

You Talking To Me?

Uwe D. Reichel¹ – Mády Katalin²
MTA Nyelvtudományi Intézet

¹uwe.reichel@nytud.mta.hu, ²mady.katalin@nytud.mta.hu

The most general aim of wh-questions is to seek for information, but they can have a wide range of other pragmatic functions. In this paper we investigate self-directed questions in dialogues that are lexicalised forms of vacillation ("how should I explain?") in cognitively demanding tasks and do not directly address the interlocutor. Their prosodic properties are compared with interlocutor-directed wh-questions. Data are based on the Hungarian version of the Columbia Game Corpus, in which two players guide each other reciprocally in an object placing game. Utterances from 5 dialogues (sum: 112 minutes) with 9 speakers were analysed. We examined the following automatically extracted prosodic variables: mean energy, f0 level in terms of absolute midline slope and intercept, f0 range in terms of the RMS between base- and topline, and the local f0 shape on the question words represented by the absolute values of third order polynomial coefficients. Based on Ohala's frequency code concept we expect higher energy, higher f0 level and range values as well as more pronounced local f0 shapes for the interlocutor-directed than for the self-directed questions. These hypotheses were confirmed for most parameters: for all energy, level and range related features as well as for two out of four local contour coefficients significantly higher values were observed in interlocutor-directed questions (Mann-Whitney test, $p < 0.05$; for the remaining two features $p < 0.1$). We conclude that in our data speakers indeed make use of the universal frequency code principles which are well captured by our prosodic parameterization. Finally, we trained random forest classifiers on the introduced feature set to predict question-directedness. A mean accuracy of 85% in 10-fold cross validation was achieved. This indicates the usefulness of the proposed parameterization for off-talk detection in dialogue systems.