



# Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

## Review of [Oscar Strik: Modelling Analogical Change. A History of Swedish and Frisian Verb Inflection; (dissertation Groningen, 2015)]

Merkuur, A.A.

### **published in**

Us Wurk. Tydskrift foar frisistyk  
2016

### **document version**

Publisher's PDF, also known as Version of record

### **document license**

CC BY-NC-SA

[Link to publication in KNAW Research Portal](#)

### **citation for published version (APA)**

Merkuur, A. A. (2016). Review of [Oscar Strik: Modelling Analogical Change. A History of Swedish and Frisian Verb Inflection; (dissertation Groningen, 2015)]. *Us Wurk. Tydskrift foar frisistyk*, 65(3-4), 190-193.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[pure@knaw.nl](mailto:pure@knaw.nl)

## BESPREK

[1330][Besprek] **Oscar Strik, Modelling Analogical Change. A History of Swedish and Frisian Verb Inflection.** Rijksuniversiteit Groningen dissertation (Groningen 2015). ISBN 9789036781527 (printed version), 978903678150 (digital version). 319 pp.

‘This book is about two things: analogy, and the past tense of verbs in Germanic languages’. Thus reads the first sentence of Oscar Strik’s 2015 dissertation *Modelling analogical change*. In short, this dissertation aims to answer the questions how and why verbs change over time, by modelling changes in the verbal inflection of Frisian and Swedish with the aid of two computer models. Strik has a background in Scandinavian and historical linguistics, two fields which he combines with a computational approach in this dissertation.

Of course, before modelling, it needs to be established which historical changes have taken place. This forms a large part (almost half) of the dissertation. In chapters 3, 4 and 5 Strik gives a clear overview of the history of verbal inflection classes in Germanic languages in general, and of Frisian and Swedish in particular. In all three chapters Strik specifically focusses on changes involving inflection classes, which he defines quite loosely as groups of verbs that inflect in rather the same way. According to Strik, inflection classes are, first and foremost, descriptive tools for grouping the verbs together, rather than psychological realities. This focus on changes in inflection (classes) is also one of the main reasons why Strik has chosen Swedish and Frisian, as both languages not only have multiple strong inflection classes but also have preserved multiple classes of weak inflection. Therefore, verbs in these languages may not only have shifted from a weak to a strong inflection class or from a strong to a weak one, but also from one strong class to another or from one weak class to another. In chapters 4 and 5, Strik maps all these changes, as well as some changes in inflectional classes themselves, and thereby provides an informative overview of the developments in both inflection systems.

After describing the historical changes, the latter half of the dissertation is reserved for accounting for these changes. Strik uses analogical computer modelling as a method because his main hypothesis is that ‘analogies, when applied systematically to the verb systems of the Germanic languages, can explain a great deal of the inflection class shifts that took place, as well as those that didn’t take place’ (p.4). Chapter 6 is dedicated solely to the concept of analogy and a description of the two models that Strik applies to the Swedish and Frisian data. A

good choice, for the notion of analogy can be quite vague and can give rise to all kinds of philosophical reflections. In line with Fertig (2013), who has dedicated a whole book to analogy, Strik considers analogy in a linguistic sense to mean ‘the capacity of speakers to produce meaningful linguistic forms that they may have never before encountered, based on patterns they discern across other forms belonging to the same linguistic system’ (2013: p.12). After arguing that inflection class shifts can be described as forms of analogical change, Strik rightly raises the question whether describing inflection class shifts as analogical change actually helps us to understand what has happened to a verb. Or in other words, does the working of analogy really explain why some verbs have shifted from one class to another?

In order to answer these questions and to find out how the concept of analogy can be formalized, he chooses to work with two computational models that operationalize the working of analogy in different manners: The Minimal Generalization Learner by Albright & Hayes (2002, 2003), which is more rule oriented, and The Analogical Modeling program (Skousen, 1989), which is more analogy based. Both of these (and other) models work as computer programs which generate output on the basis of certain hypothetical or assumed inputs. The program or learner resembles a language learner who learns and produces language (output) on the basis of the language he hears (input). The first step in modelling is to feed the programs with data: the input. In the case of verbal inflection such input may exist of present-past verb pairs: work - worked. The model extracts information from this input and can then be tested. The test data usually consists of the first part of such pairs. On the basis of pairs such as *work - worked*, the model might produce the past tense *walked* for the element *walk*.

Providing the model with input data is a crucial step in modelling, because it always forms part of the hypothesis one wants to test: based on an input X, we expect model A to generate an output Y. The aim in Strik’s case is to model the changes that took place from Old to Early Modern to Modern Swedish and from Old, to Early Modern to Modern Frisian. In both cases these are periods of at least 500 years, which Strik models in two steps. Because analogical models are said to resemble a language learner, the question is whether these steps are not too large. In an ideal situation one would model from a hypothetical language learner of generation 1 to a hypothetical language learner of generation 2 and so on.

The other element of choice when modelling language are the models themselves. Over the past decades, several learner models have been developed, each with their own underlying mechanisms. The two models Strik chooses are minimally different. Both of them use systematic comparison of all forms in the input, in order to find analogical patterns, but where the AM program derives outcomes by directly comparing a form to similar forms, the MGL describes all the patterns it can find with rules. These rules get their own reliability and confidence scores, depending on how large their scopes are and how often they apply. The MGL way

of operationalizing analogy is therefore more structural and it is not without reason that Albright & Hayes themselves consider it to be a rule-based learner program.

The results from modelling Frisian and Swedish verbs showed that, although the individual predictions do differ sometimes, on average, both models are able to predict about 60 up to 80% correctly. To interpret these results, Strik gives several baseline models to compare the results with. One of these baseline models is the no-change model which predicts that nothing changes. It turns out that in fact, this no-change model often does just as well or even better than both of the analogical models, which turn out to be quite powerful and in fact predict too many changes. A baseline model which only predicts regularization and no irregularization actually does worse than the analogical models, thereby indicating that there is more to change than regularization only and that there are effects of phonological analogy. A comment on the results is that they might have suffered from the rough modelling alignment (approximately 500 years in two steps). Future research on modelling historical changes should pay more attention to the acquisitional aspects of change, and try and model time in a more fine-grained manner by going from the output of generation 1 to the output of generation 2, and so on. Even though this would be time consuming, it would be interesting to see if it leads to better predictions.

A bonus to the dissertation is chapter 9 on 'Frequency and analogy in Early Modern Frisian verb inflection', which is based on research Strik has conducted together with Arjen Versloot, one of his supervisors. Based on the results of modelling the change from Early Modern Frisian to Modern Frisian, they show that token frequency seems to interact with analogy in various manners. Not only is a high token frequency a conservative force in general, but they find that a verb's frequency compared to the average frequency does the trick: 'the closer a verb's frequency is to the average token frequency of its current class, the more likely that an analogical pressure towards change is overridden by the conservative influence of that frequency' (p.184). The authors plan to further look into this matter in the future.

What is special about this dissertation is that the computational and the historical linguist have worked together. The book clearly shows that both fields may benefit from each other. Historical linguistics could make more use of computational models, and computational linguistics could take the historical development of languages more into account. Of course, the downside is that both disciplines have had to give up some space for the other. The element missed is the more acquisitional side of the matter. As mentioned earlier, an objection is that the two models are computational imitations of language learners, but are not used that way in this thesis, because of the rough modelling alignment. Furthermore, although notions as productivity, regularity, rules and analogy are discussed and take a central place in this research, the author takes quite a broad starting point by hypothesizing that analogy plays a large role. The reader does not get a clear view

on how the author hypothesizes verbal inflection to ‘work’. The implications of the modelling results, and especially the differences between the models in their ways of operationalizing analogy, are discussed in the concluding chapter, but this part of the thesis deserves more attention. In that regard, the additional chapter 9 also makes one curious as to how the interplay between analogy and frequency actually works and whether it indeed applies to all results. In his own recommendations for future research, Strik does not want to zoom in on the notion of analogy, but aims to zoom out as he wants to broaden the field, arguing that sociological and geographical factors should be taken into account more as well. He envisages the eventual development of a model in which ‘different layers of influence on the verb system are combined’ (p.185), such layers being not only frequency and analogy, but also demographic and social factors.

To conclude, Strik’s research provides enough information for historical, computational or other interested linguists to learn something new, not only about the developments of Frisian and Swedish verbal inflection, but also about how to use analogical models for modelling historical developments. The role of analogy in changing verbal inflections, the main subject of this research, is a tough problem to deal with. Especially because of the similar results between the two modelling programs, the question remains which place analogy as a cognitive concept takes in the variety of mechanisms related to language change and how this can be formalized best: how much analogy is enough and when and how does it play a role?

*Anne Merkuur*  
*Fryske Akademy*  
*Ljouwert*

#### References

- Albright, Adam, and Bruce Hayes. “Modeling English past tense intuitions with minimal generalization.” In *Proceedings of the ACL-02 workshop on morphological and phonological learning*, vol. 6, pp. 58-69. Association for Computational Linguistics, 2002.
- Albright, Adam, and Bruce Hayes. “Rules vs. analogy in English past tenses: A computational/experimental study.” *Cognition* 90, no. 2 (2003): 119-161.
- Fertig, David. *Analogy and morphological change*. Edinburgh University Press, 2013.
- Skousen, Royal. *Analogical modeling of language*. Springer Science & Business Media, 1989.