

3.6. Report the results of independence test using the following template:

We have found a significant association between variables Aux (doen_V / laten_V) and EPTrans (transitivity) ($p < 0.001$, $p = 0.0001553$). The odds of ratio were 2 times higher in Transitive group (-2.11883) than in Intransitive group (-1.1609). Effect size is small (Cramer's $V = 0.1745$).

3.8 Why is it not recommended to run multiple Chisq tests of independence on different variables within your dataset without adjusting for the multiplicity? (i.e. just testing all the pairs of variables one by one)

It is not recommended because this method without adjusting for the multiplicity will show dependence/independence of only two certain variables without taken into account that the third variable can simultaneously affect the first pair etc. Thus the complex interdependencies of several factors will not be found.

3.9 Provide a short text (300 words) describing the hypothesis on this study and the results of your analysis.

The hypothesis on this study suggests that there are certain linguistic and extralinguistic factors that influence the choice of synonymic auxiliary verbs 'doen' or 'laten' in Dutch language among which semantic classes of the causer and causee, transitivity of the verb used in the construction etc.

The analysis revealed the following dependency between the transitivity or non-transitivity of the notional verb and the choice of one of the two auxiliary verbs in question: doen tends to be used with intransitive verbs whereas laten goes with transitive verbs more frequently.

Judging from the Fisher test results the semantic class of the causer (animate or inanimate) as well as country and semantic domain of the caused event (mental or non-mental) are also influence the choice of the auxiliary verb.

Extra-linguistic factor that seems to have no connection to the choice of the auxiliary verbs in question is the topic of the text (economy, politics, football and games).