

# The crosslinguistic intertextuality of loanwords

Vejdemo, Susanne      Vandewinkel, Sigi  
Montero-Melis, Guillermo

7<sup>th</sup> December, 2012

## 1 Presentation and aims

The semantic meaning potential of words is to a large extent governed by their intertextual history of use (ref Linell, ref Traugott & Dasher). Loanwords are an interesting case, since they have a history in their source language, but lose some of their meaning potential when they are borrowed. We wish to examine the following topics:

- a. How much of a word's meaning changes when it is borrowed (e.g. the term 'body guard', borrowed from English into Swedish), i.e how much of its meaning potential and history is lost?
- b. After a loan word is established, how does it share the semantic space with already existing, seemingly synonymous native words (e.g. the Swedish term 'livvakt', body guard.)
- c. How do ?? and ?? compare across three European languages, i.e. Swedish, Dutch, and Spanish?

By using both quantitative experimental and corpus methods, as well as qualitative interview methods, we also wish to examine if there are marked discrepancies between the measurements of meaning and the subjective reported opinions about meanings of speakers.

## 2 Theoretical background

### 2.1 Creating semantic profiles

The increase of and access to computational power has made it possible to use large amounts of texts - corpora - to create semantic profiles for words. KOPTJEVSKAJATAMM-SAHLGREN has shown this for semantic investigations into temperature terms using a method known as Multidimensional Scaling (MDS). SIGSTRÅUM has shown how semantic profiles can be made for the Swedish term *helig* using the method Latent Semantic Analysis (LSA).

Both MDS and LSA are based on comparing collocations of a words(see REF for an overview of collocations), that is occurrences of the words with a few words of context to the left and right, with a more semantic representation of the semantic relationships of words. Semantic representations are made by measuring

how often all the words in a multimillion word text corpus co-occur with all other words in the corpus - the result is an n-dimensional space where words that co-occur often cluster together. (see DEERWEISTER, DUMAIS for general overviews of LSA, TURNEY-PANTEL <http://www.jair.org/media/2934/live-2934-4846-jair.pdf> for general overview of vector models for semantic processing).

## 2.2 Comparing semantic profiles

The semantic profiles we wish to compare can be represented as semantic vectors. Semantic vectors can be subtracted from one another, leaving a semantic difference vector - a measure of the difference between the semantic profiles. (SIKSTRÅM SEMANTIC TESTS p19) Quantitatively, all words under investigation can be ranked according to how much they change: we can thus compare an English loanword with a native language equivalent.

We also aim to develop a quantitative measurement of the semantic difference between the English word in the English corpora and the English loanword in the Swedish corpora. In addition, this comparison will also be done more qualitatively, by looking at which other words cluster with the target word.

## 2.3 What is meaning?

Meaning can be approached in a variety of ways **Koptjevskaja-Tamm 2008**

## 3 Susanne's draft

HOW TO MEASURE MEANING CHANGE As a word is borrowed from one language to another, its meaning changes - the number and kind of referents it represents can grow or shrink, and its register (in which social setting it is appropriate to use) can also vary. This project seeks to investigate this change - how large is it, and are there recurrent patterns in the kinds of changes that occur to loan words? This project aims to create both etic and emic definitions of the meaning, and meaning changes, of the loanwords. In an etic definition, the meaning of a word is defined as the set of its uses (Koptjevskaja-Tamm 2008), or the set of its situated instances (Evans 2009). In this, we follow the exemplar semantics research done by WÄLCHLI & Cysouw (2012), and will take a denotational approach to meaning where similarity in form will be assumed to represent similarity in meaning. In an emic definition, the commonalities behind the different uses are sought (Koptjevskaja-Tamm 2008), what Traugott & Dasher 2002 call a "stabilized, institutionalized, and prototypical semantic center" that can be contextually interpreted in constrained ways" and which Linell (REF) refers to as the meaning potentials of a word. The etic definitions will be investigated through the uses of the words in corpora (see section CL) and by acceptability judgments in the psycholinguistic research tradition (see section AJ). The emic definitions can partly be arrived at by a careful analysis of the commonalities in these two experiments, but the subconscious evaluations of the words by speakers will also be measured through psychological Semantic Differential experiments (see section SD).

### **3.1 Lexcial Borrowings**

**Matras2009**

## **4 Materials**

## **5 General method and phases**

## **6 Expected outcome**

## **7 Appendices**

### **7.1 Swedish loanwords**