Case study: Swedish passive

Therese Lindström Tiedemann, University of Helsinki

CLARIN workshop, 8 Dec 2017

UNIVERSITY OF HELSINKI

FACULTY OF ARTS

Project

- *Development of lexical and grammatical competences in immigrant Swedish*
- 2018–2020
- Riksbankens jubileumsfond
- Elena Volodina, Therese Lindström Tiedemann

UNIVERSITY OF HELSINKI
FACULTY OF ARTS

RIKSBANKENS JUBILEUMSFOND
STIFTELSEN FÖR HUMANISTISK OCH
SAMHÄLLSVETENSKAPLIG FORSKNING

GÖTEBORGS UNIVERSITET

Research questions

- The lexical / grammatical minimum learners at various levels demonstrate (**must-know competences**)
  - receptively (as shown in textbooks)
  - productively (as shown in essays)
- The peripheral vocabulary and grammar (**good-to-know competences**)

Material

- Two corpora (Volodina et al 2014, 2016)
    - SweLL – learner corpus
    - COCTAILL – course book corpus

Practical outcomes

- Resources and algorithms that empower L2 teachers and learners, course book writers , language assessors, etc.

  - an **inventory** of receptive and productive **vocabulary** scope / proficiency level
  - an **inventory** of receptive and productive **grammar** scope / proficiency level
  - **algorithms** for extracting lexical and grammatical features from L2 corpora

Grammar profiles

- **Positive and negative influence** from other known languages.
- **General pathways:** e.g. Pienemann's Processability theory (cf Pienemann & Håkansson 1999 in relation to Swedish)

PREVIOUS WORK:

- Swedish as a foreign language
- Inkinen-Järvi & Jokinen (2016) – **teachers' ideas** of what should be known at the **B1-B2 level**, what is taught at that level (seems to confirm PT)
- Finnish work on Swedish learner corpora (Topling, Sve2JU) – e.g. some correlation between word order and CEFR level (Lahtinen & Palvianen 2011)
- Van Dale *Grammatica Zweeds* CEFR

Combine
learner input and output

- What do learners **write**?

- What do learners **read**? / What are they expected to be able to read?

→ Bottom up approach to describing the "requirements" / CEFR-level

"PRE-PILOT": SWEDISH PASSIVE

Swedish passive

- THREE passive constructions, plus the possibility of using an impersonal construction with the impersonal pronoun "man".
- SLA questions:
  - When (and how) do **learners learn** the different constructions?
  - What is **taught** when?
  - How well does this relate to the L1 usage?
  - What kind of "errors" do they make?
  - How / when do they learn to distinguish the three / four constructions?

Need for the passive

- V2-principle

       *Vi **diskuterade** korpusar.*
       We discussed corpora.

       *Igår **diskuterade** vi korpusar.*
       Yesterday we discussed corpora

- Thematise, coherence – Swedish has very strict word order.

Swedish passive 1

- *Barnet* *sparkade* *bollen.*   (ACTIVE)
- child-DET kick-PAST ball-DET
- The child kicked the ball.

- *Bollen* *sparkades* *av barnet*.   (PASSIVE)
- ball-DET kick-PAST-PASS of child-DET
- The ball was kicked by the child.

Swedish passive 2

- *Barnet blev biten av en hund.* (PASSIVE)
- Child-DET become.PST bite-PSTPART of INDEF dog
- The child was bitten by a dog.

- *Hen var vald (av skolan) att representera eleverna.* (PASSIVE)
- 3SGPRON be.PST choose.PSTPART (by school-DET) to represent the pupils.
- (s)he had been chosen (by the school) to represent the pupils.

- *Man valde hen som elevrepresentant.*
- PRON.INDEF choose.PST 3SGPRON as pupil representative.
- One chose him/her to be the pupil representative.

So...

- **Possibilites** and **problems** there can be in studying this kind of construction and variation through a learner corpus and a course book corpus...
- What I as a linguist would need to study it better or more easily.

The different passives

	Danska	Norska	Svenska
S-passive	55%	53%	91%
Bli-passive	31%	32%	1%
Vara-passive	14%	15%	7%

Percentage of the different passive constructions in newspaper texts 2003 (Laanemets 2004: 91, cit. in Engdahl 2006)

S-forms most common

- Previous research L1 and Swedish L2/FL
    - s-passives more common
    - Check that it is also true in the corpora
        - learner corpora

        AND

        - text book corpora

L1 s-forms

- Comparison with s-forms in a Swedish newspaper Göteborgs Posten

- *finnas, kännas, hoppas, krävas, behövas, tas, lyckas, tyckas*

- Somewhat more rare *tvingas, användas*

- MOST s-verbs are deponent / medial.

Other s-forms

- **Deponent**: looks like a passive but is used actively

- *Jag **andas**.*
  =I breathe.

- **Medial**: verb where the subject is an experiencer or patient and there is no idea of an agent

- *De **förfärades** över händelserna.*
  =They were chocked by what had happened.

Material

- **(Pilot) SweLL-corpus (Språkbanken, Göteborg)**
  - SW1203 (preparatory uni course, B2)
  - Tisus (spring 2006)
  - Concordances

- **Topling corpus (Jyväskylä university)**
  - Swedish L2 (Finnish L1)
  - Upper secondary: 1/2nd course; 3rd course, 5th course
  - 5 different texts, different times of collection
  - Whole text files (concordances through Kielipankki)

Course book usage…

COCTAILL:
s-forms in course books

A2 level

- *finnas*
- *träffas*
- *ses*
- *trivas*
- *födas*
- *saknas*
- *behövas*
- *hoppas*
- *skiljas*

A few on the A1 level

- *höras*
- *fattas*
- *finnas*
- *(födas)*
- *(skiljas)*

Verbs but NOT passives!

SAG deponent verbs (incl. medial)

Method full texts (Topling data)

- **Excerpt s-forms and "potential" passives:**
  - Ex. "Min vän är mobbat i internet" (T4 OPS 3)
  - TARGET Min vän är mobbaD på internet.
  - 'my friend is being bullied on the internet'

  - Ex "Alle nya kursser har kommit, men ingenting har tagit bort" (T4 OPS 1, A2.2)
  - TARGET: Alla nya kurser har kommit, men ingenting har tagitS bort.
  - 'all new courses have come, but nothing has been removed'

- **Classification:** passive type
- **"Correct", idiomatic** – focus on passive constructions → agreement mistakes are not counted as errors:
  - Ex "… de finns <u>många fula saker </u>där som är inte menad för barn." (T4 OPS 3, A2.1, min emfas)
  - TARGET: det finns många fula saker där som inte är menade för barn

Method SweLL – s-passive

- Korp:
  - MSD: Verbs in different forms (infinitiv, imperativ, konjunktiv, presens, preteritum, supinum) + s-forms
  - POS: verb

  - →ALL s-forms, incl some words that have been misclassified *tills 'until', trots 'despite'* etc.
  - →Manual analysis of what is an s-passive
  - →CEFR Levels (evaluated) not included – separate lists
    - →How will new users be able to find out?

Results:
The task causes certain forms!

- "Studentskrivningar behövs" very common in Topling **– because of the task** – deponent / medial use

→ Information about the task is NECESSARY

  - headings
  - instructions
  - prompt texts
  - prompt vocabulary

Results – SweLL

- 1416 verbal s-forms (some mistakes: nouns, subjunctions, prepositions)
- *finnas* ('exist') – i.e. not a passive, MOST COMMON
- Many different s-forms

- Some s-forms were used "incorrectly" – target: active:
  - "Å andra sidan **känns** barnen … att de saknas någonting…" (SW1203 abc)
  - 'On the other hand the children seem to miss something'
- Task related s-forms??

Task importance

- Difficult to evaluate without access to the tasks.
- Given headings and possible prompts might include passives which might be copied (cf *Studentskrivningar behövs*)
- Some tasks will be more likely to lead to passives

- Some learners in SW1203 use A LOT of s-forms in one essay. Hard to evaluate without seeing the complete text.
  - SweLL extended context but not full texts.

Conclusions

- Passives could be better studied
  - if we had more task information
  - with semantic role annotation – search for subjects which are not agents... not only passives but... e.g. combined with *bli / vara*, verb-s
  - searchable CEFR-levels
- Can we make it possible to look at individuals?
  - 80%
  - > 3 contexts
  - whole texts, stats of passives / text
  - or without showing whole texts... but calculating on a text basis.

TACK / THANK YOU

References

- De Groot, Hans. 2014. Van Dale Grammatica Zweeds. Van Dale.

- Ekerot, L.J. 2011. *Ordföljd, tempus & bestämdhet.* 2a uppl. Gleerups.

- Engdahl , E. 2006).Semantic and syntactic patterns in Swedish Passives I Demoting the agen passive, middle and other voice phenomena. Eds. Lyngfelt & Solstad. Amsterdam: Benjamins

- Huhta, A, R. Alanen, M. Tarnanen, M. Martin, T. Hirvelä. 2014. Assessing learners' weiting skills in a SLA study: Validating the rating process across tasks, scales and languages. Language Testing 3 (3): 307-326

- Huhta, A. & G. Khushik. subm. Syntactic complexity in English as a foreign language learners' writing at different Common European Framework of Reference levels.

- Inkinen-Järvi, M. & J. Jokinen. 2016a. Språkcentrumlärarnas uppfattningar om grammatikens roll på kurserna i L2-svenska vid finländska universitet. I: T. Lehtonen & J. Vaattovaara (red.) *Näkökulmia kielenoppimisen ohjaukseen / On advisning and counselling in language learning.* Language Centre publications 6. Helsingfors: Helsingfors universitet. s. 157-176. https://helda.helsinki.fi/bitstream/handle/10138/159069/Vol_6_Nakokulmia_kielenoppimisen_ohjaukseen.pdf?sequence=2
[Hämtad: 23 nov 2016]

- Inkinen-Järvi, M. & J. Jokinen. 2016b. Grammatiken i L2-svenska på nivå B1 och B2 i Finland och Sverige. Posterpresentation Svenskans beskrivning (ms)

- Källström, R. 2012. *Svenska i kontrast – tvärspråkliga perspektiv på svensk grammatik.*

- Laanemets, A. 2004. *Dannelse og anvendelse af passiv i dansk, norsk og svensk.* Nordistica Taruensia 11. Tartu: Tartu University Press.
- Laanemets, A. 2012. *Passiv i moderne dans, norsk og svensk – Et korpusbaseret studie af tale- og skriftsprog. Tartu:* University of Tartu Press.
- Lahtinen, S. & Å. Palviainen. 2011. Omvänd ordföljd – en indikator för nivåbedömning inom CEFR? I: E. Lehtinen, S. Aaltonen, M. Koskela, E. Nevasaari & M. Skog-Södersved (red.) AFinLa-e Soveltavan kielitieteen tutkimuksi 2011. nr 3: 88-101.
- Lahtinen, S. & Toropainen, O. 2017. Jag är so sorry. Att be om ursäkt på sitt andraspråk i en mejluppgift. *Svenskans beskrivning 35*. Göteborgs universitet. https://gupea.ub.gu.se/bitstream/2077/52211/1/gupea_2077_52211_1.pdf [Hämtad 2 maj 2017]
- Manninen, S. & D. Nelson 2004. What is a passive? The case of Finnish. *Studia Linguistica* 58 (3): 212-251
- Martola, N: 2014. The interaction between passive constructions and lexical verb constructions in Swedish. *Constructions 1-6/2014.*
- Naess, Å. 2011. *Global grammatik – språktypologi för språklærere.* Gyldendal.
- Reuters, M. 2015. *Så här ska det låta.* Scriptum.
- Toropainen, O. & Lahtinen, S. (2014a). Argumentation på L2-svenska: Inlärare skriver insändare. [Argumentation in L2-Swedish: Learners write letters to the editor.] I *Svenskans beskrivning 33*. http://hdl.handle.net/10138/144498 [Hämtad 2 maj 2017]
- Toropainen, O. & Lahtinen, S. (2014b). Interrogative clauses across CEFR levels in Finnish and Swedish as an L2. *Apples, Journal of Applied Language Studies,* 71–84. http://apples.jyu.fi/
- Volodina E, Pilán I, Eide S.R, Heidarsson H. 2014. You get what you annotate: a pedagogically annotated corpus of coursebooks for Swedish as a Second Language. NLP for CALL. NEALT Proceedings Series 22
- Volodina, E, I. Pilán, I. Enström, L. Llozhi, P. Lundkvist, G. Sundberg, M. Sandell. 2016. SweLL on the rise: Swedish learner language corpus for European reference level studies. http://www.lrec-conf.org/proceedings/lrec2016/pdf/259_Paper.pdf