THE STUDY OF LINGUISTIC COMPLEXITY
AS A CHALLENGE FOR PHONOLOGY

Linguistic complexity has recently become one of the most discussed topics in linguistics. Starting with McWhorter (2001), who argued that there are more complex and less complex languages, it has been thoroughly examined in a large amount of scholarly publications, e.g. Dahl (2004), Miestamo et al. (2008), Kortmann & Szmrecsanyi (2012), to name just a few.

Phonological complexity has been extensively treated by many linguists. It is not only “envisagée en elle-même et pour elle-même,” as Saussure put it. It has also provided basis for research in other areas such as prehistoric migrations of humans (Atkinson 2011). However, the principles underlying the study of phonological complexity are far from clear.

Most studies use inventory size as the main parameter for assessing the complexity of a language. However, it is often difficult to determine the precise phonemic inventory. This topic was salient in the pioneering works on phonology (cf. Trubetzkoy 1939), but it is only briefly mentioned in modern handbooks (cf. Gussenhoven & Jacobs 2011, etc.). When comparing different languages, it is important to apply uniform criteria to all the languages in the sample.

It has always been understood that other factors are relevant as well (Maddieson 2009). These factors are much more difficult to capture and are usually neglected. This leads to unsatisfying results. For example, Haugen (1958) posits 17 consonant phonemes for Icelandic. Thus, Maddieson (2011) would classify this system as “moderately small”. In other words, the consonant system of Icelandic is relatively simple. This conclusion would surprise anyone who has the slightest knowledge of Icelandic. I argue that it is also important to consider the differences between the underlying representations (UR) and the surface representations (SR) to estimate phonological complexity. For instance, in Haugen’s analysis of Icelandic the underlying form of kjóll “dress” would be /kʰjouḷr/ which surfaces as [cʰouḷ]. The UR contains five phonemes (/kʰ/, /j/, /ou/, /l/, /r/), but only one of them (/ou/) remains unchanged in the SR. This discrepancy between the UR and the SR contributes to the overall complexity of the phonological system. If described by a rule-based theory, the higher the number of rules, the more complex the phonological system will be. In a constraint-based theory, the more complex system will be the one where markedness constraints dominate faithfulness constraints. Though it is difficult to count the rules or to list all the constraints, the similarity between URs and SRs can be measured for separate words or text samples. For example, the similarity score for the word /kʰjouḷr/ [cʰouḷ] is 20% (= 1 / 5). It might be higher for other Icelandic words, but it will be relatively low for Icelandic in general as compared to other languages. I am going to present further examples in my talk.

To sum up, the study of linguistic complexity poses some challenges for phonology. We need to assess the complexity of phonological systems taking into account not only the size of the phonemic inventory, but also the extent of the discrepancy between the underlying and the surface representations.
REFERENCES


Trubetzkoy, Nikolai. 1939. Grundzüge der Phonologie (Travaux du Cercle Linguistique de Prague 7).