Hidden variation as a result of learning: 
Is variation okay as long as you can’t hear it?

This presentation explores the possibility that variation can occur on different hidden levels of representation in the phonology of a language as a result of learning. When learning a language, children are confronted with the problem that what they hear is not what they get. Any listener/speaker of a language constructs surface structures during the comprehension and production process. While adults can fall back on the grammar they have already acquired, children need to yet construct the grammars and by that also the hidden structures from the unstructured speech stream. In the case of phonology, for instance, both the phonological surface representation of a phonetic input has to be constructed (and learned) as well as the phonological underlying representation of that input. This variation can occur on both of these levels (surface and underlying), while it does not have an effect on the communication between speakers. As long as the phonetic (or overt) forms of listener and speaker are the same, they will not be aware of the fact that they have different grammars. This kind of variation therefore has no consequence for the communication process.

The effect that there can be variation on the hidden levels of representation appeared in computer simulations on the acquisition of underlying forms (Apoussidou 2007). Virtual learners had to acquire the correct stress pattern of Modern Greek, a language where stress is mostly lexicalized, i.e. encoded in the underlying form. The virtual learners were equipped with an Optimality Theoretic grammar (Prince & Smolensky 1993) and a learning strategy, the Gradual Learning Algorithm (Boersma 1997). They had to learn both the constraint ranking that reproduced the correct stress pattern as well as the the hidden representations, underlying and surface forms, from pairs of overt forms and the corresponding meaning. The grammatical framework comprises four levels of representation shown in (1), and is based on Boersma’s (2005, 2006ab) bidirectional grammar model. The different levels of representation are connected via different kinds of constraints. The underlying and surface forms that a learner decides on are determined by the constraint ranking of the grammar and are computed anew every time a certain word is either comprehended or produced. The virtual learners were able to acquire the desired stress pattern that was contained in the overt forms they were trained on, but variation occurred on the surface level (in terms of the foot structures that the learners assigned) as well as on the level of underlying representations (in terms of lexical stress marks). For instance, the learners chose one underlying form with a certain probability when inflecting a certain root A with suffix X, while they mostly chose another underlying form when inflecting root A with another suffix Y.

The results raise questions on how much variation we want to admit in the levels of hidden representations. Is the variation just a result of imperfect learning, or is it a valid choice of analysis (as has e.g. also been suggested for palatization processes in Italian; Krämer 2007)? The presentation will provide an explanation which is also psycholinguistically plausible.
(1) The grammar model

```
  semantic representation  ‘Meaning’
                             
  phonological representations  |Underlying Form|
                             /Surface Form/

  phonetic representation  [Overt Form]
```

References


