Constructional Change Based on Cross-linguistic Analogy Building in North American Danish

Karoline Kühl

1. Introduction

Several studies of Germanic immigrant and heritage languages in North America note changes in and simplification of the passive system, e.g. Putnam & Salmons (2013) on the loss of the periphrastic passive in Moundridge Schweitzer German. Earlier, Hasselmo (2005: 2133) notes as a feature of North American Scandinavian syntax “the use of vara ['to be'] and its other Scandinavian equivalents in forming the passive of verbs denoting "transition from one state to another" ("övergangsverb"). This is corroborated by Larsson, Tingsell & Andréasson (2015: 378) who observe that North American Swedish periphrastic passives appear to be affected by English in that eventive passives are formed with the auxiliary vara ‘to be’, again at the expense of either bli ‘to become’ or a morphological passive. Hasselmo’s work appears to be the only overview paper with a focus on all three Mainland Scandinavian languages, but, to the author’s best knowledge, no contemporary publications on the passive formation in North American Danish (NAD) existed that Hasselmo could possibly have referred to. This implies that there is good reason for a systematic investigation of auxiliary choice in NAD passives. On a more general level, this study adds evidence to the gradual nature of language change, with a special focus on cross-linguistically motivated and analogy-based changes.

The paper is structured as follows: Section 2 provides an overview of auxiliary choice in Danish and defines the object of study. Section 3 describes the design of the present study and presents the rationale for choosing a Construction Grammar approach for the analysis. Section 4 describes the data and the data processing procedures, while Section 5 provides the detailed analyses. Section 6 summaries the findings, concludes and adds perspectives to the results.

2. Auxiliary choice in European and North American Danish

Auxiliary selection is central to Denmark Danish periphrastic passives: The choice of være ‘to be’ or blive ‘to become’, respectively, expresses semantic opposition of telicity. The two members of the passive auxiliaries’ paradigm constitute a grammatical and categorical opposition (Hansen & Heltoft 2011: 630; Juul Nielsen 2016: 396). The choice of auxiliary differentiates between eventive (telic or atelic) passives formed with blive and stative passives constructed with være cf. the examples (1) and (2).

(1) Soldat-en bliv-er dræb-t.
soldier-DEF become-PRES kill-SUP
‘The soldier is (being) killed.’ Eventive passive

(2) Soldat-en er dræb-t.
soldier-DEF be.PRES kill-SUP
‘The soldier is killed/has been killed.’ Stative passive

* Karoline Kühl, University of Copenhagen, karoline.kuehl@hum.ku.dk. Many thanks to Jan Heegård Petersen and Steffen Höder for discussion of earlier versions of this paper, to Joe Salmons, an anonymous reviewer and the audience at WILA 8 (Copenhagen, October 2017) for useful input.

If VÆRE ‘to be’ is combined with a telic transitional verb as kill in (2), the construction expresses the state that is reached after the transition designated by the supine (here, the state of being killed, ‘dead’) (Juul Nielsen 2016: 393-396).1 Obviously, English and Danish differ with respect to auxiliary choice in forming the periphrastic passive and the intense contact between the languages in the migration setting suggests itself as an explanation for an ‘overuse’ of VÆRE in NAD as well as the overuse of VARA in NA Swedish stated by Larsson, Tingsell & Andréasson (2015).

Yet, Kühl & Heegård Petersen (2017), a previous study on auxiliary choice in NAD, shows a not so straightforward picture: Change in auxiliary choice and use in NAD is indeed present and it favors VÆRE ‘to be’ over BLIVE ‘to become’. However, change is restricted to past tense (i.e. it favors var ‘was’ over blev ‘became’; note that verbs in Danish are not inflected for person) and it occurs mainly with the supines født ‘born’, konfirmeret ‘confirmed’ and gift ‘married’. The change is not limited to these three verbs, but these are the only ones that emerge in a substantial number of tokens allowing for generalizations (cf. Section 5 and 6). Still, the fact that other NAD verb phrases also show a non-standard auxiliary choice and use (as compared to Danish Danish) proves the general pattern of change to be potentially productive. But as the other afflicted verbs do not share semantics with født, konfirmeret and gift, there is no indication of an overall (however motivated) change in NAD. The finding that mainly født, konfirmeret and gift are subject to change is consolidated by the present study, despite integrating more data compared to Kühl & Heegård Petersen 2017. Therefore, I will limit the following analyses to these three verbs in combination with either blev ‘became’ or var ‘was’.

3. The design of the study

Methodologically, the present study is a replication of Kühl & Heegård Petersen (2017) in that it relies on corpus-based analyses following the tradition of variationist sociolinguistics in a combination of quantitative and qualitative analyses. Different from Kühl & Heegård Petersen (2017) is the amount of data which has been extended considerably (cf. Section 4) and the theoretical focus, now following a Construction Grammar (CxG) approach. Despite different schools of CxG with different aims and objectives, the assumed basic unit of language is a construction. The term denotes a symbolic unit, a pairing of form and meaning which displays structural, semantic and/or pragmatic idiosyncrasies or a high level of entrenchment.2 It is assumed that a construction is stored as a unit mentally by speakers due to its idiosyncrasies and entrenchment (see, among others, Goldberg 2006), but it is well acknowledged that constructions may change: “Constructional change seizes a conventionalized form-meaning pair of a language, altering it in terms of its form, its function, its frequency, its distribution in the linguistic community, or any combination of these” (Hilpert 2013: 3, see also Hilpert 2011).

The concept of constructional change is able to cover the selective nature of change in NAD that is described in this paper, namely the fact that almost only the transitional verbs født ‘born’, konfirmeret ‘confirmed’ and gift ‘married’ are affected by the contact with English (not the passive system nor auxiliary choice as a whole). Further, a CxG approach takes the similarities and differences of both form and function into account that would not necessarily be recognized by a strictly structural approach. The latter seems to be important with respect to the fact that the three verb phrases differ with respect to their traditional grammatical characterization as passives: konfirmere is a textbook passive where auxiliary choice denotes either the confirmation act or the state of being confirmed. With regard to født, the picture becomes more complicated: [patient + auxiliary + supine født + of agent] technically is a passive of the active [agent + føde + patient]. Yet, the passive construction with a focus on the ‘undergoer’ and no mentioning of the agent (the mother) appears to be the default (most entrenched) expression in Danish. One may ask, thus, if the verb actually represents a balanced active-passive distinction with respect to its use and, accordingly, mental entrenchment. Være gift and blive gift are not passives in the true sense,

---

1 Danish has other options of passives, too, a morphological passive and periphrastic passive construction with FÅ ‘to get’ + supine. The latter is similar to get-passives in English. However, these constructions occur rarely in the NAD data set and they do not seem to be affected by contact. Therefore, I will not address them in this paper.

2 However, see e.g. Ziegeler (2015) for critical comments on the heterogeneous nature and remarkable differences in the definition of ‘constructions’.
either, as they do not have active counterparts. Here, gift is a predicative and, accordingly, BLIVE and VÆRE must be defined as copula verbs. Nevertheless, the use of live and være gift yields the same opposition between the eventive and the stative reading as in periphrastic passives and therefore gift is included in the present paper.

To sum up: The focus of this paper will be on instances of [blev|var] + [født|konfirmeret|gift]. The reason to (tentatively) define these verb phrases as constructions despite their different characterizations in traditional grammar is a) that they form conventionalized pairs of form and meaning and, more importantly, b) that they seem to be treated as constructions by the speakers, i.e. that they are altered along similar lines (see Section 5 for an in-depth discussion). The notion of constructional change makes it possible to account for the fact that changes seem to be by and large limited to these three lexemes while at the same time allowing for the similarities in form and meaning between these verb phrases. The question whether the verb phrases should be considered instances of the same construction or three individual constructions will be discussed in Section 6. It remains to be said that constructional change and language change is a matter of degree: Changes affecting individual constructs (instantiations of constructions) might lead to changes in constructions and, ultimately, to systematic change (Hilpert 2013, Fried 2013).

4. Data and data processing

The present study relies on two corpora, namely the Corpus of North American Danish (CoNAmDa) and the LANCHART corpus of modern spoken Danish, both hosted by the University of Copenhagen. As of November 2017, the CoNAmDa contains 508,414 tokens of transcribed speech (orthographic words as well as non-word utterances) aligned with sound (cf. Kühl et al 2017 for a comprehensive account of the corpus). The corpus represents speech produced by approximately 280 immigrant or heritage speakers of Danish in the USA and Canada, born between 1871 and 1941, and recorded between 1966 and 1998 in loosely structured interviews. Most speakers in the CoNAmDa are immigrant speakers (n=151) who migrated before 1930, typically as adults with a fully acquired Danish language competence. The remaining speakers are heritage speakers born in North America into different immigrant generations. None of classic sociolinguistic variables nor factors such as immigration generation show any effect in the present study, neither do real time nor apparent time analyses.

The LANCHART corpus of modern spoken Danish consists of sociolinguistic interviews with speakers of all generations from different locations in Denmark (Gregersen & Kristiansen 2015). For want of a corpus of historical Danish covering the time of emigration of the CoNAmDa-speakers, the LANCHART corpus is used as control data. As of November 2017, the LANCHART corpus contains approximately 10.5 million tokens, with the Corpus of American Danish (North American as well as Argentine Danish) making up for approximately 1.6 million of the tokens.

Both corpora were searched for the supine forms født, konfirmeret and gift (as change in auxiliary choice appears to be widespread only with these supines, cf. section 2). Subsequently, auxiliary choice as well as the temporal inflection was annotated. In a next step, the textual environment of the VP, the co-text (Bergs & Diewald 2009: 1–14), has been rated. The co-text ultimately differentiates between an eventive or a stative reading: An eventive meaning is likely if VP occurs together with specifying adverbials as e.g. specific points of time (a date, a month, a year), specific places (in church, in a specific town), with an agent (e.g., married by a Danish pastor) or in a specific manner (e.g., confirmed in Danish). Such co-texts call for an eventive reading, i.e. coded with the use of BLIVE ‘become’ in European Danish. A stative meaning is obviously intended if the VP is combined with adverbials denoting a long-time span (e.g., married for 42 years) or other information that makes a predicative reading (e.g. married vs. unmarried) plausible.

3 The CoNAmDA contains 1st-4th immigrant generation speakers: 1st generation denotes immigrant speakers (n=151); 2nd generation represents a speaker born in North America with both parents born in Denmark (n=59), 3rd generation represents speakers with (at least) one parent born in North America (n=27) and 4th generation denotes speakers with (at least) one grandparent born in North America (n=2).
5. Analysis

The quantitative spread of auxiliary use with the supine form født, konfirmeret and gift in the LANCHART corpus and the CoNAmDa is shown in an overview in Table 1.

Table 1: Relative frequencies of auxiliaries in past tense for the supine forms født, konfirmeret and gift in North American Danish and Denmark Danish.

<table>
<thead>
<tr>
<th></th>
<th>North American Danish (CoNAmDa)</th>
<th>Modern spoken Danish (LANCHART corpus)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>var født</strong></td>
<td>264 93%</td>
<td>93 29.4%</td>
</tr>
<tr>
<td><strong>blev født</strong></td>
<td>20 7%</td>
<td>223 70.6%</td>
</tr>
<tr>
<td><strong>var konfirmeret</strong></td>
<td>28 47%</td>
<td>1 1.4%</td>
</tr>
<tr>
<td><strong>blev konfirmeret</strong></td>
<td>34 53%</td>
<td>73 98.6%</td>
</tr>
<tr>
<td><strong>var gift</strong></td>
<td>83 25.4%</td>
<td>89 21.9%</td>
</tr>
<tr>
<td><strong>blev gift</strong></td>
<td>244 74.6%</td>
<td>317 78.1%</td>
</tr>
</tbody>
</table>

Table 1 shows pronounced quantitative disparity between NAD and Denmark Danish with regard to født and konfirmeret, gift, however, shows a balanced and unremarkable frequential distribution. I will go into details regarding each verb phrase in the following analyses.

5.1. var født

The relative frequencies of var and blev født, respectively, in the CoNAmDa and the LANCHART corpus are almost diametrically opposed. Accordingly, the differences prove to be statistically significant (chi-square-test, p=<0.0001), with the expected scores differing considerably from the observed frequencies, cf. Table 2.

Table 2: Observed and expected frequencies for var født and blev født in NAD and modern spoken Danish.

<table>
<thead>
<tr>
<th></th>
<th>North American Danish (CoNAmDa)</th>
<th>Modern spoken Danish (LANCHART corpus)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>var født</strong></td>
<td>264</td>
<td>93</td>
</tr>
<tr>
<td><strong>blev født</strong></td>
<td>20</td>
<td>223</td>
</tr>
<tr>
<td><strong>var konfirmeret</strong></td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td><strong>blev konfirmeret</strong></td>
<td>34</td>
<td>73</td>
</tr>
<tr>
<td><strong>var gift</strong></td>
<td>83</td>
<td>89</td>
</tr>
<tr>
<td><strong>blev gift</strong></td>
<td>244</td>
<td>317</td>
</tr>
</tbody>
</table>

The statistical significance implies that it is very likely that the distribution represents actual differences between NAD and Denmark Danish, i.e. that NA speakers in general indeed prefer var født, while the Danish speakers prefer blev født. Yet, qualitative in-depth analyses of NAD var født that include the co-text hardly show any ‘ungrammatical’ choices (i.e. combination of var født with a co-text that beyond doubt denotes a transition, viz. the birth) and neither is var født’s function defined by a use with particular co-texts. The majority of instances of var født both in the CoNAmDa and the LANCHART corpus occur in a textual environment that is ambiguous with respect to whether the verb phrase denotes the dynamic event of the birth (Denmark Danish blev født) or the stative predicative property of being alive. (3) and (4) show such occurrences with ambiguous co-texts.
The textual ambiguity provides little restriction to auxiliary use. This seems to open up the possibility for establishing group specific language norms, i.e. a change of prototype construction (cf. Hilpert 2013: 2). A quantitative (corpus-based) approach as shown above appears to be the only way to establish such cases of noncategorical changes.

The only factor besides the shift in relative frequency that has proved to be statistically significant in the use of NAD and Denmark Danish var født is the use of verb phrase together with 1. person singular personal pronoun (jeg var født ‘I was born’) contrasted with its use with other personal pronouns (chi-square-test, p=0.0123), cf. Table 3.

Table 3: The quantitative distribution of var født in combination with personal pronouns.

<table>
<thead>
<tr>
<th></th>
<th>Modern Danish (LANCHART)</th>
<th>North American Danish (CoNAmDA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed frequencies</td>
<td>Expected frequencies</td>
<td>Observed frequencies</td>
</tr>
<tr>
<td>Observed frequencies</td>
<td>Expected frequencies</td>
<td>Expected frequencies</td>
</tr>
<tr>
<td>jeg var født ‘I was born’</td>
<td>10</td>
<td>69</td>
</tr>
<tr>
<td>83</td>
<td>74.8</td>
<td>195</td>
</tr>
<tr>
<td>X var født ‘X (NOT I) was born’</td>
<td>74.8</td>
<td>195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>212.2</td>
</tr>
</tbody>
</table>

No obvious grammatical or semantic restriction of the use of var født together with a personal pronoun of 1. ps. sg. lends itself as an explanation (note that the verb does not inflect for person in Danish), but obviously NAD and Denmark Danish have developed differently.

We may conclude that NAD and Denmark Danish do show different usage patterns of var and blev født. It is mainly a shift in frequency, except for the factor ‘1. ps. sing.’ which can be characterized as a change of function, namely a wider applicability. The qualitative analysis showed a majority of ambiguous co-texts which creates a certain freedom of choice; The ambiguity seems to provide the NA speakers with the option of choosing (and thus establishing) the diasystematically supported option in NAD, namely var født, without offending clear linguistic norms (cf. Höder 2012 for an account of Diasystematic CxG). The shift in frequency is a clear example of cross-linguistic analogy-based reorganization of a construction (Fried 2013).

5.2. var konfirmeret

The spread of frequencies of var and blev konfirmeret ‘was/became confirmed’ also shows a clear difference for NAD and Denmark Danish with respect to auxiliary choice (cf. Table 1). Whereas the auxiliaries are almost equally distributed in NAD, blev clearly dominates Denmark Danish with only a single occurrence of var konfirmeret.

As mentioned above, konfirmer is a textbook example of an active-passive opposition and differences in telicity marked by auxiliary choice. However, while the state of being confirmed does not seem to mean much in today’s Denmark, historically it implied the state of adulthood with its rights and obligations. This is relevant for the CoNAmDa speakers, many of whom were born around 1900: 6 out of 28 occurrences of var konfirmeret in the CoNAmDa can be read predicatively as ‘to be grown up’, motivating the auxiliary choice (cf. (5) below). Actually, the only occurrence of var konfirmeret in the LANCHART corpus also is used for denoting the state when a teenager is allowed to drink alcohol (cf. (6) below):
(5) Vi måtte jo tjene så snart vi var konfirmeret og det var et hårdt liv. (speaker LAR; CoAmDA)

(6) Hvornår måtte du få lov til at drikke? ... Så det var sådan cirka efter I var konfirmeret eller? at I måtte begynde at ... (speaker XMM, LANCHART)

The remaining 22 instances of var konfirmeret in the CoNAmDa, however, occur together with specifying adverbials (time, place, manner and others) that point to a dynamic reading of the construction, cf. (7) and (8) below:

(7) Jeg var konfirmeret på dansk i Ringsted, Iowa. (speaker MFN, CoNAmDa)

(8) Jeg var konfirmeret by en gammel mand. (speaker AGL, CoNAmDa)

In the LANCHART corpus blev is the default choice together with co-texts that specifies the event. This implies that the NAD use of var konfirmeret in combination with a co-text that creates a focus on the event must be considered an effect of analogy building between Danish and North American English. NAD var konfirmeret has acquired a certain degree of polysemy as it denotes both the state and the event, unlike Denmark Danish. Moreover, and also unlike Denmark Danish, var and blev occur freqeuntly balanced together with konfirmeret in NAD, implying that NAD is different from Standard Danish with regard to its default choice of construction for expressing ‘was confirmed’. Thus, the case of konfirmeret shows change both in relative frequency and semantic-grammatical function as compared to modern spoken Danish.

5.3. var gift

The relative frequency of var and blev gift in CoNAmDa and the LANCHART corpus, respectively, is unremarkable in that it shows an equal spread of auxiliaries across the two corpora (cf. Table 1). However, if the textual environment is included in the analysis, differences as well as further similarities between Denmark and NAD emerge. The similarities are that var gift is used to denote the predicative state (being married) and/or linguistic distance, i.e. pluperfect context, indirect speech or irrealis mood (Hansen & Heltoft 2011: 681–689). These unremarkable occurrences account for 56 instances of NAD var gift and for all instances in the LANCHART corpus. The remaining 23 occurrences of NAD var gift (4 instances cannot be rated due to interruptions, restarts etc.) occur together with specifying co-text, i.e. specification of a point of time or a place, as shown in (9) and (10) below:

(9) I 1946 var vi gift og min kone hør-er til den engels-k-e kirke. (speaker AOP, CoNAmDa)

(10) Jeg var gift i Ringsted, Iowa. (speaker MFN, CoNAmDa)
In (9), two facts in the co-text imply that *var gift* denotes the event, the wedding: The specifying time adverbial in the first clause (*i 1946*) might go together with a stative reading, implying that the state of being married was restricted to the year 1946. However, the use of the present tense in the second clause implies that the marriage still goes on, which is confirmed by the rest of the interview. In (10), the co-text is similarly unambiguous: The speaker mentions his participation in a wedding, he specifies the event by mentioning the place and it is not likely that he would go to look at his niece’s state of being married. Instances like (9) and (10) indicate that NAD *var gift* unlike Denmark Danish might be used to denote the dynamic transition (the wedding act), implying that NAD *var gift* has shifted to include both meanings.

### 6. Conclusion

The analyses in Section 5 have shown how cross-linguistic analogy building leads to subtle changes of conventionalized form-function pairs. The non-general nature of the changes does not mean that the observed changes are arbitrary, though: The tendency across all three verb phrases towards a higher use of *var was* follows a regular and in no way unexpected pattern towards analogical leveling in accordance with English.

Apart from this general tendency, there are remarkable differences in the way the three constructions change: *var født* changes with respect to relative frequency, thus leading to a change of construction prototype (Hilpert 2013: 2): Denmark Danish has *blev født* as a default choice, while NAD has *var født*. This in turn implies a wider applicability of *var født* in NAD with respect to the grammatic-semantic function, as compared to Denmark Danish and given the similarity of text type.

NAD *var konfirmeret* also shows change in relative frequency leading to a very different quantitative spread than Denmark Danish with *var* being frequency dominant. Further, NAD *var konfirmeret* has changed its grammatic-semantic function as it denotes both the dynamic transition (the confirmation act) and the state (being confirmed). This implies polysemy, with both the new and the original meaning present.

NAD *var gift* shows no shift in relative frequency but has changed its function in a way parallel to *var konfirmeret*: It also might be used both for denoting a state (being married) and for denoting the transition (the wedding), having acquired polysemy with both meanings present.

These changes illustrate very clearly the crucial ingredients of constructionalization as pointed out by Fried (2013: 433–434), namely (a) the co-text, (b) attraction and partial adjustment to another pattern through analogy, and (c) backward pull exerted by original meaning. The analyses in Section 5 also point out that these factors affect the three verb phrases in different ways.

One may ask whether the NAD verb phrases account for constructionalization (the emergence of new constructions, i.e. new nodes in a constructional network) or constructional change (variation in constructs that may be preceding and feeding into constructionalization) (Traugott & Trousdale 2013: 33–38). The crossing point between constructional changes and constructionalization does not seem to be settled theoretically, but, for the present study, I will argue that NAD *var født*, *var konfirmeret* and *var gift* should be considered newly emerged constructional nodes, not just micro-variation: Denmark Danish shows no recent, historical or dialectal variation that could explain the distributional or functional properties of NAD *var født/konfirmeret/gift*, respectively, meaning that these constructions with their particular features did not exist before the migration. Accordingly, they must have emerged under the conditions of language contact, more specifically based on the Danish-English bilingualism of the NA speakers that would lead to the recognition of the formal and functional analogy and, subsequently, to cross-linguistically motivated partial adjustment of individual constructs. The shift in relative frequency is important in this respect as it highlights the non-randomness of change. Apart from the general analogy-
based process, the individual differences in the change patterns seems to indicate that the NAD verb phrases are not instances of the same construction, but three individual nodes.

Two questions remain to be answered: Why can a change mainly be observed with født, konfirmeret and gift? And why is the change restricted to past tense? Both seem to be an effect of the text type ‘sociolinguistic interview’: Past tense appears to be a natural choice for elderly speakers retelling their lives. Likewise, it is a natural choice for interviewers of immigrant speakers to ask for the place and time of birth and for turning points in life as, e.g., the confirmation and marriage. Another reason for past tense might be the phonological similarity of (at least) the initial component of Danish and English auxiliaries (Danish var [va], which may be pronounced [wa] in West Jutish dialect, and English was [stressed wəz], unstressed wəz) and were [wəz]). Cross-linguistic analogy building (e.g. [w] + vowel or [w] + vowel + [r]) might also act out in phonology. Thus, a promising avenue for further research, not least with regard to bilingual data, seems to be to integrate phonology into Construction Grammar approaches (cf. Höder 2016).

References


Putnam, Michael T. & Joseph C. Salmons. 2013. Losing their (passive) voice: Syntactic neutralization in heritage German. Linguistic Approaches to Bilingualism 3(4), 476–506. Doi: 10.1075/lab.3.2.05put

