

Report on Iva Rezková: *De Story van Twee Talen*

Jacomine Nortier, Utrecht University, March 2014.

Since I am not familiar with the criteria or practices used in the Czech Republic regarding the way dissertations are judged, I will use my own criteria.

The dissertation is written with great care and I read it with pleasure. The analyses were thorough and systematic and the author managed to make an exhaustive use of only 430 sentences, which is a relatively limited corpus for this type of work. Iva is very good in explaining, both the terminology used and the way of working.

Her Dutch is very good although towards the end some errors occur.

However, in this report I will concentrate on some critical notes focusing on, in short, the following topics:

1. The MLF model
2. The neglect of loanwords
3. Written data

1. The MLF model

The choice for MLF is made, among other reasons, because of the characteristics of the corpus. But isn't that circular? First you look for what kind of cs you see in the data; then you look for a model that fits what you see. But correct me if I'm wrong!

In various studies that have been carried out in the past decades, it has been shown that the claims made by the MLF model are confirmed in language pairs in which the languages involved are typologically different. However, in work on codeswitching between typologically more similar languages the MLF model is not always able to account for the patterns found there (see for example the work by Muysken, or Deuchar). Also work by my own students showed the difficulty of explaining codeswitching between related languages or standard/dialect codeswitching.

Chapter 3.1: It is a pity that Muysken (2000) was not used in the present study. I don't think he is quoted correctly here. In *Bilingual Speech* he distinguishes three types of what he calls code-mixing: insertional, alternational and congruent lexicalization. These types are linked to communities and their characteristics. Alternation is found in stable bilingual communities with a tradition of language separation, insertion is typical for colonial settings and recent migrant communities where there is an asymmetry in language proficiency, and congruent lexicalization is associated with, among other things, bilingual speakers of closely related languages. I expect this last type to explain the instances of codeswitching in the present corpus very well. The fact that in the present study the MLF model seems to account for the codeswitching patterns found is partly because it explains the patterns indeed but partly it is a consequence of the method used. In the description of the corpus I would have taken other decisions. For example, the discussion about *samenstellingen* (compounds) is not very satisfactory. What is the difference with collocations? Why is a compound a word group instead of a word? Compounding is a morphological process, not syntactic. Furthermore I don't agree with the statement that 'fraseologie' is not part of the living language! Why? What does the author mean by 'living language'?

The level of analysis: if an English N is inserted in a larger Dutch NP with the function of a complement (Direct Object) one can't say that the N is an object. Example: in '*Hij is niet in de mood voor grapjes*' on page 98 '*mood*' is not the 'naamwoordelijk deel van het gezegde' but it is the N (complement) in a PP. Something comparable occurs in (392): *prepared to change his life/once in a lifetime*. You cannot analyze each word separately here. It is just one EL island. It makes no sense to look at an isolated *his* or *a*.

By the way: I think the function of Indirect Object is missing (meewerkend voorwerp)!

The treatment of the MLF model is very 'narrow'. Let me explain:

On page 50 the author remarks that she won't go into psycholinguistic aspects of codeswitching but the MLF model is built on a psycholinguistic production model (by Levelt). On page 51 it is mentioned that the building of a sentence frame and morpheme insertion take place in different stages in the production process. That needs to be explained. References to lemmas, the formulator, mental lexicon need to be placed in a context in order to understand the full complexity and impact of the MLF model.

MLF is one of the analytic models for codeswitching. There is a vision on how codeswitching is to be seen in the broader context of language contact. For example: is codeswitching linked to language shift and loss, to acquisition, to contact-induced change, what about proficiency levels?

Another thing that is important in the description of the premises of the MLF model is borrowing:

2. Loan words

The distinction between codeswitching and borrowing is a hot topic in the literature on grammatical aspects of codeswitching. Recently, however, most publications about codeswitching don't focus on grammar but on conversational strategies and on experimental psycholinguistic codeswitching studies.

Some linguists propose to neglect the difference and to consider all (established or not) loanwords as instances of codeswitching (such as Treffers-Daller). Myers-Scotton has elaborated on borrowing in her 1993 book *Duelling Languages* (only the reprint from 2005 is in the bibliography¹) where she introduced the MLF model. In the present study the only criterion used is the occurrence of lexical elements in WNT. That criterion may look very solid but I think it is not very reliable. It is not the list that decides, it is the speech community! Elements may be code-switching in one community and loans in another. And more importantly: they are treated as such, irrespective of their occurrence in the WNT. Unfortunately, pronunciation cannot be taken into account in this study but usually it is a very strong criterion: in general, a word pronounced according to host language rules is considered a loan word and the original donor language pronunciation indicates codeswitching. An established loan is treated as part of the host language so there is no difference between EL and ML. Since in this corpus most other language elements

¹ The same goes for Appel & Muysken who published their book in 1987. In the reprint they mention explicitly that the book is too old to be used as a single source.

are single words, the discussion on codeswitching versus borrowing is extremely relevant and should have been addressed.

The distinction between CS and borrowing in the corpus seems arbitrary: words like *bloody*, *queue*, *gesubmit(te)*, *pushen*, *spammen* (just some examples) are treated as codeswitching, probably because they don't occur in the WNT list but in practice they occur as loan words in Dutch and they are intuitively felt to be part of Dutch and therefore treated as such. If you consider them to be English (codeswitching), why would words like *process*, *offline* or *drop* be Dutch (borrowed)? *Gesubmitte* is a typical example of a morphologically integrated loan word.

Example (41) is a loan blend, would have been a nice illustration.

The fact that borrowing from English into Dutch is so widespread and common makes it difficult to analyze it with the MLF model.

3. Written data

The corpus consists of 430 sentences from the internet. Some of them are very informal chat data, others are found in more formal email exchanges. It is not clear why those different styles are used, because it is neglected in the thesis. For example we don't know whether certain phenomena only occur in the most informal (and thus least monitored) data. From the codeswitching literature we know that the level of consciousness and formality influences the amount and type of codeswitching used. Therefore it is important to distinguish between those levels when several of them are represented in the corpus.

Furthermore, I strongly disagree with the spelling criterion in defining compounds (*samenstellingen*) or the distinction between Dutch or English plural or genitive *-s*. In Dutch people usually are not aware of the spelling rules and many errors are made. As long as my (university level) students write 'huis werk' instead of 'huiswerk' the criterion is totally unreliable. And I must admit that I don't even know the rules of how to use the genitive 's in written Dutch myself. Therefore you cannot use this as a criterion to distinguish between Dutch and English.

Example (137): What is wrong with *twee autos*? The use of an apostrophe cannot be a criterion in determining whether a word is Dutch or English.

Finally there are some smaller points of discussion:

- a. Note 1 in chapter 1: what kind of adjustments have been made to the original data? In codeswitching repetitions, false starts etc. are highly interesting since they tell us a lot about the strategies the speaker uses, and about the function of codeswitching.
- b. What exactly is the definition of bilingualism? How fluent are speakers in order to be called bilinguals?
- c. In this dissertation, a switch between a main clause and a subordinate clause is analyzed as inter-sentential codeswitching. However, usually intra-sentential codeswitching includes subordinate clauses (*ondergeschikte bijzinnen*) since they fall within the grammatical scope of the main clause. In Dutch they cannot occur independently because of their deviant word order.
- d. Is a switch the same as an insertion?
- e. Both the government and Equivalence constraints have developed since their 'invention' in the seventies and early eighties. It is therefore not logical or fair to criticize (only) the oldest versions.

- f. Page 43: what is the difference between language and dialect?
- g. Spelling: Mahootian, Swahili, Schatz. Boumans (page 72) analyzed Moroccan Arabic – Dutch data. No Turkish!
- h. What is known about the background of the speakers who provided the data?
- i. Adjective –e in Dutch: explain when it is used (almost always); what are the rules?
- j. (206): *confidential bak*. *Confidential* is not an ordinary adjective. It is the name of the *bak*, probably *confidential* is written on it.
- k. 'Tangconstructie': different from the way it is used here. See <http://taaladvies.net/taal/advies/term/84/tangconstructie/>