assumption of a modern grammar in earlier times is not valid. Crucially, there was no categorical definiteness effect until the 20th century. The example in (17) is from the year 1902. It is retrieved from the IcePaHC corpus.

- (17) ... rétt eftir að farin var vöruferðin  
 ... right after that gone was **the.product.trip**  
 ‘... right after they went shopping’

The definite theme *vöruferðin* ‘the product trip’ is in a low position that is not available in current Icelandic. This leaves a subject gap that allows the participle *farin* ‘gone’ to be stylistically fronted to the first position in the sentence (Maling 1980, Holmberg 2006). Examples of this type are widely attested in the historical record and the phenomenon has been described as a “leakage” of the definiteness effect (Eythórssón 2008). However, in the early 20th century, these examples disappear and (17) is in fact the last violation of the definiteness effect in IcePaHC which does not involve a super-heavy theme.

Eythórssón (2008) suggests that the actuation of the NP change involved reanalysis of examples like (17). Under this view, the leakage of the definiteness effect was an important source of the rise

of the NP. This explanation does not seem to fit well with the attested time line. If this explanation was true, we would expect to see an increased leakage of the definiteness effect in the years leading up to the NP. This is the exact opposite of what happens. The definiteness effect seems to become a categorical constraint in the early 20th century, just before the New Passive is invented.

Under our analysis, examples of type (17) need to be ungrammatical under a Canonical Passive analysis for the New Passive to have any chance of spreading at the expense of the CP; see (9b). The NP owes its advantage to the definiteness effect. Therefore, it was the rise of the definiteness effect, not its leakage, that created favorable conditions for the spread of the New Passive. While this does not explain why the first NP speaker acquired the new grammar, it does predict that such an innovation had no chance of spreading before the 20th century. In conclusion, the model restricts the set of actuation events that have potential to successfully start a change.

## 7 Conclusion

The Icelandic New Passive is a useful test case for theories of syntactic change. It is a major syntactic change, which involves reorganization of abstract clausal syntax. In contrast to many of the best studied cases of major changes in syntax, the NP change is a change in progress. The NP is therefore a good place to develop approaches to the S-curve which go beyond fitting curves to attested data. A main goal of this study is to promote an approach to linguistic change where curve-fitting is replaced with deriving change from hypotheses about language acquisition.

We applied the Variational Model (Yang 2002) to the New Passive. The model correctly predicted that when both analyses are used in a linguistic community, the NP should gain ground at the expense of the CP. We showed how the same approach can be extended to further generations to make predictions about the rate of change. Our current estimates predict that the first speakers who do not acquire the Canonical Passive will be born around 2050. Furthermore, we showed that the evolutionary advantage of the New Passive crucially relies on a categorical definiteness effect. This result is welcome in light of the fact that the NP started spreading shortly after a strengthening of the Icelandic definiteness effect.

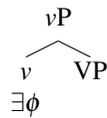
It is desirable to find out whether predictive modeling in historical linguistics is possible. The most efficient way to make progress is to make a precise formal proposal that has consequences for a wide range of empirical cases. The mechanism of change which is hypothesized in this paper can be considered for any syntactic change, although it is possibly only appropriate for types of abstract structures that have limited social impact in the sense of Labov and Harris (1986:21). We hope that the steps taken in this paper will inspire other researchers to do further work where the evolutionary trajectory of change is predicted based on facts about the environment that children grow up in.

## 8 Appendix: Syntactic Analysis

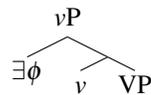
In this appendix, we provide a short outline of our syntactic analysis of the New Passive (see also Ingason et al. 2012a). The literature on the syntactic analysis of the Icelandic New Passive has focused on whether to classify the NP as an active (Maling and Sigurjónsdóttir 2002) or as a passive (Eythórsson 2008, Jónsson 2009). Empirically, the NP shares properties of active and passives. The task is therefore to find the most appropriate analysis to capture the different properties with independently motivated theoretical tools. We do this by placing the NP in a typology of passive-like constructions.

Different types of silent arguments are independently needed in the syntax. Employing these different types in the analysis of the NP accounts for the facts without construction specific machinery. Under our analysis, the properties of the New Passive follow from the subject being smaller than a silent pronoun, a weak implicit argument (WIA; Landau 2010). The WIA in the NP restricts, but does not saturate, the subject position (in the sense of Chung and Landau 2004). In this respect, we agree with Sigurðsson's analysis (2011:174), although in other respects our analysis differs from his. The typology we propose is outlined below.

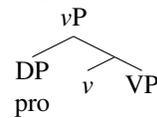
## (18) Canonical Passive



## New Passive (NP)



## Impersonal (Polish-type)



Consider (18). In the first type, a Canonical Passive, the subject position is existentially bound, [spec, *v*] is empty, and *v* bears features restricting the subject position which may be morphologically overt. (See e.g., Legate 2012 on Acehnese and Chamorro). In the second type, the New Passive (and the analogous Ukrainian: Sobin 1985), the subject position is also existentially bound, [spec, *v*] hosts features restricting the subject position, and an XP smaller than a pronoun (WIA) restricts rather than saturates the subject position. The presence of a subject in [spec, *v*] accounts for the licensing of accusative case on the theme. However, the subject not saturating the argument position accounts for the availability of *by*-phrases, the subject's inability to undergo A-movement (allowing raising across it: Jónsson 2009), and a reluctance to license depictives (Landau 2010, Jónsson 2009). In the third type, an Impersonal of the type attested in Polish, a full null pronoun in [spec, *v*] saturates the subject position.

Our analysis appropriately captures the differences between empirically attested passive-like constructions. It does so without positing new parameters and without placing any special emphasis on non-primitive labels like 'active' and 'passive'. Nevertheless, we should note that the details of the syntactic analysis are not important for the dynamics of change which are the main topic of the paper. Our model of change operates on empirical generalizations and is not sensitive to the inner workings of the two incompatible analyses, as long as there is conflicting evidence in the linguistic environment.

## References

- Árnadóttir, Hlíf, Thórhallur Eythórsson, and Einar Freyr Sigurðsson. 2011. The passive of reflexive verbs in Icelandic. *Nordlyd* 37:39–97.
- Barðdal, Jóhanna, and Valéria Molnár. 2003. The passive in Icelandic: Compared to mainland Scandinavian. *Structures of Focus and Grammatical Relations* 65:231–260. Ed. Jorunn Hetland and Valéria Molnár. Tübingen: Max Niemeyer Verlag.
- Bell, Allan. 1984. Language style as audience design. *Language in society* 13:145–204.
- Bush, Robert R, and Frederick Mosteller. 1951. A mathematical model for simple learning. *Psychological Review* 58:313.
- Chung, Sandra, and William Allen Ladusaw. 2004. *Restriction and saturation*, volume 42. MIT press.
- Eythórsson, Thórhallur. 2008. The New Passive in Icelandic really is a passive. *The Rosendal Papers* 173–219. Ed. Thórhallur Eythórsson. Amsterdam and Philadelphia: Benjamins.
- Holmberg, Anders. 2006. Stylistic fronting. *The Blackwell companion to syntax* 532–565.
- Ingason, Anton Karl. 2010. Explaining case variation. A Death Rattle Hypothesis for Minority Rules. A talk given at Scandiasyn/N'CLAV Grand Meeting, Sommaroya, Tromsø, June 8, 2010.
- Ingason, Anton Karl, Julie Anne Legate, and Charles Yang. 2012a. Structural and Evolutionary Basis of the Icelandic New Impersonal Passive. A paper presented at Non-Canonically Case-Marked Subjects within and across Languages and Language Families, Iceland, June 4-8, 2012.
- Ingason, Anton Karl, Einar Freyr Sigurðsson, and Joel Wallenberg. 2012b. Antisocial Syntax. Disentangling the Icelandic VO/OV parameter and its lexical remains. A paper presented at DIGS 14, University of Lisbon, July 4-6, 2012.
- Jónsson, Jóhannes G. 2009. The new impersonal as a true passive. *Advances in comparative Germanic syntax* 281–306.
- Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.
- Labov, William. 1972. *Sociolinguistic patterns*, volume 4. Univ of Pennsylvania Press.
- Labov, William, and Wendell A Harris. 1986. De facto segregation of black and white vernaculars. *Diversity and diachrony* 53.
- Landau, Idan. 2010. The explicit syntax of implicit arguments. *Linguistic Inquiry* 41:357–388.

- Legate, Julie Anne. 2012. Subjects in Acehnese and the nature of the passive. *Language* 88:495–525.
- Maling, Joan. 1980. Inversion in embedded clauses in Modern Icelandic. *Íslenskt mál og almenn málfræði* 2:175–193.
- Maling, Joan, and Sigríður Sigurjónsdóttir. 2002. The new impersonal construction in Icelandic. *The Journal of Comparative Germanic Linguistics* 5:97–142.
- Rögnvaldsson, Eiríkur, Anton Karl Ingason, and Einar Freyr Sigurðsson. 2011a. Coping with variation in the Icelandic Diachronic Treebank. *Oslo Studies in Language* 3.
- Rögnvaldsson, Eiríkur, Anton Karl Ingason, Einar Freyr Sigurðsson, and Joel Wallenberg. 2011b. Creating a dual-purpose treebank. *JLCL* 26:139–150.
- Rögnvaldsson, Eiríkur, Anton Karl Ingason, Einar Freyr Sigurðsson, and Joel Wallenberg. 2012. The Icelandic Parsed Historical Corpus (IcePaHC). In *LREC, 1977–1984*.
- Sigurðsson, Einar Freyr. 2012. Germynd en samt þolmynd. Um nýju þolmyndina í íslensku [Active but still passive: On the New Passive in Icelandic.]. Master's thesis, University of Iceland.
- Sigurðsson, Halldór Á. 2011. On the new passive. *Syntax* 14:148–178.
- Sigurðsson, Einar Freyr, and Brynhildur Stefánsdóttir. 2013. 'By'-phrases in the Icelandic New Impersonal Passive. Paper presented at the 37th Penn Linguistics Colloquium.
- Sobin, Nicholas J. 1985. Case assignment in Ukrainian morphological passive constructions. *Linguistic Inquiry* 16:649–662.
- Thráinsson, Höskuldur, Sigríður Sigurjónsdóttir, Hlíf Árnadóttir, and Thórhallur Eythórsson. Forthcoming. Um þolmynd, germynd og það. [On passive, active and there]. *Tilbrigði í íslenskri setningagerð II: Helstu niðurstöður. Tölfræðilegt yfirlit með skýringum*, ed. H. Thráinsson, Á. Angantýsson and E. F. Sigurðsson. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson, and Eiríkur Rögnvaldsson. 2011. Icelandic Parsed Historical Corpus (IcePaHC). Version 0.9. Size: 1 million words. URL [http://www.linguist.is/icelandic\\_treebank](http://www.linguist.is/icelandic_treebank).
- Weinreich, Uriel, William Labov, and Marvin I Herzog. 1968. *Empirical foundations for a theory of language change*. University of Texas Press.
- Wood, Jim. 2013. The Syntactic Underpinnings of Passive Voice. A talk given at the Speaker Series of the University of Pennsylvania.
- Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

Department of Linguistics  
 University of Pennsylvania  
 Philadelphia, PA 19104–6305  
[ingason@ling.upenn.edu](mailto:ingason@ling.upenn.edu)  
[jlegate@ling.upenn.edu](mailto:jlegate@ling.upenn.edu)  
[charles.yang@ling.upenn.edu](mailto:charles.yang@ling.upenn.edu)