In this talk, we introduce a model to capture two strategies of reference tracking: continuing topic and topic shift in the discourse-configurational language (É. Kiss, 1995), Hungarian and the non-configurational language, Lakhota (Ullrich, 2016). For modeling the discourse-semantics of the given phenomena, we propose a dynamic (update) model based on frame-semantic representations (Kallemeyer & Osswald, 2013, Petersen, 2015, Löbner, 2017). The syntactic representations are built upon the framework of Role and Reference Grammar (Van Valin, 2005).

We propose a frame-based model of the cognitive process of discourse interpretation, where next to the information updates reference tracking is of central importance. Certain morphosyntactic choices depend on whether the given NP expresses a continuing topic or a topic shift. In pro-drop languages, it often determines the choice between a zero morpheme and a noun phrase or demonstrative pronoun. Consider the following Hungarian examples based on the ‘Frog Story’ (Mayer, 1967).

(1) A kisfiú kergette a béka-t,  
the boy chased the frog-ACC  
‘[The boy]TOP chased the frog,’
1. aztán el-ugrott egy faág-ra.  
then PRT-jumped a branch-SUBL  
‘and then he jumped away to a branch.’
2. #aztán a kisfiú el-ugrott egy faágra.  
then the boy PRT-jumped a branch-SUBL  
‘and then the boy jumped away to a branch.’
3. aztán a béka/az el-ugrott egy faágra.  
then the frog/that one jumped away to a branch.
   then the frog/that PRT-jumped a branch-SUBL  
   ‘and then the frog/that one jumped away to a branch.’

In (1), the sentence topic is the referent of the boy. The continuations illustrate different morpho-syntactic choices. In (1a), the subject NP is dropped, coded as zero, expressing the continuation of the sentence topic. In this case, a full NP is out (1b). In (1c), the sentence topic is shifted to the referent of the frog and this is signaled by either the demonstrative az ‘that’ or the NP a béka ‘the frog’ in topic position.

Similar examples can be found in Lakhota, a head-marking language. In Lakhota, pro-drop is used extensively, and the inflected verb alone can form a complete sentence. Lakhota has two definite articles: k(ŋ) ‘the’, which can be used to refer to referents in the immediate discourse context (physical environment, old referents, associative anaphora), while k’uŋ can only be used in a noun phrase referring to an explicitly mentioned discourse antecedent.

(2) a. Wičháša waŋ1 wa-náse-Ø1-I man  
w-aNU-hunt-3SGA-go and  
w-aØ1-pȟáta-haŋ yunŋkáŋ winyan waŋ2  
NU-3SGA -butcher-CONT and.here woman a
él hi-ná<Ø₂>žíŋ k’eyaš átayaš a<Ø₂>yúta-šní (...) 
there arrive-stand<3SGA> but entirely look.at<3SGU>-NEG

‘A man₁ went hunting, and he₁ was cutting up [his game], and here a woman₂ arrived and stood there, but he₁ didn’t look at her₂ at all (...)’

b. (...) naŋ waŋná Ö₁-khi-gniŋ-kta
and now 3sgA-arrive-go.back-POT
čhaŋké wíŋyaŋ k’uŋ₂ ...lo-Ø₂-čhiŋ ki and.so woman the.AF ... food-3sgA-want the
úŋ a<Ø>yúhel na<Ø₂>žíŋ-haŋ (...) on.acct.of wait<3sgU> stand<3sgA>CONT

‘(...) and now he₁ was going to go back; and so the aforementioned woman₂ was standing waiting around on account of wanting food (...)’

(Boas & Deloria 1942:160)

The two protagonists in the story (‘The Stingy Hunter’) are introduced as indefinite NPs; the man (the hunter) is zero coded from the second clause on and is the continuing topic until the second clause in (2b) in which the woman is coded by N k’uŋ, signaling that she is the new sentence topic. When the hunter returns to being the sentence topic seven clauses later, he is referred to by N k’uŋ.

In our talk, we introduce the details of our frame-based model providing the sentence-level representation and a model of the immediate discourse context. The frames used in our representations are defined as base-labelled typed feature structures following the approach of Kallmeyer & Osswald (2013).

The sentence-level representation (see example (3)) registers both the newly introduced (N) and the anaphoric discourse referents (A), and the content of the sentence, represented by a frame (Fₖ). The discourse referents are anchors for reference in the given discourse. All discourse referents are linked to base-labelled nodes in the corresponding frame, the link represented by cross-labelling. In the example below, for example, the discourse referent dₓ is linked to the node labelled x in the frame.

(3) The boy chased the dog.

\[ \langle \mathcal{N}, \mathcal{A}, F_\delta \rangle = \langle \{d^e\}, \{d^x, d^y\}, \rangle \]

The local context is formally modeled as a pair \( \langle \mathcal{R}, F_c \rangle \) of the set of discourse referents \( \mathcal{R} \) and a frame representation \( F_c \) of the information given at that point of the discourse. Updating the context with an utterance, \( \langle \mathcal{R}, F_c \rangle | \langle \mathcal{N}, \mathcal{A}, F_\delta \rangle \), adds the newly introduced referents to the discourse \( \mathcal{R} \cup \mathcal{N} \) and the context-frame is extended by the sentence-frame via frame unification \( F_c \sqcup_{\delta} F_\delta \). Anaphoric referents are constrained as
their corresponding base-labelled nodes in the sentence-frame must be unified with a base-labelled node in the context-frame, hereby expressing anaphora resolution.

Topic selection is represented both at the sentence-level and in the immediate discourse context as a distinguished element in the set $N$ or $A$ and in $R$. The process of continuing topic can be captured similarly to anaphora resolution together with a constraint that it must be resolved as the previous topic. The process of topic switch explicitly states a change of the topic selection during the update process. In our talk, we will introduce the details of the processes sketched above together with the syntax-semantic interface of the morpho-syntactic strategies. We will discuss language specific issues, examples from Hungarian and Lakhota, as well as generalizations on the phenomena of continuing topic and topic switch.


