Udi Person Markers and Lexical Integrity

Michael T. Wescoat

1 The problem. Harris (2002) observes that Udi person markers (PMs) like *ne* in (1) may occur (a) between morphemes in a verb, (b) inside a monomorphemic verb, e.g. in *ef* 'keep,' and (c) on a focus constituent (FocC) before the verb. She proposes alignment principles like (2) as OT constraints, ranked (a) ≫ (b) ≫ (c). Harris reasons that since the analysis of Udi PMs requires reference to both syntactic constructs (FocC) and word-internal entities (incorporated element, verb stem), it challenges the Lexical Integrity Hypothesis advocated by Bresnan and Mchombo (1995), among others. In this paper, I argue that the notion of *lexical sharing* (Wescoat 2002) leads to a view of the lexicon in which all of the information necessary to predict the placement of Udi PMs is available in the lexicon, thus preserving the lexicalist separation between syntax and morphology.

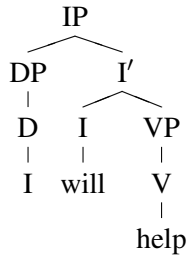
(1) a. aš-**ne**-b-sa
work-3SG-*do*-PRES
'She works.'

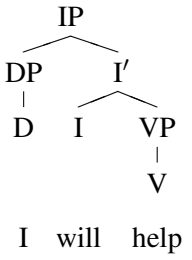
b. e-**ne**-f-sa
keep$_1$-3SG-*keep*$_2$-PRES
'She keeps (it).'

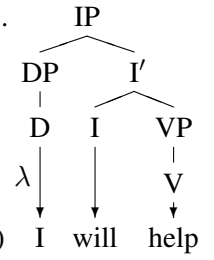
c. merab-en ayt-**ne** ef-sa
Merab-ERG *word.*ABS-3SG *keep*-PRES
'Merab keeps **his word**.'

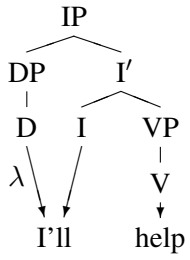
(2) a. Align the left edge of the PM with the right edge of a FocC. [see (1c)]
b. Align the left edge of the PM with the right edge of an incorporated element. [see (1a)]
c. Align the right edge of the PM with the right edge of the verb stem.[1] [see (1b)]

2 Lexical sharing. Wescoat proposes lexical sharing for cases where one word created in the lexicon corresponds to two elements of c-structure. Examples include some English monosyllabic pronoun-auxiliary contractions, like *I'll* [ajl/ɑl], exhibiting various lexical idiosyncrasies. One may approach lexical sharing in four conceptual steps, illustrated in (3): (a) Start with a traditional c-structure; (b) Sever the word-labeled terminal nodes from the tree, and arrange the words in a separate, linearly ordered representation, which one may call *l*(*exical*)-*structure*; (c) Establish a mapping, called λ, between the new terminals of c-structure (the former preterminals) and the words of l-structure; (d) Since λ is independent of the c-structure relation of *domination*, it is possible to have λ map two or more terminal nodes to the same word, in which case the former are said to *share* the latter, whence the name *lexical sharing*.

(3) a.


b.


c.


d.


Lexical sharing extends the information listed in the lexicon. *Lexical exponence rules* encode syntactic categories (and ultimately feature information) for an ordered sequence of one or more adjacent terminals. Notationally, "I'll ← D I" indicates that λ can map adjacent D and I into the same word, *I'll*.

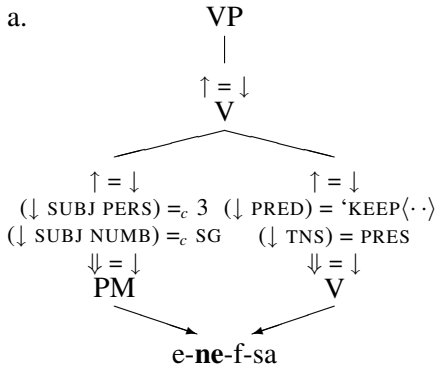
Harris argues that Udi PMs are clitics. Wescoat (2002, 2005) shows that lexical sharing is a useful tool for analyzing some varieties of clitics, and I propose to extend this approach to Udi PMs. The PM and its host form a word that is shared by two terminals in c-structure, much as *I'll* is shared in (3d).

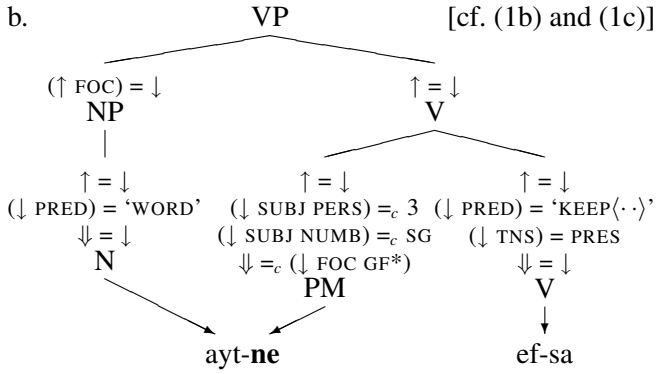
3 LFG with lexical sharing. To put lexical sharing into LFG, one may assume three things: (a) Words in l-structure map into f-structures; (b) The metavariable ⇓ means 'f-structure of the word below'; (c) The right-hand sides of lexical exponence rules are given annotations, as in (4), to be associated with the c-structure terminals instantiated by the word on the rule's left-hand side, as in (5).

(4) a. e-**ne**-f-sa ← PM V
(↓ SUBJ PERS) =$_c$ 3 (↓ PRED) = 'KEEP⟨· ·⟩'
(↓ SUBJ NUMB) =$_c$ SG (↓ TNS) = PRES
⇓ = ↓ ⇓ = ↓

b. ayt-**ne** ← N PM
(↓ PRED) = 'WORD' (↓ SUBJ PERS) =$_c$ 3
⇓ = ↓ (↓ SUBJ NUMB) =$_c$ SG
⇓ =$_c$ (↓ FOC GF*)

[1] The optimal alignment is one segment from the end of the verb stem; only this segment separates the aligned right edges.

(5) a. 

b. 

[cf. (1b) and (1c)]

The VPs in (5) reflect the fact that Udi is verb-final, with a position for FocC immediately before the verb complex. The annotation "$\Downarrow =_c$ (↓ FOC GF*)" on PM in (4b) implements a constraint requiring that N be part of a FocC. Note that this crucially exploits the mechanisms furnished by an LFG with lexical sharing. In the absence of "$\Downarrow =_c$ (↓ FOC GF*)," the equation "$\Downarrow = \downarrow$" would occur by default, and that would effectively limit the choice of host for the PM to the verb, lest functional inconsistency result.

4 Optimality theory. I assume defeasible constraints, as in OT-LFG (Bresnan 2000). I eliminate (2a), replacing it with (6a) and (6b); the constraints are ranked (6a) ≫ (2b) ≫ (2c) ≫ (6b). I follow Harris in assessing one violation for every phonological segment separating aligned edge points. One can calculate violations per lexical item, as seen in (7). I assume that words containing no PM, like *ayt* 'word' and *ef-sa* 'keep-PRES,' do not violate any of the constraints under consideration. Likewise, I assume that items do not violate a constraint when a morphological reference point mentioned in the constraint (incorporated element, verb stem) is lacking. In evaluating the sentence, the violations for each word are summed, as in tableau (8), which is left partial to save space.

(6) a. Mark the PM for focus (i.e. annotate with "$\Downarrow =_c$ (↓ FOC GF*)").
b. Align the left edge of the PM with the right edge of the stem.

(7) a.

	(6a)	(2b)	(2c)	(6b)
(4a) e-**ne**-f-sa	*		* [f]	***** [nefsa]

b.

	(6a)	(2b)	(2c)	(6b)
(4b) ayt-**ne**				

(8)

	(6a)	(2b)	(2c)	(6b)
☞ (1c) ayt-**ne** ef-sa				
ayt e-**ne**-f-sa	*!		*	*****

This scheme favors marking the PM for focus, requiring the PM to be associated with FocC if one is present. In the absence of a FocC, however, a sentence like (1c) could not even be generated; consequently a sentence with the PM associated with the verb would become the most harmonic by default, the position of the PM morpheme within the verb being determined by the alignment constraints. Note that in determining the host of the PM, it is not necessary to consult syntactic structure; lexical sharing puts enough information within the grasp of the lexicon to achieve the right effects.

5 Conclusion. Adopting lexical sharing in OT-LFG expands the information available within the lexicon in a way that makes functional constraints on the choice of clitic hosts a lexical matter. Thus, this work demonstrates the viability of the lexicalist approach in the face of a new empirical challenge.

Bresnan, Joan (2000) 'Optimal syntax.' In Joost Dekkers, Frank van der Leeuw, & Jeroen van de Weijer, eds., *Optimality Theory: Phonology, Syntax, and Acquisition*. Oxford: Oxford University Press, 334–385.

Bresnan, Joan, & Sam A. Mchombo (1995) 'The lexical integrity principle: Evidence from Bantu.' *Natural Language and Linguistic Theory* 13, 181–254.

Harris, Alice C. (2002) *Endoclitics and the Origins of Udi Morphosyntax*. Oxford University Press.

Wescoat, Michael T. (2002) *On Lexical Sharing*. Ph.D. dissertation, Stanford University.

(2005) 'English nonsyllabic auxiliary contractions: An analysis in LFG with lexical sharing.' In Miriam Butt, & Tracy Holloway King, eds., *Proceedings of the LFG05 Conference*. Stanford: CSLI, 468–486.