
Wao Terero lexical suffixes: Bridging the lexicon and discourse

XXX
YYY

In this talk I discuss a formal treatment of a subcomponent of the Wao Terero lexical suffix system, adjectival classifiers. Wao Terero (Glottocode waor1240) is a linguistic isolate spoken in the Ecuadorian Amazon. Data is from ongoing fieldwork. Lexical suffixes are bound elements that provide nominal meanings to their host constructions (Sapir 1911). In the context of my fieldwork, I utilize a fragment methodology, where grammatical models encode hypotheses that can be tested through elicitation. To support this effort, I have developed Lexical Proof Morphology (LP), a theoretical framework embedded in a constructive logic (Coquand & Huet 1988). The theory is particularly concerned with interface issues and is compatible with the tenets of Word and Paradigm Morphology (WP) (Robins 1959). It seamlessly interfaces with existent, broad coverage theories of syntax (Pollard & Worth 2015) and dynamic semantics (Martin & Pollard 2012). This integration is necessary for a treatment of Wao Terero lexical suffixes. Lexical suffixes have predictable but polysemous lexical semantic meanings. Some may be used as classifiers, where they play a role in anaphoric constructions. In these cases, the polysemy offered by the lexicon must be narrowed to include only those qualities compatible with a referent. Therefore, the dynamic semantic context must limit lexical variability. At the same time, discourse provides another domain where classifier constructions receive multiple interpretations. These are driven by the dichotomy created by information introduction and information reference. This means that morphological forms sit at the intersection of multiple interpretations in two semantic domains. On the one hand, this relationship evokes classic realizational assumptions. On the other, as will be made clear, the dynamic context recruits lexical meaning into diverse interpretative contexts, such that realization also behaves as a conduit between two semantic domains. LP makes the relationship between intrinsic, lexical semantic meanings and extrinsic, dynamic meanings explicit using a unique, proof-theoretic realizational architecture.

The Wao Terero lexical suffix system is complex. There are roughly 35 suffixes. These may occur with nearly every part of speech – including demonstratives, nouns, adjectives, verbs, question words, and others. Each construction type comes with particular quirks and semi-productive nuance. For the sake of simplicity, I focus on productive, transparent adjectival constructions but it is helpful to get a taste of the broader system. As examples (1) and (2) show, where lexical suffix glosses are in bold, lexical suffixes may be used in constructions that have compound and incorporation-like meanings. This is despite the fact that the suffixes do not correspond to the stems of free words. Meanings may also be classifier-like (Peeke 1968), as in (2) and (3), depending on the construction type and context. Though the classifier constructions in these examples may resemble grammatical agreement, classifiers are optional. They are acceptable in many contexts but may occur only occasionally in spontaneous speech.

- | | | |
|--|---|--|
| (1) <i>kewe-ñabo</i>
cassava- leaf
'cassava leaf' | (2) <i>Onom-po kem-po-tabopa.</i>
body- hand cut- hand -1.past
'I cut my hand.' | (3) <i>Ñene-po wipo impa.</i>
big- canoe canoe copula
'The canoe is big.' |
|--|---|--|

No comprehensive formal treatments of lexical suffixes exist in the literature but there are some outline proposals. Wiltschko (2009) sketches a Distributed Morphology (DM) (Halle & Marantz 1993) treatment of a similar system in a Salishan language. Her proposal takes

advantage of the hybrid nature of DM to treat lexical suffixes in an item and arrangement manner. Specifically, the suffix and its host are roots below a root node: [$\sqrt{\text{root}}$ [$\sqrt{\text{host}}$] [$\sqrt{\text{suffix}}$]]. As (2) and (3) show, lexical suffixes do not inhabit syntactic argument positions, otherwise they would block the occurrence of overt nominal arguments. Placing lexical suffixes in sub-root non-argument positions allows for this quality of lexical suffix behavior according to DM-like assumptions.

Wiltschko's representation of lexical suffixes as roots in hierarchical configurations is inadequate for Wao Terero and likely other languages. There are two reasons for this. One is lexical semantic and the other is due to dynamic semantics. The lexical issue is easily explained. Wao Terero lexical suffix constructions are highly polysemous. The form-meaning correspondences of DM roots assume meaning underspecification, which allows for some quasi-polysemous meaning variation. Wao Terero patterns exhibit true polysemy, multiple meanings (Copestake & Briscoe 1995), rather than underspecification.

- | | | |
|--|--|---|
| (4) <i>ñene-we</i>
big-plant/tree/pole
'big (plant/tree/pole)' | (5) <i>ñene-mo</i>
big-eye/fruit/face
'big (eye/fruit/face)' | (6) <i>ñene-mpo</i>
big-canoe/hand/finger
'big (canoe/hand/finger)' |
|--|--|---|

Polysemy is particularly clear in adjectival classifier constructions, where the lexical suffix imposes a selectional restriction on an argument, which may be explicit in syntax or supplied by discourse. This is the case in (4), (5) and (6), where the selectional restriction varies. This means that (6) is an appropriate answer to the questions, "What is her hand like?" or "What is her canoe like?" but not a question about a plant.

The second problem posed by a root-configuration approach requires consideration of the discourse context. The hierarchical root schema is far too simplistic to predict the diversity of lexical suffix construction interpretations in discourse, in particular their role in introducing information and their role as anaphora. Adjectival classifier constructions may play (at least) three discourse roles. In role 1 the construction modifies a noun when a referent is introduced into discourse (see (3)), analogous to *A short boat exists*. In both role 2 and 3, the construction is anaphoric. In role 2 only the adjectival information is new information, analogous to *The boat is short*. In role 3 both the adjectival and classifier content are part of the descriptive content of the anaphor, *It is the short boat*.

The diagnostic for establishing that adjectival constructions can serve each of these roles involves negation. In formal pragmatics it has been observed that the descriptive content of an anaphoric expression is not offered for acceptance or rejection in discourse (Roberts 2010). The interlocutors presuppose its validity. This means that if one negates an expression containing an anaphor, the descriptive content of that anaphor will not be negated. For example, the 'boat' meaning in the anaphoric expression *the boat* falls outside the scope of negation in *It isn't the boat*. One can felicitously follow up with *The boat is black*. This would not be the case in the non-anaphoric case *It isn't a boat*. Presenting a full paradigm of these diagnostics for Wao Terero constructions will not fit within the confines of this abstract. Role 2 is perhaps the most interesting case because the diagnostic predicts a split interpretation of meaning components, where the adjectival meaning is negated but not the classifier's descriptive content. This can be seen in (7), which is an answer to the question "Is the canoe short and red?" In it *okampo*, 'short', is negated, but not the descriptive content of the classifier.

- (7) *Obatawe wii oka-mpo inamai impa.*
red not short-canoe not is
'It is red but not short.'

The proposal of a particular phrase-like structure cannot speak to these discourse phenomena. The DM proposal fails to predict anything like it, as would similar schemata proposed in popular construction morphologies (Booij 2010). Truly accounting for form-meaning correspondences requires a linkage to some engine of interpretation, such as dynamic semantics, where meanings are composed and their entailments can be verified.

LP is a multi-paradigm theory, roughly in the mold of (Sadler & Spencer 2001). A morphological paradigm space exists, which is a non-symmetrical taxonomic space of triples (mc, mf, lx) , called form entries, where mc is a purely morphological category, called an m-cat, mf is a morphological form, and lx is a lexeme identifier. LP provides a declarative system for defining this space that does not make direct reference to syntactic or semantic categories. An mc tends to be named after formal (phonological) characteristics of a form. So $\tilde{n}enempo$, ‘big (canoe, hand, fore-paw etc.)’, has a category of po . This category may look redundant here but in inflectional systems with a high degree of allomorphy, these categories span diverse forms. It is also important to note that the category is for the whole of the form, not just the suffix. A form entry for $\tilde{n}enempo$ is $(po, \tilde{n}enempo, \tilde{N}ENE)$. The lexical entries, called signs, of the system constitute the syntax-facing paradigm. There is no intermediary notion of cells but there is a family of relations between form-entries and signs called form-sign mappings that provide generalizations over paradigmatic structure. These are declarative, natural deduction-style rules. A simplified example is below. It lacks important but distracting technical details.

$$\frac{(mc, mf, lx) \quad class(lx) \leq adj \quad meaning(s, lx, mc)}{(mf, RefAdj, (\lambda P_1 P_2. P_1(x)/P_2(x))(\pi_1 s)(\pi_2 s))}$$

This says that given a form-entry with a lexeme of a particular class – in this case adjectival – when there is a meaning s , corresponding to the lexeme and m-cat, there is a sign with the corresponding form-entry’s morphological form, the syntactic category RefAdj, for referential adjective, and a meaning derived from s . This form-sign mapping could be seen as the analog of a *cell* for adjectives that have a referential sub-part. The sign it is used to prove can be seen as the realization of that cell for a lexeme.

The meaning of the resulting sign is complex. For adjectives of this type, s is a pair of an adjectival predicate and classifier predicate. The πs are projection functions for accessing these elements. The slash notation, $P_1(x)/P_2(x)$, is essentially an annotation to avoid going into details of the dynamic semantic theory. It is intended to communicate that the two predicates exist in different scopal relationships to logical operators, such as negation. P_2 , the classifier meaning, should be outside negation’s scope. The value of the variable x is supplied by discourse.

An important part of the form-sign mapping is the predicate *meaning*, which is true when there is a meaning s for lx and mc . That is to say, it is true when a lexeme of a particular category has a meaning. The existence and value of s is determined by proof. In order to provide such a proof further axioms and theorems must be provided.

The lexical semantic system assumed here is minimal. This is because elaborating that system is a matter of ongoing research into Wao Terero lexical patterns. In the system assumed here, intrinsic meanings are associate with a lexeme as pairs $(\tilde{N}ENE, big)$, $(\tilde{N}ENE, fat)$, etc. These are axioms of the lexical semantic theory. The intrinsic meanings associated with categories are likewise given as axioms, $(po, hand)$, $(po, canoe)$. Then rules are provided to describe the licit combinations of these meanings.

$$\frac{(lx, P_1) \quad (mc, P_2) \quad class(lx) \leq adj \wedge mc \leq classifier}{meaning((P_1, P_2), lx, mc)}$$

The first elements of the premise in the example rule above refer to the previously mentioned pairs. The notion that lexemes belong to classes has already been introduced. It is also

the case that morphological categories are hierarchically ordered, as indicated by the use of $mc \leq classifier$. In this case the m-cat hierarchy ensures that only the lexical suffixes that distribute like classifiers take part in the rule.

An example proof using the rule would be:

$$\frac{(\tilde{N}ENE, \text{big}) \quad (po, \text{canoe}) \quad class(\tilde{N}ENE) \leq adj \wedge po \leq classifier}{meaning((\text{big}, \text{canoe}), \tilde{N}ENE, po)}$$

The result of this proof can be used to provide the meaning to the form-sign mapping above.

$$\frac{(po, \tilde{n}enempo, \tilde{N}ENE) \quad class(ene) \leq adjectival \quad meaning((\text{big}, \text{canoe}), \tilde{N}ENE, po)}{(ene, \text{RefAdj}, \text{big}(x)/\text{canoe}(x))}$$

The point of all this is that when this sign is composed with other signs, the restriction on the referent is clear. Lexical ambiguity is still available based on the multiplicity of homophonous signs that may be proven. This captures the notion that when a speaker makes an assertion about some previously mentioned *canoe*, they are not also vaguely making a statement equally applicable to some *hand* or other compatible referent.

Realizational mechanisms have been provided in two domains, with the lexical semantics directly feeding the dynamic semantics. This is the bridge between the intrinsic and extrinsic. Polysemy is navigated by the *meaning* predicate, while the multiplicity of discourse interpretations is handled by form-sign mappings. This realization apparatus utilizes meaning representations that are plausible, justified and interpretable based on practices in state of the art semantic theories.

References

- Booij, Geert. 2010. *Construction morphology*. Oxford University Press.
- Copestake, Ann & Ted Briscoe. 1995. Semi-productive polysemy and sense extension. *Journal of Semantics* 12. 15–67.
- Coquand, Thierry & Gérard Huet. 1988. The calculus of constructions. *Information and Computation* 76(2). 95–120. [https://doi.org/10.1016/0890-5401\(88\)90005-3](https://doi.org/10.1016/0890-5401(88)90005-3).
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth L. Hale & Samuel J. Keyser (eds.), *The View from Building 20*, 111–176. Cambridge, MA: MIT Press.
- Martin, Scott & Carl Pollard. 2012. A higher-order theory of presupposition. *Studia Logica: An International Journal of Symbolic Logic* 100. 727–751. <https://doi.org/10.1007/s11225-012-9427-6>.
- Peeke, M. Catherine. 1968. *Preliminary grammar of auca*. Indiana University PhD.
- Pollard, Carl & Chris Worth. 2015. Coordination in linear categorial grammar with phenogrammatical subtyping. In *Proceedings for the esslli 2015 workshop on empirical advances in categorial grammar (cg 2015)*, 162–182.
- Roberts, Craige. 2010. Retrieval and definite noun phrases. In *Chicago workshop on semantics and philosophy*.
- Robins, Robert H. 1959. In defence of WP. *Transactions of the Philological Society* 58(1). 116–144.
- Sadler, Louisa & Andrew Spencer. 2001. Syntax as an exponent of morphological features. In *Yearbook of Morphology 2000* (YOMO (Yearbook of Morphology)), 71–96. Springer.
- Sapir, Edward. 1911. The problem of noun incorporation in American languages. *American Anthropologist* 13(2). 250–282.
- Wiltschko, Martina. 2009. $\sqrt{\text{root}}$ incorporation: Evidence from lexical suffixes in Halkomelem Salish. *Lingua* 119(2). 199–223.