

# Abstract

Over the past five decades, the phenomenon of ‘*Urban Youth Languages*’ has been observed across Europe where young people in multi-ethnic urban areas follow specific linguistic practices. One example is Kiezdeutsch (‘hood German’), which is a German-language variety spoken primarily by teenagers from multi-ethnic urban neighborhoods. Since its emergence some 30 years ago, Kiezdeutsch has developed systematic linguistic structures that identify it as an independent variety of German. On the syntactic level, variants found in Kiezdeutsch reflect changes in functional categories. Examples of such changes include: bare noun phrases (NPs) lacking determiners or prepositions lack of copula verbs, Verb-first declaratives (V1), and preserved subject-verb-object (SVO) word order in sentences beginning with an adverb (ADV SVO).

The research to date has focused on either qualitative analyses or small-scale quantitative analyses of Kiezdeutsch to provide evidence. This thesis bridges the gap in the literature by providing empirical evidence from a large-scale logistic regression analysis of Kiezdeutsch. At its heart, the thesis exploits generalized linear models (GLMs) to learn which syntactic constructions found in Kiezdeutsch are characteristic of it in comparison to standard German. This is done on both the word and the phrase level using part-of-speech n-grams. The thesis identified several POS n-gram types which support the following phenomena: bare NPs, ADV SVO, and V1. Moreover, significant associations between Kiezdeutsch and POS trigrams with negation were identified. Furthermore, the thesis found limited evidence to show that lack of relative clauses is linked to Kiezdeutsch.

The contributions of this thesis are not limited to providing evidence as it also outlines, in a methodological manner, a robust approach for model and parameter selection in logistic regression analysis for Kiezdeutsch which can be applied to data from different sources like the Internet and for other fields like sociolinguistics, psycholinguistics and code-switching.