

***"De story van Twee Talen."* Grammatical Description and Analysis of
Intrasentential Codeswitching between Dutch and English – PhD
Dissertation by Iva Rezková**

Opponent's report: Dr Roel Vismans, University of Sheffield

Introduction

This dissertation is a study of Dutch-English code switching, more particularly the grammatical parameters within which code switching between Dutch and English can occur. The study is based on the analysis of a small corpus of sentences obtained from four different internet media: chats, emails (both private and business), travelogues, and social networks and discussion fora. The theoretical framework for the study is the Matrix Language Frame Model (MLF model) developed by Myers-Scotton in several publications over the last 20 years or so.

I have a general interest in code switching but I am not a specialist. I know that another external opponent for this dissertation is Dr Jacomine Nortier of the University of Utrecht who has studied code switching (in Dutch) intensively, and in any difference of opinion between myself and Dr Nortier I would bow to her superior knowledge in this field.

In my report I will comment briefly on three specific areas of the dissertation: the model, the data and the analysis. I finish with suggestions for questions during the defence of the dissertation.

The model

Chapters 2-4 together form the theoretical background to the study and necessarily entail a review of the relevant literature. The candidate presents a thorough explanation of the MLF model in chapter 4, having briefly outlined alternative models in chapter 2, which in effect offers a brief historical outline of theoretical developments in the field, whilst chapter 3 presents a useful discussion of the differences between code switching and various kinds of borrowing. The candidate demonstrates a clear grasp of the main issues outlined in these chapters. Nevertheless, there are three ways in which the theoretical discussion could have been improved.

1. The candidate's analysis and discussion rely to a great extent on a relatively small number of references. This is evidenced in the modesty of the bibliography and the fact that chapters 2-4 regularly refer to works that are quoted indirectly from other authors. Such secondary referencing leaves the candidate at the mercy of those other authors and their interpretation. That is why scholarly studies of this kind require that a candidate goes back to original sources as much as possible in order to arrive at their own interpretations and conclusions.
2. The MLF model presented in chapter 4 appears to be largely based on the work of Myers-Scotton (2002,¹ 2005a and 2005b), which the candidate critiques on the basis of Boumans (1998). That is problematic: Boumans criticism of Myers-Scotton must be based on her earlier work, because it pre-dates the relevant studies of Myers-Scotton's. This may seem largely a matter of presentation, but could also have repercussions for the candidate's appreciation of the model.
3. Perhaps as a result of 1 and 2, the candidate does not really seem to engage in a critical discussion of the literature and develop her own point of view. There are

¹ Or 2003? This is not clear. The bibliography says 2002, but on for example p. 60 the date is 2003.

places in the theoretical discussion where further theorising would be appropriate, for example on p. 45 where the candidate asserts that only bilingual speakers can code-switch, whereas all monolingual speakers can do is use loan words. But in a situation (as in the Dutch language area) where strict monolingualism is almost nonexistent, it is not clear what the boundary between the two is. The theoretical chapters of a doctoral dissertation are an appropriate place for such fundamental discussions.

The data

The data for this study come from a modest corpus that the candidate herself has compiled for the purpose. It is entirely drawn from computer-mediated communication (CMC) sources and consists of four subsets: (1) chats, (2) business emails, (3) personal emails and travelogues, and (4) messages on social media and discussion fora.

This is an admirable venture. Corpus building is a useful and often necessary research skill and this experience will stand the candidate in good stead in the future. However, she should have provided more details about the composition of the corpus and of her methodology for its annotation and detailed study. There is no information, for example, about the extent of regional variation within the corpus. We gain some insight from the list of internet sources in the bibliography (p. 193), but this only concerns (parts of) subsets (3) and (4) of the corpus. We do not have any information about the location of the multinational company whose chats and emails provided the data for subsets (1) and (2). Is it in the Netherlands or Flanders (where the candidate apparently spent four months)? Nor do we have information about the language users' characteristics, such as gender and age. This is admittedly not a sociolinguistic study, but can linguistic analysis and speaker characteristics really be separated so strictly? The numerical analysis of the data in this thesis would certainly have warranted some reflection on the social background of the speakers involved.

Similarly, some reflection would have been appropriate on the fact that the source of the corpus is CMC. On p. 74 we read without further comment that this is a corpus of written as opposed to spoken Dutch. However, the last 20 years or so have seen a great deal of discussion about the nature of CMC and the fact that much of it is a hybrid of spoken and written communication. This corpus in particular consists of one subset, business emails, that probably approaches formal written language, whereas the other three subsets are to various degrees more akin to spoken language. It is, therefore, not entirely homogeneous. Moreover, its subsets ('registers') in effect constitute different text types whose linguistic characteristics are likely to vary. And finally, 'register' 3 consists of two different text types (personal emails and travelogues) that are not necessarily very similar.²

The analysis

The analysis in the dissertation is presented in two chapters: chapter 6 offers a detailed description of code switching in the corpus, whilst chapter 7 then relates the analysis of the corpus back to the MLF model. This is solid work that lays bare the fundamental parameters within which Dutch-English code switching can take place, which is groundbreaking work. It also demonstrates convincingly the soundness of the model.

There is a certain amount of numerical analysis in this part of the dissertation. It is not always clear how significant this quantitative work really is. What do we learn from the fact that, for example, the majority of code switches occurs in NPs? Also, in some cases statistical significance testing could have enhanced the findings, esp. those concerning the Uniform

² Incidentally, the *Corpus Gesproken Nederlands* is an extensive corpus of spoken Dutch, freely accessible to all researchers. The candidate could have reflected in the dissertation on whether or not it would have been an appropriate source.

Structure Principle. Yet, at the same time an important conclusion about deverbal adjectives is based on just two instances.

The detail in the analysis sometimes hinders a more profound engagement with theoretical aspects of code switching. For example, a number of compounds in the corpus seem to be limited to the vocabulary of the specific company that is the source of some of the data (e.g. the word *breach*). That raises questions about jargon, and about the boundary between borrowing and code switching. Also, spelling issues, which figure in several places in chapter 6, ignore the fact that non-adherence to spelling rules is a significant characteristic of CMC. This may throw significant doubt on the dissertation's assertions about spelling as an indication of code switching.

At the same time a number of details about Dutch and English grammar seem to have been ignored in the discussion. For example, §6.1.1 fails to take into account the role of accentuation in determining whether a combination of (Dutch or English) words is a compound or not; §6.2.2.2.2 (p. 117) does not acknowledge the complexities of the formation of comparatives in English; and §6.2.1.6, where diminutives are discussed, does not seem to appreciate that systematic diminutive suffixation is absent in English. What does the MLF model say about this? Where do diminutives fit into figures 3 and 4 (pp. 62 and 76 respectively)?

Conclusion and recommendation

This dissertation presents some innovative work in its description of Dutch-English code switching and is embedded in an appropriate theoretical framework. The research for the dissertation is based on a credible corpus. However, I recommend that the dissertation be revised prior to its defence in order for it to improve in a number of respects, as detailed above. The revised version should:

1. Demonstrate profound critical engagement with the literature on code switching (cf. section on 'The model', above);
2. Reflect critically on the limitations of the corpus in the light of existing research into CMC (cf. section on 'The data', above).

Questions

As requested, I finish with a number of possible questions for the candidate during the dissertation's defence. Questions 1-3 are specific to the work in the dissertation, and question 4 invites the candidate to speculate on related issues. Question 5 is of a more abstract nature.

1. Why has the candidate chosen the LFM model as the theoretical basis for her analysis? What are its strengths and weaknesses? How could the work she has done for this thesis help to address those weaknesses and enhance the model further?
2. Given the provenance of the data for this dissertation (a number of different 'registers' or text types collected from Computer-Mediated Communication), to what extent does the candidate think her findings are applicable across other text types and registers, and across both written and spoken language?
3. To what extent does the candidate think that the *Woordenlijst Nederlandse taal* (2005) is an appropriate instrument for determining whether a word belongs to the Dutch lexicon? What alternatives has she considered? What would have been their possible advantages over the *Woordenlijst*, and why has she rejected them?
4. The dissertation explicitly takes a grammatical angle. This is an entirely appropriate delineation for a study of this kind. However, the candidate also writes in her first chapter (pp. 13-14) that she 'does not consider the social factors, that influence the use of code switching, as irrelevant. The sociolinguistic "why" questions are a crucial

- part of code switching. There is no doubt that code switching is a linguistic phenomenon with a clear social function and motivation.' I would like to invite her to go beyond the descriptively grammatical. What are her reflections on the 'why' of Dutch-English code switching, especially in relation to her corpus and its 'registers'?
5. This dissertation highlights a number of grey areas or fuzzy boundaries. For example, the boundaries between different kinds of bilingualism are unclear. Drawing a clear line between borrowing and code switching is impossible (§3.3.1). The precise distinction between compounds and free combinations of words is difficult to explain (§6.1.1). Even the difference between a phrase and an EL Island does not always seem to be clear. To what extent does the candidate think such grey areas are characteristic of (human) language and why do they exist?

28 March 2014