Short-Term Scientific Missions (STSMs) to the team at the Utrecht Institute of Linguistics by Derya Cokal

The purpose of the Short-Term Scientific Missions (STSMs) visit with Prof. Sanders and his team at the Utrecht Institute of Linguistics was to: (1) perform a cross-linguistic comparison between Dutch and Turkish causal connectives; (2) determine the underlying principles of schemes for annotating causal discourse relations; (3) annotate the objective and subjective relations these devices convey, and (4) explain the cognitive mechanisms that cause writers/speakers to choose one causality marker instead of another.

During the meetings in Utrecht, the three objectives mentioned above were fulfilled and new research questions were developed. Dr. Cokal, Prof. Sanders, and his team analyzed Turkish tokens retrieved from Middle East Technical University (METU) data. This report summarizes the main issues that arose during the meetings.

A. Annotation

Initially, the group decided to use a common annotation methodology. As in (1) each argument (i.e., token and connective) is placed within brackets:

1. [hayatı seveceksin]$_{ARG1}$ çünkü.[kim sevmez hayatı?] $_{ARGU}$

Secondly, ambiguous cases that Prof. Zeyrek and Dr. Cokal found in the METU corpus were discussed. In order to determine whether the causal relation is subjective or objective each sentence is algebraically paraphrased or expressed. For example, if the sentence is put into the following form: Argument (X) leads to the fact (Y), then the causality relation between the sentences is objective. On the other hand, if the sentence is read, "Argument (X) leads me to conclude or surmise," then the causality relation is subjective. During the meeting, each ambiguous token was read and paraphrases developed. Naturally, there are exceptional tokens, where such paraphrasing does not resolve or uncover subjective/objective causal relations. In the following section, these exceptional cases and the annotation criteria set for such cases are given.

B. Types of narration and its role in the determination of objective and subjective cases:

In (2), the narration is in the first person, but the narrator tells what her brothers would have thought in that situation. Therefore, instead of a subjective causal relation, an objective causal relation is annotated.

2. [Beni almaktan yana değillerdi.]$_{ARG1}$ [Bebeği açıklayamazdık] $_{ARG2}$ çünkü.

   [They (my brothers) were not willing to take me back.] [We can't tell the baby there (in the village)] $_{ARG2}$ because.

In (3), the narration is again in the first person, but a real time situation is given in Argument 2. The narrator was on the night train between Istanbul and Ankara when the meeting time came.

3. Çünkü [randevu saati geldiğinde], [ben, Ankara'ya hareket eden gece treniyle İstanbul'u terk ediyordum.]

    Because [when the appointment time came] ARG1 [I was leaving Istanbul on the night train to Ankara.] ARG2

C. A token has a nucleus and satellite

4. Sürekli yenilenen bir birliktelikti amacımız. Annelerimiz , amcalarımız , ağabeylerimiz gibi olmamalıydık. [Oysa evlilik üzerine hiçbir şey bilmiyorduk]ARG1 Çünkü [hiç evlenmemiştik] ARG2, [yaşanmadan nasıl bilinirdi ki.] ARG3

    Our aim was always to renew our relationship. We should have not been like our aunts, uncles, and brothers. [However, we did not know anything about the marriage.] ARG1 Because [we were not get married,] ARG2 [how could we know without experiencing it?] ARG3

In (4), there is a nucleus and satellite. The relation between ARG1 and ARG2 is implicit and a subjective causal one. On the other hand, the relation between ARG2 and ARG3 is implicit and an objective relation.

Similarly, the same nucleus and satellite relation is seen in (5). Again, the relation between ARG1 and ARG2 is a subjective relation, because the writer uses "I guess" in ARG2 to state that ARG1 "the ruins being from Byzantine" is the writer's assumption. On the other hand, the relation between ARG1 and ARG3 is an objective causal one.

5. Bunun eski bir sultan sarayı olduğu söylenirdi ama bir taş yığıntısından farksız olan kalıntı besbelli bizim sultanlarımızdan çok daha eskiydi ve bir [Bizans yapısının harabesi] ARG1 olsa gerekti çünkü [taşları eşelediğinizde renk renk mozaik parçaları bulabiliyordunuz.] ARG2

    It was said that it was an old sultan palace, but it was not different from a pile of stones, the ruins were obviously older than our sultans and [there were Byzantine ruins] ARG1, [I guess] ARG2, because [when you scrabbled the stones, you could find coloured mosaic pieces.] ARG3

D. Other causal connectives (zira and bundan dolayı) in subject causal relations

Though there were not many tokens with "bundan dolayi" in the METU corpus (N=16), Prof. Zeyrek and Dr. Cokal agreed to handle causality connectives with a deictic expression (because of this) separately from explicit connectives. This methodological decision is illustrated in (6), in which the first person narration and the wording "in favour of" make the causal relation subjective.

6. [Kesinlikle bu işin barışçıl yollarla çözümünden yanayız ve BM silah denetçileri şu ana kadar herhangi bir nükleer ya da kimyasal silahla ilgili Bulguya ulaşmış değil.] Bundan dolayı da [bir yıl daha bu işin uzatılmasını talep ettiler.]

   [We are definitely in favour of a peaceful resolution and the United Nations weapon inspectors have not found any nuclear and chemical weapons]. Because of this, [they demanded a one-year extension.]

The annotations performed by Prof. Zeyrek and Dr. Cokal have shown the retrieved tokens from the METU corpus with "zira" (because) were only used in subjective causal relations. For example, in (7) below, the narrator uses "for me" to state the reason why he runs away from the place.

7. Fırtınanın yaklaşmakta olduğunu sezinlediğim an [kaçardım], zira [bana göre dayak yemek çok alçaltıcı bir muamele sayılırdı.]The storm was approaching, [I ran] because, I think [being beaten was a very degrading attitude.]

E. Volitional and non-volitional marked Turkish connectives

During the meeting at Utrecht, an examination of the initial distribution of için and çünkü in agreed annotation cases revealed "çünkü" was used in a subjective relation in 90% of them. On the other hand, in 70% of için cases, the relation was objective. This led to the group questioning whether Turkish discourse connectives have such volitional and non-volitional distinctions. Specifically, "Does the high use of için in objective relation express volitional act? Are the 30% of icin cases with a subjective relation non-volitional and/or express speech acts? In an attempt to answer these questions, the annotation will be examined again and tagged regarding its volitionality/non- volitionality status.

F. Acceptability of the use of çünkü

During the meeting another important issue arose. In (8), the use of "çünkü" seems to be ungrammatical instead of için. Besides, in order to be able to use çünkü, the argument structure in the token needs to be changed. Therefore, the group decided to retrieve similar cases and find out the commonalities in the tokens and question why için is grammatical rather than çünkü.

8. [Top ve para sahibi oldukları] $_{ARG1}$ için [iktidarlarını sürdüreceklerdir] $_{ARG2}$ ama ne zamana kadar?

   [Since they have money and guns,] $_{ARG1}$ [they'll always control the government,] $_{ARG2}$ But how long will they control it?

As a result of the week's meetings, the group decided to undertake the following additional data analysis: (1) go over the annotations and tag

regarding (a) volitionality/non-volitionality, and (b) annotate narrator perspective according to methodology outlined in Sanders & Spooren (2015). The initial results will be presented at the Discourse Relational Devices, LPTS 2016 Conference, January 22, in Valencia, Spain.

I. REFERENCES

Sanders, T. J. M., & Spooren, W. P. M. S. (2015). Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text. *Linguistics*, *53*(1), 53-92.