Processing canonical and non-canonical sentences in Turkish within a context
Barış Kahraman & Yuki Hirose (The University of Tokyo)
kahraman@phiz.c.u-tokyo.ac.jp

Previous studies have shown that canonical sentences are generally easier to process than non-canonical sentences. Keiser and Trueswell (2004) argued that old information tends to occur early, and new information tends to occur later in the sentences [1]. They showed that when the subject and object were mentioned in a supportive context, non-canonical sentences (OVS) were not more difficult to process than canonical ones (SVO) in Finnish. Keiser and Trueswell concluded that the processing difficulty of non-canonical sentences in the previous studies is due to violations of discourse factors, rather than the complexity/frequency based factors. However, in establishing discourse relationships, Keiser and Trueswell repeated nouns from the preceding contexts in their test sentences. Therefore, we are not sure whether their results are due to repeated noun benefit or givenness. We also want to test whether the impact of the givenness can be generalized to other languages having different word orders.

We conducted a self-paced reading experiment in Turkish (N=35) to examine the impact of the information status of the subject and object nouns on the processing of canonical (SVO) and non-canonical (OSV) sentences by manipulating the information status of the subject and the object. The target sentences were embedded in the beginning of a complement clause, and always preceded by a context such as "the ticket seller’s name was John.", as shown in (1a)-(1d). To avoid unnaturalness and name repetition effects, we used pronouns when referring to given referents in the target sentences. The complexity/frequency based accounts predict that the verbs of canonical sentences should be read faster than the verbs of non-canonical sentences because canonical sentences have higher frequency and less complexity in Turkish (Demiral, 2007) [2]. On the other hand, discourse function based accounts predict that the processing difficulty of the verb will be affected by information status of the nouns [1].

We found only a main effect of information status of the nouns ($F_1 (1,32) = 5.78, p=0.022; F_2 (1,21) = 5.51, p=0.029$), with verbs after new-given noun order being read faster than the verbs after given-new noun order. This may be due to fact that pre-verbal position is focus position in Turkish, and given nouns are easier to be focused than the discourse-new nouns. In the first two regions, there was also a main effect of the information status of the nouns, and in the first region given nouns (pronouns) were read faster than the discourse-new nouns ($F_1 (1,32) = 3.84, p=0.059; F_2 (1,21) = 5.33, p=0.031$). In the second region, on the other hand, discourse-new nouns were read faster than the given nouns ($F_1 (1,32) = 12.26, p=0.001; F_2 (1,21) = 40.84, p=0.0001$). This confirms that old-new information order is also easier to process than new-old information order in Turkish regardless of the surface word order, as in Finnish. Overall, this study suggests that non-canonical sentences are not always more difficult to process than the canonical ones, and that the processing difficulty of non-canonical sentences in the previous studies may indeed be due to violations of discourse demands. More importantly, complexity/frequency based accounts cannot explain the present results in Turkish, while discourse based factors can – thus extending Keiser and Trueswell’s results for Finnish.

(1a) [He-NOM_{given} Mary-ACC_{new} cheats] that the stationmaster said.
(1b) [Mary-NOM_{new} he-ACC_{given} cheats] that the stationmaster said.
(1c) [He-ACC_{given} Mary-NOM_{new} cheats] that the stationmaster said.
(1d) [Mary-ACC_{new} he-NOM_{given} cheats] that the stationmaster said.

References