Katharina Korecky-Kröll

Scientific report: New experimental methods in research on morphonotactic processing

Purpose of the visit:

The theoretical background of the project is the Strong Morphonotactic Hypothesis (Dressler & Dziubalska-Kołaczyk 2006) which claims that phonotactics helps in decomposition of words into morphemes: If a sequence occurs only or by default over a morpheme boundary and is thus a prototypical morphonotactic sequence, it should be acquired earlier and processed faster and more accurately than a purely phonotactic sequence. The computational model PHACTS simulates the processing of ortho-phonological words and generalizes the resulting phonotactic knowledge to novel stimuli.

Our Italian-Austrian project team developed and conducted several computational and psycholinguistic processing experiments concerning German phonotactic and morphonotactic consonant clusters.

First results of the computational experiments show that the PHACTS model is sensitive to the phonotactic or morphonotactic status of n-grams.

First results of a psycholinguistic "split clusters" experiment demonstrate that Viennese teenagers aged 13 -15 show more significant processing differences between phonotactic and morphonotactic consonant clusters than Viennese adults, but the sample of teenagers was small.

Therefore an improved version of the same “split clusters” experiment and a new sequence monitoring experiment have been conducted in winter and spring 2013 with a bigger sample of Viennese high-school and university students.

The short visit was planned to focus on the preparation of two further publications and of two further experiments which investigate the processing of German-like pseudowords in comparison to existing German words both containing typical phonotactic and morphonotactic consonant clusters, including vowel length as an additional factor.

The minor part of the short visit should concern the planning of a European project on the linguistic input and the morphological skills of immigrant children with different socio-economic backgrounds in their second language, which is also the majority language of the respective country.

Description of the work during the visit:

I arrived in Pisa on Monday, Sept. 23rd, 2013, after a night on the train and immediately went to the Scuola Normale. In a first meeting with Chiara Celata and Irene Ricci, it turned out that there was a problem with the reaction times of our first “split clusters” experiments. The reaction times had been automatically measured by the Presentation Software, but unfortunately these measures were not always correct (i.e. if a test person took a louder breath before answering or if there was some background noise, this was recorded as the beginning of the answer). We decided to go back to the original data and to check and correct the reaction times before working on the statistical analyses and the publications. Chiara Celata and Irene Ricci spent the major part of their time during my visit to correct these data. Of course it was not planned like that and we would have preferred working already on the analyses and the publications, but in my view, it was necessary to solve this

unforeseen problem first in order to avoid running statistical analyses on data which are not entirely correct.

After further meetings with Chiara Celata, Irene Ricci, and Pier-Marco Bertinetto (the latter one stressed that it would be especially important to find out whether phonemes or phoneme sequences after or before a morpheme boundary are easier to detect for participants), I mainly worked on the preparation of two further "sequence monitoring" experiments on German which should include vowel length as an additional factor and involve not only consonant clusters but also vowels (e.g. e-schwa or -en which are frequent inflectional suffixes in German). I was not able to finish these preparations during my stay, but I will have still some time to finish them because these experiments can be conducted in summer term 2014 at the earliest when I hope to teach a certain course at the University of Vienna which would allow me to get students' help for data collection.

On Friday, September 27th, 2013, when the colleagues from Scuola Normale had to prepare a public lecture, I visited Vito Pirrelli and Claudia Marzi at the Istituto di Linguistica Computazionale "Antonio Zampolli". In a very fruitful meeting we discussed two major areas of future collaboration:

1. The European project on the linguistic input and the morphological skills of immigrant children with different socio-economic backgrounds in their second language (the majority language of the respective country): While our research group in Vienna has already been working on the INPUT project ("Investigating Parental and Other Caretakers' Utterances to Kindergarten children", in which we collect longitudinal spontaneous speech data and test data of 61 three- to four-and-a-half year old children) for 1 ½ years, a first common project proposal of several countries (Norway, Austria, Germany, Sweden) in the NORFACE framework has unfortunately been rejected. But Vito Pirrelli had just learned that his Italian project proposal had been accepted by the Italian Research Council. It is smaller than the Viennese INPUT project, but it will nevertheless give very important results for Italy and may also serve as a very important preparatory work for another common European project proposal involving more countries. Vito Pirrelli would like to start his project soon and will contact me and I will give him detailed information about my experience with the INPUT project. As soon as the new European Horizon 2020 framework will publish its first calls, we will think about submitting a joint proposal.
2. Claudia Marzi is currently working on the learning of German noun plurals by Self-Organizing Maps and showed me a very interesting presentation she gave in Dubrovnik. As I have been working on acquisition of plurals in German-speaking children together with my colleagues, it will be an interesting area of cooperation for both of us to compare SOM data with "real" children's data.

On Friday afternoon I came back to the Scuola Normale in order to work on the preparation of the new experiments for some more hours. Then I took a train to Florence in the early evening and from there the night train back to Vienna.

Description of the main results obtained:

Due to the problem with the incorrect reaction times, the statistical analyses of our two experiments could not yet be finished. A preliminary χ^2 test on the quality of the adults' responses in the ist/sit experiment reveals a gender difference: Although "ist" is preferred over "sit" by most adults, it is

significantly more preferred by women than by men, which is an unexpected result. We must definitely run other analyses as well (preferably a mixed-effects model).

As to the preparation of the future visual sequence targeting experiments, I found already many real words for the following types of phonemes and phoneme sequences (always with short or long vowels before or after): T, ST, E, EN, S, N

e.g.: T
TEAM (phonotactic, long vowel, initial t)
TIPP (phonotactic, short vowel, initial t)
RINNT (morphonotactic, short vowel, final t, morpheme boundary before t)
SPRINT (phonotactic, short vowel, final t)
PINNTE (morphonotactic, short vowel, medial t, morpheme boundary before t)
QUINTEN (morphonotactic, short vowel, medial t, morpheme boundary after t)
FLINTE (phonotactic, short vowel, medial t)
SCHIENT (morphonotactic, long vowel, final t, morpheme boundary before t)
DIENTE (morphonotactic, long vowel, medial t, morpheme boundary after t)
[no phonotactic cluster with long vowel]

The vowels and the consonants before the target consonant or cluster must also be varied, and of course half of the test items will be distractors and not contain the target consonant or cluster at all.

The pseudoword experiment will be designed according to the real word experiment. There is a strong tendency of long vowels to combine with morphonotactic clusters in German (although there are also short vowels combining with morphonotactic clusters), but in many cases there are just no existing words in which long vowels combine with phonotactic clusters (see example above). Within the pseudoword experiment, we may be able to close many gaps that exist in the real words experiment due to the non-existence of certain combinations. We hypothesize that participants might treat pseudowords with long vowels before consonant clusters as (pseudo-)words containing morpheme boundaries.

Future collaboration with the host institution:

As already mentioned, the work on phonotactic and morphonotactic processing between the Scuola Normale Superiore di Pisa (Chiara Celata, Irene Ricci, Pier Marco Bertinetto) and the Department of Linguistics of the University of Vienna (Katharina Korecky-Kröll, Wolfgang U. Dressler) will definitely continue and go on even in a still more intense way as before. We will continue working on the statistics of the two experiments and give a joint presentation on this topic at the "Second International Workshop on Phonotactics and Phonotactic Modeling" which will take place in Pisa from November 21 to 22, 2013. Probably there will also be a joint volume of the workshop papers (as for the papers of the first workshop in 2012 which are now under review for a joint volume of the *Language Sciences* journal). If our results are as convincing as we hope them to be, we will also try to submit a joint paper to an A journal. Further steps are the two new experiments which will be conducted in 2014.

Furthermore, a fruitful cooperation is planned with Vito Pirrelli (concerning the project on input to bilingual and monolingual children which will hopefully result in a big European project involving many different countries) and Claudia Marzi (concerning the acquisition of noun plurals).

Projected publications/articles:

1. A joint paper on the “split clusters” and the “sequence monitoring” experiment in the workshop proceedings
2. Probably a joint paper on the same topic in an A journal (if the results are convincing)
3. At least one further paper on the planned “real words” and “pseudowords” experiments (2014 or 2015)

Other comments:

I am very grateful for this opportunity to pass five days in Pisa with very intense and fruitful work. I think that the ESF Networds short visits program is extremely useful, especially for younger researchers who in general do not have many other possibilities to get money for travelling.

Comments on the travel costs: My train ticket was a bit cheaper than expected because for the train ride from Florence to Vienna, there was no more free place in the sleeping car, but only in the couchette coach.