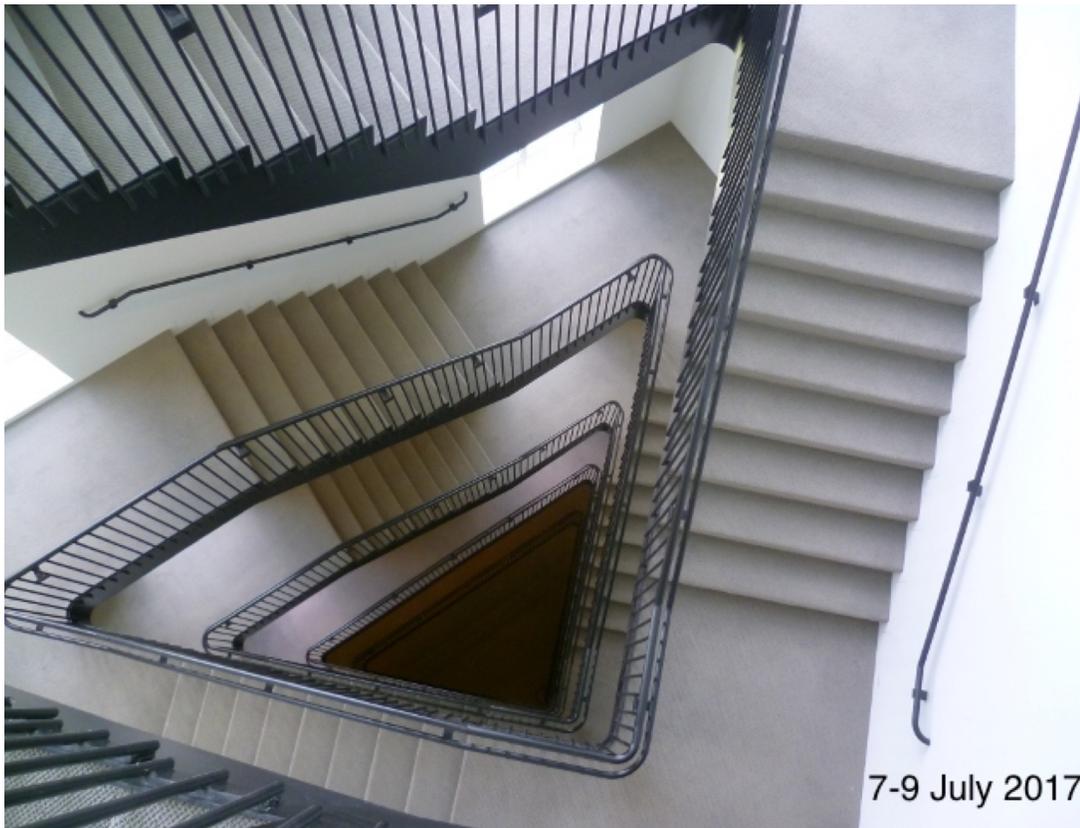


HPSG 2017



24th International Conference on Head Driven Phrase
Structure Grammar

Abstracts



Office of the Vice President
for Research



<http://hpsg.uky.edu/>



We are pleased to welcome you at the 24th HPSG Conference in Lexington, Kentucky. This work could not have been carried out without the insights, advice, and material support of a number of people. Our sincere thanks go to the members of the program committee for their helpful insights:

- Anne Abeillé (Université Paris Diderot and CNRS)
- Doug Arnold (University of Essex)
- Emily Bender (University of Washington)
- Francis Bond (Nanyang Technological University)
- Gosse Bouma (University of Groningen)
- George Broadwell (University of Florida)
- Philippa Cook (Goethe-Universität Frankfurt am Main)
- Berthold Crysmann (Université Paris Diderot and CNRS)
- Kordula De Kuthy (Universität Tübingen)
- Daniel Flickinger (Stanford University)
- Antske Fokkens (Vrije Universiteit Amsterdam)
- Petter Haugereid (University of Bergen)
- Fabiola Henri (University of Kentucky)
- Anke Holler (Universität Göttingen)
- Jong-Bok Kim (Kyung Hee University)
- Jean-Pierre Koenig (SUNY at Buffalo)
- Robert D. Levine (Ohio State University)
- Nurit Melnik (The Open University of Israel)
- Philip Miller (Université Paris Diderot)
- Stefan Müller (Humboldt Universität zu Berlin)
- Tsuneko Nakazawa (The University of Tokyo)
- Joanna Nykiel (University of Silesia)
- Gerald Penn (University of Toronto)
- Manfred Sailer (Goethe-Universität Frankfurt am Main)
- Pollet Samvellian (*Université Sorbonne Nouvelle*)
- Sanghoun Song (Incheon National University)
- Stephen Wechsler (University of Texas at Austin)
- Shûichi Yatabe (University of Tokyo)
- Eun-Jung Yoo (Seoul National University)

In addition to the scientific support of the program committee, we appreciate the following institutions and people for their material support:

- Department of Linguistics, Ohio State University
- Department of Modern and Classical Languages, Literatures and Cultures, University of Kentucky
- Office of the Vice President for Research, University of Kentucky
- Stefan Müller, Institut für deutsche Sprache und Linguistik
- VisitLex

Finally, we would also like to thank Sahar Taghipour and Koorosh Ariyae, graduate students at University of Kentucky's Master's program in Linguistic Theory and Typology.

Rui Chaves, Organizing Committee Chair
Fabiola Henri, Local Organizing Committee Chair

The HPSG community supports academic dialog and collaboration between people of all nations and languages. We are excited to inform you that on top of the main session of the conference, HPSG 2017 will include papers focusing on languages spoken in Seven countries. For this reason, we are going to have a special session, called Languages of the Seven. The goal of this special session is to promote free academic dialog in spite of policies which try to restrict it. It focuses on languages from Iran, Iraq, Libya, Somalia, Sudan, Syria and Yemen, and is not restricted to any particular theoretic framework.

Invited Talk: Laura Michaelis & Paul Kay	Partial Inversion in English	1
Tali Arad Greshler & Nurit Melnik	Backward Control in Modern Standard Arabic	4
Abeer Alsulami, Doug Arnold & Robert Borsley	Simple and Complex Comparatives in Modern Standard Arabic	9
Berthold Crysmann	Resumption and Case: a new take on Modern Standard Arabic	14
Poster Session		
Sahar Taghipour	Subject agreement in Laki	19
Ali Salehi	Rule conflation of definite nouns in Kurdish	24
Elodie Winckel and Stefan Müller	French dont-relative clauses (and their gap)	32
Kei Tanigawa	NP-internal Coordination and Agreement in French	37
Maksymilian Dabkowski	Yucatecan Control and Lexical Categories in SBCG	43
Justin Bai, Maksymilian Dabkowski, Kalinda Pride & Nicholas Tomlin	Yucatec Maya in SBCG: A Fragment	48
Takafumi Maekawa	English Prepositional Numeral Constructions	53
Aixiu An & Anne Abeillé	Agreement and interpretation of binominals in French	58
Douglas Ball	'VP' Adverbs without a VP: The Syntax of Adverbs in Tongan	63
Berthold Crysmann & Olivier Bonami	Atomistic and holistic exponence in Underspecified Realisational Morphology	68
Thomas Diaz, Jean-Pierre Koenig & Karin Michelson	Oneida prepronominal prefixes in Information-Based Morphology	73
Robert Hepburn-Gray	Against Split Morphology: 'inflection', 'derivation', and the structure of the SBCG lexicon	78
George Broadwell	Parallel affix blocks in Choctaw	83
Petter Haugereid	An incremental approach to gapping constructions	88
Sarah Hyeyeon Lee	The Syntax of the not only...but also... Construction: A Linearization Account	93
Invited Talk: Robert Borsley	Explanations and 'engineering solutions'? Aspects of the relation between Minimalism and HPSG	99

Languages of the Seven session

Partial Inversion in English

Paul Kay (Stanford/UC Berkeley) and Laura A. Michaelis (University of Colorado Boulder)

This study concerns the SBCG representation of five English constructions that combine special grammatical form with a special discourse pragmatics. The **grammatical features** include agreement between an intransitive verb and an argument that appears in what is usually thought of as direct object position. While the postverbal nominal argument has certain subject properties (in particular, controlling verb agreement), the preverbal argument has others (including undergoing raising and occupying preverbal position). We refer to this ‘split’ in subject properties as *partial inversion*. The **discourse-pragmatic features** are presentational illocutionary force and a postverbal focused argument. The constructions are:

- Presentational-*there*: *The earth was now dry, and there grew a tree in the middle of the earth.*
- Deictic Inversion (DI): *Here comes the bus.*
- Existential-*there*: *There’s a big problem here.*
- Presentational Inversion (PI): *On the porch stood marble pillars.*
- Reversed Equative-*be*: *The only thing we’ve taken back recently are plants.* (Corbett 2006: 63-64)

Although some of these constructions have been treated extensively in the literature (especially PI and Existential-*there*), to our knowledge no previous analyses have attempted to account for both the formal and interpretational similarities among all five, or the details of their formal and interpretive differences. Our observations, based on web attestations of the relevant patterns, include the following:

1. The post-verbal, focal NP cannot be characterized as having a particular (structural or inherent) case; both nominative and accusative variants are attested.
2. DI is an aspectually sensitive construction, yielding progressive construals of simple-tense dynamic verbs (e.g., *Here came the waitress*).
3. Some of the maximal constructions specify verbal and adverbial LID values, while others do not.
4. The ‘setting’ argument of PI is not as constrained as it is typically said to be: it need not fill the discourse-pragmatic role of topic (*pace* Webelhuth 2011 *inter alia*), it need not be a location (Postal 2004) and it need not be a PP (contra Postal 2004), but it might be appropriately characterized in notional terms, as a reference point or ‘origo’ (cf. Partee and Borschev’s 2007 notion of ‘perspectival center’).
5. As also noticed by Salzman (2013), the ‘setting’ argument of PI need not be a selected argument of the verb (*pace* Bresnan 1994); this is shown by attested tokens like (5a-b):
 - a. And there, in the midst of the swirl, was smiling Sam Waksal—a reedy, charming bachelor biotech entrepreneur
 - b. From the mast flapped the banner of King Aurelius and of Braime.
6. The argument-realization properties associated with PI are found in a variety of structural configurations: Raising, Cleft, Nonsubject *Wh*-interrogative and Auxiliary Initial (including *do*-support contexts, *pace* Bruening 2010); this is shown by attested tokens like (6a-b):
 - a. Out of the woodwork during their show seemed to emerge all of these really big meat-head type guys and they started moshing hard. (raising)
 - b. Did out of this impasse come existentialism, a belief that man could define his temperament and impose meaning on the world? (polar interrogative)

Observations (4-5) suggest that PI is not, as some analysts (including Bresnan 1994) have contended, a kind of topicalization (filler-gap) structure: the fact that PI permits verbs, like *smile* and *flap* in (5a-b), that assign no locative argument damages the case for topicalization, since in such examples there is no gap within the clause to share f-structure attributes with the pre-clausal oblique expression. Observation (6) suggests that the PI pattern is a fact of argument structure rather than phrase structure. Accordingly, we see the partial-inversion pattern as the product of a lexical derivation. No phrasal construction is required to license any of the examples that is not also required to license clauses exhibiting canonical subjects and canonical agreement. Thus, for example, the PI sentence *Down came the rain* is analyzed as a Subject-Predicate construct whose head daughter is the Head-Complement construct *came the rain*.

To account for the unity among the five cases, we propose a type hierarchy in which each pattern of partial inversion inherits syntactic and discourse-pragmatic constraints from a non-maximal type, the **Split Subject (SS) construct type**. As a derivational construction, SS licenses a unary branching structure, that is, a mother with a unique daughter. Constructions of this kind implement the concept of lexical rule (Müller and Wechsler 2014). The single daughter of SS is of the type *intransitive-verb-lexeme*. The mother is of the type *split-subject-intransitive-verb-lexeme*. Unselected ‘setting’ arguments are readily accommodated by this analysis: to rely on a derivational construction as the licenser of PI lexemes is in fact to anticipate a valence mismatch between the two lexemes mediated by the rule. The verbs and verb classes permitted in each partial-inversion pattern are represented in the mother’s FRAMES set, where semantic frames, like syntactic patterns, also participate in the type hierarchy (Davis and Koenig 2000).

In the mother of a SS construct, the agreement trigger is not the XARG; it is an internal argument that is identified with the XARG of the daughter. In creating an external argument that is not the agreement trigger, SS effectively splits canonical subject properties between the ‘new’ and ‘old’ XARGs. Since in each case the preverbal, non-agreeing constituent appears in subject position in a simple declarative clause, this analysis contrasts with many approaches to PI (e.g., Bresnan 1994, Postal 2004, Kim 2003, Bruening 2010), which analyze the preverbal constituent of PI as occupying a filler position in an extraction structure. One virtue of our approach is that it permits us to separate the XARG role from that of agreement trigger. While agreement features are included in referential indices (as per Pollard and Sag 1994: CH 2), subject behaviors like raising and control of a tag subject are a function of XARG status—a status occupied, e.g., by the ‘setting’ argument in PI.

This account also decouples partial inversion from the ‘setting subject’ phenomenon; it applies as well to those cases in which the preverbal, non-agreeing constituent is a non-oblique nominal expression (i.e. predications with *Reversed Equative-be*). In so doing, it captures agreement variability—attestation of both canonical and ‘backwards agreement’ patterns in equative clauses, the latter being licensed by a split-subject derivational construction.

Unlike the lexical rules that modulate between usual valence patterns, we find that lexical rules involved in marked phenomena like partial inversion tend to produce derived lexical items with properties not usually found in listemes, for example, an external argument that does not control agreement. This should not be surprising, since such facts define the marked phenomena that call for a lexical analysis in the first place.

References

- Bresnan, J. (1994). Locative Inversion and the Architecture of Universal Grammar. *Language* 70: 72-131.
- Bruening, B. (2010). Language-Particular Syntactic Rules and Constraints: English Locative Inversion and *do*-support. *Language* 86: 43-84.
- Corbett, G. (2006). *Agreement*. Cambridge: Cambridge University Press.
- Kim, J.B. (2003). English Locative Inversion: A Constraint-based Approach. *Korean Journal of Linguistics* 28: 207-235.
- Davis, A.R. and J.P. Koenig. (2000). Linking as Constraints on Word Classes in a Hierarchical Lexicon. *Language* 76: 56-91.
- Keenan, E. (1976). Towards a Universal Definition of Subject. In C.N. Li, (ed.), *Subject and Topic*. New York: Academic Press. 303-333.
- Müller, S., and S. Wechsler. (2014). Lexical Approaches to Argument Structure. *Theoretical Linguistics* 40: 1-76.
- Partee, Barbara and Vladimir Borschev. (2007). Existential Sentences, BE and the Genitive of Negation in Russian. In K. von Stechow and I. Comorovski, (eds.), *Existence: Semantics and Syntax*. Berlin: Springer Verlag. 147-190.
- Pollard, C. and I.A. Sag. (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Postal, P. M. (2004). A Paradox in English Syntax. *Skeptical Linguistic Essays*. Oxford: Oxford University Press. 15-82.
- Sag, I.A. (2012). Sign Based Construction Grammar: An Informal Synopsis. In Boas, H. C., & I.A. Sag, (eds.), *Sign-Based Construction Grammar*. Stanford: CSLI Publications. 69-202.
- Salzmann, M. (2013). Repair-driven Verb Movement in English Locative Inversion. In P. Brandt and E. Fuss, (eds.), *Repairs: The Added Value of Being Wrong*. Mouton: De Gruyter. 155-206.
- Van Valin, R.D., and R.J. LaPolla. (1997). *Syntax: Structure Meaning, and Function*. Cambridge: Cambridge University Press.
- Webelhuth, G. (2011). Motivating Non-canonicity in Construction Grammar: The Case of Locative Inversion. *Cognitive Linguistics* 22: 81-105.

Backward control in Modern Standard Arabic

Tali Arad Greshler and Nurit Melnik

1 Introduction

Subjunctive complement clauses in Modern Standard Arabic (MSA) are used in contexts where English (and other languages) uses the infinitives. However, unlike English infinitivals, subjunctives in MSA exhibit agreement. Moreover, they alternate between control interpretations, where the matrix subject and the embedded subject share reference, and no-control interpretations, where they have disjoint reference. A corpus-based investigation conducted by Arad Greshler et al. (2016) (henceforth AHMW) revealed that there are no obligatory control (OC) predicates in MSA. Consequently, they propose a uniform *pro*-drop based analysis of control and no-control. The analysis accounts for all patterns, except one – backward control – which exhibits a surprising agreement pattern. AHMW explain the discrepancy by suggesting that this construction may involve extra-grammatical factors.

This study proposes an alternative account of the backward control construction, which builds on new corpus findings regarding the types of predicates which are licensed in this construction. We propose that these predicates can optionally form complex predicates with the embedded subjunctives. When this occurs, the complex predicate exhibits the regular agreement patterns associated with VSO and SVO clauses in MSA.

2 Background

MSA is a *pro*-drop language whose unmarked word order is VSO, yet SVO order is also available. While the two word orders are possible, each is associated with a different agreement pattern. Post-verbal subjects trigger partial agreement, which only involves gender, while the number feature is invariably singular. Pre-verbal subjects trigger full agreement on the verb.

MSA subjunctive clauses are preceded by the particle *?an* and are obligatorily verb-initial. They typically resemble control constructions in English (and other languages), where an unexpressed subject is controlled by a matrix argument. However, the agreement marking on the subjunctive verb reveals the agreement properties of the intended subject. In (1a) the subjunctive *yaktuba* ‘write’ exhibits 3SM agreement. Consequently, the understood embedded subject can be construed as the matrix subject (*control*) or as a different singular–masculine referent (*no-control*). The control and no-control interpretations are also possible in the backward pattern illustrated in (1b), where the subject appears in the embedded clause. In addition, since there is no OC, embedded subjunctives may exhibit agreement properties distinct from the matrix predicate. In (2) the embedded subject is optional but control is impossible due to the agreement mismatch (matrix 3SM and embedded 3SF).

- (1) a. *ħaawala muħammad-un [?an yaktuba maqaal-an.]*
tried.3SM Muhammad-NOM(M) AN write.3SM.SBJ article.ACC
 ‘Muhammad tried to write an article.’
 ‘Muhammad_i tried that he_j would write an article.’
- b. *ħaawala [?an yaktuba muħammad-un maqaal-an.]*
tried.3SM(M) AN write.3SM.SBJ Muhammad-NOM(M) article.ACC
 ‘Muhammad tried to write an article.’
 ‘He tried that Muhammad would write an article.’
- (2) *ħaawala muħammad-un [?an taktuba (hind-un) maqaal-an.]*
tried.3SM Muhammad-NOM(M) AN write.3SF.SBJ (Hind-NOM(F)) article.ACC
 ‘Muhammad tried that Hind/she would write an article.’

3 Subjunctive reference patterns in MSA

AHMW conducted a corpus-based investigation with the goal of finding whether all *?an*-clause selecting predicates allow for both control and no-control between the two subjects, or whether there are OC predicates. The corpus that was used is the 115-million-token sample of the *arTenTen* corpus of Arabic (Arts et al., 2014).

The corpus investigation led AHMW to conclude that there are no OC predicates in MSA. They found evidence for control and no-control with various types of predicates: volitionals, implicatives, manipulatives, modals, and aspectuals. These findings echo Habib (2009), who assumes that there are no “real” control predicates in MSA. They do constitute, however, counterexamples to the generalization made by Landau (2013, p.106), who predicts that “[t]here cannot be a language where modal, aspectual and implicative verbs or evaluative adjectives allow an uncontrolled complement subject”, provided that the embedded predicate exhibit morphological agreement.

Under the assumption that there is no OC in MSA, AHMW argue for one structure for all cases, namely, a no-control structure. Constructions with *?an* complement clauses are simply structures with two independent subjects. The omission of a subject in either clause is due to the *pro*-drop property of MSA; each of the clauses, the matrix clause and the embedded clause, can either have an overt subject or a *pro*-dropped subject. There are no constraints on the agreement relations between the two predicates, and therefore they do not need to match. What resembles subject control is in actuality co-indexation at the semantico-pragmatic level.

One pattern proved problematic for this analysis. The simple example of the backward pattern in (1b) masks a more complex agreement pattern which is only discernable with plural human subjects, for which agreement varies depending on the position of the subject relative to the verb. AHMW find that when the embedded subjects are both human and plural the matrix verb exhibits partial agreement (i.e., only in gender and person) with the subject (3).

- (3) *haawalat* [*?an taktuba_i* *l-banaat-u_i* *maqaal-an*].
tried.3SF AN write.3SF.SBJ the-girls-NOM article-ACC
‘The girls tried to write an article.’

This is unexpected under the *pro*-drop analysis. *Pro* subjects are assumed to trigger full agreement on their predicates. If so then it is not clear how a 3SF *pro* can co-refer with the plural embedded subject.

AHMW conclude that there is no evidence for the existence of OC predicates in MSA. A one-structure *pro*-drop analysis accounts for most of the data, with the exception of the agreement pattern attested in the backward construction (3). They suggest that the use of partial agreement in this pattern is motivated by analogy to the partial subject–verb agreement found in simple VSO clauses, and that the integration of this construction into the theory requires some additional assumptions, which may involve extra-grammatical factors, possibly related to the non-native status of MSA.¹

4 The distribution of backward control

The goal of this study is to propose an alternative account of the backward control construction illustrated in (3). It begins by conducting more focused corpus investigations of the backward pattern. First, we ask whether it is indeed the case that there are no instances of full agreement when the subject is expressed in the embedded clause. Moreover, we extend the range of predicates examined by AHMW to investigate whether all predicates are compatible with backward control.

¹MSA is the literary standard of the Arab world, but it is acquired in school. The mother tongue of its speakers is some regional dialect of Arabic.

A corpus investigation limited to cases with plural animate subjects revealed instances of backward control with volitions, implicatives, modals and aspectuals. No instances of backward control were found with the following predicates : *qarrara* ‘decide’, *xafiya* ‘fear’, *rafad^fa* ‘refuse’, *tarradada* ‘hesitate’, and *ʔiqtaraha* ‘propose’.

With regards to agreement, contrary to AHMW, we found instances of full agreement on the matrix predicate. However, unlike a similar raising construction discussed by Wurmbrand & Haddad (2016), whose matrix predicates alternate between full and partial agreement, the difference in the agreement marking was found to affect the interpretation of the two sentences. When the embedded subject is plural and human and the matrix predicate exhibits partial agreement with it the sentence is ambiguous (4a). The unexpressed matrix subject can be construed as the embedded subject (control) or as a singular-feminine referent (no-control). When the matrix predicate is plural, there is only one no-control interpretation (4b).

- (4) a. *haawalat_{i/j} [ʔan taktuba_i l-banaat-u_i maqaal-an].*
tried.3SF AN write.3SF.SBJ the-girls-NOM article-ACC
 ‘The girls tried to write an article.’ / ‘She_j tried that the girls_i would write an article.’
- b. *haawalna_{*i/j} [ʔan taktuba_i l-banaat-u_i maqaal-an].*
tried.3PF AN write.3SF.SBJ the-girls-NOM article-ACC
 ‘They_j tried that the girls_i would write an article.’ / Not: ‘The girls tried to write an article.’

This asymmetry does not occur in the forward pattern (1a); the two interpretations (control/no-control) are always possible and the embedded predicate exhibits full agreement with its construed subject.

A similar phenomenon is found in Modern Greek (MG), a language which shares a number of syntactic properties with MSA. Subjunctive complement clauses in MG fall into two categories: controlled-subjunctives (C-subjunctives), which enforce control between the matrix and embedded subject, and free-subjunctives (F-subjunctives), which, like in MSA, allow for both control and no-control. Also similarly to MSA, in both types of constructions the subject can be expressed either in the matrix clause or the embedded clause, yet the backward pattern with F-subjunctives is more restricted.

With C-subjunctives a control interpretation is the only option regardless of the position of the subject. With F-subjunctives, on the other hand, the forward pattern in (5a) is ambiguous between control and no-control, but in the backward pattern (5b) the embedded subject cannot be controlled by the matrix subject (Alexiadou et al., 2010, ex. 39). This is similar to the MSA data in (4b).

- (5) a. *o Janis_i elpizi [na fai pro_{i/j} to tiri]*
John-NOM hopes subj eats pro the cheese
 ‘John_i hopes that he_{i/j} will eat the cheese.’
- b. *pro_{*i/j} elpizi [na fai o Janis_i to tiri]*
pro hopes subj eats John-NOM the cheese
 ‘He hopes that John will eat the cheese.’

Alexiadou et al. (2010) propose a *pro*-drop analysis for F-subjunctives, similar in spirit to the one proposed by AHMW. Consequently, they attribute the impossibility of coreference in (5b) to Principle C. The embedded referential subject, *Janis*, cannot be bound by the matrix *pro* subject. The fact that there is no Principle C effect in the case of C-subjunctives is taken by Alexiadou et al. (2010) as evidence that control with these predicates does not involve a *pro*-dropped subject.

The similarity between MSA and MG is even greater when the types of predicates which are licensed by the different constructions are considered. Landau (2004) argues that the predicates which select C-subjunctives in MG belong to a category of predicates which cross-linguistically select *semantically untensed* complements, and include the implicatives, aspectuals, modals, and evaluative adjectives. Predicates which select F-subjunctives, on the other hand, are those which select *semantically tensed*

In a nutshell, we propose that *?an* and the subjunctive combine in a head-marker structure. The matrix predicate selects as its complement a marked subjunctive with matching agreement properties, and inherits all its arguments. Moreover, the referential index of the shared subject is structure-shared with the values of the semantic roles of the relations denoted by each of the predicates, as is expected in a control construction (not shown in the sketch due to lack of space).

In the full version of the paper we show that AHMW's analysis together with the current proposal account for all the patterns attested with *?an* complement clauses. This analysis is given support by cross-linguistic findings regarding parallel predicate types and their syntactic behavior. Consequently, what AHMW consider a surprising pattern can be explained with known syntactic mechanisms.

References

- Abeillé, Anne, Danièle Godard & Ivan A. Sag. 1998. Two kinds of composition in French complex predicates. In E. Hinrichs, A. Kathol & T. Nakazawa (eds.), *Complex predicates in nonderivational syntax*, vol. 30 Syntax and Semantics, San Diego: Academic Press.
- Alexiadou, Artemis, Elena Anagnostopoulou, Gianina Iordachioaia & Mihaela Marchis. 2010. No objections to backward control. In M. Polinsky & N. Hornstein (eds.), *Movement theory of control*, 89–117. Amsterdam: John Benjamins.
- Arad Greshler, Tali, Livnat Herzig Sheinflux, Nurit Melnik & Shuly Wintner. 2016. Reference patterns in subjunctive complement clauses of modern standard arabic. In D. Arnold, M. Butt, B. Crysmann, T. H. King & S. Müller (eds.), *Proceedings of the joint 2016 conference on head-driven phrase structure grammar and lexical functional grammar, polish academy of sciences, warsaw, poland*, 4–22. Stanford, CA: CSLI Publications.
- Arts, Tressy, Yonatan Belinkov, Nizar Habash, Adam Kilgarriff & Vit Suchomel. 2014. arTenTen: Arabic corpus and word sketches. *Journal of King Saud University - Computer and Information Sciences* 26(4). 357 – 371. doi:<http://dx.doi.org/10.1016/j.jksuci.2014.06.009>.
- Grano, Thomas. 2015. *Control and restructuring*. Oxford: Oxford University Press. (Oxford Studies in Theoretical Linguistics series).
- Habib, Rania. 2009. The syntax of the Standard Arabic *?an* and *?anna*. In K. K. Grohmann & P. Panagiotidis (eds.), *Selected papers from the 2006 Cyprus syntaxfest*, 159–194.
- Hinrichs, Erhard & Tsuneko Nakazawa. 1990. Subcategorization and vp structure in german. In S. Hughes & J. Salmons (eds.), *Proceedings of the third symposium on germanic linguistics*, Amsterdam: Benjamins.
- Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22(4). 811–877.
- Landau, Idan. 2013. *Control in generative grammar: A research companion*. Cambridge University Press.
- Monachesi, Paola. 1998. Italian restructuring verbs: A lexical analysis. In E. Hinrichs, A. Kathol & T. Nakazawa (eds.), *Complex predicates in nonderivational syntax*, vol. 30 Syntax and Semantics, 313–368. San Diego: Academic Press.
- Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford: CSLI Publications.
- Roussou, Anna. 2009. In the mood for control. *Lingua* 119(12). 1811–1836.
- Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.
- Wurmbrand, Susi & Youssef A Haddad. 2016. Cyclic spell-out derived agreement in Arabic raising constructions. In Y. A. Haddad & E. Potsdam (eds.), *Perspectives on Arabic linguistics XXVIII: Papers from the annual symposium on Arabic linguistics, Gainesville, Florida, 2014*, 193–228. Amsterdam/Philadelphia: John Benjamins Publishing Company.

SIMPLE AND COMPLEX COMPARATIVES IN MODERN STANDARD ARABIC

Abeer Alsulami, King Saud University, Riyadh, University of Essex
 Doug Arnold, University of Essex
 Robert D. Borsley, University Essex

1. Introduction

Like many languages, Modern Standard Arabic (MSA) has simple comparatives with a comparative form of an adjective and complex comparatives with two separate elements:

- (1) a. heya ʔaTwal-u min xalid-in
 she taller.MASC.SG.NOM from Khalid-GEN
 ‘She is taller than Khalid.’
 b. ʔanaa ʔakthar-u thakaʔ-an min ali-in
 I.1SG.MASC more-NOM intelligence.ACC from Ali-GEN
 ‘I am more intelligent than Ali.’

(These are verbless sentences. Past tense counterparts would contain a past tense copula.) Superficially, these examples are much like their English translations and like simple and complex comparatives in other languages., but as the translation of (1b) makes clear, *thakaʔ-an* is not an adjective but what we will call an adjectival noun. (In traditional Arabic grammar it is known as *masdar*.) An adjective is not possible:

- (2) *ʔanaa ʔakthar-u thakay-an min ali-in
 I.1SG.MASC more-NOM intelligent.ACC from Ali-GEN
 ‘I am more intelligent than Ali.’

MSA simple comparatives are much like their counterparts in other languages. However, there is evidence that MSA complex comparatives are a quite different construction from the complex comparatives of many other languages and that they require a rather different analysis.

2. Basic data

Simple adjectival comparatives involve what is known as the relative form of an adjective and a PP headed by *min* ‘from’ (a *min*-phrase) expressing the standard of comparison. As one might expect, *min* may have either an NP, as in (1a), or an elliptical clause, as in (3), as its complement:

- (3) kamal-un ʔakbar-u mi-maa kaan ʔab-uu-hu
 kamal-NOM older-NOM from-what was father-NOM-his
 ‘Kamal is older than his father was.’

Simple comparatives can be used predicatively, as in (1a) and (3), or attributively, as in the following:

- (4) kamal-un rajul-un ʔakbar-u min ali-in
 kamal-NOM man-NOM older-NOM from Ali-GEN
 ‘Kamal is an older man than Ali.’

Some MSA adjectives do not have a comparative form, and for them a comparative meaning can only be expressed by a complex comparative involving one of a small number of general comparative words and an accusative adjectival noun. As one would expect, these constructions also have attributive uses:

- (5) qaabal-tu rajul-an ʔakthar-a thakaʔ-*an* min xalid-*in*
 met-1SGM man-ACC more-ACC intelligence-ACC from Khalid-GEN
 ‘I met a man more intelligent than Khalid.’

As one might also expect, there are similar examples with *ʔaqall* ‘less’ and an adjectival noun:

- (6) ʔanaa ʔaqall-u thakaʔ-*an* min ali-*in*
 I.1SG.MASC less-NOM intelligence-ACC from Ali-GEN
 ‘I am less intelligent than Ali.’
 (7) qaabal-tu rajul-an ʔaqall-a thakaʔ-*an* min xalid-*in*
 met-1SGM man-ACC less-ACC intelligence-ACC from Khalid-GEN
 ‘I met a man less intelligent than Khalid.’

The adjectival noun in a complex comparative is always accusative, but the case of the comparative word reflects the position of the construction. Thus, it appears that the complex comparative is a head-complement structure, in which the comparative word is a head and the adjectival noun and the *min*-phrase its complements. It seems surprising that it contains an adjectival noun and not an adjective. However, there is evidence that this is not at all surprising. It is also possible to have an ordinary noun instead of the adjectival noun, as in the following:

- (8) ʔanaa ʔakthar-u maal-*an* min-ka
 I more-NOM money-ACC from-you
 ‘I have more money than you.’
 (9) qaabal-tu rajul-an ʔakthar-a kutub-*an* min ali-*in*
 met.1SGM man-ACC more.ACC books-ACC from Ali-GEN
 ‘I met a man with more books than Ali.’

It is clear that these examples involve the same construction as the examples with an adjectival noun. Thus, we have a construction in which a comparative adjective takes an accusative nominal complement, which may be either an ordinary noun or an adjectival noun. We have translated the examples with an ordinary noun with ‘have’ when used predicatively and with ‘with’ when used attributively. Examples with an adjectival noun could be translated in the same way. That is, we could have ‘He has more intelligence’ and ‘a man with more intelligence’ rather than ‘he is more intelligent’ and ‘a more intelligent man’.

Essentially any comparative can combine with an ordinary noun in a complex comparative. Here are a few relevant examples:

- (10) a. ʔanaa ʔakbar-u sinn-*an* min ali-*in*
 I.1SG.MASC older-NOM age-ACC from Ali-GEN
 ‘I am older in age than Ali.’
 b. ʔanaa ʔafSaH-u lisaan-*an* min ali-*in*
 I.1SG.MASC more fluent-NOM tongue-ACC from Ali-GEN
 ‘I have a more fluent tongue than Ali.’

Only a small number of comparatives can combine with an adjectival noun, but we assume this is just a matter of semantics, of what makes sense.

MSA has another construction in which an adjective takes a nominal complement, the adjectival construct construction (Ryding 200: 253-4, Al-Sharifi and Sadler 2009), illustrated by the following:

- (11) 'anta azīm-u -l-hazz-i
 you great-NOM the-fortune-GEN
 'You have great luck'/'You are very lucky'
- (12) imra-at-un ḡamīl-at-u -l-waḡh-i
 woman-F-NOM beautiful-F-NOM the-face-GEN
 'a woman with a beautiful face'

These have a non-comparative adjective and the nominal complement is genitive and definite, but they seem to have the same basic structure and essentially the same type of meaning, 'have' when used predicatively and 'with' when used attributively.

3. HPSG analyses

An analysis of adjectives needs to accommodate both predicative and attributive uses. Assuming that predicative adjectives have a non-empty SUBJ value reflecting the first member of the ARG-ST list and are [MOD *none*] and that attributive adjectives have a value for the MOD feature coindexed with the first member of the ARG-ST list and are [SUBJ <>], the following constraint can be proposed:

(13)

$$\left[\begin{array}{l} \textit{word} \\ \text{HEAD } \textit{adj} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{HEAD [MOD } \textit{none}] \\ \text{SUBJ } \langle [1] \rangle \\ \text{ARG - ST } \langle [1] \rangle \oplus \text{L} \end{array} \right] \vee \left[\begin{array}{l} \text{HEAD [MOD N'}_{[i]} \\ \text{SUBJ } \langle \rangle \\ \text{ARG - ST } \langle [i] \rangle \oplus \text{L} \end{array} \right]$$

This will apply both to basic adjectives and to adjectives which are the product of a lexical rule.

We propose that simple comparative adjectives are the product of the following lexical rule:

(14) Simple comparative lexical rule

$$\left[\begin{array}{l} \text{HEAD } \left[\begin{array}{l} \textit{adj} \\ \text{AFORM } \textit{pos} \end{array} \right] \\ \text{ARG - ST } [1] \end{array} \right] \Rightarrow \left[\begin{array}{l} \text{HEAD } \left[\begin{array}{l} \textit{adj} \\ \text{AFORM } \textit{comp} \end{array} \right] \\ \text{ARG - ST } [1] \oplus \langle \text{PP}[\textit{min}] \rangle \end{array} \right]$$

This changes the value of AFORM from *pos(itive)* to *comp(arative)* and adds PP[*min*] to the end of the ARG-ST list.

We assume that complex comparatives involve an adjective with a nominal complement, which may be an adjectival noun or an ordinary noun and must be accusative and indefinite. The complement has essentially the same role as the first argument of a basic comparative. We assume, therefore, that adjectives in a complex comparative have an extra argument as the first member of their ARG-ST list, which is the subject if it is predicative or is coindexed with the

modified NP if it is attributive. Given these assumptions, we will have lexical descriptions of the following form:

(15)

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{adj} \\ \textit{AFORM comp} \end{array} \right] \\ \text{ARG - ST} \langle \text{NP}, \text{NP}[\text{DEF} -, \text{CASE} \textit{acc}], \dots \text{PP}[\textit{min}] \rangle \end{array} \right]$$

Descriptions like this can be derived from descriptions for simple comparative adjectives by the following lexical rule:

(16) Complex comparative lexical rule

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{adj} \\ \textit{AFORM comp} \end{array} \right] \\ \text{ARG - ST} \langle \text{[1]} \rangle \end{array} \right] \Rightarrow \left[\text{ARG - ST} \langle \text{NP} \rangle \oplus \langle \text{NP}[\text{DEF} -, \text{CASE} \textit{acc}]_{[1]} \rangle \oplus [1] \right]$$

This adds an extra argument to the beginning of the ARG-ST list and marks the original initial argument, which is now the second argument, as [DEF -] and [CASE *acc*]. [1] will often contain just PP[*min*], but where the basic non-comparative adjective has a complement, there will be another member.

Adjectival constructs also involve an adjective with a nominal complement, which must be genitive and definite. As in complex comparatives, the complement has essentially the same role as the first argument of a basic adjective, and the adjective has an extra argument as the first member of its ARG-ST list. This means lexical descriptions of the following form:

(17)

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{adj} \\ \textit{AFORM pos} \end{array} \right] \\ \text{ARG - ST} \langle \text{NP}, [\text{DEF} +, \text{CASE} \textit{gen}], \dots \rangle \end{array} \right]$$

Descriptions like this can be derived from descriptions like for simple adjectives by the following lexical rule:

(18) Construct adjective lexical rule

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{adj} \\ \textit{AFORM pos} \end{array} \right] \\ \text{ARG - ST} \langle \text{[1]} \rangle \end{array} \right] \Rightarrow \left[\text{ARG - ST} \langle \text{NP} \rangle \oplus \langle \text{NP}[\text{DEF} +, \text{CASE} \textit{gen}]_{[1]} \rangle \oplus [1] \right]$$

This adds an extra argument to the beginning of the ARG-ST list and marks the original initial member as [DEF +] and [CASE *gen*]. [1] will often but not always be the empty list.

The construct adjective lexical rule and the complex comparative lexical rule are obviously quite similar. Both add an extra argument to the beginning of an ARG-ST list and turn the original initial member into the second member so that it is realized as a complement. They differ in whether they apply to [AFORM *pos*] or [AFORM *comp*] adjectives and in whether

they require the original initial member the ARG-ST list be indefinite and accusative or definite and genitive. It is natural to ask whether the two lexical rules could be combined. We can do this as follows:

(19)

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{adj} \\ \text{AFORM [1]} \end{array} \right] \\ \text{ARG - ST} \langle \text{[i]} \rangle \oplus [2] \end{array} \right] \Rightarrow [\text{ARG - ST} \langle \text{NP} \rangle \oplus \langle \text{NP}[\text{DEF} [3] \text{CASE} [4]][\text{i}] \rangle \oplus [2]]$$

$$([1] = \textit{comp} \ \& \ [3] = - \ \& \ [4] = \textit{acc}) \vee ([1] = \textit{pos} \ \& \ [3] = + \ \& \ [4] = \textit{gen})$$

Here we have a rule with an attached disjunctive statement of the possible values of the features AFORM, DEF and CASE. The first disjunct specifies the values for complex comparatives and the second gives the value for construct adjectives. This is quite complex, but it does capture the similarity between the two sets of words.

4. A further issue

Where the noun in a complex comparative takes a PP complement, this may follow the *min*-phrase. Here is an example with an ordinary noun:

- (20) ?anaa ?akthar-u Kutub-an min ali-in fi n-naHw-i
 1SG.MAS more-NOM books-ACC from Ali-GEN at the-syntax-GEN
 ‘I have more books than Ali about syntax.’

There are similar examples with adjectival nouns. Such examples might suggest that an argument composition analysis would be appropriate for complex comparatives. However, there is evidence from examples like the following that this is a more general phenomenon not specifically connected with complex comparatives:

- (21) ?a-ʕTaa kamal-un kitaab-an ?ila ali-in fi n-naHw-i
 3SG.MAS-gave Kamal-NOM book-ACC to Ali-GEN at the-syntax-GEN
 ‘Kamal gave a book to Ali about about syntax.’

Both types of example can be analysed with the EXTRA mechanism applied by Kay and Sag (2012) to a variety of English types of extraposition.

REFERENCES

Al-Sharifi, B. and L. Sadler (2009), The adjectival construct in Arabic, in M. Butt and T. H. King, (eds.) *Proceedings of the LFG09 Conference*. CSLI Publications, Stanford, CA, pp. 26-43.

Ryding, K. (2005), *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Kay, P. and I. A. Sag (2012), Cleaning up the big mess: Discontinuous dependencies and complex determiners, in Boas and Sag (eds.), *Sign-Based Construction Grammar*, Stanford: CSLI Publications, 229-256.

Resumption and Case: a new take on Modern Standard Arabic

BERTHOLD CRYSMANN
Laboratoire de linguistique formelle, CNRS

Over the past few years, there has been renewed interest in the treatment of resumption in HPSG: despite areas of convergence, e.g. the recognition of resumptive dependencies as SLASH dependencies, as motivated by Across-the-Board (ATB) extraction, there is no unified theory to date, with differences pertaining, e.g., to the exact formulation of SLASH amalgamation (Ginzburg and Sag, 2000), or the place of island constraints in grammar. While Borsley (2010) and Alotaibi and Borsley (2013) relegate the difference in locality of gap and resumptive dependencies to the performance system, Crysmann (2016) captures insensitivity to strong islands as part of the grammar. Harmonising existing proposals becomes even more acute, if we consider the cross-linguistic similarity of the phenomenon, in particular, if we compare languages like Hausa and Arabic, which both feature island insensitivity to some degree, as well as bound pronominal resumptive objects and zero pronominal resumptive subjects, to name just a few of the parallels.

In this paper, I shall reexamine resumption (and extraction) in Modern Standard Arabic (henceforth: MSA) and propose a reanalysis that improves on Alotaibi and Borsley (2013) in several areas: first, it resolves problems with case assignment in ATB extraction where there is a gap in one conjunct, and a resumptive in the other. Second, I shall show how the weight-based theory of islands developed in Crysmann (2012) can be parameterised to account for the MSA data. Furthermore, by adopting the underspecification approach advanced in Crysmann (2016), I shall reinstate a deterministic version of SLASH amalgamation, as well as provide identical semantics for gaps and resumptives. Finally, the underspecification approach takes on board Borsley’s insight (Borsley, 2010; Alotaibi and Borsley, 2013) that decision on resumptive vs. non-resumptive function of pronominals should be associated with the governing head, thus addressing McCloskey’s generalisation (McCloskey, 2002).

1 Gaps and resumptives in MSA

Unbounded dependency constructions in MSA provide evidence for both gap and resumptive strategies in the grammar of extraction: as shown in (1) arguments of prepositions, as well as possessor arguments of nouns may only extract with a resumptive element in situ (a bound pronominal affix).¹

¹For reasons of space, I shall gloss over the possibility of pied piping which is immaterial for the points made here.

- (1) a. ?ayy -u/*-i đ̣aamiʕat-in ðahaba
which -NOM/-GEN university-GEN went.3SM
Ahmad-u ?ilai -ha / *∅ ?
Ahmad-NOM to -it
‘Which university did Ahmad go to?’ (A&B 2013, p. 7)
- b. ?ayy -u/*-i mu?allif-in gara?a
which -NOM/-GEN author-GEN read.3SM
Ahmad-u kitaab-a -hu / *∅ ?
Ahmad-NOM book-ACC -his
‘Which author’s book has Ahmad read?’ (A&B 2013, p. 7)

By contrast, non-nominal complements, e.g. PP-complements of verbs or adjectives may only extract by means of a gap strategy (see Alotaibi and Borsley, 2013, p. 11, for data and discussion).

Direct objects, however, witness overlap between the two strategies: while it is possible to extract by means of a gap strategy in certain constructions, cf. (2), others feature the presence of a bound pronominal affix on the governing verb (3).

- (2) ?ayy-a T-tullaab-i qaabala l-qaa?id-u ∅ ?
which-ACC the-students-GEN met.3SM the-leader-NOM
‘Which of the students has the leader met?’ (A&B 2013, p. 8)
- (3) ?ayy-u T-tullaab-i qaabala-**hum**
which-NOM the-students-GEN met.3SM-them
l-qaa?id-u ?
the-leader-NOM
‘Which of the students has the leader met?’ (A&B 2013, p. 8)

Choice between the two strategies depends on several factors: first, while both strategies are available with wh-extraction and relatives with a definite antecedent, only resumption is an option with indefinite antecedents (4). Furthermore, extraction out of strong islands, e.g. relative clauses make use of a resumptive obligatory (5).

- (4) qaabaltu rajul-an [ʔaʕrifu -hu / *∅] ?
met.1SM man-ACC knew.1SM -him
‘I met a man that I knew’ (A&B 2013, p. 9)
- (5) ?ayy -u/*-a bint-in ra?aita l-ʔasad-a llaðii
which -NOM/-ACC girl-GEN saw.2SM the-lion-ACC that
ʔakala -ha / ∅
ate.3SM -her
‘Which girl did you see the lion that ate?’ (A&B 2013, p. 12)

Case marking of fillers correlates with the choice of extraction strategy: while gaps feature a matching effect, giving accusative case on the filler, the fronted constituent bears nominative case in the event of a resumptive. Note that nominative case marking of fronted possessors as in (1b) or complements of prepositions (1a), which appear in the genitive when in situ, is congruent with this observation.

As for subject extraction, Alotaibi and Borsley (2013) observe that the subject-agreement pattern (full agreement in person, number, and gender) in relativisation and wh-fronting (6) parallels that of topicalised (7) and pro-dropped (8) subjects, in contradistinction to post-verbal subjects (9) in situ (partial agreement in person and gender).

- (6) ʔayy-u Tullaab-in ʕaraf-uu / *ʕarafa
 which-NOM students-GEN knew.3PM knew.3SM
 l-ʔijaabat-a?
 the-answer-ACC
 ‘Which students knew the answer?’ (A&B 2013, p. 10)
- (7) T-tullaab-u qaabaluu / *qaabala Ahmad-a
 the-students-NOM met.3PM met.3SM Ahmad-ACC
 ‘The students met Ahmad’ (A&B 2013, p. 9)
- (8) a. laqad qaabala Ahmad-a
 indeed met.3SM Ahmad-ACC
 ‘He met Ahmad.’ (A&B 2013, p. 10)
 b. laqad qaabaluu Ahmad-a
 indeed met.3PM Ahmad-ACC
 ‘They met Ahmad.’ (A&B 2013, p. 10)
- (9) qaabala / *qaabaluu T-tullaab-u Ahmad-a
 met.3SM met.3PM the-students-NOM Ahmad-ACC
 ‘The students met Ahmad’ (A&B 2013, p. 9)

Alotaibi and Borsley (2013) therefore correlate fronting with the null subject property and conclude that subject extraction involves a zero resumptive, rather than a gap.

2 Alotaibi and Borsley (2013)

In order to capture both the distribution of gaps vs. resumptives, Alotaibi and Borsley (2013) suggest that gap dependencies involve full reentrancy between an argument’s LOC value with a member of SLASH, whereas resumptives, which are treated as ordinary pronominals in the spirit of McCloskey (2002) and Borsley (2010), give rise to an optional NP member on SLASH where reentrancy with the pronominal argument is limited to INDEX.

Now given that the slashed NP specification exhibits only very limited reentrancy with properties of the resumptive element, exempting most of CONT (leaving alone INDEX) and all of CAT, the exceptional assignment of nominal case to the filler of resumptive unbounded dependencies is finally accounted for by means of restricting this specific member of SLASH to nominative case. To this end, they propose an implicational constraint on words where a pronominal argument has its INDEX shared with an element in SLASH, see (10).

$$(10) \left[\begin{array}{l} \text{word} \\ \text{SLASH } \left\{ \boxed{1} [\text{INDEX } \boxed{2}] \right\} \\ \text{ARG-ST } \left\langle \dots \left[\begin{array}{l} \text{pro} \\ \text{INDEX } \boxed{2} \end{array} \right] \dots \right\rangle \\ \rightarrow \\ \text{SLASH } \left\{ \boxed{1} [\text{CASE } \textit{nom}] \right\} \end{array} \right]$$

Assignment is thus uniformly fixed at the bottom of the dependency, including resumptive and gap dependencies. Given that case properties are imposed on SLASH elements, either by reentrancy (gap) or stipulation (resumptives), they inevitably percolate up, ensuring nominative fillers for resumptives and matching fillers for gaps.

Alotaibi and Borsley (2013) further propose that case can be used to control the distribution of gaps and resumptives in a more fine-grained way. While definite relatives marked by complementiser *llaði* license both gaps and resumptives for direct objects, indefinite relatives, which are headed by a zero complementiser according to Alqurashi and Borsley (2012), only permit a resumptive. Alotaibi and Borsley (2013) suggest that this difference can be captured by the following lexical entries for *llaði* (11) and the zero relative complementiser 12:

$$(11) \left[\begin{array}{l} \text{PH } \langle \textit{llaði} \rangle \\ \text{HD } \left[\begin{array}{l} \textit{comp} \\ \text{MOD NP } \left[\begin{array}{l} \text{DEF } + \\ \text{IND } \boxed{i} \end{array} \right] \end{array} \right] \\ \text{COMPS } \left\langle \text{S } \left[\text{SLASH } \left\{ \text{NP } \left[\text{IND } \boxed{i} \right] \right\} \right] \right\rangle \end{array} \right]$$

$$(12) \left[\begin{array}{l} \text{PH } \langle \rangle \\ \text{HD } \left[\begin{array}{l} \textit{comp} \\ \text{MOD NP } \left[\begin{array}{l} \text{DEF } - \\ \text{IND } \boxed{i} \end{array} \right] \end{array} \right] \\ \text{COMPS } \left\langle \text{S } \left[\text{SLASH } \left\{ \text{NP } \left[\begin{array}{l} \text{CASE } \textit{nom} \\ \text{IND } \boxed{i} \end{array} \right] \right\} \right] \right\rangle \end{array} \right]$$

3 ATB and Case

The idea to exploit case properties in order to regulate the distribution of resumptives and gaps runs into quite some serious problems once we consider ATB extraction.

In MSA, like in many other languages that offer both gap and resumptive strategies, mixing of gap and resumptives is possible, as shown, e.g. in (13): while the ATB constraint can be shown to be operative in the language, it apparently treats gap and resumptive dependencies alike.

- (13) a. * man [tuhibu \emptyset wa tušaḩiṣu Aḩmad-a fii
 who like.2SM and support.2SM Aḩmad-ACC in
 nafs-i l-waqt-iṣ]
 same-GEN the-time-GEN
 ‘Who do you like and support Aḩmad at the same
 time?’ (A&B 2013, p. 13)
- b. man [tuhibu \emptyset wa tušaḩiṣu \emptyset fii nafs-i
 who like.2SM and support.2SM in same-GEN
 l-waqt-iṣ]
 the-time-GEN
 ‘Who do you like and support at the same time?’
 (A&B 2013, p. 13)
- c. man [tuhibu \emptyset wa tušaḩiṣu -hu fii
 who like.2SM and support.2SM -him in
 nafs-i l-waqt-iṣ]
 same-GEN the-time-GEN
 ‘Who do you like and support at the same time?’
 (A&B 2013, p. 14)’

It is precisely for this reason that almost all approaches to resumption in HPSG treat both dependencies via SLASH.

As discussed by Alotaibi and Borsley (2013), mixing of resumptives and gaps leads to a conflict of case specifications on SLASH: if nominative case is assigned at the bottom of a resumptive dependency, yet standard accusative is assigned to object gaps, unification of SLASH values must fail. However, mixing is not only possible with case-ambiguous fillers, as in (13), but also with unambiguously case-marked fillers. Speakers find resolution to the gap’s accusative case requirement perfectly acceptable, whereas judgements degrade for nominative: “[t]hey find examples like [(14b)] with nominative case less acceptable, but do not generally reject them” (Alotaibi and Borsley, 2013, p. 21).

- (14) a. ʔayy -a Tullaab-in [qaabalta \emptyset
 which -ACC students-GEN met.2SM and
 wa taḩadda θ ta ʔilai-**hum**?
 talked.2SM to-them
 ‘Which students have you met and talked to?’ (A&B
 2013, p. 21)
- b. ʔayy -u Tullaab-in [qaabalta \emptyset
 which -NOM students-GEN met.2SM and
 wa taḩadda θ ta ʔilai-**hum**?
 talked.2SM to-them
 ‘Which students have you met and talked to?’ (A&B
 2013, p. 21)

As admitted by the authors, both the perfectly well-formed accusative variant and the marginal nominative one are erroneously ruled out as ungrammatical by their account. This analysis of MSA resumption therefore contradicts the standard account of the ATB effect (Pollard and Sag, 1994), which derives the constraint quite elegantly by simple unification of the SLASH sets of the conjunct daughters.

I shall argue, however, that case conflict in mixed gap/resumptive ATB constructions is not an inherent problem of MSA, but rather constitutes an artefact of the way Alotaibi and Borsley (2013) implement case assignment with resumptive dependencies, namely assignment to an element in SLASH at the bottom of the dependency and therefore

transmission of this case requirement along the extraction path.

It is of note that MSA provides no evidence at all that case transmission is required in resumptive dependencies, owing to the absence of a matching effect: with wh extraction, the stipulated nominative case assignment at the bottom hardly ever corresponds to what case would normally be assigned here, which is either accusative (for direct objects) or genitive (prepositions and possessed nouns). Similarly, in *ʔanna* clauses, which obligatorily involve resumption, the accusative requirement for the fronted NP is locally assigned by the complementiser *ʔanna*, both for subject (zero pronoun) and object (pronominal affix) resumptives. Finally, the definite relative complementiser *llaḩi* indeed does inflect for case (in addition to number and gender), but case matching uniformly targets the antecedent, not the relativised argument (Alqurashi and Borsley, 2012, p. 29). Owing to the absence of a matching effect with resumptive dependencies, we can conclude that nominative case assignment at the bottom of the dependency is empirically undermotivated. It so happens that this is indeed the assumption that is at the root of the problem with ATB constructions.

4 A reanalysis

In order to overcome the problems with case assignment in ATB extraction with mixed gap/resumptive dependencies, I shall propose that nominative case assignment to the filler in constructions with resumptives is fixed directly at the top of the dependency. As for exerting more precise control over the distribution of gaps and resumptives, I shall rely instead on the weight-based theory of resumption and extraction developed in Crysmann (2012, 2016).

4.1 A weight-based theory of extraction and resumption (Crysmann, 2016)

The weight-based theory of resumption and extraction implements a distinction of SLASH elements in terms of the amount of information that is minimally or maximally transmitted. As illustrated by the type hierarchy in (15), *local* values are differentiated according to the amount of information they carry: While *weak-local* contains no CAT, and only INDEX features in CONT, *full-local* has both CAT and CONT features, including semantic relations on RELS. The value of the LOC attribute of *synsem* objects therefore is of the later type, cf. (16). As a consequence, *weak-local* values essentially live on non-local features, such as SLASH sets.² Reentrancy of an element with a LOC feature, as in the filler-head schema or for the type *gap-synsem* automatically coerces the element into the full type. The crucial point of this theory now is that resumptives by themselves are generally underspecified as to the local type on their SLASH set: all they require is minimal sharing of INDEX.

²See Crysmann (2013) for a similar proposal regarding locality constraints on complement clause vs. relative clause extraposition.

with selection for a *weak-local* UDC, effectively ruling out a gap at the bottom. As for selection of resumptive vs. gap strategies, I have replaced problematic case percolation with weight-based selection, a move which also opens up for the possibility to address island-sensitivity within the grammar, rather than delegating it to performance. Now that there is no case assignment at the bottom of resumptive dependencies, no conflict can ever arise between case requirements for gaps and resumptives: in essence, it is always the case of the gap that wins out. This not only reflects the absence of any observable matching effect with resumption, but it also corresponds to the perceived degradation associated with nominative as compared to accusative case.³ Coordination of slashed constituents will always enforce the stronger constraint on the mother and both conjunct daughters, i.e. in coordination involving gaps and resumptives, the SLASH values of the mother and both daughters will be coerced to *full-local*, a specification that is compatible at the bottom of the dependency with both gaps (trivial) and resumptives (by underspecification). At the top, however, only the standard filler-head schema, but not the parochial “resumptive” filler-head structure, is compatible with this specification, thereby deriving the preference for case matching with the gap. Moreover, the present proposal does away with the rather stipulative nature of using case to regulate the distribution of resumptives vs. gaps, and replaces it with a more principled concept relating to the necessary amount of percolated information.

Applying the theory of resumption proposed in Crysmann (2016) has several advantages: first, it provides a unified analysis of resumption for Hausa and Arabic, a property that is highly desirable given the amount of parallelism in the constraints that both languages observe. The crucial differences between Hausa and MSA is actually encapsulated in a single place: MSA has a parochial filler-head rule that is compatible with resumptive dependencies, whereas Hausa does not. This is sufficient to derive the fact that in Hausa, only relativisation can escape strong islands (since the standard filler-head rule enforces *full-local*), whereas in MSA, this option is open to wh-fillers as well. The successful application of Crysmann (2016) to MSA further solves two somewhat more technical issues: SLASH amalgamation can be stated in a deterministic fashion and gaps and resumptives will receive identical semantics.

5 Conclusion

In this paper, I proposed an analysis of resumption and ATB extraction in Modern Standard Arabic that builds on previous work on resumption in Hausa (Crysmann, 2016). In addition to providing a more unified theory of the phenomenon in the two languages, the weight-based model of locality permits fine-grained control over the distribution of gaps and resumptives in a more principled way than what is offered

³I specifically rule out the marginal example in (14b) as ungrammatical. Speakers clearly perceive the degradation, as stated by Alotaibi and Borsley (2013), and it is an often-made observation that second-language speakers (MSA is used as a vehicular language by speakers of different Arabic dialects) tend to have less sharp judgements, being more inclined to accept ungrammatical sentences.

by the case-based approach of Alotaibi and Borsley (2013). Not only does it generalise better from case languages like Modern Standard Arabic to case-less languages like Hausa or vernacular Arabic varieties, but it also avoids case conflict in ATB constructions with mixed gap/resumptive strategies. Postulating a parochial “resumptive” filler-head construction for Modern Standard Arabic not only solves the case issue, but it also derives why wh-fillers can escape strong islands, in contrast to Hausa, which only features standard filler-head structures with full local reentrancy.

References

- Alotaibi, M. and Borsley, R. D. (2013). ‘Gaps and resumptive pronouns in Modern Standard Arabic’. In S. Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*. 6–26.
- Alqurashi, A. and Borsley, R. D. (2012). ‘Arabic relative clauses in HPSG’. In S. Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon*. Stanford, CA: CSLI Publications, 26–44.
- Borsley, R. D. (2010). ‘An HPSG approach to Welsh unbounded dependencies’. In S. Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot, Paris 7, France*. Stanford, CA: CSLI Publications, 80–100.
- Crysmann, B. (2012). ‘Resumption and island-hood in Hausa’. In P. de Groote and M.-J. Nederhof (eds.), *Formal Grammar. 15th and 16th International Conference on Formal Grammar, FG 2010 Copenhagen, Denmark, August 2010, FG 2011 Lubljana, Slovenia, August 2011*, vol. 7395 of *Lecture Notes in Computer Science*. Springer, 50–65.
- (2013). ‘On the locality of complement clause and relative clause extraposition’. In G. Webelhuth, M. Sailer, and H. Walker (eds.), *Rightward Movement in a Comparative Perspective*. Amsterdam: John Benjamins, 369–396.
- (2016). ‘An underspecification approach to Hausa resumption’. In D. Arnold, M. Butt, B. Crysmann, T. H. King, and S. Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar, Polish Academy of Sciences, Warsaw, Poland*. Stanford, CA: CSLI Publications, 194–214.
- Ginzburg, J. and Sag, I. A. (2000). *Interrogative Investigations. The Form, Meaning, and Use of English Interrogatives*. Stanford: CSLI Publications.
- McCloskey, J. (2002). ‘Resumptives, successive cyclicity, and the locality of operations’. In S. D. Epstein and T. D. Seely (eds.), *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell, 184–226.
- Pollard, C. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Subject Agreement in Laki
 Sahar Taghipour
 University of Kentucky

This study investigates the inflection of subject agreement in Laki verbal paradigms. This language based on Windfuhr (2009) and Anonby (2004) belongs to Kurdish Language, as one of the Northwestern Iranian languages. In this language, verbs inflect for subject agreement by different sets of markers. Tafakkori and Omidi (2014: 43-44) and Moradi (2015: 7-8) classify Laki verbal paradigms into three different groups based on the subject agreement patterns they exhibit. Moradi (2015) discusses that this classification is based on two properties; one is transitivity and the other is tense. Based on this assumption she considers the following classes of verbs in this language¹:

Group 1: suffixes for the present tense

1sg	2sg	3sg	1pl	2pl	3pl
em/m	in/n	i/e	im/men	inon/non	en/n

Group 2: suffixes for the intransitive past tense

1sg	2sg	3sg	1pl	2pl	3pl
em/m	in/n	-	im/men	inon/non	en/n

Group 3: enclitics for the transitive past tense

1sg	2sg	3sg	1pl	2pl	3pl
im/m	it/t	te/e	imon/mon	iton/ton	won/on

Moradi explains that the suffixes listed in group 1, mark subject agreement in present verbs. In the paradigm that follows, they are marking subject agreement in the conjugation of *warden* ‘to eat’ in the present paradigm. In this paradigm habitual and progressive present conjugations are extracted.

Table 1. Habitual and progressive present conjugation of *warden* ‘to eat’

		Present	
		Habitual	Progressive
SG	1	<i>marem</i>	<i>derem marem</i>
	2	<i>marin</i>	<i>derin marin</i>
	3	<i>mari</i>	<i>deri mari</i>
PL	1	<i>marimen</i>	<i>derim marimen</i>
	2	<i>marinan</i>	<i>derin marinan</i>
	3	<i>maren</i>	<i>deren maren</i>

¹ You may notice a slight difference in the markers introduced in Tafakkori and Omidi and Moradi’s research and the ones listed in my paradigms. The mid back vowels of these markers are the short back vowels in my data.

Moradi explains that the suffixes listed in group 2, mark subject agreement in past intransitive verbs. In the paradigm that follows, they are marking subject agreement in the conjugation of *haten* ‘to come’. In this paradigm the simple past conjugation is extracted.

Table 2. Simple Past Conjugation of *haten* ‘to come’

		Simple Past
SG	1	<i>hatem</i>
	2	<i>hatin</i>
	3	<i>hat</i>
PL	1	<i>hatimen</i>
	2	<i>hatinan</i>
	3	<i>haten</i>

suffixes listed in group 3, mark subject agreement in past transitive verbs. The simple conjugation of *warden* ‘to eat’ is illustrated in Table 3.

Table 3. The conjugation of simple past of *warden* ‘to eat’

Simple Past	
<i>wardem</i>	‘I ate’
<i>wardet</i>	‘You ate’
<i>wardi</i>	‘He/She ate’
<i>wardman</i>	‘We ate’
<i>wardtan</i>	‘You ate’
<i>wardan</i>	‘They ate’

Based on 6 tests suggested by Pullum and Zwicky (1983), Moradi considers all of these markers as enclitics, except the marker of {3 sg}, which she believes based on distributional evidence should be regarded as a suffix. I follow her reasoning, and consider the subject marker in group 3 as clitics except the one for {3sg}, which is regarded as a suffix. However, it seems that her classification does not explain everything in Laki verbal paradigms in terms of subject agreement. The problematic case in Moradi’s classification is how verbs are conjugated in the present and future perfect paradigms in contrast with the conjugation of other verbs in present and future paradigm. All present and future verbs choose class one markers presented in Table 2, while verbs in the present perfect and future perfect, choose markers presented in group 2 in the case of being intransitive, and markers presented in group 3 in the case of being transitive.

Table 4. Present habitual and present perfect conjugation of *warden* ‘to eat’

		Present Habitual	Present Perfect Future Perfect
SG	1	<i>marem</i>	<i>wardema</i>
	2	<i>marin</i>	<i>wardeta</i>
	3	<i>mari</i>	<i>wardiyasi</i>
PL	1	<i>marimen</i>	<i>wardmana</i>
	2	<i>marinan</i>	<i>wardetana</i>
	3	<i>maren</i>	<i>wardana</i>

I propose three possible solutions to this problem. The first solution is borrowed from what Tafakkori and Omid (2014) proposed. They classify verbal inflection in Laki based on the form of the stem. Tafakkori and Omid proposed that subject agreement in Laki is sensitive to the verbal stem and transitivity. As it is evident based on the data presented so far (Table 1, 3 and 4 in particular), past and perfect verbs share the same stem form, and all the present verbs (excluding present perfect) and future verbs (excluding future perfect) have the same stem form throughout the paradigm. Tafakkori and Omid suggested a binary distinction between past and present stem, and they call them past and present stem. But this analysis poses a problem. The problem is, the past stem is displayed throughout the present and future perfect paradigms, despite the semantics of non-past in these paradigms. So it seems that there is a morphomic¹ mismatch in form-content association of stems in present and future paradigms that is not explained in their study. This idea is what that has been addressed in other studies on Iranian languages. Haig (2008) explains that alignment pattern in Iranian languages is sensitive to stem forms (past and present), and the correlation between the form of the stem and its semantics does not match in some Iranian languages. If I accept what Tafakkori and Omid came up with, I should consider Laki as a language where the association of the form and the content of the stem does not display a canonical pattern, which means there is a mismatch between the form of the stem and its semantic interpretation.

To solve the problem of subject agreement in present and future perfect verbs, the second solution would be to suggest the possibility that the subject agreement in Laki is sensitive to perfectivity. This solution helps to account for the same pattern of subject agreement in future and present perfect verbs, and some verbs in past paradigm, but not all verbs in past paradigm. Because not all past verbs have the semantics of perfectivity. For instance, past habitual and past progressive verbs have the semantics of past but not perfect. Yet they inflect for subject agreement the same as the future and present perfect verbs do, as it is displayed in table 5.

¹ This term was originally introduced by Aronoff (1994) in morphology. What he basically discussed in 1994 was that, there are some cells in the paradigms of languages that share the same inflectional forms with distinct contents, such that they do not form a natural class. In other words, he believes that in these instances, the inflectional morphology of the language provides wrong data for the semantic interpretation and the distribution of words in syntax. This should be considered as a mismatch between form and function. Later, other people such as Maiden et al. (2011) defined it more precisely. they think of morphemes as patterns in paradigms that don't make any sense syntactically or semantically, but which morphology reinforces. One instance of that is the so-called L pattern in a lot of Romance languages. There is a particular stem that is used throughout the singular and in the third person plural, and then for the first and second person plural there is a different stem. From the point of view of the content, there is no reason that why they should be that way. Yet, it seems to be a regular pattern that happens in all Romance languages. So morpheme is an abstract pattern in the paradigm. In other works like Stump (2016), morphemes have been addressed as elements that guide morphology, but not anything else in the grammar of languages. This phenomenon is an important piece of evidence for the autonomy of morphology from other components of the grammar.

Table 5. Past habitual and past progressive conjugation of *warden* ‘to eat’

		Habitual	Progressive
SG	1	<i>mawardem</i>	<i>daftem mawardem</i>
	2	<i>mawardet</i>	<i>daftet mawardet</i>
	3	<i>mawardi</i>	<i>dafti mawardi</i>
PL	1	<i>mawardman</i>	<i>daftman mawardman</i>
	2	<i>mawardtan</i>	<i>daftan mawardtan</i>
	3	<i>mawardan</i>	<i>daftan mawardan</i>

It is evident that perfectivity cannot account for the subject agreement in Laki verbal paradigms. Then the third solution is to say that apart from transitivity, the subject agreement is sensitive to preterite property. What I am going to argue is, all the past and perfect verbs (present and future) should be considered as preterite and the present and future verbs (excluding the perfect paradigms) should be regarded as non-preterite. The argument that I am going to support this idea is borrowed from what the philosopher Hans Reichenbach proposed in 1947, and it is fully explained in Michaelis’s paper (2006). Reichenbach proposed that each tense can be represented as the sequence of three time points: speech time **S**, event time **E**, and reference time **R**. Reference time is a time potentially distinct from speech time and event time about which a tensed sentence makes a claim. In absolute tenses, reference time is the same as event time; thus, the simple present, simple past and simple future (all absolute tenses) might be represented as follows (where ‘→’ means ‘precedes’ and a comma means ‘is simultaneous with’):

Simple present: E,R,S

Simple past: E,R → S

Simple future: S → E,R

But in relative tenses (present perfect, past perfect, future perfect), event time precedes reference time:

Present perfect: E → S,R

Past perfect: E → R → S

Future perfect: S → E → R

This seems relevant to my analysis, since the simple past forms a natural class with all three relative tenses: these are exactly the tenses in which E precedes something (either R or S or both). So I think “preterite” means “E → X” in Laki, and “nonpreterite” means “E ↗ X” (where $X \in \{S, R\}$). If so, then the preterite property which is expressed by stem alternation can account for the subject agreement in the future perfect and present perfect verbs. Moreover, the stem alternation does not represent a form-content mismatch, like what I discussed in the previous proposed stem classifications (past~present and perfect~imperfect). On the other hand, preterite property seems to be not the only determining property in alignment pattern of Laki. Transitivity is the other determining factor in subject agreement pattern of this language; otherwise we will not be able to explain why verbs in the paradigm below inflect for subject agreement differently:

Table 6. Simple past conjugation of *haten* ‘to come’ and *warden* ‘to eat’

		<i>haten</i> ‘to come’	<i>warden</i> ‘to eat’
SG	1	<i>hatem</i> ‘I came’	<i>wardem</i> ‘I ate’
	2	<i>hatin</i> ‘You came’	<i>wardet</i> ‘You ate’
	3	<i>hat</i> ‘He/She came’	<i>wardi</i> ‘He/She ate’
PL	1	<i>hatimen</i> ‘We came’	<i>wardman</i> ‘We ate’
	2	<i>hatinan</i> ‘You came’	<i>wardtan</i> ‘You ate’
	3	<i>haten</i> ‘They came’	<i>wardan</i> ‘They ate’

As the conclusion, I classify subject agreement in Laki verbal paradigm as follows:

1. Subject agreement of preterite transitive verbs

=em	=man
=et	=tan
-i	=an

2. Subject agreement of preterite intransitive verbs

-em	-imen
-in	-inan
-	-en

3. Subject agreement of non-preterite verbs

-em	-imen
-in	-inan
-i	-en

References

- Anonby, Erik. 2004. “Kurdish or Luri. Laki’s Disputed Identity in the Luristan Province of Iran”. *Kurdische Studien*, 4(5), 7-22.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. MIT press.
- Haig, Geoffrey. 2008. *Alignment change in Iranian languages: a construction grammar approach*. Vol. 37. Walter de Gruyter.
- Maiden, Martin, John Charles Smith, Maria Goldbach, and Marc-Olivier Hinzelin (eds.). 2011. *Morphological autonomy: Perspectives from Romance inflectional morphology*. Oxford University Press.
- Michaelis, Laura B. 2006. Tense in English. In Bas Aarts and April McMahon, (eds.), *The Handbook of English Linguistics*. Oxford. Blackwell.
- Moradi, Sedigheh. 2015. *Laki Verbal Morphosyntax*. Masters thesis. Linguistics Department of University of Kentucky.
- Pullum Geoffrey, and Arnold Zwicky 1983. “Cliticization vs. inflection: English n’t”.
- Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. New York: Macmillan.
- Stump, Gregory. 2016. *Inflectional paradigms: Content and form at the syntax-morphology interface*. Cambridge University Press.
- Tafakkori, Shoja, and Abbas Omidi 2014. Laki verbal suffixes: ergative construction, subject agreement or pronoun? *Iranian Studies* 4.1, 37-55.
- Windfuhr, Gernot. 2009. *The Iranian languages*. London: Routledge.

Rule Conflation of Noun Markers in Kurdish

Ali Salehi

University of Kentucky

1.1 Introduction

This study investigates various exponents of noun markers, specifically definiteness, number, and person in various syntactic environments in two Kurdish dialects. Kurdish is one of the northwestern Iranian languages spoken by Kurd people in Western Asia. Its three main dialects are Kurmanji (Northern Kurdish), Sorani (Central Kurdish) and Kalthori (Southern Kurdish). In Kurdish, like lots of other languages, NPs that code definiteness, person and number are marked by affixation and demonstratives as morphological markers to express specificity and definiteness. The markers for definiteness, number and person are diverse and show different behaviors in terms of form and placement in various noun phrases. The position of the markers in the noun phrase is seen to be dependent not just on morphology, but on the morphology and the syntax of the sentence. For instance, in Southern Kurdish (Kalthori), the common markers that code definiteness are *-ægæ* and *-ægan*. They are undetachable cumulative suffixes that denote two morphosyntactic features (definiteness and number): *-ægæ* represents both singularity and definiteness and to represent the plural definite, the singular definite marker *-ægæ* is changed to *-ægan*. Neither marker can be segmented to represent each morphosyntactic property separately. In Sorani, the definite state is represented by *-ækæ* and *-ækan*; these follow similar rules for number but go through some phonological changes when placed after vowels at the end of the words. In Kurmanji Kurdish, a noun in the absolute state represents both the generic sense of the noun and the definite sense. For example, a noun like *miróv* (man) depending on the context of the language can mean ‘man’ (in general) or ‘the man’. In Kurmanji, nouns are inflected in four cases, nominative, oblique, construct and vocative.

To investigate and exemplify the behavior of markers, different noun phrases are categorized and glossed based on Abbott’s (2004) classification of NP types. A list of NPs with markers and demonstratives will be provided to compare the distribution of definite markers in various morphotactic environments. The analysis of the definite articles is included in (1) nouns as NPs and nouns in coordination, (2) NPs with adjectival modifiers (3) some other adjuncts of an NP’s head noun e.g. possessive constructions, relative clauses, object relative and subject relative clauses.

2. Lit review: Previous research on topic

Dryer (2013) defines a definite article as “a morpheme which accompanies nouns and which codes definiteness or specificity”. Edmonds (1995), McKenzie (1961), and Yarmoradi (2005) state that *-ægæ* and *-i* are used as affixes to represent definiteness and indefiniteness in Southern Kurdish. While definiteness depends on the unambiguous identification of the participants of the discourse, specificity depends on the knowledge of the speaker. In a sentence like “We don’t know who the president will be.”, the participants have a background knowledge of the discourse, but the referent is not identified. Choubsaz and Rezai (2014) have investigated the definite marker and argue that the morpheme *-ægæ* represents specificity in Kermanshahi Kurdish (a dialect of Southern Kurdish), not definiteness. They focus on the difference between specificity and definiteness in Kermanshahi Kurdish. They draw upon the definition of specificity by Enç (1999) and Karimi (1999, 2003) to investigate *-ægæ* as a morpheme. Based on Lyon’s classification, Kermanshahi Kurdish is placed among languages that only mark indefiniteness. The authors introduce the Kurdish morpheme *-ægæ* as the morphological marker of noun phrases in various syntactic positions. It is notable that *-ægæ* appears at the end of definite and some indefinite noun phrases, so it is considered as a clitic. Using collected samples and data, it is

indicated that the presence of this morpheme is obligatory not only in definite noun phrases also in indefinite noun phrases. The article concludes that unlike what has been claimed in describing this morpheme, it mainly indicates specificity and not definiteness and the morpheme *-æga* is not a mark of definiteness but a mark for specificity. Finally, based on Lyons (1999) view, the authors conclude that Kermanshahi Kurdish only entails the indefinite morphemes.

3. Data Analysis (selected NPs for the sake of the length of the abstract)

3.1.1 Demonstratives

The demonstratives in Sorani Kurdish cover the noun on the edges like a circumfix. ‘this’ is represented by *am... (y)á*, and ‘that’ is represented by *aw... (y)á*. The nouns that these demonstratives surround are absolute singulars and the indefinite plurals.

<i>am</i>	<i>pyâw á</i>	<i>am</i>	<i>ktâwî á</i>	<i>am</i>	<i>dargâ yá</i>	<i>aw</i>	<i>nâma yá</i>
this	man	this	book	this	door	that	letter
‘this man’		‘this student’		‘this door’		‘that letter’	

<i>am</i>	<i>pyâw-ân á</i>	<i>am</i>	<i>ktâwî á</i>	<i>am</i>	<i>dargâ-yân á</i>	<i>aw</i>	<i>nâm-ân á</i>
this	man.PL	this	student	this	door.PL	that	letter.PL
‘these men’		‘this student’		‘these doors’		‘those letters’	

3.1.2 Demonstratives in the coordination of nouns in an NP

<i>am</i>	<i>ktâw o</i>	<i>dáftar á</i>	<i>am</i>	<i>káníjk o</i>	<i>kor á</i>		
this	book	and	notebook	this	girl	and	boy
‘this book & notebook’				‘this girl & boy’			

3.1.3 Demonstratives in larger phrases with adjectives

<i>am</i>	<i>ktâw á</i>	<i>qoy á</i>	<i>am</i>	<i>ktâw á</i>	<i>qoy á</i>	<i>rash á</i>				
this	book	-	big	-	this	book	-	big	-	black
‘this big book’				‘this big black book’						

<i>am</i>	<i>hotel á</i>	<i>gowra</i>	<i>xas á</i>		
this	hotel	-	big	good	-
‘this big good hotel’					

am kaniʃk o kor a balabarz a boor a
 this girl and boy -- tall -- blond --
 ‘this tall, blond girl & boy’

3.1.4 Adjectival modifiers

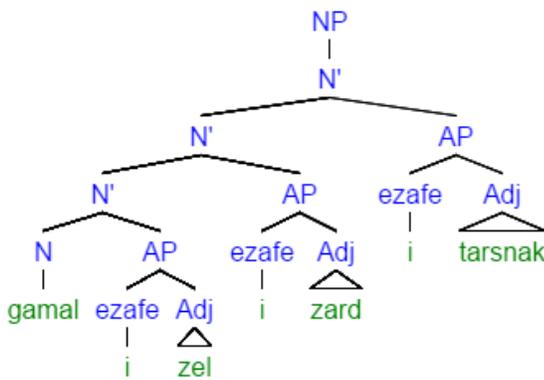
The attributive adjective is placed after the noun and is linked to the noun by the unstressed vowel *i* (the ezafe vowel).

<i>hotel-ân</i>	<i>i</i>	<i>bash</i>		<i>nâm-ân</i>	<i>i</i>	<i>drezh</i>
hotel PL.	EZF	good		letter PL.	EZF	long
‘(some) good hotels’				‘(some) long letters’		

gamal i fera zel
 Dog a very big
 ‘a very big dog’

gamal i zel i zard i tarsnak
 dog a big - yellow - fierce
 ‘a big, fierce, yellow dog’

In this structure, the ezafe vowel precedes all the adjectives to modify them. The distribution of the ezafe vowel in a syntactic tree will be as follows.



The *i* vowel is indeed the ezafe which appears in the indefinite NPs. It's a proclitic and it precedes the head in an NP. The ezafe *i* vowel seem to select attributive adjectival heads. On the other hand, *a* seems to be an enclitic appearing after the head and by default it selects a predicative adjective. In the phrase, ‘this big, yellow dog’, *i* is changed to *a* as an enclitic.

Am- gamal a zel a zard -a
 this dog EZF big yellow --
 ‘this big, yellow dog’

The indefinite singular and plural markers (-*èk*, -*ân*) are attached to the head noun in phrases with adjectival modifiers.

3.1.5 Adjectives with demonstratives and definites

when both the adjective and the noun are covered by the demonstrative, the linking vowel between the two words changes to *a*. Since the distribution of the vowel is very similar to the ezafe vowel *i*, the *a* seems to be the same ezafe vowel *i* which is changed due to vowel harmony in the phrase.

Phrases with definite markers:

Hotel a bash -aká
 hotel EZF good SG.DEF
 ‘the good hotel’

Pyâw a amarîkî -aká
 man EZF American SG.DEF
 ‘the American man’

dars -ân a âsân -aká
 lesson PL. EZF easy SG.DEF
 ‘the easy lessons’

dars -ân a sakht -aká
 lesson PL. EZF hard SG.DEF
 ‘the hard lessons’

Phrases with demonstratives:

Am- hotel a bash -á
 this hotel EZF good -
 ‘this good hotel’

aw- pyâw a amarîkî -á
 that man EZF American -
 ‘that American man’

am- dars -ân a âsân -á
 that lesson PL. EZF easy -
 ‘these easy lessons’

aw- dars -ân a sakht -á
 that lesson PL. EZF hard -
 ‘those hard lessons’

In singular forms, the definite marker is placed after the adjective while the plural marker is after the head noun in the phrase. With definite nouns in plural forms, the noun and the adjective are linked with the ezafe ‘*i*’ vowel again but the placement of the definite suffix, in both is flexible while all the phrases express the same meaning semantically. All the following phrases mean the same:

Rule A Realizes definiteness $\{\alpha\}$ By means of	Rule B Realizes plurality $\{\beta\}$ By means of	Rule [A © B] Realizes $\{\alpha\} \cup \{\beta\}$ By means of
<i>-aká prefixation</i>	<i>-ân suffixation</i>	<i>=akân suffixation</i>

Rule [A © B] $\{\alpha\} \sqcup \{\beta\}$: suffix *-a-b*

In the first expression, both blocks apply in sequence to realize the inflection of the head noun for both definiteness and plurality; in the second, one block applies to the head noun as the plural marker but the other block applies to the word at the NP's right periphery to express definiteness. And in the third, both blocks apply to the peripheral word indicating both plurality and definiteness.

4. Nouns Southern Kurdish: In southern and Sorani Kurdish, the definite marker is postnominal for singular and plural nouns.

4.1.1. singular definite: Southern Kurdish marks definiteness with *-ægæ* which also represents singularity. The marker is attached postnominally to the end of the nouns.

ketaw	-ægæ	dæftær	-ægæ
book	-DEF.SG	notebook	-DEF.SG
'the book'		'the notebook'	

In Sorani, definiteness is marked by *-ækæ* which both represents singularity and definiteness.

4.1.2 Plural definite: The marker *-ægan* marks both definiteness and plurality for the nouns.

ketaw	-ægan	dæftær	-ægan
book	-DEF.PL	notebook	-DEF.PL
'the books'		'the notebooks'	

4.1.4 Noun phrases with adjectival modifiers

When there exists a modifier as an adjective in a noun phrase, the marker attaches to the adjective. Since the adjectives are the end of the phrase in SK, the markers again behave as edge clitics.

4.1.5. Singular noun as head: In noun phrases with adjectives as modifiers, the definite marker behaves as an edge clitic and attaches to the last adjective at the end of the NP.

kænişk	hyl	-ægæ	kor	juwan	-ægæ
girl	blonde	-DEF.SG	boy	young	-DEF.SG
'the blonde girl'			'the young boy'		

4.1.6. with a plural noun as head: both plurality and definiteness is represented by the marker *-ægan*. Behaving like an edge clitic, the marker is again attached to the edge of the phrase after the adjective.

kor	juwan	-ægan	kænişk	hyl	-ægan
boy	young	-DEF.PL	girl	blonde	-DEF.PL
'the young boys'			'the blonde girls'		

4.1.7. Possessive Constructions: Kurdish possessive markers are represented as inflections represented at the edge of the phrases. The possessive markers in Kurdish are *-m*, *-æt*, *-ey*, *-eman*, *-etan*, *-ian*. Behaving like an edge clitic, the possessive markers are placed at the end of the phrase. They behave like the “-’s” in English. This clitic in English appears at the end of the noun to mark the possessive case. In possessive constructions, the definite marker, whether singular or plural is placed on the thing possessed. In this case, the definite markers do not behave as edge clitics. They seem to be morphemes attached to the thing possessed. In the examples below, the definite marker is not placed NP finally. In this particular instance, it’s placed before the possessive marker for person in the NP.

Bawg rafiq -ægæ -m
 Father friend -DEF.SG my
 ‘The father of my friend’

Bawg rafiq -ægan -em
 Father friend -DEF.PL my
 ‘The father of my friends’

saw mal -ægæ -m
 owner house -DEF.SG my
 ‘my landlord’ (the owner of my house)

Saw mal -ægan -eman
 owners house -DEF.PL our
 ‘our landlords’ (the owners of our houses)

rafiq bawg -em
 friend father my
 ‘a friend of my father’

rafiq -æg -ei bawg -em
 friend DEF ezafe father my
 ‘The friend of my father’

4.1.8. Demonstratives: When definiteness is determined by demonstratives in a phrase, both the singular and plural definite markers exhibit an unpredictable behavior. In case of demonstratives, the definite markers can be segmented. The definite marker is attached to the head noun though there is an adjective in the predicate and the definite word is distinct from the demonstrative. Here, the definite marker for both plurals and singulars is *-æ* and plurality is marked by *-æl*.

i kænijk -æ qafæng -æ
 this girl -DEF.SG beautiful is
 ‘This girl is beautiful.’

i kænijk -æl -æ qafæng -en
 this girl -PL DEF beautiful are
 ‘These girls are beautiful.’

i ketaw -æ qoi o kow- æ.
 this book -DEF big and blue is
 ‘This book is big and blue.’

i ketaw -æl -æ qoi o ow -en.
 this book -PL DEF big and blue are
 ‘These books are big and blue.’

i kænijk qafæng -æ
 this girl beautiful DEF.SG
 ‘this beautiful girl’

i kænijk qafæng -æl -æ
 this girl beautiful -PL -DEF
 ‘These beautiful girls’

4.1.9. Object relative & subject relative phrases: A complex NP involving a relative clause does not license a definite marker at the edge of the NP. In this case, the definite marker goes with the head noun in the phrase.

kæniʃk	-ægæ	dyæʃo	di:m	qæʃæŋg	-æ
Girl	-DEF.SG	last night	see.PST	beautiful	is

‘The girl I saw last night is beautiful.’ (object related)

Kæniʃk	-ægan	dyæʃo	haten	qæʃæŋgen
Girl	-DEF.SG	last night	come.PST	beautiful.are

‘The girls who came last night are beautiful.’ (subject related)

5. References

- Abbott, B. (2004) “Definiteness and indefiniteness”, in Laurence R. HORN & Gregory WARD (eds.), *Handbook of Pragmatics*, Oxford: Blackwell.
- Edmonds, C.J. (1955). “Propositional and Personal Affixes in Southern Kurdish”, In *Bulletin of the School Oriental and African Studies*. University of London. Vol. 17. No. 3. Pp. 490-502.
- Edmonds, C.J. (1955). “Propositional and Personal Affixes in Southern Kurdish”. In *Bulletin of the School Oriental and African Studies*. University of London. Vol. 17. No. 3. Pp. 490-502 [In Persian].
- Enç, M. (1991). “The Semantics of Specificity”. *Linguistic Inquiry*. Pp. 22: 1-25.
- Gadilia, K. (2013). A typological study of (in)definiteness in the Iranian languages. Retrieved from <https://www.researchgate.net/publication/236587678>
- Gadilia, K. (2013). “A Typological Study Of (In)Definiteness in The Iranian Languages”. Retrieved from: https://www.researchgate.net/publication/236587678_A_TYPOLOGICAL_STUDY_OF_INDEFINITENESS_IN_THE_IRANIAN_LANGUAGES
- Hedberg, N., & Görgülü, E. & Mamani, M. (2009). On Definiteness and Specificity in Turkish and Persian Proceedings of the 2009 annual conference of the Canadian Linguistic Association.
- Rezai, S. & Choubasaz, Y. (2014). Specificity in Kermanshahi Kurdish. *Language Related Research*. Vol.4, No.4 (Tome 16).
- Zahedi, K., & Mehrzmay, R. (2011). Definiteness in Sorani Kurdish and English. *Dialectologia* 7, 129-157. ISSN: 2013-2247

French dont-relative clauses (and their gap)

Elodie Winckel and Stefan Müller

24th International Conference on HPSG

French relative clauses have been addressed in many generative works (a.o. Kayne, 1976; Godard, 1988), and have been the core of many HPSG analyses (a.o. Sag and Godard, 1994; Kolliakou, 1999; Abeillé and Godard, 2007). We will concentrate in this paper on the relativization of *de*-phrases.

De-phrases (i.e. PPs – or genitive NPs – introduced by a *de*) can be expansions of nouns (without subject island constraint, see example (1)¹), verb complements or adjective complements.² *De*-complements can be relativized (a.o.) with *dont*.

- (1) Cette cuvée très particulière, [**dont** le lancement a eu lieu en this vintage very particular DONT the launching has taken place in 1995], est une reprise de vinification très ancienne encore pratiquée en 1995 is a revival of winemaking very old still practiced in 1935. (FR-WAC)
1935

If *dont* is the expansion of a noun, Abeillé and Godard (2007) assume some restrictions. The noun modified by the relative clause cannot be itself inside a PP:

- (2) *Voilà le pays dont Paul revient [de la capitale -].
that's the country DONT Paul is-returning from the capital
(Abeillé et al., 2006:4)

‘Here is the country from the capital city of which Paul is returning.’

In example (2), the *de*-phrase [*de la capitale du pays*] (‘from the capital city of the country’) is considered as a PP. However, not all *de*-phrases are PPs. Abeillé et al. (2006) have defined two types of *de*: in its oblique uses, *de* is a preposition, projecting a PP; the nonoblique *de* is a weak head which adopts the head value of its complement. The difference between PP[*de*] and NP[*de*] was already present in Sag and Godard (1994) and Godard and Sag (1996), based on the fact that extraction out of an NP is impossible in certain cases, that we can only briefly list here:

1. if the NP is embedded in a PP (see above)

¹from the FRENCH Web As Corpus corpus (FR-WAC)

²We ignore the latter in this paper.

2. if the N is a verb or adjective derivate and the *de*-phrase can be related to an indirect object or modifier of the verb or the adjective
3. when *de* indicates a (local) origin (Sportiche, 1981)
4. in case of multiple *de*-phrases: (Sag and Godard, 1994; Godard and Sag, 1996)
 - a *de*-phrase expressing Agent or Theme cannot be extracted if a *de*-phrase expressing Possessor is also present
 - a *de*-phrase expressing Theme cannot be extracted if a *de*-phrase expressing Agent is also present

Based on point 2, a distinction has been made between *de*-phrases being arguments (genitive NPs) and *de*-phrases being adjuncts. The point 4 has been assumed to follow from the fact that only the first *de*-argument in the ARG-ST of a noun can be extracted, iff this *de*-phrase is indeed a NP[de] and not a PP[de].³

Following *de*-phrases are considered PPs[de]:

- some expansions of the noun which cannot be extracted under any circumstance: local origin, but also compounds-like combinations as in example (3)⁴:
- (3) a. Le trésor de guerre s'épuise.
 the treasure of war declines
 'The stock declines.'
- b. *la guerre dont le trésor s'épuise
 the war DONT the treasure declines
- *de*-phrases as non-sentential argument of verbs, like the one in example (2)
 - expansions of deverbal nouns, when this expansion can be related to the *de*-argument of a verb

To the best of our knowledge, no one has ever taken into consideration the relativization out of nominal *de*-complements of verbs like the ones illustrated in 4:

- (4) a. Il s' occupe [de la carrière de Mickaël Landreau]
 he REFL cares DE the career DE Mickaël Landreau
 'He takes care of Mickaël Landreau's career.'

³Sag and Godard (1994) assume for that a thematic hierarchy: Possessor > Agent > Theme. If several *de*-phrases are realized, only the highest in the hierarchy can be extracted. Kolliakou (1999) recaptures Godard's analysis but her basis is the distinction between IDPs (individual-denoting phrases), i.e. phrases denoting individuals that refer to an entity in discourse, and PDPs (property-denoting phrases), i.e. phrases denoting properties that determine a type of entity. There can only be one IDP in an NP. If two *de*-phrases are realized in the NP, only one can be a IDP, the other one is necessarily a PDP. This distinction offers an alternative to the distinction on the thematic roles of *de*-phrases, which are very hard to determine with certainty.

⁴in which the compound *trésor de guerre* (lit. 'treasure of war') refers to any kind of capital or savings that are being stocked

- b. Je me suis assuré [de la vérité de cette déclaration].
 I REFL have ensured DE the truth DE this statement
 ‘I checked the truth of this statement.’

It is however possible to extract from those complements, as examples in 5 can show:

- (5) a. Proche de l’ ex-gardien du PSG Mickaël Landreau, dont friend of the ex-goalkeeper of-the PSG Mickaël Landreau DONT il s’ occupe [de la carrière _ sur le plan juridique], l’ he REFL cares of the career on the level legal the avocat Didier Domat a déjà défendu avec succès les lawyer Didier Domat has already defended with success the intérêts du Racing et du Red Star par le passé.⁵ intererests of-the Racing and of-the Red Star in the past
 ‘Friend of the PSG ex-goalkeep Mickaël Landreau, whose career he takes care of on the legal level, the lawyer Didier Domat has already defended successfully in the past the Racing’s and the Red Star’s interests.’
- b. D’après cette déclaration, dont je me suis assuré [de la following this statement DONT I REFL have ensured DE the vérité _], j’ai rédigé le présent acte [...].⁶ truth I wrote the present act
 ‘Following this statement, whose truth I checked, I wrote this act’

If PPs are islands to extraction, these examples are in contradiction with the delimitation of PPs[de] mentionned previously. We conduced a study of *dont* relative clauses on a web corpus which shows that three kinds of relativizations out of a *de*-phrase can be observed :

- relativization of a *de*-phrase which is itself embedded in a *de*-phrase⁷
- relativization out of sentential complements of verbs, adjectives and nouns⁸
- relativization out of the indirect *de*-object of a verb (similar to example (5)); It is a rare phenomenon,⁹ but we couldn’t identify any specific characteristic of the verbs allowing this kind of relativization, so that there is

⁵ source: <http://www.leparisien.fr/abo-paris/l-homme-qui-fait-trembler-la-fff-11-06-2010-959398.php>

⁶ source: Chateaubriand, Mémoires d’outre-tombe (1ère partie, livre 4), 1848

⁷All the embedding *de*-phrases however respect the constraints mentionned previously, so that they all may be NP[de]. In this case, it is a violation of the prepositional constraint to extract from them.

⁸Analysing *de* as a weak head, Abeillé et al. (2006:151) account for the relativization out of sentential complements: the *de*-phrase is verbal.

⁹not absolutely restricted to indirect NP objects introduced by *de*:

- (6) l’ eau d’ irrigation dont il plaide pour la rationalisation de l’ usage
 the water of irrigation DONT he argues for the rationalization of the use
 ‘the irrigation water, whose usage he argues for the rationalization of’
 (source:
<http://www.lesoirdalgerie.com/articles/2017/02/04/article.php?sid=208772&cid=2>)

Unlike *de*, *pour* (‘for’) has always only been analysed as a preposition.

no evidence that some of these indirect objects introduced by *de* would be NPs[de] whereas the others would be PP[de].

A straightforward analysis is therefore to consider that *de*-complements of verbs are not PPs.

In our HPSG-analysis, as proposed by Abeillé et al. (2006), the *de* building a PP[de] is a preposition, whereas *de* building an NP[de] is a weak head, and therefore has a HEAD-feature value *noun* if it subcategorizes for a noun. Both have a marking *de*.

Following Müller (1999), we assume that clauses of the type *relative-clause* are non-headed phrases and their relative word is a relative pronoun. Relative clauses introduced by a complementizer however are not a subtype of *relative-clause*, but are of the type *head-complement-phrase*, like other clauses introduced by a complementizer. *Dont* is a complementizer: like *que*, another French complementizer, and unlike relative pronouns, *dont* is invariable, does not constrain its antecedent, constrains the mode of the relative which has to be finite and cannot be used in a pied-piped construction; these are all characteristics of complementizers (Abeillé and Godard, 2007). The *dont* in the relative clauses we have mentioned so far fills a gap in its relative clause. This is not always the case.¹⁰

Dont is a complementizer and we will assume that it has the following lexical entry:

(8) Lexical entry for *dont*:¹¹

$$\left[\begin{array}{l} \text{PHON} \\ \text{LOC|CAT} \\ \text{NONLOC} \end{array} \left[\begin{array}{l} \langle dont \rangle \\ \left[\begin{array}{l} \text{HEAD} \\ \text{MARK} \\ \text{COMPS} \\ \text{SLASH} \end{array} \right] \\ \left[\begin{array}{l} \text{INHER|TO-BIND} \end{array} \right] \end{array} \right] \right]$$

$\left[\begin{array}{l} \left[\begin{array}{l} \text{compl} \\ \text{MOD } \bar{N}_{[1]} \end{array} \right] \\ dont \\ \left\langle \text{VP} \left[\begin{array}{l} \text{VFORM } finite \\ \text{SLASH } \left\langle [2] \text{NP} \left[\begin{array}{l} \text{MARK } de \\ \text{INDEX } [1] \end{array} \right] \right\rangle \right] \right\rangle \end{array} \right]$

Unlike the complementizer *que*, *dont* can only build relative clauses. It is therefore always modifying a noun. *dont* subcategorizes for a complement which is a verbal projection with all argument saturated (the COMPS and SUBJ lists must be empty). This complement also has an element in its SLASH list.¹²

¹⁰Sometimes, *dont* is only co-indexed with a resumptive pronoun inside of the relative clause. This resumptive pronoun must not necessarily correspond to a *de*-phrase.

(7) une difficulté dont_i Paul est certain qu' il la_i résoudra
 a difficulty DONT P. is sure that he it.ACC will-resolve
 (Abeillé and Godard, 2007)

'a difficulty that Paul is sure to resolve'

¹¹simplified representation, based on Abeillé and Godard (2007:30)

¹²As pointed out in footnote 10, *dont* relative clauses without gap also exist. When *dont* builds a relative clause without gap, the antecedent of the relative clause must be coindexed with a relative pronoun in the relative clause. Relative clauses in Hausa exhibit a similar pattern (Crysmann, 2016). Crysmann's analysis defines resumptive pronouns as having an element in their SLASH list (coindexed with themselves). The lexical entry above accounts therefore for both kinds of relative clauses.

In all *dont* relative clauses, the antecedent of the relative clause is either coindexed with a gap or with a resumptive pronoun. The first case is the one we are dealing with in this poster. The gap must correspond to an NP[de] (this is ensured in the lexical entry by the feature TO-BIND).

Dont will fill any kind of gap corresponding to the NP[de] of its verbal complement, even a gap embedded in the complements of the verb: following Godard and Sag (1996), the SLASH element of any nominal complement of the verb is amalgamated into the SLASH list of the verb.

In the poster – here only briefly sketched – we have two goals: (i) gather and make compatible with each other (and partially update) the previous HPSG works on *de*-phrases and (ii) present new corpus data what motivate us to reconsider some delimitations that have been stated about *de*-phrases and their relativization.

References

- Anne Abeillé and Danièle Godard. Les relatives sans pronom relatif. In M. Abecassis, editor, *Le Français Parlé, Normes et Variations*, pages 37–60. L’Harmattan, 2007.
- Anne Abeillé, Olivier Bonami, Danièle Godard, and Jesse Tseng. The Syntax of French *à* and *de*: An HPSG Analysis. In Patrick Saint-Dizier, editor, *Syntax and Semantics of Prepositions*, Text, Speech and Language Technology, pages 147–162. Springer Verlag, 2006.
- Berthold Crysmann. An underspecification approach to Hausa resumption. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King, and Stefan Müller, editors, *Proceedings of the Joint 2016 Conference on Head-Driven Phrase Structure Grammar and Lexical Functional Grammar, Polish Academy of Sciences, Warsaw, Poland*, pages 194–214, Stanford, CA, 2016. CSLI Publications.
- Danièle Godard. *La syntaxe des relatives en français*. Science du langage. Eddu Centre national de la Recherche Scientifique, Paris, 1988.
- Danièle Godard. Extraction out of NP in French. *Natural Language and Linguistic Theory*, 10:233–277, 1992.
- Danièle Godard and Ivan A. Sag. Quels compléments de nom peut-on extraire en français ? *Langages*, 122:60–79, 1996.
- Richard S. Kayne. French relative *que*. pages 255–299, Washington, 1976. Georgetown University Press.
- Dimitra Kolliakou. *De*-phrase extractability and individual/property denotation. *Natural Language and Linguistic Theory*, 17:713–781, 1999.
- Stefan Müller. *Deutsche Syntax Deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Linguistische Arbeiten. Max Niemeyer Verlag, Tübingen, 1999.
- Ivan A. Sag and Danièle Godard. Extraction of *de*-phrases from the French NP. In *Proceedings of the North East Linguistic Society 24*, pages 519–541, Department of Linguistics, University of Massachusetts, 1994. GLSA.
- Dominique Sportiche. Bounding nodes in French. *The Linguistic Review*, 1(2):219–246, 1981.

NP-internal Coordination and Agreement in French

Kei TANIGAWA

University of Tokyo

kei.tanigawa@gmail.com

The aim of this study is to examine the syntactic structure of coordination within noun phrases (NPs) in French. When French speakers coordinate two phrases headed by nouns, they usually coordinate two NPs, as shown in (1).

- (1) Mon chien et mon chat sont obéissants.
my.M.SG dog and my.M.SG cat be. PL obedient
‘My dog and my cat are obedient.’

Several studies (e.g. Blanche-Benveniste 1995) have reported that the determiner of the second conjunct can be omitted in conversation, but in literary works, the omission of the determiner of the second conjunct is rare except when the two coordinate NPs have the same index as shown in (2).

- (2) Un mathématicien et physicien anglais, Lord Kelvin, suggéra[...]
a mathematician and physician English, Lord Kelvin, suggested[...]
‘An English mathematician and physician, Lord Kevin, suggested[...]
(Berkeley, Cerveaux Geants, 1957)

In (2), *mathématicien* and *physicien* have the same index (*Lord Kelvin*).

Over the past few decades, several studies have focused on coordination and grammatical agreement within NPs in Portuguese (Villavincencio et al. 2005, Arnold et al. 2007) and agreement within NPs in other languages (King and Dalrymple 2004 among others). As for French, Abeillé and Mouret (2010), Shiraïshi and Abeillé (2016) have studied the relationship between coordination and the agreement, and Abeillé (2006) also treats coordination within NPs, but there seems to be no established theory that captures the relationship between the coordination and the agreement within NPs.

In this study, a questionnaire was conducted in which 15 native speakers of French who live in Paris were asked to judge the acceptability of French sentences involving NP-internal coordination on a 4-point scale. One purpose of this questionnaire was to

compare the acceptability of NP-internal coordination involving several different determiners: a possessive form, a demonstrative form, the indefinite article, and the definite article. When the sentences contained a definite article, the informants were asked to judge the acceptability both with and without a given context, because definite articles in French allow generic interpretation, and that interpretation needs to be distinguished from their more canonical interpretation.

[given context: I have a desk and a chair.]

- (3) ?*Mes/ ?*Ces/*Des/?*Les chien et chat sont obéissants.
 my. PL/these/INDEF.PL/DEF.PL dog and cat be. PL obedient
- (4) *Mon/*Ce/*Un/*Le chien et chat sont obéissants.
 my.SG/this/INDEF.SG/DEF.SG dog and cat be.PL obedient
- (5) *Mon/*Ce/*Un/*Le chien et chat est obéissant.
 my.SG/this/INDEF.SG/DEF.SG dog and cat be.SG obedient

(intended) ‘My dog and my cat are obedient.’

The questionnaire results concerning the sentences in (3)-(5) clearly showed that all of these sentences are unacceptable or at least very difficult to accept.¹ No other pair of words in this study showed different results. This result, especially the low acceptability of the sentence as shown in (4), is consistent with the data as pointed out and discussed in Heycock and Zamparelli (2005)².

Unlike sentences like (3), involving coordination of singular nominals³, sentences such as (6), involving a plural determiner and a plural nominal, are naturally acceptable.

- (6) Mes/Ces/Des/Les animaux sont obéissants.
 my.PL/these/INDEF.PL/DEF.PL animal.PL be.PL obedient
 ‘My/these/∅/the animals are obedient.’

Therefore, it can be concluded that plural determiners distinguish a plural nominal from a coordinate structure consisting of singular nominals.

¹ An and Abeillé (2017) give some data that singular nominal whose referents differ can be coordinated. This is the further issue to examine and analyse.

² Heycock and Zamparelli (2005) suggest a semantic analysis, but in this study, it is supposed that the agreement phenomena are syntactic issues.

³ Within an NP, the whole part except determiners will be called ‘nominal(s)’ in this study.

Sentences like (7), involving a singular determiner and a singular nominal, are naturally also acceptable.

- (7) Mon/Ce/Un/Le chien est obéissant.
 my.SG/this/INDEF.SG/DEF.SG dog be.SG obedient.

‘My/this/a/the dog is obedient.’

Comparison between the sentences like (7) and ones involving coordination of singular nominals like (4) and in (5) shows that singular determiners distinguish a singular nominal and a coordinate structure consisting of singular nominals, accepting the former but the latter as its complement.

To recapitulate; regarding the coordination of nominals in French, it can be said

- (i) that determiners in plural distinguish a plural nominal from a coordinate structure consisting of singular nominals (they take only a plural nominal as its complement), and
- (ii) that determiners in singular distinguish a singular nominal from a coordinate structure consisting of singular nominals (except for cases in which singular NPs with the same index are coordinated, singular determiners take only a non-coordinate singular nominal as its complement).

The analysis suggested here is roughly based on the analysis of subject-verb agreement in Yatabe (2004). In order to capture the phenomena shown above, we propose the following hypotheses.

Hypothesis 1: There are at least two lexical entries for the word *et* ‘and’ in French, and both can conjoin NPs. One lexical entry is to form an NP with a plural index (*plural-and*) and another one is to form an NP with a singular index (*singular-and*).

Hypothesis 2: Both determiners and nouns have a HEAD feature which is called AGR. The AGR value of determiners is a feature structure that consists of NUM(ber), GEN(der), and the value of this feature on nouns is a feature structure that has NUM, GEN, and PER(son).

Hypothesis 3: As defined in Yatabe (2004: 342), when a coordinate structure *M* is made up of *n* conjunct daughters ($d_1 \dots d_n$) and the HEAD value of *M* will be described as below.

The HEAD value of *M* is

$$\left[\begin{array}{l} \text{CONJ } \boxed{0} \\ \text{ARGS } \langle \boxed{a_1}, \dots, \boxed{a_n} \rangle \end{array} \right]$$

where $\boxed{a_1} \dots \boxed{a_n}$ are the HEAD values of $d_1 \dots d_n$ respectively, and $\boxed{0}$ is the SYNSEM|CONT|RELN value of M .

In previous works, NP-internal agreement has been characterized in terms of closest conjunct agreement and resolution of the feature values on coordinated elements. That type of analysis works for most cases, but not general enough to cover the entirety of complex agreement phenomena.

With Hypotheses 2 and 3, it is possible to cover all the agreement patterns of coordination within NPs in French and other languages. The *number* agreement relation, which is required to hold between a determiner and the immediately following nominal, is defined as follows. $\boxed{1}$ is the HEAD value of a determiner and $\boxed{2}$ is the HEAD value of a nominal. The value k means an arbitrary number.

$$\begin{aligned}
 &\text{number agreement } (\boxed{1}, \boxed{2}) \equiv \boxed{1} [\text{AGR} \mid \text{NUM } \boxed{3}k] \wedge \boxed{2} [\text{AGR} \mid \text{NUM } \boxed{3}] \\
 &V(\boxed{2}: \left[\begin{array}{l} \text{CONJ } et \\ \text{ARGS } \langle \boxed{a_1}, \dots, \boxed{a_n} \rangle \end{array} \right] \\
 &\quad \wedge \text{ number agreement } (\boxed{1} [\text{AGR} \mid \text{NUM } pl], \boxed{a_1}) \\
 &\quad \wedge \dots \\
 &\quad \wedge \text{ number agreement } (\boxed{1} [\text{AGR} \mid \text{NUM } pl], \boxed{a_n})) \\
 &V(\boxed{2}: \left[\begin{array}{l} \text{CONJ } ou \\ \text{ARGS } \langle \boxed{a_1}, \dots, \boxed{a_n} \rangle \end{array} \right] \\
 &\quad \wedge \text{ number agreement } (\boxed{1}, \boxed{a_1})) \\
 &V(\boxed{1}: \left[\begin{array}{l} \text{CONJ } ou \\ \text{ARGS } \langle \boxed{b_1}, \dots, \boxed{b_n} \rangle \end{array} \right] \\
 &\quad \wedge \text{ number agreement } (\boxed{b_n}, \boxed{2})) \\
 &V(\boxed{2}: \left[\begin{array}{l} \text{CONJ } et \\ \text{ARGS } \langle \boxed{a_1}, \dots, \boxed{a_n} \rangle \end{array} \right] \\
 &\quad \wedge \text{ number agreement } (\boxed{1} [\text{AGR} \mid \text{NUM } sg], \boxed{a_1}) \\
 &\quad \wedge \dots \\
 &\quad \wedge \text{ number agreement } (\boxed{1} [\text{AGR} \mid \text{NUM } sg], \boxed{a_n}) \\
 &\quad \wedge (\boxed{a_1} [\text{INDEX } \boxed{4}x] \wedge \boxed{a_2} [\text{INDEX } \boxed{4}] \wedge \dots \wedge \boxed{a_n} [\text{INDEX } \boxed{4}])) \\
 &\dots
 \end{aligned}$$

The relation above might seem to be overly complicated, but it will be argued that this complexity is necessary to cover the agreement phenomena in French and other languages. For example, NP-internal coordinate structure in English shows an agreement strategy that is slightly different from that of French, as shown in (9).

(9) This man and woman are in love. (From Heycock and Zamparelli 2005)

In (9), the determiner *this* is in singular and the noun-verb agreement is in plural. King and Dalrymple (2004) suggest that the properties of phrases like (9) can be explained by using the distributive CONCORD feature and non-distributive INDEX feature, and give an analysis along similar lines for determiner-noun agreement across the languages.

Heycock and Zamparelli (2005) propose a semantic analysis of the distribution of the split reading and the joint reading across the languages. It will be argued that the empirical coverage of these theories is not as comprehensive as that of the theory above.

To summarize, questionnaire results indicate that, within NPs in French, plural determiners distinguish a non-coordinate plural nominal form from a coordinate structure consisting of singular nominals, and singular determiners distinguish a non-coordinate singular nominal from a coordinate structure consisting of singular nominals. A strategy was proposed that successfully characterizes these observations and some other agreement phenomena in several languages.

References

- Abeillé, Anne. 2006. In defense of lexical coordination. In O. Bonami and P. Cabredo Hofherr (eds.), *Empirical Issues in Syntax and Semantics* 6, Paris: 7-36.
- Abeillé, Anne and François Mouret. 2010. Quelques contraintes sur les coordinations elliptiques en français. *Revue de Sémantique et de Pragmatique* 24,177-207.
- An, Aixiu and Anne Abeillé. 2017. Agreement and interpretation of binominals in French. In 4th European Workshop on HPSG. University of Chicago Center, Paris.
- Arnold, Doug, Louisa Sadler, and Aline Villavicencio. 2007. Portuguese : corpora, coordination and agreement. In S. Featherston and W. Sternfeld (eds.), *Roots : Linguistics in Search of its Evidential Base*, Mouton de Gruyter, Studies in Generative Grammar 96, The Hague : 9-28.
- Blanche-Benveniste, Claire. 1995. Répéter ou ne pas répéter. In H.B. Shyldkrot and L. Kupferman (eds.), *Tendances Récentes en Linguistique Française et Générale*. John Benjamins, Amsterdam: 55-74.
- Heycock, Caroline and Roberto Zamparelli. 2005. Friends and colleagues: Plurality, coordination, and the structure of DP. In I. Heim and A. Kratzer (eds.), *Natural Language Semantics* 13, Springer, Berlin: 201-270.
- King, Tracy H. and Mary Dalrymple. 2004. Determiner agreement and noun conjunction. *Journal of Linguistics* 40(1), Cambridge University Press, Cambridge: 69-104.
- Shiraïshi, Aoi and Anne Abeillé. 2016. Peripheral ellipsis and verb mismatch. *Proceedings of the Joint 2016 Conference on HPSG and LFG*, CSLI Publications, Stanford: 662-680.
- Villavicencio, Aline, Louisa Sadler and Doug Arnold. 2005. An HPSG account of closest conjunct agreement in NP coordination in Portuguese. In S. Müller (ed.), *Proceedings of the 12th International Conference on Head-driven Phrase Structure Grammar*, CSLI Publications, Stanford: 427-447.
- Yatabe, Shûichi. 2004. A Comprehensive theory of coordination of unlikes. In Stefan Müller (Ed.), *Proceedings of the HPSG 04 Conference*, CSLI Publications, Stanford: 335-355.

Yucatecan Control and Lexical Categories in SBCG: An Abstract

Maksymilian Dąbkowski

*Department of Cognitive, Linguistic & Psychological Sciences,
Brown University, Providence*

Yucatec Maya has two different types of complement control constructions, where *control construction* is understood as one where the implicit subject of a given verb is determined by some other expression in the clause. In this paper, the constructions will be referred to as *incompletive control* and *subjunctive control*, since what distinguishes them is the subjunctive or incompletive status marking on the embedded verb. *Status* is a traditional term employed in Mayanist literature for verbal suffixes whose choice is subject to the aspect, mood and transitivity of the verb.¹

- (1) In k'aat in ts'iib-t-∅-∅ le kàarta-o'.
A1SG wish A1SG write-APP-SBJ-B3 DEF letter-D2
"I want to write the letter." (lit. "To write the letter is my wish.")
- (2) Táan in bin-∅ in ts'iib-t-ik-∅ kàarta-o'ob.
PROG A1SG going-INC A1SG write-APP-INC-B3 letter-B3PL
"I am going (around while) writing letters."

Here, sentence (1) demonstrates a construction where the embedded verb *ts'iibt* takes on subjunctive status.² Subjunctive status is required by desideratives as the above *k'aat* and verbs such as *kanik* "learn," *ojel* "know" or *sajak* "fear." Sentence (2) demonstrates an incompletive control construction where the embedded verb is explicitly marked with the suffix *-ik*. Other matrix clauses selecting for incompletive complements include verbs such as *chúunsik* "begin," *k'a'ajsik* "remind" or *k'a'ajal* "remember." Motion verbs select for either status, but their semantics then vary: the subjunctive status indicates a purpose while the incompletive indicates a "manner" or simultaneity [3]. Notice the overt agreement between the matrix clause and the embedded clause, in this case both expressly marked for 1st person singular in (1) and in (2).

Main status markers are: *completive*, *incompletive* and *subjunctive*, the first one having a considerably restricted distribution. Despite their labels, evocative of aspectual information, their semantic import is frequently negligible. Clauses with active verbs begin with one of several words indicating aspectual or modal information, known in the Mayanist literature as Aspect-Mood markers (AM markers). In most constructions, the status of the verb is strictly governed by an AM marker, so its meaning, whatever it be, is subsumed under the AM marker's much stronger semantics.

	TRANSITIVE	INTRANSITIVE
INC	...- <i>ik</i>	...- <i>Vl</i> , -∅
SBJ	...- <i>ej</i> , ...-∅	...- <i>Vk</i> , ...- <i>ak</i>
CPL	...- <i>aj</i>	...- <i>aj</i> , ...-∅

Table 1: Status morphology

¹Abbreviations for glosses: 1: first person, 2: second person, 3: third person, A: set A, APP: applicative voice, B: set B, CAUS: causative voice, CPL: completive status, D2: distal clitic, DEF: definite article, IMP: imperfective AM marker, INC: incompletive status, ONGL: onglide, PL: plural, PROG: progressive AM marker, PRV: perfective AM marker, REC: recent past AM marker, REL: relational, SBJ: subjunctive status, SG: singular, TERM: terminative AM marker.

²In this case, the subjunctive status is morphologically empty, but that does not have to be the case. Status morphology is sensitive to voice, transitivity and type of verb. Table 1, adapted from AnderBois and Armstrong (unpublished manuscript), gives status suffixes for active verbs [3].

- (3) Táan in páan-ik-∅ u y-okom-al.
 PROG A1SG dig.out-INC-B3SG A3 ONGL-pillar-REL
 “I am digging out (holes) for the pillars.” (E447) [7]

The presence of the status suffix in (3) is, in some sense, semantically redundant, since the progressive aspect marker already has an “incompletive” sense. Other times, the “meaning” of the status suffix is entirely contradicted by the AM marker. In (4), the terminative AM marker *ts’o’ok* does not conflict with the incompletive status. To the contrary, it demands it and overrides its meaning.

- (4) Ts’o’ok a took-ik-en ti’ le kim-il-o’.
 TERM A2 wrest-INC-B1SG PREP DEF die-NML-D2.
 “You have wrested me from death.” (73a) [3]

The main preoccupation of this paper is discussing, understanding and accounting for an unexpected property of subjunctive control. While incompletive control easily generalizes from transitive verbs to intransitive ones, subjunctive control is not as well-behaved. In fact, the most obvious approach to generating the subjunctive control with intransitive verbs (i.e. one employing intransitive verbs with subjunctive morphology) is ungrammatical (5). The proper intransitive equivalent to (1) is expressed via an *incompletive* verb stripped of the appropriate agreement marking, as displayed below (6).

- (5) *J tãal-∅-en na’ak-ak-en. (6) J tãal-∅-en na’ak-al.
 PRV come-CPL-B1SG ascend-SBJ-B1SG PRV come-CPL-B1SG ascend-INC
 intended: “I came to ascend.” “I came to ascend.”

This paper is indebted to AnderBois and Armstrong (unpublished manuscript, henceforth A&A), whose analyses are here frequently evoked and examples cited. A&A argue, in my opinion convincingly, that Yucatecan control constructions are best analyzed as essentially verbal. Their approach stands in juxtaposition to Coon (2013) who argues that the syntax of analogous constructions in Chol is nominal [9]. Yet, A&A do not sever all ties to Coon’s original proposal; just like her, they classify the intransitive subjunctive control (henceforth ISC) as nominal in an attempt to explain its unusual properties. This paper evaluates the arguments supporting that position, points at syntactic similarities between the categories of noun and verb in Yucatec Maya, ultimately rejects the nominal treatment of ISC, and proposes an alternative approach carried out in SBCG, a framework incorporating insights from both Berkeley Construction Grammar and Head Driven Phrase Construction Grammar [6]. I argue that a constructionist approach to constructions such as Yucatecan control is the simplest and most accurate in generating analyses which eschew unjustified, from a language-internal perspective, stipulations. Finally, I suggest that ISC’s properties might be viewed as a disambiguation mechanism which arose in a historical process.

One of two fundamental arguments presented by A&A in favor of the nominal analysis for ISC rests on the observation that the incompletive status, one of two main quirks of ISC, and the instances of unambiguous nominalization use the exact same set of suffixes. The claim is then that nominalizations are used instead of verbal arguments if the verbal arguments fail to fulfill specific syntactic requirements of the matrix clause. In this case, the requirement is that the embedded verb’s set-A morphemes agree with the subject of the matrix clause.

	SINGULAR	PLURAL
1ST	...-en	...-o’on
2ND	...-ech	...-e’ex
3RD	...-∅	...-o’ob

Table 2: Set-B morphology

	SINGULAR	PLURAL
1ST	<i>in (w-)</i> ...	<i>k</i> ... <i>in (w-)</i> ...-o’on
2ND	<i>a (w-)</i> ...	<i>a (w-)</i> ...-e’ex
3RD	<i>u (y-)</i> ...	<i>u (y-)</i> ...-o’ob

Table 3: Set-A morphology

Set-A and set-B are two sets of agreement markers which appear on Yucatecan verbs and nouns. The somewhat uninformative names are traditionally used in Mayanist literature, for the functions of those morphemes do not perfectly correspond to any categories for which Latinate nomenclature has been established. Nonetheless, some approximations can be made. Set-A, broadly understood as ergative-genitive, cross-references subjects of transitive verbs, subjects of incompletive intransitive verbs and pos-

sessors of nouns. Set-B, broadly understood as absolutive, cross-references subjects of stative predicates (nouns and adjectives), objects of transitive verbs, and subjects of intransitive verbs marked for subjunctive or completive status. Tables 2 and 3 have been adapted from Lehmann (2002) [15].

The other argument invoked by A&A pertains to the lack of agreement morphology on the embedded verb in ISC constructions. The subject of an incomplete intransitive is normally expressed through set-A morphology. That is syntactically identical to how possessors are expressed in noun phrases.

(7) in wakax A1 cow “my cow”	(8) in na’ak-al A1 ascend-INC “I ascend” perhaps: “my ascending”	(9) na’ak-al ascend-INC perhaps: “ascending”
--	---	--

Even so, set-A morphology does not appear in ISC. A&A claim that nominalizations can be expected to have a looser syntactic relation to their arguments. The claim seems plausible on the surface. After all, both “his studying” and “studying” are viable noun phrases, but “he studies” does not licence “studies” as a grammatical sentence.

Nonetheless, a closer look reveals that neither of the two proposed tests can be used to successfully argue for the nominal status of ISC complements in Yucatec Maya. While it is undoubtedly the case that nominalizations use the same morphology as the incomplete status, it must be noted that they can be definite, which means that they are accompanied by the definite article *le* and a distal clitic such as *’o*. Definite morphology might be indeed preferred. During elicitations, sentences with nominalizations unmarked for definiteness tend to be repeated with a distal clitic, which disambiguates the syntactic category they belong to (10-11) [2]. But that is not the case with incomplete subjunctive control, where neither the definite article *le* nor distal clitics can appear. Were we to postulate that ISC truly involves nominalizations, we would then replace one idiosyncrasy of Yucatec Maya with another, and start scratching our heads all over again, pondering the new question: if ISC arguments are verbs which had undergone nominalization, why do they shun definite morphology? Nominalization as explicans brings forth only more explananda.

(10) ?Uts-∅ xook-∅. good-B3GS study-INC/NML “It’s good to study.” / “Studying is good.” (?)	(11) Uts-∅ xook-∅-o’. good-B3GS study-INC/NML-D2 “It’s good to study.” (?) / “Studying is good.”
---	--

The argument from lack of set morphology does not fare much better. First of all, applying facts about how nominalization works in English might not bear much fruit in Yucatec Maya. As argued by Lehmann, Yucatecan nouns are quite rigidly divided into two categories: those which must be possessed (inalienably possessed nouns) and those which cannot (absolute nouns) [15]. Absolutive uses of inalienably possessed nouns are possible, as are possessed uses of alienable nouns, but not without undergoing some derivational morphology first.³ Thus, the facts about of possessors’ optionality, so familiar from English, do not apply in Yucatec Maya.

One possible stance in light of the above facts is that ISC arguments are absolute nominalizations, i.e. nominalizations which, akin to absolute nouns, do not admit possessors. Of course, this explanation would hardly be an explanation at all: there is no a priori reason why verbs should *necessarily* lose their syntactic arguments when nominalized, especially given the facility with which they could be expressed in Yucatec Maya. If anything, we would expect the nominalizations to be necessarily possessed, as that would match their semantics more closely.

Nonetheless, there is an even stronger reason to reject the link between a lack of set-A morphology in ISC and nominalization. Yucatecan nouns and verbs display a wide range of syntactic parallelisms. In fact, the numerous similarities make some posit that all dynamic verbs should be analyzed as “nominal.” Sentence (12) could be then thought of as “your watching of a cow is ongoing,” instead of its more natural-sounding translation [7].

³Lehmann includes in his analysis a third category: neutral nouns [15]. Nonetheless, he himself notices that even neutral nouns are subject to tone lowering when possessed. It is therefore perhaps better to analyze them as a separate subgroup of absolute nouns.

- (12) Tàan a w-il-ik-∅ wakax.
 PROG A2 ONGL-watch-INC-B3 cow
 “You are watching a cow.”
- (13) wakax-o’ob
 cow-B3PL
 “they are cows (a cow)” / “cows”

Now, that is not to say there is no difference between Yucatecan nouns and verbs. Even though that distinction in Mayan languages is not as fundamental as, let’s say, in Indo-European, they most surely constitute two discrete categories; *wakax* “cow” could never be used as an active verb, at least not without undergoing some derivational morphology first. But the point I want to make is different: the distinction between nouns and verbs does not translate into a distinction in set morphology. As shown in examples (7) and (8), set-A morphology can play both an ergative and a genitive function. The ambiguity can be even more radical when set-B is involved. In addition to its verb-restricted usages, noun phrases with suffixed with *-o’ob* can have a predicative and plural readings, as demonstrated in example (13).

From the syntactic perspective, nonetheless, set-A and set-B morphemes are agnostic about what they attach to. Most importantly, the parallelisms between set-A and set-B on nouns and verbs go deeper than what first meets the eye. As was already mentioned, the requirements of both nouns and verbs with respect the presence of set morphology (especially set-A morphology) are strict. That is, the set morphology must or must not appear, and optionality is generally not an option. In addition, both nouns and verb participate in syntactically identical head-complement constructions, where the complement follows the noun. For nouns, the complement is the possessor while for verbs it is the agent.

- (14) u tàataj le ts’akya_j=o’
 A3 father DEF doctor=D2
 “the doctor’s father”
- (15) u y-il-ik-∅ le tsakya_j=o’
 A3 ONGL-watch-INC-B3SG DEF doctor=D2
 “the doctor is watching”

All the above examples are presented to motivate a unified treatment of set-A and set-B morphology where set morphology consists of one group of words/morphological functions applicable to both nouns and verbs, a claim easy to model in a multiple-inheritance hierarchy such as SBCG. Given these parallelisms between nominal and verbal set morphology, it seems exceedingly unlikely that morphological categories generally indifferent to the lexical status of their hosts would make that distinction for one construction only. My proposal then is to preserve the elegance and simplicity of a unified set morphological account, and stipulate the restrictions on control constructions separately.

Arbitrary stipulations are generally dispreferred in syntactic theory. Nonetheless, syntactic idiosyncrasies are not unattested and when encountered, they should be recognized as such. Fortunately, another construction from Yucatec Maya gives us a hint that this approach is correct. This construction is agent focus and verbs bereft of set-A morphology are its cornerstone. In contrast to ISC, in agent focus the “defective” verb is necessarily transitive.

- (16) K-u w-il-ik-∅ polok wakax Maruch.
 IMP-A1SG ONGL-watch-INC-B3SG fat cow Mary
 “Mary is watching a fat cow.”
- (17) Maruch il-ik-∅ polok wakax.
 Mary watch-INC-B3SG fat cow
 “MARY is watching a fat cow.” / “It is Mary who is watching a fat cow.”

In essence, the agent focus construction is characterized by fronting the transitive verb’s agent and removing the associated AM marker alongside with the set-A morpheme. Even though the quirks of agent focus are very different from subjunctive control, agent focus points at a precedence in Yucatecan grammar of verbs without the otherwise expected set-A. Interestingly, it has been suggested that this way of marking agent focus emerged in Yucatec Maya to disambiguate between agent and patient focus after all morphology associated with agent focus was lost [17]. Were that true, one could look at intransitive subjunctive control in a similar way—here too it is set-A whose presence or lack disambiguates between two otherwise identical constructions: the intransitive incomplete control and the intransitive subjunctive control.

References

- [1] Scott AnderBois. Focus and uninformativity in Yucatec Maya questions. *Natural language semantics*, 20(4):349–390, 2012.
- [2] Scott AnderBois. Personal communication, January-February 2017.
- [3] Scott AnderBois and Grant Armstrong. *On a transitivity-based split in Yucatec Maya control*. Brown University and University of Wisconsin, unpublished.
- [4] Grant Armstrong. On Copular Sentences in Yucatec Maya. In *Proceedings of the Conference on Indigenous Languages of Latin America-IV.*, 2009.
- [5] Grant Armstrong. Vocabulary & Grammar Packet — Level II. In *Yucatec Maya Summer Institute*. UNC-Duke Consortium for Latin American Studies, 2014.
- [6] Hans C. Boas and Ivan A. Sag, editors. *Sign Based Construction Grammar*. Center for the Study of Language and Information Publications, 2012.
- [7] Jürgen Bohnemeyer. *The Grammar of Time Reference in Yucatec Maya*. Lincom Europa, 2002.
- [8] Jessica Coon. When Ergative = Genitive: Nominals and Split Ergativity. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, pages 99–107, 2008.
- [9] Jessica Coon. *Aspects of split ergativity*. Oxford University Press, 2013.
- [10] Marshall Durbin and Fernando Ojeda. Basic word order in Yucatec Maya. *Papers in Mayan linguistics*, pages 69–77, 1978.
- [11] Colette Grinevald and Marc Peake. Ergativity and voice in Mayan: A functional-typological approach. *Ergativity, valency and voice*, pages 15–49, 2012.
- [12] Rodrigo Gutierrez-Bravo and Jorge Monforte y Madera. On the nature of word order in Yucatec Maya. *Information structure in indigenous languages of the Americas*, pages 139–170, 2010.
- [13] Javier Abelardo Gómez Navarrete. *Diccionario Introductorio. Español-Maya, Maya-Español*. Universidad de Quintana Roo, 2009.
- [14] Edward Keenan and Bernard Comrie. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1):63–99, 1977.
- [15] Christian Lehmann. *Possession in Yucatec Maya. Second, revised edition*. Universität Erfurt, 2002.
- [16] Philip H. Miller and Ivan A. Sag. *French Clitic Movement without Clitics or Movement*. 1995.
- [17] Elisabeth Norcliffe. *Head Marking in Usage and Grammar: A Study of Variation and Change in Yucatec Maya*. Stanford University, 2009.
- [18] Elisabeth Norcliffe. *Revisiting Agent Focus in Yucatec*. Stanford University, 2009.
- [19] Mike Reape. Domain union and word order variation in German. In J. Nerbonne, K. Netter, and C. J. Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, page 151–197. Stanford University: CSLI Publications, 1994.
- [20] Elisabeth Verhoeven. *Experiential constructions in Yucatec Maya*. John Benjamins Publishing Company, 2007.

Yucatec Maya in SBCG: A Fragment

Justin Bai, Maksymilian Dabkowski,
Kalinda Pride, Nicholas Tomlin

Yucatec Maya (YM) is a Mesoamerican language spoken by about a million speakers primarily in the Yucatan Peninsula. In this paper, we present parts of a grammatical fragment with particular emphasis on the pronominal system, verbal morphosyntax, focus, deictic clitics and topicalization through the lens of Sign-Based Construction Grammar (Boas & Sag 2012), which allows us to account elegantly for disparate constructions which share many characteristics.

1 Predicativity and agreement

In our fragment, we adhere to the conventional analyses of Yucatecan word order and treat it as an underlyingly verb (or perhaps better: predicate) initial language. As for copular predicates then, the word order of the following sentences will be taken as most basic.¹

(1) Koolnáal-en (tèen).
farmer-B.1.SG I
'I am a farmer.'

(2) Polok-ech (tèech).
fat-B.1.SG you
'You are fat.'

	SINGULAR	PLURAL
1ST PERSON	in (w) <i>stem</i>	k <i>stem</i> in (w) <i>stem-o'on</i>
2ND PERSON	a (w) <i>stem</i>	a (w) <i>stem-e'ex</i>
3RD PERSON	u (y) <i>stem</i>	u (y) <i>stem-o'ob</i>

Figure 1: Set-A morphology

	SINGULAR	PLURAL
1ST PERSON	<i>stem-en</i>	<i>stem-o'on</i>
2ND PERSON	<i>stem-ech</i>	<i>stem-e'ex</i>
3RD PERSON	<i>stem-Ø</i>	<i>stem-o'ob</i>

Figure 2: Set-B morphology

There are two categories of agreement morphemes called Set-A and Set-B, terminology employed in the traditional Mayanist literature. The atypical nomenclature is used as the two categories do not quite completely overlap with other, typologically more common categories. Set-A, for example, perhaps best approximated as ergative-genitive, cross-references agents on transitive verbs, and subjects on intransitive incomplete verbs, but also functions as a possessor on nouns. Set-B, best approximated as absolutive, cross-references patients on transitive verbs and subjects on intransitive ones, and also functions as plural marking on nouns. Set-A and Set-B morphemes can attach to both nouns and verbs, but their behavior is unaffected by the lexical class of their host. That motivates a unified treatment, which we provide. We recast Set-B morphology via one morphological function and Set-A morphology using the same function accompanied by 3 additional lexemes.

All basic non-copular sentences in Yucatec Maya—that is, sentences containing active verbs—start with one of fifteen Aspect-Mood (AM) markers. Here, we adapt basic analytical insights from Bohnemeyer (1998) and Bohnemeyer (2002) and postulate that AM markers are in fact the predicates of those sentences. Example (3) would then be more literally translated as ‘my reading of the newspaper is/was/will be near,’ as *ta'itak* is the predicate. This allows for a simpler analysis and preserves coherence of VOS word order throughout the language.

All basic non-copular sentences in Yucatec Maya—that is, sentences containing active verbs—start with one of fifteen Aspect-Mood (AM) markers. Here, we adapt basic analytical insights from Bohnemeyer (1998) and Bohnemeyer (2002) and postulate that AM markers are in fact the predicates of those sentences. Example (3) would then be more literally translated as ‘my reading of the newspaper is/was/will be near,’ as *ta'itak* is the predicate. This allows for a simpler analysis and preserves coherence of VOS word order throughout the language.

¹Glossing abbreviations: 1: first person, 2: second person, 3: third person, A: set A, APP: applicative voice, B: set B, CPL: completive status, D2: a deictic clitic, D3: a deictic clitic, DET: determiner, IMP: imperfective AM marker, INC: incomplete status, NEG: negation, OBL: obligative AM marker, ONGL: onglide, PL: plural, PROG: progressive AM marker, PROX: proximal AM marker, PRV: perfective AM marker, REC: recent past AM marker, SBJ: subjunctive status, SG: singular, TERM: terminative AM marker, TOP: topical.

- (3) Ta'itak in xok-ik-Ø le periyòodiko-o'.
 PROX A.1.SG read-INC-B.3.SG DET newspaper-D2
 'I /have/had/will have/ almost read the newspaper,'
 'I /am/was/will be/ about to read the newspaper.' (taken from Bohmeyer 2002)

2 Focus Constructions

YM sentences often contain a syntactically focused constituent that appears to the left of the other material in the sentence. The focused constituent is gapped in the remainder of the sentence. For example, in (4), *Maruch* is focused, creating a sentence identical in truth-conditional meaning to (5), but very different in emphasis. Focus constructions are also possible in copular sentences, where the predicate is an adjective or a predicative noun instead of an AM marker. For example, in (6), the argument of the predicative noun *Maruchech*, *tèech*, is focused. The difference between the two is, again, purely pragmatic.

- | | |
|---|---|
| <p>(4) Maruch sáam u yil-Ø-Ø Jwàan.
 Maria REC A.3.SG see-SUBJ-B.3.SG Juan
 'Juan saw MARIA,' 'It's Maria that Juan saw.'</p> <p>(5) Sáam u yil-Ø-Ø Maruch Jwàan.
 REC A.3.SG see-SUBJ-B.3.SG Maria Juan
 'Juan saw Maria.'</p> | <p>(6) Tèech Maruch-ech.
 you Maria-B.2.SG
 'YOU are Maria.'</p> <p>(7) Maruch-ech (tèech).
 Maria-B.2.SG you
 'You are Maria.'</p> |
|---|---|

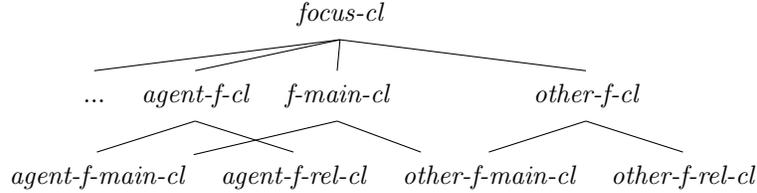
Bohmeyer (2002) argues that the focused constituent can also be treated as a sort of predicate. Under that analysis, the second translation given for (4) reflects its meaning more literally, although sentences like (4) should not be confused with syntactically distinct Yucatecan clefting. We account for this by identifying the left daughter with the head daughter. The right daughter (everything following the focused element) may be any type of predicate, so long as it has an appropriate gap. The mother is also stipulated to be predicative, despite the fact that the head daughter is not.

What is the proof for the predicativeness of Yucatecan focus constructions, though? There have been several tests proposed in the past, many of which are not unproblematic. Bohmeyer (2002)'s *wáaj* test, for example, fails to recognize that *wáaj* can appear in more environments than first meets the eye. Nonetheless, we identify a new test: just like regular predicates, focused constituents are ungrammatical with definite morphology. We then suggest that the definite article *le* selects for [PRED -] items and thus explain the ungrammaticality of definite focus.

Relative clauses in YM are nearly identical to the focus constructions above. The only difference is that instead of standing alone as sentences, they have the same distribution as NPs. Importantly, this means they may be determined by *le*:

- (8) wakax sáam u yil-Ø-Ø Maruch
 cow REC A.3.SG see-SUBJ-B.3.SG Maria
 'a cow that Maria saw'
- (9) le wakax sáam u yil-Ø-Ø Maruch-o'
 DET cow REC A.3.SG see-SUBJ-B.3.SG Maria-D2
 'the/that cow that Maria saw'

Due to the similarities between focus constructions and relative clauses, we define a multiple inheritance hierarchy where these two constructions differ only in the PRED values of the mother.



We have extended this account to include coverage of agent focus, a distinct but related construction. Notably, our treatment of focus also captures the syntax of *wh*- questions. In YM, these kinds of questions are distinguished from declarative statements only by the presence of a focused element which is pragmatically uninformative, such as a focused *wh*-word.

3 Deictic Clitics

Deictic clitics is a term used traditionally in Mayanist literature to refer to four clause-final morphemes, whose occurrence must always be licensed by a clitic-triggering constituent. Their prototypical purpose is to orient the speaker deictically with regard to a constituent present in the clause. There is at most one clitic per clause, and it always appears in a clause final position.

The four clitics *a'*, *o'*, *e'*, and *i'* are glossed as D1, D2, D3, and D4, respectively. The first two clitics relate to the speech situation and are triggered by the definite article *le*. *A'* occurs if a referent is near the speaker; *o'* occurs otherwise. (They can be likened to English “this” and “that.”) The semantics of *e'* and *i'* are somewhat harder to pin down. For example, *e'* is triggered by *way* “here,” and serves the purpose of a topical marker. Negation and locatives frequently require *i'*. As already mentioned, deictic clitics appear not at the end of the phrase that triggers them, but rather at the end of the clause. In (10), we see the clitic appearing on *nool* even though it has been triggered by the previous constituent.

- (10) Yàan u bon-ik-Ø le naj in nool-o'.
 OBL A.3.SG paint-INC-B.3.SG DET house A.1.SG grandfather-D2
 ‘My grandfather has to paint the (that) house.’

When there are multiple clitic-triggering elements in a clause, the most “highly-ranked” one wins. Speech-situation clitics are the most highly ranked, followed by *e'*, and right after *e'* comes *i'*. Trivially, a clitic always outranks the absence of a clitic. The hierarchy can therefore be stated as follows:

$$a', o' > e' > i' > \emptyset$$

In (11), for example, the clitic *i'* (triggered by *ma'*) is outranked by *e'*.

- (11) Le paax-e' ma' t-u yu'ub-a'al-Ø way-e'
 DET music-D3/TOP NEG IMPF-A.3.SG feel-PASS.INC here-D3
 The music is not heard over here. [7]

All this data is straightforwardly accounted for via a non-maximal construction, from which most maximal constructions inherit.

$$deictic-q-cxt \Rightarrow \left[\begin{array}{l} \text{MTR} \left[\begin{array}{l} \text{ENQ-D } F_{\max}(\boxed{1}, \boxed{2}, \dots, \boxed{n-1}, \boxed{n}) \\ \text{DEQ-D } \boxed{m} \end{array} \right] \\ \text{DTRS} \left\langle \left[\begin{array}{l} \text{ENQ-D } \boxed{1} \\ \text{DEQ-D } \textit{none} \end{array} \right], \dots, \left[\begin{array}{l} \text{ENQ-D } \boxed{n-1} \\ \text{DEQ-D } \textit{none} \end{array} \right], \left[\begin{array}{l} \text{ENQ-D } \boxed{n} \\ \text{DEQ-D } \boxed{m} \end{array} \right] \right\rangle \end{array} \right]$$

The above construction formalizes all the relevant insights of the current section. It introduces two new features: DEQUEUE-DEICTIC (or DEQ-D) and ENQUEUE-DEICTIC (or ENQ-D). ENQ-D contains the information of which clitic is triggered by each word or, in other words, which clitic is “enqueued” by a given word. DEQ-D

contains information about which clitic is actually attached to a given word, or which word “dequeues” it. The ENQ-D of the mother is the highest-ranked clitic of all those which are triggered by the daughters. The hierarchy of clitics is formalized with the F_{\max} function.

Of course, all that is to little avail unless we also specify what the top node looks like. All this machinery is set up to ensure the ENQ-D and the DEQ-D of the top node are identical. Intuitively, all it says is that the highest-ranked clitic which some constituent requires is also the one which the sentence eventually realizes. Thus, we arrive at the sentential node presented on the right.

$$S = \left[\text{SYN} \left[\begin{array}{l} \text{CAT} \quad \left[\text{PRED} \quad + \right] \\ \text{VAL} \quad \langle \rangle \\ \text{GAP} \quad \langle \rangle \\ \text{ENQ-D} \quad \boxed{1} \\ \text{DEQ-D} \quad \boxed{1} \end{array} \right] \right]$$

4 Topicalization

YM sentences are frequently prefaced by one or more topic phrases, which introduce background information and behave idiosyncratically with respect to clitics. Topics may precede any grammatically independent utterance of YM, including AM phrases, sentences consisting of a focus and AM phrase, and predicates. The topicalized constituent is typically gapped from the main clause. It necessarily occurs with the enclitic e' , or with another higher-ranked deictic clitic if one is appropriate. Unlike with focus, there is no requirement for the presence or absence of the determiner *le* in the topic phrase. Topics always appear to the left of a focus if there is one, which is captured through the MRKG feature. Specifically, focus clauses prevent their daughters from being marked *topical*, while the topic construction ensures a topical marking on the mother node. Since relative clauses are focus clauses, this also correctly predicts that topics cannot occur within relative clauses.

- (12) Ten- e' in k'àaba'- e' Maks-Ø.
 I-D3 A.1.SG name-D3 Maks-B.3.SG
 ‘As for me, as for my name, it is Maks.’ (or, a little less awkwardly: ‘My name is Maks.’)
- (13) Le ts'akyah-o' u k'àaba'- e' Pedro-Ø.
 DET doctor-D2 A.3.SG name-D3 Pedro-B.3.SG
 ‘As for that doctor, as for his name, it is Pedro.’ (or: ‘The doctor’s name is Pedro.’)

In the above examples, three of the the four topic-accompanying clitics are e' . Only in the first topic in (13), e' is outranked by o' . Significantly, those clitics appear within a sentence, so the construction responsible for topics cannot inherit from *deictic-queue-ext*. We need a new one, which stipulates the clitic-related mechanisms somewhat differently:

$$\text{topical-cl} \Rightarrow \left[\begin{array}{l} \text{MTR} \quad \left[\begin{array}{l} \text{MRKG} \quad \textit{topical} \\ \text{ENQ-D} \quad \boxed{1} \\ \text{DEQ-D} \quad \boxed{1} \end{array} \right] \\ \text{DTRS} \quad \left\langle \left[\begin{array}{l} \text{ENQ-D} \quad \boxed{3} \\ \text{DEQ-D} \quad \mathbf{F_{\max}(e', \boxed{3})} \end{array} \right], \boxed{4} \left[\begin{array}{l} \text{CAT} \quad \left[\text{PRED} \quad + \right] \\ \text{VAL} \quad \langle \rangle \\ \text{GAP} \quad \langle \boxed{2} \rangle \oplus L \\ \text{MRKG} \quad \textit{mrk} \\ \text{ENQ-D} \quad \boxed{1} \\ \text{DEQ-D} \quad \boxed{1} \end{array} \right] \right\rangle \\ \text{HD-DTR} \quad \boxed{4} \end{array} \right]$$

The most important part of the construction is in boldface. The clitic dequeued by the topic has to be e' or a higher ranked clitic, if one is enqueued by a constituent of the topic phrase. To formalize that, we use the F_{\max} function again. The ENQ-D and DEQ-D of the mother are identified with those of the head.

Selected References

- [1] Scott AnderBois. Focus and un informativity in Yucatec Maya questions. *Natural language semantics*, 20(4):349–390, 2012.
- [2] Grant Armstrong. On Copular Sentences in Yucatec Maya. In *Proceedings of the Conference on Indigenous Languages of Latin America-IV.*, 2009.
- [3] Hans C. Boas and Ivan A. Sag, editors. *Sign Based Construction Grammar*. Center for the Study of Language and Information Publications, 2012.
- [4] Jürgen Bohnemeyer. *Time Relations in Discourse: Evidence from a Comparative Approach to Yucatec Maya*. PhD Dissertation, Center for Language Studies Nijmegen-Tilburg, 1998.
- [5] Jürgen Bohnemeyer. *The Grammar of Time Reference in Yucatec Maya*. Lincom Europa, 2002.
- [6] Jessica Coon. When Ergative = Genitive: Nominals and Split Ergativity. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, pages 99–107, 2008.
- [7] Yucatec Maya example N^o 110. (Available online at <http://valpal.info/languages/yucatec-maya/examples/110>. Accessed on 2016-12-28). In: Lehmann, Christian. *Yucatec Maya Valency Patterns*. In: Hartmann, Iren & Haspelmath, Martin & Taylor, Bradley (eds.). *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013.
- [8] Colette Grinevald and Marc Peake. Ergativity and voice in Mayan: A functional-typological approach. *Ergativity, valency and voice*, pages 15–49, 2012.
- [9] Rodrigo Gutierrez-Bravo and Jorge Monforte y Madera. On the nature of word order in Yucatec Maya. *Information structure in indigenous languages of the Americas*, pages 139–170, 2010.
- [10] Edward Keenan and Bernard Comrie. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1):63–99, 1977.
- [11] Christian Lehmann. *Possession in Yucatec Maya. Second, revised edition*. Universität Erfurt, 2002.
- [12] Philip H. Miller and Ivan A. Sag. *French Clitic Movement without Clitics or Movement*. 1995.
- [13] Elisabeth Norcliffe. *Head Marking in Usage and Grammar: A Study of Variation and Change in Yucatec Maya*. Stanford University, 2009.
- [14] Elisabeth Norcliffe. *Revisiting Agent Focus in Yucatec*. Stanford University, 2009.
- [15] Mike Reape. Domain union and word order variation in German. In J. Nerbonne, K. Netter, and C. J. Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, page 151–197. Stanford University: CSLI Publications, 1994.
- [16] Elisabeth Verhoeven. *Experiential constructions in Yucatec Maya*. John Benjamins Publishing Company, 2007.

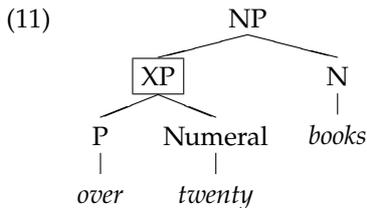
- (9) a. *from ten to fifteen* judges
 b. *in excess of ninety* delegates
 c. *up to twenty* minutes

(Huddleston & Pullum 2002:357)

As one would expect, they can be modified by *somewhere*.

- (10) a. **somewhere** *from 500,000 to 650,000* people
 (http://www.memphismagazine.com/December-2006/The-Return-of-the-Spanish-Lady-pt-II/)
 b. **somewhere** *in excess of 50* scuds (COCA: 1991 SPOK ABC_Nightline)
 c. **somewhere** *up to 100,000* people
 (http://www.huffingtonpost.ca/glen-pearson/sudan-indepencece_b_873072.html)

The above points indicate that a PNC has something like the following structure.



The following example shows that this is a right analysis.

- (12) [over thirty] but [not more than forty] students (Bob Borsley, p.c., March 2017)

What is the prenominal constituent (XP in (11)), then? A possible analysis might be that it is a PP, composed of a prepositional head and a numeral as its complement. In the next section we will see that there are some objections to this analysis.

2 PP analysis of the prenominal phrase

Aarts (2011) states that the prenominal constituent is a PP.

- (13) [NP [PP over twenty] Iranians] (Aarts 2011:119)

Corver & Zwarts (2006) also argue that the prenominal constituent is a PP. They claim that the N and the prenominal phrase are merged inside the NP and make a small clause. The prepositional numeral then moves up to Spec NumP for checking its cardinality feature with the Num head.

- (14) [NumP [PP around 20]_i [Num' NUM [NP children t_i]]] (Corver & Zwarts 2006:828)

However, the PP analysis of the prenominal element is not without problems. First, unlike a normal PP, it is in the prenominal position.

- (15) a. *[on the desk] books⁴
 b. books [on the desk]
- (16) a. [over thirty] books
 b. *books [over thirty]

The second problem is related to the following generalisation: modifiers with complements are systematically excluded from the prenominal position in English (Sadler & Arnold 1994:190).

- (17) a. a child [grateful [for the present]]
 b. *a [grateful [for the present]] child (Sadler & Arnold 1994:189)

The PP analysis would pose a serious challenge for the above generalisation: a PP contains a complement and should be excluded from the prenominal position.

It seems, then, that the PP analysis of the prenominal element of PNCs is unsatisfactory.

3 The prenominal phrase is a numeral

In this section we will see some pieces of evidence that the prenominal phrase of PNCs is headed by the numeral, not the preposition. First, it occurs in the prenominal position like normal numerals.

⁴The italicised phrases in the following examples are PPs, but we follow Sadler & Arnold (1994:189) in assuming that they are the result of some word formation process.

- (i) an *on board* entertainment console
 (ii) an *up-to-the-minute* new report

(Sadler & Arnold 1994:189)

- (18) a. [thirty] books
b. *books [thirty]

- (19) a. [over thirty] books
b. *books [over thirty]

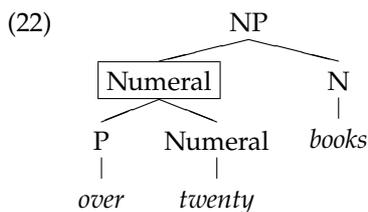
Second, it can appear in the position which is typically filled by a numeral.

- (20) a. an amazing [over fifty] years
b. a negligible [under ten] people
c. an estimated [around 10,000] students
d. an additional [about 100] jobs

Third, the prenominal phrase involving *one* can function as a determiner, like the numeral *one*.

- (21) a. *(one) year
b. *(around one) year

The above pieces of data show that the prenominal phrase of a PNC behaves like a numeral. This means that the numeral heads the prenominal phrase. What we want is roughly structures like (22).



4 HPSG Analyses

It is important to note that only a limited variety of prepositions can appear in PNCs. With their spatial meaning, the pairs of prepositions in (23) are almost interchangeable.

- (23) a. The water came up *above/over* our knees. (Swan 2005:3)
b. I'd like to travel *around/round* the world. (ibid.:50)
c. Look in the cupboard *below/under* the sink. (ibid.:85)

However, only one of each pair is available in PNCs.

- (24) a. She had *over/*above* thirty pairs of shoes. (Sinclair 2004:5)
b. He owns *around/*round* 200 acres. (ibid.:39)
c. There were *under/*below* twenty people at the lecture. (Swan 2005:86)

We need a framework which provides representations detailed enough to differentiate *over*, *around* and *under* from *above*, *round* and *below*, respectively, and to capture the idiosyncratic properties of the former type of prepositions. HPSG is such a framework.

The lexical description of a normal preposition which takes a noun as its complement is something like the following.

- (25)
$$\left[\begin{array}{l} \text{HEAD } \textit{preposition} \\ \text{COMPS } \langle \left[\text{HEAD } \textit{noun} \right] \rangle \end{array} \right]$$

(25) says that normal prepositions take a noun as their complement. It is clear that the prepositions in PNCs have quite different properties from those of normal prepositions. They do not form a prepositional phrase with the following numeral. Rather, the numeral functions as a head and the phrase behaves as a numeral.

In the rest of this section we will look at three possible HPSG analyses of the prepositions in English PNCs. The first and second analyses appear to be unsatisfactory, but the third seems to give a satisfactory account of the facts.

4.1 Weak head analyses

Przepiórkowski (2013) gives an analysis of Polish constructions which are very similar to English PNCs.

- (26) W pokojach będą po dwa fotele. [Polish]
in rooms be-FUT.PL DISTR TWO-NOM.PL armchair-NOM.PL
'There will be two armchairs in each room.'
(Przepiórkowski 2013:166)

The word *po* is a preposition encoding distance distributivity: it ‘attaches to the noun phrase denoting the distributed quantity and looks elsewhere in the sentence for the set to distribute over’ (Przepiórkowski 2013:162). He claims that *po* in (26) is a weak head (Tseng 2002, Abeillé et al. 2006) taking *dwa fotele* ‘two armchairs’ as its complement. A weak head inherits most of syntactic and semantic properties of its complement and those properties are passed on to the phrasal level. This propagation of information from non-heads to phrases can account for the fact that the prepositional phrase can act as a nominative noun phrase and induces number agreement with the verb: *po* preserves the grammatical case and number of the complement on the phrase.

It seems that Przepiórkowski’s (2013) weak head analysis works for the Polish data, but the examples that we observed in section 1 pose a problem for analysing English PNCs along these lines. In that section we claimed that the PNC is an NP headed by the noun which is preceded by a combination of the preposition and the numeral, as described in (22). In the weak head analysis the PNC is an NP, but it is headed by a (weak head) preposition which takes a combination of the numeral and the noun as its complement. This analysis is incompatible with our conclusion about the constituent structure of PNCs (22).

One might argue for an analysis in which the weak head preposition takes a numeral as its complement and the resulting phrase combines with the head noun. This analysis would produce structures like (22). However, there is an objection to this analysis. As discussed in section 2 modifiers with complements are systematically excluded from the prenominal position in English. This analysis is contrary to this generalisation.

We conclude, then, that the approaches employing weak heads are unsatisfactory.

4.2 Functor analysis

We will turn to an analysis which we think provides a satisfactory analysis of the data. In this analysis prenominal elements, such as adjectives and determiners, are uniformly treated as ‘functors’ (Van Eynde 2006, 2007, Allegranza 1998), which select a head. The combination of the functor and its head (called ‘head-functor phrase’) is subject to the following constraint (Van Eynde 2006:164).

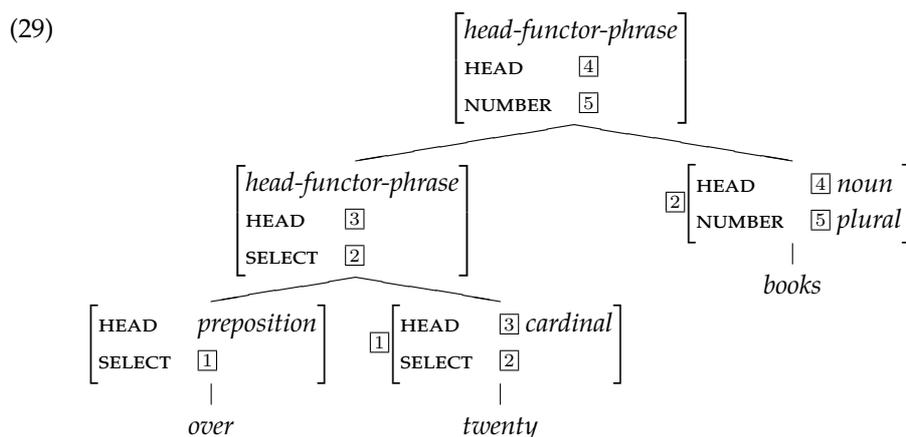
$$(27) \text{ head-functor-phrase} \rightarrow \left[\begin{array}{l} \text{DAUGHTERS} \quad \langle [\text{SELECT } \boxed{1}], \boxed{2} [\text{SYNSEM } \boxed{1}] \rangle \\ \text{HEAD-DAUGHTER } \boxed{2} \end{array} \right]$$

Constraint (27) states that in a phrase of type *head-functor-phrase* the non-head daughter selects a head daughter. With these assumptions the lexical description of prepositions in PNCs is something like the following.

$$(28) \left[\begin{array}{l} \text{HEAD } \textit{preposition} \\ \text{SELECT } \left[\begin{array}{l} \text{HEAD } \textit{cardinal} \end{array} \right] \end{array} \right]$$

They select a cardinal numeral. It is a sort of functor, which selects a head.

The internal structure of PNCs can be analysed as in (29).



The preposition *over* combines with *twenty* to form a head-functor phrase, utilising the SELECT specification (1). The head daughter’s SELECT value is propagated to the mother node (2). The phrase *over twenty* combines with the head noun *books* to form another head-functor phrase, utilizing the SELECT value (2) inherited from *twenty*. The head daughter’s HEAD and NUMBER values are the same as that of the mother node (3, 4 and 5).

The functor analysis can handle the problems we noted with the previous analyses in section 2: the combinations of the preposition and the numeral are not PPs but phrases headed by the numeral so it is natural that they occur in the prenominal position in the same way as bare numerals; the numeral in the PNCs is not a complement of the preposition so it does not contradict the generalisation that prenominal modifiers do not take a complement; and we can make a clear distinction between the prepositions which can occur in PNCs and those which cannot.

Moreover, the functor analysis is more satisfactory than the weak head analyses because it can accommodate all the data observed in section 1 and does not contradict the generalisation that prenominal modifiers do not take a complement.

5 Conclusion

We provided a detailed description of English PNCs and especially of the prepositions employed in the constructions. We then considered how PNCs should be analysed within the framework of HPSG. We looked at three different analyses: two in terms of weak heads and one in terms of functor daughter, and showed that the functor analysis provides a satisfactory account of the data. We employed only existing and independently motivated theoretical apparatus.

References

- Aarts, B. (2011). *Oxford Modern English Grammar*. Oxford: Oxford University Press.
- Abeillé, A., Bonami, O., Godard, D. & Tseng, J. (2006). The syntax of French *à* and *de*: An HPSG analysis. In Saint-Dizier, P. (ed.), *Syntax and Semantics of Prepositions*, Springer Verlag. Number 29 in Text, Speech and Language Technology, 147–162.
- Allegranza, V. (1998). Determiners as functors: NP structure in Italian. In Balari, S. & Dini, L. (eds.), *Romance in HPSG*, Stanford: CSLI Publications. 55–108.
- Corver, N. & Zwarts, J. (2006). Prepositional numerals. *Lingua* **116**. 811–835.
- Davies, M. (2004–). BYU-BNC (Based on the British National Corpus from Oxford University Press). <http://corpus.byu.edu/bnc/>.
- Davies, M. (2008–). The Corpus of Contemporary American English: 450 Million Words, 1990–Present. <http://corpus.byu.edu/coca/>.
- Huddleston, R. & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Kayne, R. S. (2010). *Several, few and many*. In *Comparisons and Contrasts*, Oxford University Press. 29–56.
- Nouwen, R. (2010). Two kinds of modified numerals. *Semantics & Pragmatics* **3**. 1–41.
- Pollard, C. J. & Sag, I. A. (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Przepiórkowski, A. (2013). The syntax of distance distributivity in Polish: Preserving generalisations with weak heads. In Müller, S. (ed.), *Proceedings of the 20th international conference on head-driven phrase structure grammar, freie universität berlin*. 161–181. <http://csli-publications.stanford.edu/HPSG/2013/przepiorkowski.pdf>.
- Sadler, L. & Arnold, D. J. (1994). Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* **30**. 187–226.
- Sinclair, J. (ed.) (2004). *Collins COBUILD English Usage*. London: HarperCollins.
- Swan, M. (2005). *Practical English Usage: Third Edition*. Oxford: Oxford University Press.
- Tseng, J. L. (2002). Remarks on marking. In Van Eynde, F., Hellan, L. & Beermann, D. (eds.), *Proceedings of the 8th International Conference on Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications. 267–283. <http://csli-publications.stanford.edu/HPSG/2/>.
- Van Eynde, F. (2006). NP-internal agreement and the structure of the noun phrase. *Journal of Linguistics* **42**. 139–186.
- Van Eynde, F. (2007). The big mess construction. In Müller, S. (ed.), *The Proceedings of the 14th International Conference on Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications. 415–433. <http://csli-publications.stanford.edu/HPSG/8/>.

Agreement and interpretation of binominals in French

Aixiu An, Anne Abeillé

Laboratoire de Linguistique Formelle
University of Paris Diderot –Paris 7

aixiu0806@gmail.com

abeille@linguist.univ-paris-diderot.fr

Abstract

This paper investigates the agreement of coordinated binominals in French in the form *Det N1 et N2*. We provide corpus data and experimental data to show that different agreement strategies exist in French and that they challenge previous syntactic analyses of binominals. We then propose an HPSG analysis to account for French binominals.¹

1 Introduction

Binomial coordinations such as *Det N1 and N2* have raised many discussions (Heycock & Zamparelli 2005; Le Bruyn & de Swart 2014; King and Dalrymple 2004, Arnold et al. 2006). Semantically, binominals can have two distinct readings: a joint reading (1a, colleague and friend are co-referent) and a split reading (1b, with two distinct individuals: boy and girl).

- (1) a. This friend and colleague is a nice person.
b. This boy and girl are coming tonight.

The split binominals are an instance of natural coordination (e.g., Haiman 1983, Wälchli 2005...), in which the coordinated parts express semantically closely associated concepts. Not all bare coordinations are equally felicitous (2a, b).

- (2) a. this boy and girl
b. *this boy and cat

Syntactically, two structures have been proposed for split reading, either [*Det [N1 and N2]*] (King and Dalrymple 2004) or [*Det N1 and N2*] (Le Bruyn and de Swart 2014).

As for determiner agreement, languages may have different strategies: a singular determiner requires the conjuncts to be singular in English, and conjuncts with different number cannot be coordinated (3a) (King and Dalrymple 2004),

whereas Spanish exhibits closest conjunct agreement (3b) (Demonte et al. 2012)

- (3) a. This boy and girl.
*This boy and girls
* These boys and girl
b. [{E1/*Los} abdomen y
the.M.SG/M.PL abdomen.M.SG and
pecho] aparecen relativamente
chest.M.SG appear.PRS.3PL relatively
abultados.
swollen
'The abdomen and chest look relatively swollen.'

This paper will make two points: on the one hand, we provide new data for French binominals, which challenge Le Bruyn and de Swart (2014)' syntactic analysis; on the other hand, we propose a preliminary HPSG analysis for French binominals.

2 Empirical evidence of binominal agreement in French

In French, for singular joint reading, like in many languages, only the singular determiner is allowed (4). For split reading, Heycock & Zamparelli (H&Z)(2005) and Le Bruyn & de Swart (B&S)(2014) assume that French is an exception, as singular nouns are infelicitous (5a), while only plurals are allowed (5b):

- (4) Le/*les collègue et ami
the.SG/PL colleague.SG and friend.SG
de Jean est venu hier
of Jean PAST come.SG yesterday.
'The colleague and friend of Jean came yesterday.'
- (5) a. *Ce/*Ces marin et
this.MSG/PL sailor.MSG and
soldat sont souvent ensemble.
soldier.MSG be.PL often together
b. Ces marins et soldats
this.PL sailor.MPL and soldier.MPL
sont souvent ensemble
be.PL often together.
'These sailors and soldiers are often together'

¹ This work was supported by Labex Empirical Foundations of Linguistics (ANR-10-LABX-0083) and China Scholarship Council. We would like to thank Barbara Hemforth, Olivier Bonami and engineers: Johan Ferguth, Doriane Gras and Loïc Liégeois. for their help. We also thank the audiences of 4th European HPSG workshop.

The examples in H&Z (2005) and B&S (2014) only consist of animate nouns. We will present a corpus study (corpus frWac) and an experiment of acceptability judgments challenging these data for singular nouns. For plural nouns, we assume that both joint and split readings are allowed: example (5b) is ambiguous between joint and split reading.

2.1 Corpus data

In frWac (1.6 billion words, Baroni et al. 2009), we found 371.000 tokens (96612 types) for the construction *Det N1 et N2*. We annotate the number of Det and Nouns with *Flemm* (Namer 2000). There are 51711 tokens (31412 types) for *Dsg N1sg et N2sg* with either joint reading (6a for animate nouns and 6b for inanimate nouns) or split reading (7), 5137 tokens (1308 types) for *Dpl N1sg et N2sg* with only split reading (8).

- (6) a. Le chanteur et poète québécois
 the.SG singer.SG and poet.SG Quebec
 Gilles Vigneault publie en France
 Gilles Vigneault publish.3SG in France
 un livre d' entretiens
 a.MSG book of interviews
 'The singer and poet of Quebec, Gilles Vigneault, publishes a book of interviews in France'
- b. Le restaurant et bar Starlight
 the.SG restaurant.SG and bar.SG Starlight
 propose un menu international.
 offer.3SG a menu international
 'The restaurant and bar, Starlight, offers an international menu.'
- (7) Présentez-vous à la date et lieu
 introduce you at the.FSG date.FSG and place.MSG
 indiqué pour suivre votre formation
 indicated.MSG to follow your training.
 'Introduce yourself at the date and place indicated to follow your training.'
- (8) Les lieu et programme
 the.PL place.MSG and program.MSG
 seront précisés sur le bulletin.
 be.FUT.PL specified.PL on the bulletin
 'The places and programs will be specified on the bulletin'

We extracted the binominals with more than five occurrences and removed the errors. We annotated noun animacy with an external dictionary (Bonami pc.) and the joint or split reading manually.

The results (fig. 1) show that for the joint reading, only the Dsg is allowed, whereas both Dsg and Dpl are allowed for the split reading: 3084 tokens (60 types) for Dpl, 7563 tokens (456 types) for Dsg.

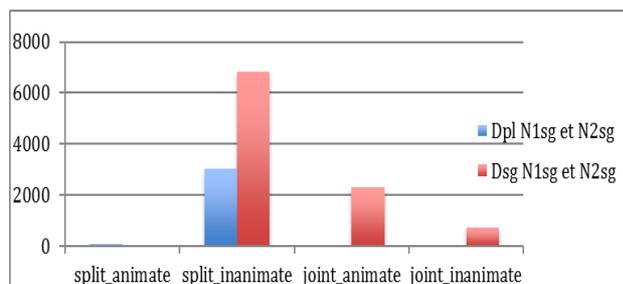


Figure 1. Animacy and readings in frWac binominals

Furthermore, there is an interaction with animacy: the joint reading is more frequent with animate than inanimate nouns. For the split reading, there is also an interaction between Det agreement and animacy: for split animate binominals, Dpl is preferred in a two-tailed binomial test ($p < .001$), whereas Dsg is preferred ($p < .001$) for split inanimate binominals.

Our hypothesis is that Dsg is dispreferred for split animates in order to avoid the joint reading.

2.2 Acceptability judgment experiment

We designed an experiment of acceptability judgment to test the acceptability of Dsg/Dpl for split singular binominals. We had 12 singular animate binominals (9) and 12 singular inanimate binominals (10), as well as 6 control items (11). These items were inspired by corpus data. Participants could only see one D number for each binomial, the number of which was counterbalanced across participants. The binominals are in subject position and the predicate is plural and collective, in order to force the split reading:

- (9) Le/Les directeur et
 the.SG/PL director.SG and
 sous-directeur du secteur se
 under-director.SG of.MSG sector REFL
 sont mis en accord sur le projet
 PAST put in agreement on the project
 'The director and vice-director of the sector agreed on the project.'
- (10) Il arrive souvent que votre/vos
 it happens often that your.SG/PL
 identifiant et mot de passe ne
 username.SG and password.SG NEG
 soient pas reconnus par le site.
 be_SBJV.PL NEG recognized.PL by the site
 'It often happens that your username and password are not recognized by the site.'
- (11) La tête dans le/les genoux,
 the.FSG head.FSG in the.MSG/PL knee.PL,
 je dormirais peut-être deux heures.
 I sleep.COND.1SG perhaps two hour.PL

‘With the head in the knee, I would sleep perhaps two hours.’

We had 42 participants, all French native speakers, recruited on the website RISC. The results (Fig.2) shows that the acceptability of experimental items slightly lower than good controls (without coordination) in green, but much higher than bad controls (in yellow)

We analyze the data using a mixed-effect linear regression model. Our dependent variable is participant’s acceptability judgment (range from 1-10), independent predictors are Noun animacy and Det number. We find significant effects for both animacy and Det number and there is no interaction between these factors. Dpl is better judged than Dsg both for animate nouns and inanimate nouns. Meanwhile, we also find that animacy plays a role, since inanimate nouns are more acceptable than animate nouns.

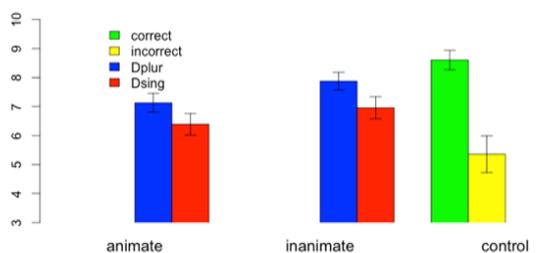


Figure2 Acceptability judgement for D Nsg et Nsg for split reading

3 Syntactic structures of binominals in French

Le Bruyn and de Swart (2014) propose two different syntactic structures depending on the meaning. For joint reading, Det lives in a position above the joint Coord phrase (12a). For split reading, Det combines with the first conjunct only (12b), predicting thus ungrammaticality of Dpl when followed by two coordinated Nsg, as in English (13a), Spanish (13b) and supposedly for French (13c).

- (12) a. joint reading: [DP D [CoordP NP and NP]]
 b. split reading: [CoordP [DP D NP] and NP]

- (13) a. *These boy and girl are eating a pizza
 b. *Los abdomen y pecho
 the.M.PL abdomen.MSG and chest.MSG
 c. *Les homme et femme sont venus
 the.PLman andwoman PAST come.PL
 ‘The man and woman have come.’ (B&S 2014)

According to B&S (2014), (12b) is also compatible with bare binomials N1 et N2 (Roodenburg 2005), which only have the split reading (*Nom et prénom doivent être écrits en noir*. ‘Name and first name must-pl be written in black’)

We agree with B&S’s syntactic structure for joint reading. However, for split reading, we argue that there is a Det above coordinated nouns for split reading as well. On the one hand, the data presented above show that Dpl is acceptable in French, as long as the two N form a natural pair (14a). Moreover, plural numerals may be used: example (14b) refers to a pair, one brother and one sister.

- (14) a. Les mari et femme sont
 the.PL husband.SG and wife.SG be.PL
 d’accord sur le partage des biens.
 of agreement on the division of.PL property.PL
 ‘The husband and wife agreed on the division of these property.’

- b. [Mes deux [frère et sœur]]
 my.PL two brother.SG and sister.SG

Our data show that Dsg is also possible for split reading in French, at least with inanimates (524 tokens for *vos nom et prénom* (‘your.PL name.SG et surname.SG’), 383 for *votre nom et prénom* (‘your.SG name.SG et surname.SG’)). When there is a mismatch of number, the determiner agrees with the closest conjunct.

- (15) a. La plupart de nos établissements sont
 the most of our facilities be.PL
 ouverts tous les jours y compris
 open.PL all the.PL days including
 le dimanche et jours fériés
 the.SG Sunday.SG and day.PL holiday.PL
 ‘Most of our facilities are open every day including Sunday and public holidays.’ (frWac)

- b. Chacun essaye de trouver sa place
 everyone tries to find his place.SG
 en fonction de ses dons et
 in accordance of his.PL gift.PL and
 charisme.
 charisma.SG

‘Everyone tries to find his place according to his gifts and charisma.’ (frWac)

We thus assume that Det agreement for split binominals does not involve an abstract structure but a fairly superficial structure (Borsley 2009), where the Det number is influenced by the adjacent conjunct.

We assume that Dsing is disfavored for split singular animates since Dsing favors joint reading, and most animate nouns are ambiguous between a referential use (*Un collègue est venu*. ‘A colleague has come’) and a predicative use (*Jean est un collègue*. ‘Jean is a colleague’)

4 HPSG analysis

4.1 Previous work

As is generally accepted in HPSG, we use two distinct agreement features, CONCORD for morphosyntactic agreement and INDEX for semantic agreement (Pollard & Sag 1994, Wechsler & Zlatic 2000). Nouns, determiners, and attributive adjectives carry a CONCORD feature, closely related to inflection. INDEX agreement is more semantic, whose value is related to the referential/semantic possibilities of the associated nominal.

King & Dalrymple (2004) propose an LFG analysis where CONCORD features are distributive. The conjuncts require the Det to have as the same CONCORD value as the conjuncts. INDEX features are non-distributive, representing the set formed by the coordinating structure and triggering verb agreement.

Arnold et al. (2006) show that in Portuguese, the determiner always agrees in gender with the first conjunct, and in number either with the first conjunct (16a) or with the coordinated structure (16b).

- (16)a. No on.the.M.SG povo population.M.SG gente hebreia and e

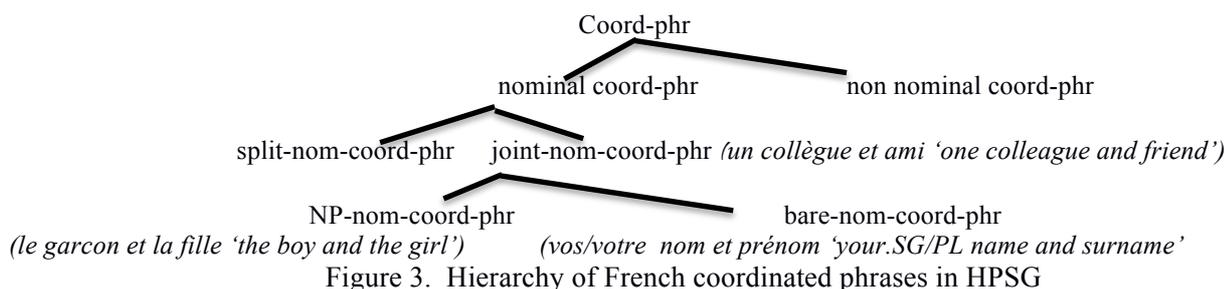
people.F.SG Hebrew.F.SG (Arnold et al. 2006)

- b. Os provaveis director the.MPL probable.PL director.MSG e ator principal and actor.MSG principal.MSG ‘the likely director and main actor’

In addition to CONCORD and INDEX, Arnold et al. (2006) propose two new features: L-AGR for the leftmost conjunct, R-AGR for the rightmost conjunct. In closest conjunct agreement, Det agrees with the first N via L-ARG. L-ARG and R-ARG are head features; the value of L-ARG of the coordination structure comes from the L-AGR of the leftmost daughter, while the CONCORD value reflects the resolved agreement features of the coordinate structure, with identical values of INDEX.

4.2 Our analysis

To deal with the agreement of binominals, we propose a hierarchy of nominal-coordinate-phrases (Fig. 3)



We follow Borsley (2005) who argues that coordinated phrases are unheaded. Disregarding conjunction features, SLASH features are shared between the conjuncts and the coordinate phrase (Mouret 2007) and VALENCE features are shared by default (\setminus) (17).

- (17) Coord-phr =>

$$\left[\begin{array}{l} \text{VAL } \setminus [2] \\ \text{SLASH } [3] \\ \text{DTRS } \langle \left[\begin{array}{l} \text{VAL } \setminus [2] \\ \text{SLASH } [3] \end{array} \right], \dots, \left[\begin{array}{l} \text{VAL } \setminus [2] \\ \text{SLASH } [3] \end{array} \right] \rangle \end{array} \right]$$

In this paper, we only deal with binominals whose head is a noun.

- (18) nom-coord-phr => [HEAD noun]

For joint nominal coordinated phrases, CONCORD and INDEX features are also shared (19):

- (19) joint-nom-coord-phr =>

$$\left[\begin{array}{l} \text{HEAD } [\text{CONCORD } [1]] \\ \text{INDEX } i \\ \text{DTRS } \text{list } \left(\left[\begin{array}{l} \text{CONCORD } [1] \\ \text{INDEX } i \end{array} \right] \right) \end{array} \right]$$

For split nominal coordination phrases, the coordinated phrase has a different INDEX value than the conjuncts and a plural CONCORD value (20).

- (20) split-nom-coord-phr =>

$$\left[\begin{array}{l} \text{HEAD } [\text{CONCORD } [\text{NUM } pl]] \\ \text{INDEX } k = i + \dots n \\ \text{DTRS } \langle [\text{INDEX } i], \dots, [\text{INDEX } n] \rangle \end{array} \right]$$

For NP coordination (*le frère et la soeur* ‘the brother and the sister’), the VALENCE of the daughters is saturated and the number is plural.

We then consider bare binominals, which can only have a split reading (Roodenburg 2005).

- (21) a. Nom et prénom doivent être

name.SG and surname.SG must be écrits en noir.
written.PL in black
'Name and surname must be written in black.'

b. Votre/Vos nom et prénom
your.SG/PL name.SG and surname.SG
doivent être écrits en noir.
must be written.PL in black
'Your name and surname must be written in black.'

For them, we assume the SPR value to be optional. We propose that a split bare coordinated phrase does not necessarily share its CONCORD feature with the conjuncts. It can either inherit the CONCORD value of the first conjunct, expecting a singular Det if it is singular, a plural Det if it is plural, or have a resolved CONCORD value, expecting a plural Det (22).

(22) bare-nom coord-phr=>

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{CONCORD [NUM [0]]} \\ \text{LAGR [NUM [1]]} \\ \text{RAGR [NUM [2]]} \end{array} \right] \\ \text{VAL} \left[\text{SPR} < (D[\text{NUM [0]} \vee [1]]) > \right] \\ \text{DTRS} \left\langle \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{CONCORD NUM [1]} \\ \text{LAGR NUM [1]} \end{array} \right] \right], \dots, \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{CONCORD NUM [2]} \\ \text{LAGR NUM [2]} \end{array} \right] \right] \right\rangle \\ \text{VAL} \left[\text{SPR} < D[\text{NUM [1]}] > \right], \dots, \left[\text{VAL} \left[\text{SPR} < D[\text{NUM [2]}] > \right] \right] \end{array} \right]$$

Selected references

Arnold, D, L Sadler, A Villavicencio. 2005. *An HPSG account of closest conjunct agreement in NP coordination in Portuguese*. In 12th International Conference on Head-driven Phrase Structure Grammar, ed. Stefan Müller, 427–447. Stanford, CA: CSLI Publications.

Borsley, R. 2005. Against ConjP, *Lingua* 115, 461-482

Borsley, R. 2009. *On the superficiality of Welsh agreement*. *NLLT*, 27: 225–265.

de Swart, H, Le Bruyn, B. 2014. *Bare coordination: the semantic shift*. *NLLT* 32 (4): 205-1246

Demonte, V, I Pérez-Jiménez. 2012. *Closest conjunct agreement in Spanish DPs. Syntax and beyond*. *Folia Linguistica* 46(1): 21–74.

Haiman, John. 1983. *Iconic and economic motivation*. *Language* 59(4): 781–819.

5. Conclusion

On the basis of large corpus data, we argue that singular split binominals do exist in French, and both singular and plural determiners are possible. Our experimental data further show that animacy play a role in the acceptability judgments: inanimate binominals are better accepted than animate binominals. The fact that the determiner can agree with the coordinated phrase suggest that the determiner is placed above the coordinated nouns.

We also propose an HPSG analysis with different subtypes for split reading and joint reading. For split reading, following Arnold et al (2006), we use L-ARG feature to capture the closest conjunct agreement fact.

This paper presents on-going work, which leaves a number of questions open. We have identified the role of animacy for the determiner agreement, whether there are other factors should also be explored.

Heycock, C, R Zamparelli. 2005. *Friends and colleagues. Plurality, coordination and the structure of DP*. *Natural Language Semantics* 13: 201–270.

King, T, M Dalrymple. 2004. *Determiner agreement and noun conjunction*. *J of Linguistics* 40 (1): 69–104.

Pollard, C, Sag, I. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press

Roodenburg, J, 2005, *Une coordination particulière : les syntagmes N Conj N en français*. *Langages* 160 : 93-109.

Wälchli, Bernard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press

Wechsler, S, Zlatić L. 2000. *A theory of agreement and its application to Serbo-Croatian*. *Language* 76:799–832.

‘VP’ Adverbs without a VP: The Syntax of Adverbs in Tongan

Douglas Ball
Truman State University

1 Introduction

The work of Jackendoff 1972 introduced a two-way distinction for adverbs: S-adverbs and VP-adverbs. While these terms could be taken as syntactic stand-ins for semantic notions (such as ‘sentence operator’ and ‘predicate operator’, utilized by Thomason and Stalnaker (1973)), they also can be taken as approximations of their syntax (and they are in most analyses of English adverbs; see, e.g., the fairly concrete and non-complex analysis in Potsdam 1998). As the latter, they raise an interesting cross-linguistic question: how do ‘VP-adverbs’ behave in languages ostensibly without VPs, such as verb-initial languages? This paper investigates the matter of ‘VP-Adverbs’ in one such language, the verb-initial language, Tongan (a member of [from largest to smallest] the Austronesian, Oceanic, and Polynesian language families). The data reveal that Tongan does have a regular syntactic location for predicate-centered adverbs: immediately following the predicate. I propose an HPSG account for this location that relies not on the presence of some phrasal structure, but on the presence of a particular level of saturation. After proposing this analysis, I consider its interaction with an unlikely further collection of the data: the syntax of the dependent pronouns.

2 The Basics of the Tongan Clause Structure and Tongan Adverb Location

A basic sentence in Tongan, from which some characteristics of Tongan clauses can be explicated, is given in (1):

- (1) Na‘e tāmata‘i ‘e Tēvita ‘a Kōlaiate.
PST kill.TR ERG David ABS Golaith
‘David killed Golaith.’ (Churchward 1953, 67)

The sentence begins with a word that indicates temporal-aspectual (and sometimes modal) information; I henceforth will call these TAM words. The TAM word is most often followed by the word that denotes the main predication in the sentence, what I will call the predicate. After the predicate come the arguments (and adjuncts) of that predication. As (1) illustrates, the principal indication of predicate-argument relationships is through what are arguably prepositions at the left edge of the argumental expressions. As is further evident from (1), this system of argument flagging is ergatively-aligned. Nevertheless, when the subject (of a predicate of any transitivity) is pronominal in nature (and is realized as what I will call a dependent pronoun), it appears after the temporal-aspectual word; non-subject pronouns are after the predicate, as in (2):

- (2) Na‘á ku manuatu‘i ia.
PST 1SG.SUBJ remember.TR 3SG
‘I remembered him.’ (Churchward 1953, 66)

Following proposals surrounding the structure of the Tongan clause given by Dukes 2001 and Ball 2008, I will assume that the example in (1) has the structure shown in Figure 1 (only essential features included). As Figure 1 indicates, this analysis puts forth that there are two constituents of note within a Tongan clause.

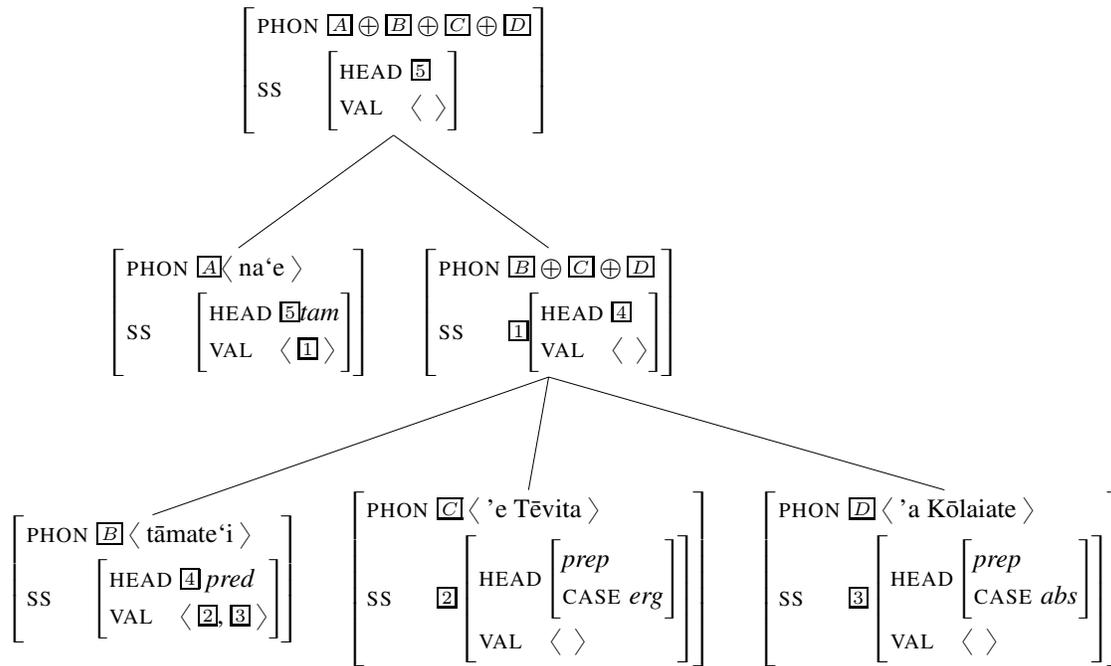


Figure 1: Structure of An Example Tongan Sentence

The lower – a fully saturated phrase headed by the predicate (containing the tag 1 in Figure 1) – is put together by a rule that combines a head and all its valents together (identical, or nearly so, to Schema 3 from Pollard and Sag 1994 and *sai-ph* from Ginzburg and Sag 2000). The higher – a fully saturated phrase headed by the TAM word – is likewise made up of a head and its valent.

As many languages, the broad collection of Tongan adverbs can appear in a number of different locations in the clause. Some adverbs appear between the TAM word and the predicate, as in (3) [adverbs here and in the following examples are italicized]:

- (3) Na'e *toutou* fakama'a 'e Tēvita e faliki.
 PST repeatedly clean ERG (name) ABS.DET floor
 'Tēvita cleaned the floor repeatedly.' (own data)

Others appear clause-finally, as in (4):

- (4) Na'á ne fai eni 'aneafi.
 PST 3SG.SUBJ do PROX.DEM yesterday
 'He did this yesterday' (Churchward 1953, 66)

However, they seem to most frequently appear immediately after the main predicate, as shown in examples (5)–(7):

- (5) Na'e fakama'a *fakalelei* 'e Pita e faliki.
 PST clean well ERG (name) ABS.DET floor
 'Pita cleaned the floor well.' (own data)

(6) Na'e tali *totoka* 'a Mele ki he pasi.
 PST wait calmly ABS (name) to DET bus
 'Mele waited calmly for the bus.' (own data)

(7) Na'e tō *'anefē* 'e Sione 'a e manioke?
 PST plant when.PST ERG (name) ABS DET cassava
 'When did Sione plant some cassava?' (own data)

This location is unsurprising for Tongan; in a different domain, adnominal modifiers in the language likewise follow their nominal heads. In the post-predicate location, adverbs can also stack, as illustrated with *fakalelei* 'well' and *ma'u pē* 'always' in (8):

(8) Na'e fakama'a *fakalelei ma'u pē* 'e Sione 'a e faliki.
 PST clean well always ERG (name) ABS DET floor
 'Sione always cleaned the floor properly.' (own data)

With these two adverbs, the order depicted in (8) is the only possible order: reversing the order to *ma'u pē fakalelei* results in unacceptability. This datum thus suggests that right-to-left order corresponds to the scopal interpretation of the adverbial elements.

Having laid out the basic kinds of patterns in Tongan, let me turn to an account of the 'VP adverbs' – the ones that appear immediately after the predicate.

3 A Sketch of An Analysis of Post-Predicate Adverbs

A central idea of most generative accounts of the syntax of the adverbs is that there is some structure present – be it a VP, an S, a V' , or something else – for the adverb to attach to. Indeed, in textbook X-bar theory, adjuncts, such as adverbs, are sister to a X' . In the case of the post-predicate adverbs of Tongan, this assumption immediately forces the analyst to appeal to some sort of covert structure: minimally, a covert V' above the V (assuming that, in fact, the predicates are actually Vs) to which an adverb could attach. For analysts working in the transformational tradition, further assumptions might lead them to proposals that the order is due to movement of heads around the adverb (in the style of Pollock 1989) or even more elaborate movement-based proposals. Massam (2010) discusses the syntax of adverbs of Tongan's sibling language Niuean and notes how, assuming a fixed order to adverbs (cf. Cinque 1999), one might derive the order found in Niuean (and presumably other Polynesian languages) with so-called 'roll-up movement' (Cinque 2005). Although Massam's paper is not a critique of 'roll-up movement' *per se*, it does offer a picture of some of the problems that the 'roll-up movement' approach brings, even confined to the Minimalist framework.

In contrast to all of this, HPSG, with its more flexible approach to syntactic combinatorics, offers a much more transparent solution. Instead of needing to assume additional covert structure or elaborate phrasal manipulations, an HPSG approach can, in fact, enforce a fairly direct combination of predicates and adverbs. On such an approach, post-predicate adverbs would thus be specified as in (9):

(9)
$$\left[\text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \left[\text{adv} \left[\text{MOD} \left[\text{HEAD} \quad \textit{pred} \right] \left[\text{VAL} \quad \textit{ne-list} \right] \right] \right] \right]$$

The value of the non-empty list (*ne-list*) for the modified predicate interacts crucially with the view that the predicate phrase is composed by a rule that puts together a head and all its valents. The presence of what could be termed the head-all-valents rule results in just two possible kinds of predicate constituent: fully saturated and fully unsaturated. The adverb, with its MOD specification in (9), must combine with a fully unsaturated verb, forcing it to be verb-adjacent. A tree illustrating this result, analyzing example (6) but showing only the predicate phrase of that clause, is shown in Figure 2.

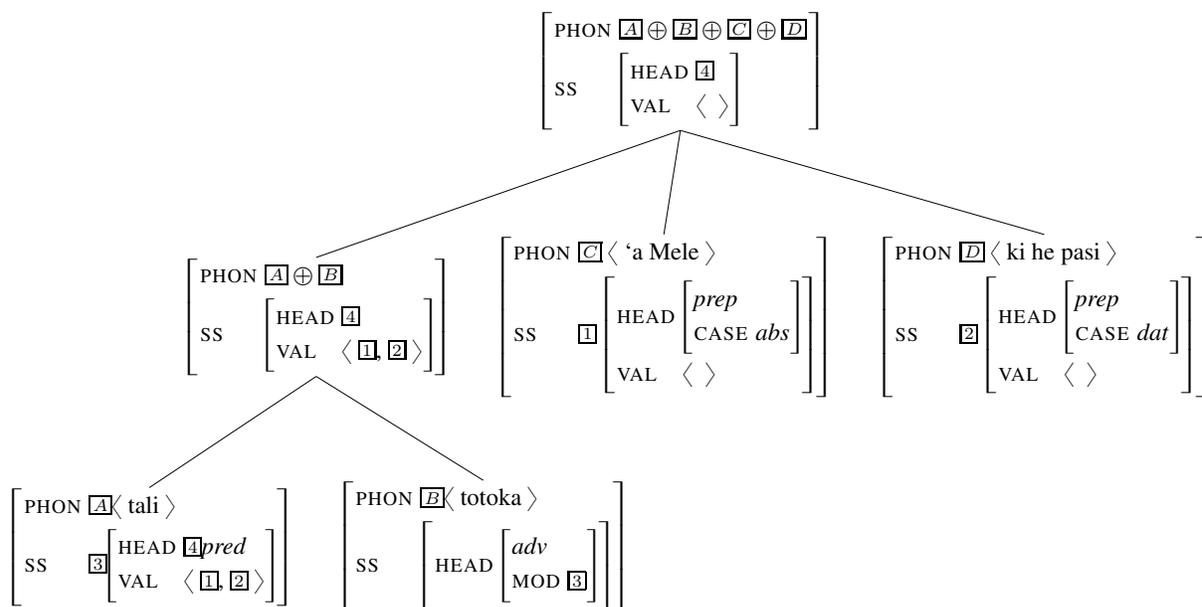


Figure 2: Structure With a Post-Predicate Adverb

4 Consequences and Challenges

The specification in (9) has no difficulty capturing the data (such as (5)–(8)) where an adverb co-occurs with arguments after the predicate. However, what happens when there are no arguments after the predicate, such as when there is an intransitive predicate with a pre-predicate dependent pronoun? Post-predicate adverbs are indeed possible in this case, as the example in (10) shows [the adverb is again italicized]:

- (10) Na'a nau *ō* *leva*.
 PST 3PL.SUBJ go.PL at.once
 'They went at once.' (Churchward 1953, 196)

Examples like (10) pose a potential challenge to the specification in (9). As Tongan independently allows for a fair amount of “zero anaphora”, it seems plausible to assume that a verb like *ō* in (10) is [VAL < >]; that is, all “zero anaphors” – even those that co-refer with pre-predicate dependent pronouns – are ARG-ST list items that do not appear on the corresponding VAL list (this is the analysis provided of this phenomenon in Ball 2008, ch. 3). However, such a proposal would lead to the misprediction that (10) would be impossible, as the specification of [VAL < >] is not consistent with the MOD specification in (9). This, thus, suggests that, to preserve the analysis embodied in (9), a pre-verbal argument would best be treated as a VAL list item on the predicate itself.

The potential issues arising from the interaction of the syntax of adverbials and the syntax of pre-predicate dependent pronouns also raise the question of whether an adjuncts-as-complements approach to the adverbs (see Bouma, Malouf, and Sag 2001; Bonami and Godard 2007; Sato and Tam 2008; among others, for this approach) would be preferable for Tongan adverbs. Preliminarily, while the adjuncts-as-complements approach could be implemented for the Tongan adverbs and would allow the adverbs to be analyzed in a way consistent with previous analyses, it also does not offer a particularly well-motivated analysis of adverb location (the post-predicate location would have to be forced by special rule or linear precedence constraint, while on the MOD-based approach, it falls out from the fact that adverb seek the minimal relevant verbal constituent) and would less straightforwardly handle the scopal properties evident from

(8). So even if an adjuncts-as-complements approach might ultimately be pursued, hopefully these positive features of the MOD-based analysis would be preserved within it.

Overall, though, it is clear that an HPSG approach can readily account for the location of adverbs in Tongan and can do so without recourse to covert structure or covert operations on structure. Thus, this kind of analysis seems to be a firmer foundation for building an even more complete analysis of these elements in Tongan.

References

- Ball, Douglas L. 2008. "Clause Structure and Argument Realization in Tongan." Ph.D. diss., Stanford University.
- Bonami, Olivier, and Danièle Godard. 2007. "Integrating Linguistic Dimensions: The Scope of Adverbs." In *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford Department of Linguistics and CSLI's LinGO Lab*, edited by Stefan Müller, 25–45. Stanford, CA: CSLI Publications.
- Bouma, Gosse, Rob Malouf, and Ivan A. Sag. 2001. "Satisfying Constraints on Extraction and Adjunction." *Natural Language and Linguistic Theory* 19 (1): 1–65.
- Churchward, C. Maxwell. 1953. *Tongan Grammar*. London: Oxford University Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- . 2005. "Deriving Greenberg's Universal 20 and Its Exceptions." *Linguistic Inquiry* 36 (3): 315–332.
- Dukes, Michael. 2001. "The Morphosyntax of 2P Pronouns in Tongan." In *The Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, edited by Dan Flickinger and Andreas Kathol, 63–80. Stanford, CA: CSLI Publications.
- Ginzburg, Jonathan, and Ivan A. Sag. 2000. *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Stanford, CA: CSLI Publications.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Massam, Diane. 2010. "Deriving Inverse Order: The Issue of Arguments." In *Austronesian and Theoretical Linguistics*, edited by Raphael Mercado, Eric Potsdam, and Lisa deMena Travis, Volume 167 of *Linguistik Aktuell/Linguistics Today*, 271–296. Amsterdam/Philadelphia: John Benjamins.
- Pollard, Carl J., and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Pollock, Jean-Yves. 1989. "Verb Movement, Universal Grammar, and the Structure of IP." *Linguistic Inquiry* 20 (3): 365–424.
- Potsdam, Eric. 1998. "A Syntax for Adverbs." *The Proceedings of the Twenty-Seventh Western Conference on Linguistics*. Fresno, CA: Department of Linguistics, California State University-Fresno., 397–411.
- Sato, Yo, and Wai Lok Tam. 2008. "Towards a Unified Account of Adjuncts." Edited by Stefan Müller, *The Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications, 173–188.
- Thomason, Richmond H., and Robert C. Stalnaker. 1973. "A Semantic Theory of Adverbs." *Linguistic Inquiry* 4 (2): 195–220.

Atomistic and holistic exponence in Underspecified Realisational Morphology

BERTHOLD CRYSMANN & OLIVIER BONAMI
Laboratoire de linguistique formelle, CNRS & U Paris-Diderot

A core concern of any theory of inflectional morphology is to capture the fact that the same exponents may be used in different ways in different contexts. Relevant phenomena are both wide-spread and varied. In this paper we shall examine the following two cases: (i) parallel exponence, where the same shapes in different positions realise related but distinct property sets (see Table 2 from Swahili), and (ii) ‘gestalt exponence’ (Blevins et al., in press), where the cooccurrence of two (or more) exponents in a word realises a property that neither realises in isolation (see Table 1 from Estonian).¹ As we shall show, the first case is best conceived in terms of an *atomistic* view of exponence, which establishes correspondences between function and minimal segmentable morphs, whereas the second one is best understood in *holistic* terms, where form and function are established rather at the level of the whole word. Building on previous work in Information-based Morphology (Crysmann and Bonami, 2016), we suggest that both views can be reconciled under a single formal approach to morphology that relies crucially on underspecification in inheritance hierarchies of typed feature structures, and show how this conception improves over other realisational approaches to inflection.

1 Data

1.1 Estonian

Noun declension in Estonian has served as the primary piece of evidence to argue that form-function correspondences are better understood in holistic terms, i.e. established in terms of relations between fully inflected words, rather than in atomistic terms, involving the combination of sub-word units (Blevins, 2005, 2006).

As illustrated in Table 1, morphological marking of number (SG/PL) and core cases (NOM/GEN/PART) clearly provides distinct forms for all six paradigm cells (modulo syncretism between two cell in the SEMINAR class), but the individual devices used to express the distinctions do not align well with the functional distinctions they are supposed to express. On the side of pure exponence, we find several devices: presence vs. absence of an inflection class-specific theme vowel (-a/-u/-i), which segregates the nominative singular from all other forms, suffixation of case/number markers, which is sometimes identical across inflection classes (e.g. NOM.PL -d), and sometimes not (e.g. GEN.PL -d/-t). Similarly, while one might be tempted to further decompose e.g. the genitive plural marker -de/-te there is no constant plural form or corresponding singular form on which this decomposition could be modelled.

¹Other phenomena that exhibit the very same general properties include polyfunctionality, where identical forms express different function (Spencer and Stump, 2013; Ackerman and Bonami, in press), variable placement, where one exponent realising one property set occurs in different linear positions depending on the morphosyntactic context (Stump, 1993; Crysmann and Bonami, 2016); and exuberant exponence, where some property is marked over and over again by the same forms (Harris, 2009; Crysmann, 2014).

NOKK ‘beak’			ÕPIK ‘workbook’		
	SG	PL		SG	PL
NOM	nokk	nokad	NOM	õpik	õpikud
GEN	noka	nokkade	GEN	õpiku	õpikute
PART	nokka	nokkasid	PART	õpikut	õpikuid

SEMINAR ‘seminar’		
	SG	PL
NOM	seminar	seminarid
GEN	seminari	seminaride
PART	seminari	seminarisid

Table 1: Partial paradigms exemplifying three Estonian noun declensions (core cases; Blevins, 2005)

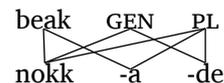


Figure 1: *m:n* relations in Estonian

Finally, the NOKK class displays an alternation between geminated and non-geminated stems, which witnesses an alignment with case that is the exact opposite in the singular and the plural. As summarised by the diagram in Figure 1 for *nokk-a-de* ‘beak.GEN.PL’ no form exclusively expresses a single property, and conversely, no property is exclusively expressed in a single place.

Thus, while individual formal devices can clearly be identified, association with function must be established at a level that involves combinations of forms. For Estonian, a holistic, or constructional (Gurevich, 2006) view appears therefore inevitable.

1.2 Swahili

In contrast to Estonian core cases, which are encoded in a highly opaque fashion, Swahili is much more transparent, thereby being compatible with an atomistic view that associates function more directly with individual exponents that serve to express them. However, if a holistic view can shed light on systems like Estonian that are not fully amenable to an atomistic analysis, one might wonder whether a more elegant model of morphology might not be arrived at by generalising all form/function relations to the level of the morphological word. Such an approach has been pursued, e.g. by Koenig (1999) who has proposed an essentially word-based constructional analysis of (part of) the Swahili position class system.

The phenomenon of parallel exponence, however, resists such a mode of analysis. Swahili verbs can inflect for both subject and object agreement, inserting exponents into different templatic slots (slots 2 and 5, respectively). While choice of morphosyntactic properties, and therefore, forms,

PER	GEN	SUBJECT		OBJECT	
		SG	PL	SG	PL
1		ni	tu	ni	tu
2		u	m	ku	wa
3	M/WA	a	wa	m	wa
	M/MI	u	i	u	i
	KI/VI	ki	vi	ki	vi
	JI/MA	li	ya	li	ya
	N/N	i	zi	i	zi
	U	u	—	u	—
	U/N	u	zi	u	zi
	KU	ku	—	ku	—

Table 2: Swahili person markers (Stump, 1993)

are independent for both functions, it is clear from Table 2 that pairings of form and function draw largely on the same inventories. As a result, an analysis that wants to capture this generalisation must permit the reuse of the same morphological resources for different purposes within the same word, which necessitates reifying correspondences between shapes and partial morphosyntactic description—precisely what a holistic approach sets out not to do.

To summarise our presentation of the basic data, we can conclude that a credible morphological theory must afford ways to accommodate both atomistic and holistic analyses within the same formal system, rather than enforce one view or another.

2 Analysis

In extant inferential-realisation models of morphology, and most prominently in Paradigm Function Morphology, variable use of exponents has typically been addressed by supplementing the core system of ordered rules of exponence by various extensions, including metarules (Stump, 1993, 2001), rules of referral (Stump, 2001; Spencer and Stump, 2013), meta-operators used in the statement of rules and/or paradigm functions (Stump, 2012a,b), rules of paradigm linkage (Spencer and Stump, 2013), and exponence schemata (Stump, forthcoming). We submit that such a ‘divide and conquer’ approach misses the essential unity of the phenomena at hand.

We build on Crysmann and Bonami (2016) and offer an alternative formal conceptualisation of exponence that builds almost exclusively on underspecification techniques fully in the spirit of constraint-based approaches to syntax and semantics. Under this view, every exponent is seen as an association between a shape and a position in the word and introduced by a rule specifying which property set this exponent realises. The crucial difference with respect to a classical realisation approach is that rules of exponence are organised in a monotonic inheritance hierarchy. Hence it is always possible to capture what two exponents have in common by referring to the appropriate super-type in the hierarchy. For example in the case of Swahili person marking, as illustrated in Figure 2, common properties are captured further up the hierarchy, whereas specific instantiations in terms of combination of more abstract rule types are found towards the bottom of the hierarchy.

An interesting feature of the formal device of underspecification is that it is largely agnostic as to the distinction between what Blevins (2006) calls a *constructive* view of morphology, where words are derived from minimal elements, and what he calls an *abstractive* view, where words are taken as prior, and entities such as stems and affixes, to the extent that they are useful analytic devices, are higher-level

abstractions over words.² Nodes in the inheritance hierarchy are nothing more than generalisations on the distribution of recurrent partials, i.e. useful abstractions from surface word-sized Saussurean signs. Because inheritance is monotonic — there are no defaults, unlike what happens in Network Morphology (Brown and Hippisley, 2012) and Construction Morphology (Booij, 2010) —, the hierarchy can be seen both from a top-down point of view, as a way of encoding optimally constraints on exponence, and from a bottom-up point of view, as an explicit representation of relations of similarity and difference between words.

From the constructive point of view, it is important that the full hierarchy can be deduced from a partial hierarchy through the use of online type construction (Koenig, 1999): this means that only those realisation rules that include some constraint not inherited from supertypes need to be explicitly listed, rather than inferred from the shape of the system, by means of systematic intersection of leaf types from each dimension (boxed). Such inferrable types are indicated by dashed lines, as shown in Figure 2. From an abstractive point of view, on the other hand, the leaf types in the hierarchy are ontologically prior, as they constitute the directly observable associations between content and form.

The double interpretation of rule inheritance hierarchies as abstractive vs. constructive views of morphology is crucial to permitting both atomistic and holistic analyses, as we will see below.

The discussion up to now is compatible with various ways of conceptualising realisation rules. In the remainder of the paper we adopt the view, proposed by Crysmann and Bonami (2016), that they express relations between m properties to be expressed and n exponents, which may even be discontinuous. Hence the simple case of a 1:1 relation between a feature and an affix is just a particular case of that general situation. In this talk we use this general framework to highlight differences between inflection systems that help explain why they rather call for atomistic or holistic analyses.

The crucial contrast between the two systems is that rules for Swahili express 1:1 relations between morphs and partial property sets, while rules for Estonian express globally a word-level $m:n$ relation between a sequence of exponents and a property set. We contend that these are necessary features of adequate analyses of these two systems. In the case of Swahili, no word-level constraint can capture the fact that the same affixes play double duty as subject and object markers — hence the word-level analysis proposed by Koenig (1999) is sub-optimal, and abstraction of realisation rules of sub-word relevance is crucial. At the other end of the spectrum, in Estonian, simultaneous introduction of all morphs is the formal rendering of the idea of ‘gestalt exponence’ — words are segmentable, but content is attributed to combinations of morphs rather than individual morphs. Note that adopting such a ‘gestalt’ view in no way precludes identifying generalisations across words where they are relevant. For instance, the fact that plural marking is always manifested at the right edge of the word in the Estonian dataset is captured by a general type linking the expression of plural to position 2 without constraining its shape. In this sense the approach is close in spirit to Berkeley Construction Grammar, where generalisations hold at variable levels of granularity.

In the following two subsections, we shall provide walk-through sample analyses of Swahili, as an example of a nec-

²Blevins introduces the notion of an abstractive approach in the context of the study of the implicative structure of paradigms (Wurzel, 1984), arguing that segmentation is of little help to study that structure. Most work claiming the label ‘abstractive’ pursues the same agenda (e.g. Ackerman and Malouf 2013; Bonami and Beniamine 2016; Sims 2015). We contend however that the idea of an abstractive approach to morphology applies beyond the domain of implicative structure, and is relevant even for the analysis of agglutinative systems where segmentation is not disputed.

essarily atomistic analysis, and, in more detail, Estonian, as a representative that lends itself more to a holistic approach.

2.1 Swahili

The analysis in this subsection essentially rehearses the proposal in Crysmann and Bonami (2016), mainly serving the purpose of contrasting the advantages of an atomistic analysis of this system, compared to the holistic approach required by Estonian.

Crysmann and Bonami (2016) draw a fundamental distinction between hierarchies of rules of exponence, representing recipes, and general principles of morphological completeness and coherence that regulate rule combination. While formally, this approach shares a great degree of similarity with Koenig (1999), in particular reliance on Online Type Construction (Koenig and Jurafsky, 1994), it is the shift of perspective from word-level or lexeme-level hierarchies to hierarchies of morphological rules that establishes its versatility in the domain of inflectional morphology.

By way of illustration, Figure 2 provides a partial description of parallel exponence: types in the **[EXPONENCE]** dimension on the right pair shapes (phonology of morphs) with person, number and gender properties, whereas the two types in the **[MORPHOTACTICS]** dimension specify position class information for subject vs. object agreement. Systematic intersection of leaf types (one each from either dimension) yields the fully expanded set of rules, effectively distributing positional marking of grammatical function over the exponents. In order to derive a morphologically wellformed, fully inflected word, every element of the morphological property set must be realised by some realisation rule, i.e. each member of the property set must be “consumed” by some MUD element of exactly one rule. As a result, rule type hierarchies constitute a repository of recipes that can be referred to more than once, e.g. for subject and object agreement.

2.2 Estonian

Owing to the interdependence of stem alternation, theme marking and number/case suffixation, the hierarchy of Estonian core case rules is certainly more baffling at first sight. We establish a system of three dimensions that derives a set of $m:n$ rules, the leaves of which simultaneously express lexical identity (stem selection), as well as inflection for number and case. From the start, distinct formal devices can be clearly identified, and this is how the hierarchy in Figure 4 is primarily organised: at the top, we find three dimensions corresponding to stem selection (**[STEM]**), selection of the theme vowel (**[THEME]**), as well as selection of suffixal exponents. Since the latter exhibit a major split between the singular and a suffixal plural, we have labeled this dimension **[PLU]**.

Within the **[PLU]** dimension, the first major split is between plural categories and the singular. The former are uniformly expressed by a morphotactic combination of stem, theme vowel and suffix, a property that is abstracted out by way of the MPH feature, specifying three morphs in adjacent position classes, associated with the appropriate number value. Immediate subtypes capture partial generalisations over exponent shapes, paired with case properties. At the level of the leaves, both form and function get fully specified. The singular rules, by contrast, are more heterogeneous morphotactically: the supertype only states there be minimally a stem ([PC 0]). Leaf types capture the difference on morphotactic complexity, ranging from monomorphic, stem-only nominatives, via bi-morphic genitives, being expressed by stem and theme vowel, to tri-morphic partitives, which show a combination of stem, theme vowel and the suffix *-t*.

Within the **[PLU]** dimension, the rule type for nominative singular is special, as it is pre-linked to the **[THEME]** dimension as well. All other rule types are not, so they will be intersected with the only open type in the **[THEME]** dimension, thereby fixing the theme vowel, based on the inflection class of the stem (TV).

Finally, the **[STEM]** dimension on the left determines stem selection from the inventory of lexemically specified stem spaces (Bonami and Boyé, 2006): as detailed in Figure 3, Estonian lexemes minimally specify a stem (ST) and a theme vowel (TV), the latter being fixed per inflection class. In the case of the *nokk* inflection class³ (*n-lid*), lexemes provide an alternative weak stem WK-ST, in addition to the standard (strong) stem (ST). Stem selection rule types in the **[STEM]** dimension of Figure 4 pick out the lexemically specified stem phonology under ST and associate it with the phonology of a morph in position class 0. An alternative rule type, constrained to inflection class *n-lid* accesses the weak stem, with the subtypes extensionally enumerating the two cells in which this stem is licit. All other cells are filled by the default rule, as determined by Pāṇinian closure on the type hierarchy (Crysmann and Bonami, 2016).

Note that both our analysis of Swahili and the one of Estonian build on essentially the same formal devices, namely underspecification and Online Type Construction. The main difference between the holistic analysis we formally develop for Estonian, and the atomistic analysis we suggest for Swahili is the cardinality of MPH (form) and MUD (function), which is 1 for both in Swahili, while it goes up to 3 for Estonian, providing true $m:n$ relations at the fundamental level of realisation rules.

3 Conclusion

Although these two analyses purposefully showcase (sub)systems lying at extreme ends of a gradient of morphological opacity, nothing in the formal setup we assume entails that a system may contain only word-level or only morph-level rules: indeed, outside the domain of core cases, Finno-Ugric Estonian is rather of the agglutinative type. This opens up the possibility of capturing appropriately diverse combinations of opaque and transparent corners of an inflection systems, and hence helps provide a formally sound typological characterisation of exponence systems, rather than assume a ‘one size fits all’ view of morphological modeling that masks diversity. In particular, it is notable that the framework allows for the definition of a classical morpheme — a 1:1 association between a morph and a property set — where it is useful, without forcing its universal adoption, even in the analysis of the same system. In contrast to morpheme-based theories, this 1 : 1 relation does not enjoy any special formal status compared to $m : n$: it just happens to have a very simple and straightforward specification.

The main claim advanced here relates to the unique importance of ontological abstraction over rules: given monotonic type hierarchies, graded differences between opaque and transparent morphological systems can be given a unified analysis, representing the scale of opacity in terms of degrees of informational abstraction and analytical complexity.

References

Ackerman, F. and Bonami, O. (inpress). ‘Systemic polyfunctionality and morphology-syntax interdependencies’. In A. Hippisley

³We label the inflection class after the initial of the examples in Table 1.

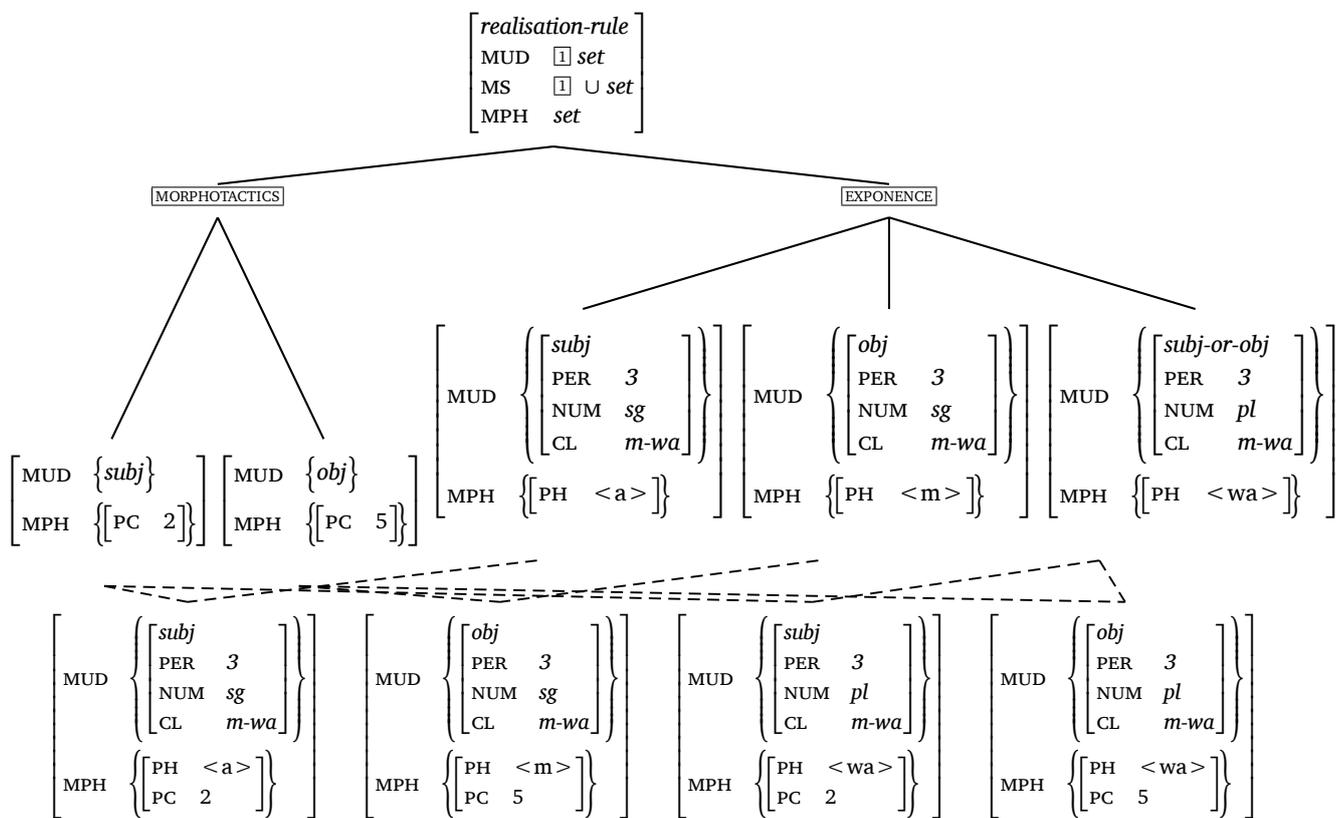


Figure 2: Rule type hierarchy for Swahili parallel position classes (Crysmann and Bonami, 2016)

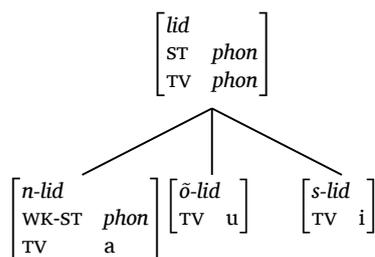


Figure 3: Lexical class type hierarchy for Estonian

and N. Gisborne (eds.), *Defaults in Morphological Theory*. Oxford: Oxford University Press.

Ackerman, F. and Malouf, R. (2013). ‘Morphological organization: the low conditional entropy conjecture’. *Language*, 89:429–464.

Blevins, J. P. (2005). ‘Word-based declensions in Estonian’. In G. E. Booij and J. v. Marle (eds.), *Yearbook of Morphology 2005*. Springer, 1–25.

——— (2006). ‘Word-based morphology’. *Journal of Linguistics*, 42:531–573.

Blevins, J. P., Ackerman, F., and Malouf, R. (in press). ‘Morphology as an adaptive discriminative system’. In H. Harley and D. Siddiqi (eds.), *Morphological Metatheory*. Amsterdam: John Benjamins, 35pp.

Bonami, O. and Beniamine, S. (2016). ‘Joint predictiveness in inflectional paradigms’. *Word Structure*, 9:156–182.

Bonami, O. and Boyé, G. (2006). ‘Deriving inflectional irregularity’. In *Proceedings of the 13th International Conference on HPSG*. Stanford: CSLI Publications, 39–59.

Booij, G. (2010). *Construction morphology*. Oxford: Oxford University Press.

Brown, D. and Hippiisley, A. (2012). *Network Morphology: a defaults based theory of word structure*. Cambridge: Cambridge University Press.

Crysmann, B. (2014). ‘Deconstructing exuberant exponence in Batsbi’. Paper presented at the 16th International Morphology

Meeting, Budapest, May 29–June 1.

Crysmann, B. and Bonami, O. (2016). ‘Variable morphotactics in information-based morphology’. *Journal of Linguistics*, 52:311–374.

Gurevich, O. I. (2006). *Construction Morphology: the Georgian version*. Ph.D. thesis, University of California, Berkeley.

Harris, A. C. (2009). ‘Exuberant exponence in Batsbi’. *Natural Language and Linguistic Theory*, 27:267–303.

Koenig, J.-P. (1999). *Lexical relations*. Stanford: CSLI Publications.

Koenig, J.-P. and Jurafsky, D. (1994). ‘Type underspecification and on-line type construction’. In *Proceedings of WCCFL XIII*. 270–285.

Sims, A. (2015). *Inflectional defectiveness*. Cambridge: Cambridge University Press.

Spencer, A. and Stump, G. T. (2013). ‘Hungarian pronominal case and the dichotomy of content and form in inflectional morphology’. *Natural Language and Linguistic Theory*, 31:1207–1248.

Stump, G. T. (1993). ‘Position classes and morphological theory’. In G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 1992*. Kluwer, 129–180.

——— (2001). *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

——— (2012a). ‘Overriding default orderings in inflectional morphology’. Paper presented at the Conference on Defaults in Morphological Theory, University of Kentucky, Lexington, May 21–22, 2012.

——— (2012b). ‘Variable affix order and the amorphousness hypothesis’. Paper presented at the First American International Morphology Meeting (AIMM), University of Massachusetts, Amherst, September 22, 2012.

——— (forthcoming). ‘Morphotactically conditioned overrides of default affix ordering’. In A. Hippiisley and G. T. Stump (eds.), *Cambridge Handbook of Morphology*. Cambridge: Cambridge University Press.

Wurzel, W. U. (1984). *Flexionsmorphologie und Natürlichkeit. Ein Beitrag zur morphologischen Theoriebildung*. Berlin: Akademie-Verlag. Translated as Wurzel (1989).

——— (1989). *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer.

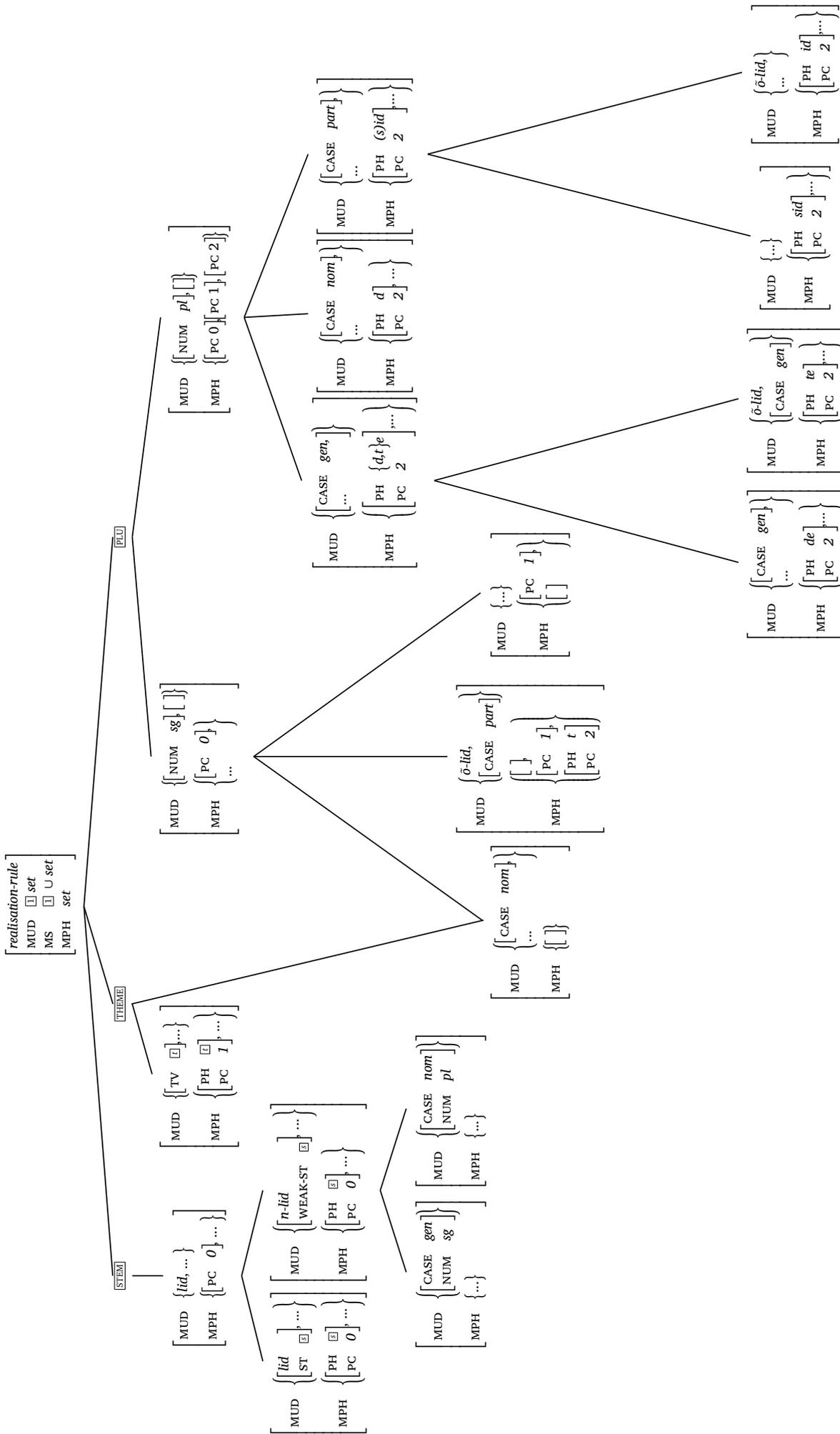


Figure 4: Gestalt exponence in Estonian

Oneida prepronominal prefixes in Information-Based Morphology

Thomas Diaz, Jean-Pierre Koenig and Karin Michelson

The inflectional morphology of Northern Iroquoian, which consists of aspect suffixes and prepronominal and pronominal prefixes (as shown in (1)), is known to be particularly complex. It makes large sets of inflectional distinctions (e.g., 58 pronominal prefixes in a single position class; 11 prepronominal prefixes that can combine into 43 attested prepronominal prefix combinations). And allomorphy is robust and complicated/unnatural (at times) (see Koenig and Michelson 2015).

$$(1) \quad [\text{WORD} \textbf{prepro-pro}_s \text{-} [\text{STEM} \text{BASE-asp}]]$$

The particularly complex allomorphy that the combination of prepronominal prefixes induces has led many Iroquoianists to abandon Lounsbury’s (1953) analysis of prefixes into a position class system. Instead, they treat prepronominal prefixes as a single block with at least 436 allomorphs for the 43 attested prefix combinations (see, e.g., Abbott, Christjohn and Hinton 1996 and Michelson and Doxtator 2002 for Oneida and Woodbury 2003 for Onondaga).

In this paper, we argue that an analysis of prepronominal prefixes into a position class system is possible within Information-Based Morphology (Crysmann and Bonami 2016; IBM, hereafter). The position class system we argue for is shown in Figure 1. We show that this analysis reduces the number of exponent rules needed by a factor of ten; allows us to model well-known generalizations that a single-block model would miss; and provides a simple account of the variable positional realization of the factual and optative, which can occur in four and three positions, respectively. Figure 1 assumes that prepronominal prefixes realize, overall, five distinct bundles of morphosyntactic features (members of the MS set in IBM parlance), as shown in (2). Importantly, since there is an interaction between prepronominals, pronominals, and stem when it comes to prepronominal prefix allomorphy, the value of the MS attribute also includes members for pronominal prefixes and the stem and thus has 7 members, as shown in (2).

$$(2) \quad \left\{ \begin{array}{l} \left[\begin{array}{l} \text{COIN} \ /- \\ \text{PART} \ /- \\ \text{NEG} \ /- \\ \text{CONTR} \ /- \end{array} \right], \left[\text{DUALIC} \ /- \right], \left[\text{MOOD} \ \textit{fut, opt, fact/none} \right], \left[\text{TRANS} \ /- \right], \left[\text{REP} \ /- \right] \\ \left[\text{AGR} \ \textit{list}(\phi\text{-features}) \right], \textit{lid} \end{array} \right\}$$

In the rest of this abstract, we present some of the most interesting and challenging facts concerning the morphology of Oneida prepronominal prefixes, and discuss how IBM allows for a relatively easy model of these facts. One of the hallmarks of position class systems is that morphosyntactic properties assigned to the same position compete for realization, and only one of them can ever be realized at a time. In the context of Oneida prepronominal prefixes, this means that the negative, contrastive, coincident, and partitive compete for realization in position 1, and the cislocative and repetitive for realization in position 5. The various mood values in positions 4 and 6 are not in competition for realization, as they are values of the

1	2	3	4	5	6	7	8
Negative	Translocative	Dualic	Factual	Cislocative	Factual	Pronominal	Stem
Contrastive	Factual		Optative	Repetitive	Optative	Factual	
Coincident			Future			Optative	
Partitive							

Figure 1: Position classes for Oneida inflectional prefixes

same attribute, and a word cannot therefore be simultaneously factual and optative (or future). Mood and pronominal prefixes do not compete for realization in position 7 either, as forms occurring in position 7 (the usual position for pronominal prefixes) that mark mood are portemanteau morphs, as we show below. Finally, the translocative and factual in position 2 do not compete for realization either, as the exponent of the factual that occurs in position 2 only does so when the factual is word initial, i.e. when there is no prefix occurring in position 1, 3, and 5, and when the value of the translocative attribute is –.

Because Oneida prepronominal prefixes, while inflectional, add semantic information, competition for realization in positions 1 and 5 raise an expressibility question: What happens when the semantic concept expressed by the coincident (*When ... or ... same ...*) and the negative (*... not ...*), or the repetitive (*... again ...*) and the cislocative (*... (to) here ...*), must both be conveyed? Oneida uses two distinct solutions to solve the potential expressibility issues that positions 1 and 5 pose. In the case of position 1, the COIN/CONTR/NEG/PART Morphological Feature Combination includes a hierarchy of *realization* (see (3)), which obviates any issue of semantic expressibility. If the value of a feature higher in the hierarchy is +, then the realization of the MFS takes the form reserved for that feature (i.e. the form when all other features have – values), as illustrated in examples (4).

(3) COIN > CONTR > NEG > PART

- (4) a. né tshá'-ka-t
 assertion COIN-3Z/N.SG.A-same[STV]
 'it's the same'
- b. yah né tshá'-ka-t
 not assertion COIN-3Z/N.SG.A-same[STV]
 'it's not the same'

We implement the hierarchy in (3) in terms of a hierarchy of realizational rules as shown in Figure 2. Such a hierarchy of realizational rules is one way to overcome expressibility limitations that can occur in a position class system.

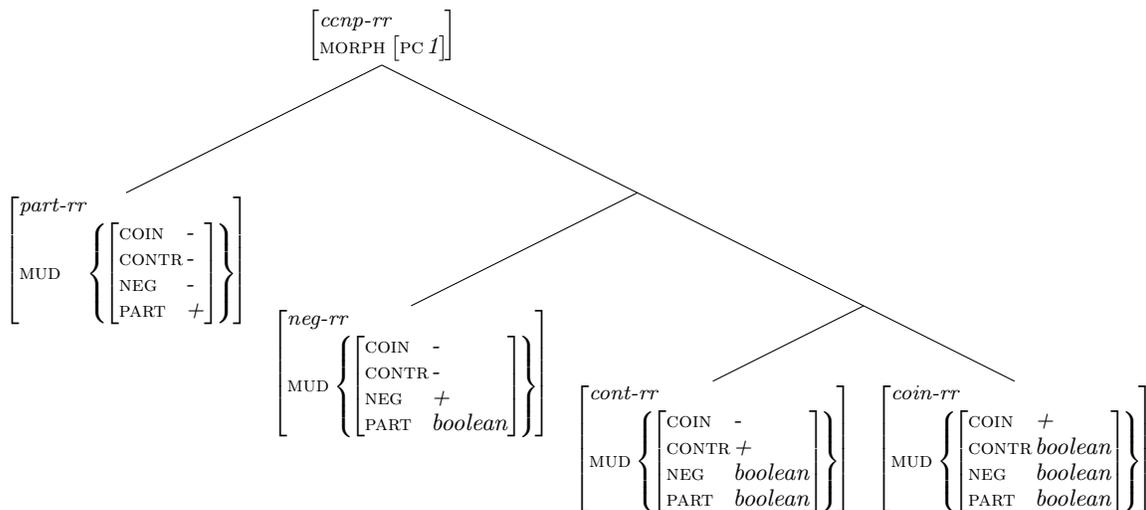


Figure 2: A portion of the hierarchy of position 1 rules

Oneida deals with the challenge of expressibility of the occurrence of the repetitive and cislocative in the same position class in a quite distinct way. In this case, the dualic prefix must occur, whether the dualic meaning is conveyed or not, as illustrated in (5)-(7). (5), which includes the translocative prefix in position 2, does not raise an expressibility challenge and no dualic is present. The corresponding example with the cislocative in (6) does, and the dualic must occur. The required presence of the dualic in position 3, when the meaning of both the cislocative and repetitive must be expressed, means that examples like

(7) can convey three possible sets of concepts: The cislocative and dualic concepts, the cislocative and repetitive concepts (the dualic prefix being required formally, but without contributing any meaning), and the cislocative, repetitive, and dualic concepts. To the extent the last set of concepts is sufficiently distinct from the first set, this seems to be empirically correct. We model this fact through a constraint on MS membership as shown in (8).

- (5) y-usa-yakw-ahtΛ tí
 TRL-FACT:REP-1EX.PL.A-leave:PNC
 ‘we left for there (to go home) again’
- (6) t-uta-yakw-ahtΛ tí
 DL-FACT:CSL-1EX.PL.A-leave:PNC
 ‘we left from here again’
- (7) t-Λ-t-hati táne?
 DL-FUT-CSL-3M.PL.A-stand,stop:PNC
 ‘they will stop there (again)’ (*possible ambiguity*)
- (8) $\left(\left[\text{MS}\{[\text{CSL } +]\} \cup \text{set} \right] \wedge \left[\text{MS}\{[\text{REP } +]\} \cup \text{set} \right] \right) \Rightarrow \left[\text{MS}\{[\text{DUALIC } +]\} \cup \text{set} \right]$

The use of constraints on MS membership is needed to model another case of “prefix substitution,” namely another case where the realization of a prefix does not semantically convey what this prefix typically conveys, but what another prefix conveys. This case involves the negative and contrastive prefixes. The presence of a negative prefix in Oneida precludes dualic, translocative, or mood prefixes. If any of these prefixes occur, the contrastive prefix is used. In other words, when any prefix occurs in positions 2 through 4, the contrastive substitutes for the negative, as shown in (9). The constraint in (10) models this “prefix substitution,” given the hierarchy in Figure 2 which ensures that when both the negative and contrastive features have + values (and the coincident is –), the contrastive prefix occurs.

- (9) a. tho y-a?te-yakw-atlást-a?
 there TRL-DL-1EX.PL.A-meet-HAB
 ‘we meet, get together there’
- b. yah tho th-y-a?te-yakw-atlást-a?
 not there CONTR-TRL-DL-1EX.PL.A-meet-HAB
 ‘we don’t meet, get together there’
- c. *yah tho te?-y-a?te-yakw-atlásta?
 not there NEG-TRL-DL-1EX.PL.A-meet-HAB

$$(10) \left(\left[\text{MS}\{[\text{NEG } +]\} \cup \text{set} \right] \wedge \left(\begin{array}{l} \left[\text{MS}\{[\text{DUALIC } +]\} \cup \text{set} \right] \\ \vee \left[\text{MS}\{[\text{TRANS } +]\} \cup \text{set} \right] \vee \left[\text{MS}\{[\text{MOOD } \textit{potent}]\} \cup \text{set} \right] \end{array} \right) \right) \Rightarrow \left[\text{MS}\{[\text{CONTR } +]\} \cup \text{set} \right]$$

The allomorphy of the translocative prefix is sensitive to the presence of a phonologically-determined accent on the prefix. To simplify, we must distinguish between the simple allomorphs *y-*, *ye-*, and *ya?*- and the complex forms *yehe-*, *yΛhΛ-*, and *yaha?*-. The latter are used whenever the translocative prefix is accented, i.e., when the pronominal prefix and stem do not form more than a single syllable (given the rules governing accent in Oneida, see Michelson 1981). We model this kind of allomorphy by a constraint on verb forms that conditions the use of any member of the set of realizational rules for the complex allomorphs of the translocative (subtypes of the *trl-acct-rr* type) on the form of the pronominal prefix and the stem, as shown in (11). The presence in IBM of the set of realizational rules a word exemplifies in the representation of that word makes it easy to model this otherwise rather unusual condition on the allomorphy of the translocative.

$$(11) \left[\begin{array}{l} \textit{word} \\ \text{RR} \quad \mathbb{I} \left\{ \dots \left[\begin{array}{l} \text{MUD}\{lid\} \\ \text{PH} \langle \sigma \rangle \end{array} \right], \left[\begin{array}{l} \text{MS}[\text{AGR } list(\phi)] \\ \text{PH } list(C) \end{array} \right], \dots \right\} \\ \text{MS} \quad \{ \dots [\text{TRANS } +] \dots \} \end{array} \right] \Rightarrow \textit{member}(\textit{trl-acct-rr}, \mathbb{I})$$

As shown in Figure 1, the factual (and optative) can be realized in several positions. In some cases, the two positions correspond to distinct realizations of the same morph (positions 2 and 4 or 6 and 7, respectively). In other cases, the two positions correspond to what Lounsbury (1953) calls “partials,” meaning that two distinct morphs in two distinct positions together realize the factual: The repetitive (*s-*) together with the factual, for example, surfaces as *-usa-* and the cislocative (*t-*) and factual as *-uta-*. Partials are easy to model in Information-Based Morphology without resorting to empty morphs (as Lounsbury does), as on-line typing (Koenig 1999) for realizational rules allows the expression of complex regularities like those involved in the exponence of the factual (and optative). We represent informally, the exponence rules for the factual in Figure 3. We assume a default null realization rule for each dimension (omitted for space considerations). The first dimension realizes the first part of the factual partials in positions 2 or 4, depending on whether the repetitive/cislocative occurs or nothing occurs before the factual (i.e., the factual is word initial). The second position realizes the second part of the factual in position 6 or 7.

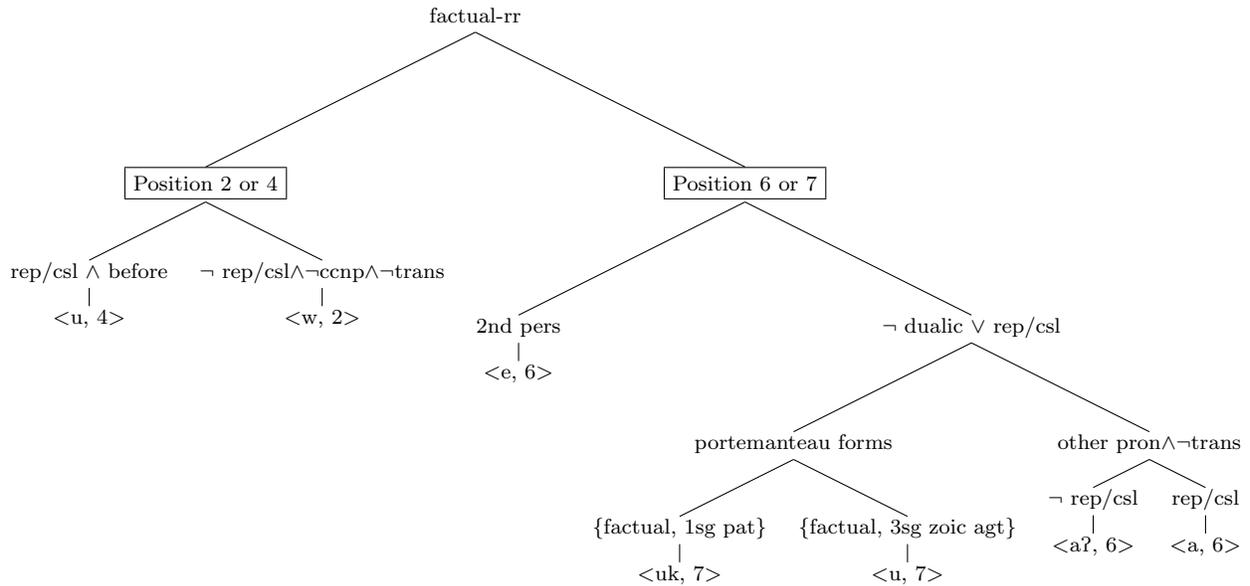


Figure 3: The hierarchy of realizational rules for the factual

The realization of the factual partial in positions 6 or 7 are interesting, as they are conditioned by the pronominal prefix realized in position 7 (for position 6) or are portemanteau morphs realizing both the factual partial and the pronominal prefix in position 7. We concentrate on position 7 here as it illustrates best how IBM helps model a particularly idiosyncrastic case of allomorphy. Consider the verb forms *uki táwe?* and *ustáthΛ?* in (12) and (13), respectively. *U-* in the first case realizes both the factual and the *wa* part of the pronominal prefix for the first singular Patient intransitive prefix (ordinarily, *wak-*). *U-* in the second case realizes both the factual, the third person feminine-zoic/neuter Agent intransitive prefix (ordinarily, *w-*), and the stem initial *a-* of so-called *a-*stems (see Koenig and Michelson 2015), as shown in the “underlying forms” represented below the actual forms in (12) and (13).

- (12) *uki·táwe?*
wa?-*wak-ita?**w-e?*^{UR}
 FACT-1SG.P-sleep-PNC
 ‘I fell’
- (13) *ustáthΛ?*
wa?-*w-astathv-?*^{UR}
 FACT-3Z/N.SG.A-become.dry-PNC
 ‘it became dry’

The realizational rules for both portemanteau morphs are provided below. We use ‘!’ to indicate that our constraint is an abbreviation of the actual constraint, and [*cr*, F +] indicates that either the cislocative

or repetitive attribute have a + value. Additionally, we use *1sg patient* and *3rd zoic agent* as abbreviations for the relevant AVM for space considerations. As customary within IBM, the presence of two members in the value of the MUD attribute (for Morph Under Discussion) indicates that *uk-* and *u-* are portemanteau morphs. Finally, the *u-* form requires the stem realization rule in (15), which realizes the *a*-stem without its initial vowel. What looks like rather crazy allomorphy rules (*wa-wak-* > *uk-*; *wa-w-a...* > *u...*) receive a comparative treatment within IBM because of its assumption that all prepronominal prefixes, pronominal prefixes, and the stem are realized simultaneously (therefore licensing portemanteau of prepronominal and pronominal prefixes).

$$(14) \left[\begin{array}{l} \text{MUD} \mathbb{1} \{ [\text{MOOD } \textit{factual}], \textit{1sg pat} \} \\ \text{MS} \mathbb{1} \cup ! \left([\text{DUALIC-}] \vee \left[\begin{array}{l} \textit{cr} \\ \textit{F} \end{array} \right] \right) ! \cup \textit{set} \\ \text{MPH} \left\{ \left[\begin{array}{l} \text{PHON} \langle \textit{uk} \rangle \\ \text{PC} \quad \gamma \end{array} \right] \right\} \end{array} \right] \quad \left[\begin{array}{l} \text{MUD} \mathbb{1} \{ [\text{MOOD } \textit{factual}], \textit{3rd zoic agent} \} \\ \text{MS} \mathbb{1} \cup ! \left([\text{DUALIC-}] \vee \left[\begin{array}{l} \textit{cr} \\ \textit{F} \end{array} \right] \right) ! \cup \textit{set} \\ \text{MPH} \left\{ \left[\begin{array}{l} \text{PHON} \langle \textit{u} \rangle \\ \text{PC} \quad \gamma \end{array} \right] \right\} \end{array} \right]$$

$$(15) \left[\begin{array}{l} \text{MUD} \quad \mathbb{s} \left\{ \left[\begin{array}{l} \textit{lid} \\ \text{STEM} \mathbb{1} \end{array} \right] \right\} \\ \text{MS} \quad \mathbb{s} \cup \{ [\text{MOOD } \textit{opt/fac}], \textit{3rd zoic agent} \} \cup \textit{set} \\ \text{MPH} \quad \left\{ \left[\begin{array}{l} \text{PH} \langle \mathbb{3} \rangle \\ \text{PC} \textit{11} \end{array} \right] \right\} \\ \text{S-PHON} \left\{ \left[\begin{array}{l} \text{PHON} \mathbb{1} \left(\langle \mathbb{2} \textit{a} \rangle \oplus \mathbb{3} \right) \\ \text{CLASS} \mathbb{2} \end{array} \right] \right\} \end{array} \right]$$

Our paper shows that, contra prevailing Iroquoianists' wisdom, an IBM analysis of Oneida prepronominals into individual prefixes is not only possible, but more intuitive and insightful. In fact, it reduces the number of allomorphs by a factor of ten. Moreover, IBM makes the kind of allomorphy Oneida prefixal inflectional morphology exhibits rather easy to model, partly because of the ability to include other prepronominal prefixes *and* pronominal prefixes and stem-type as conditioning environment for allomorph selection and partly, because of its use of on-line type construction to model partial exponents of morphosyntactic features. But our analysis also shows that there are some unique aspects of Oneida inflectional prefixes that still make it complex. First, conditioning environments for exponence is particularly complex as some of the rules we provided show. Second, Oneida uses two distinct strategies for overcoming limitations on expressibility of position class systems: a hierarchy of exponence (for prefixes that occur in the first position) and prefix "stand-in" (whereby the realization of the dualic indirectly expresses the presence of both the repetitive and cislocative meanings). Third, prepronominal prefix allomorphy is often conditioned by what precedes (or does not precede) the exponent, something that is difficult to directly model within IBM and requires the rather cumbersome listing of all morphosyntactic feature bundles that have – values (and are thus realized by a default zero realization rule).

Against Split Morphology

Robert Hepburn-Gray
 University at Buffalo (SUNY)
 rchepbur@buffalo.edu

In this paper I present data from several Niger Congo languages, illustrating how the paradigms which make up the noun class systems of these languages are problematic to analyze within traditional morphosyntactic frameworks. I outline possible solutions to this problem, and argue for the introduction of an exemplar-based Word and Paradigm (Blevins 2006) approach to morphology within SBCG. I then outline the consequences of this approach for the structure of the SBCG lexicon.

Most morphological theory, including that currently employed within SBCG (Sag and Boas 2012), often differentiates between inflectional and derivational processes. This division is referred to as “split morphology” by Bauer (1997). In SBCG, this distinction is explicitly represented in the hierarchy of construct(ion) types, as shown here:

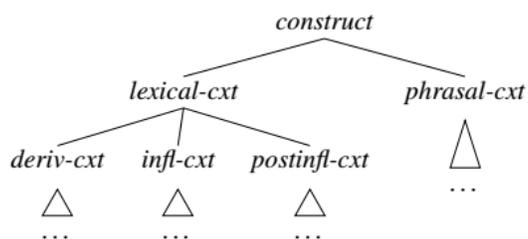


Figure 1: SBCG Construct Types (Sag 2012)

This traditional division of inflection and derivation has been defined and supported by a number of different criteria, outlined in Stump (2005) and adapted here:

- A. Derivation can change part-of-speech class, while inflection cannot
- B. Inflection applies to a category without exception; derivation applies sporadically
- C. Inflection is semantically regular; derivation is frequently less than fully semantically regular
- D. Inflection is syntactically determined; derivation is not
- E. Derivational processes apply before inflectional processes

Stump notes that none of these criteria are unproblematic. Indeed it is not hard to find contradictions even in English; the gerund for instance, changes category (V > N), but applies without exception and is semantically regular, a contradiction between criterion A, and B and C. Despite the problems that he discusses, Stump maintains that there is indeed a “conceptual” difference between inflection and derivation, reiterating the common analysis that inflection relates all the words related to a single lexeme, while derivation creates new lexemes from other lexemes.¹ This analysis is also assumed in SBCG, and is achieved through the type signatures of *infl-cxt* and the *deriv-cxt*:

¹ Contra Stump (2001:253) , “both inflectional paradigms and derivational paradigms are inventories projected from a single lexeme...”

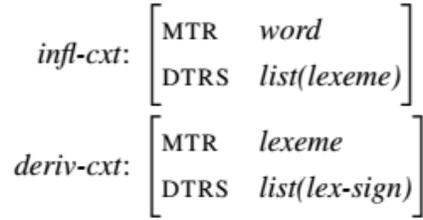


Figure 2: Inflectional & Derivational Construction Type Signatures

Sag does note, however, that the inflection/derivation dichotomy may not be suitable for some languages, and that a more fine-grained distinction between morphological types may be necessary for the analysis of these languages (2012:112). I argue, based on the problems of the definitional criteria already noted by Stump and the Niger-Congo data to be presented below, that exceptions to this dichotomy are not marginal, and that a split-morphology analysis should not be assumed when analyzing newly documented languages.

Turning now to the Niger-Congo data, consider the noun class system of the Kordofanian language Otoro, described in Stevenson (2009).

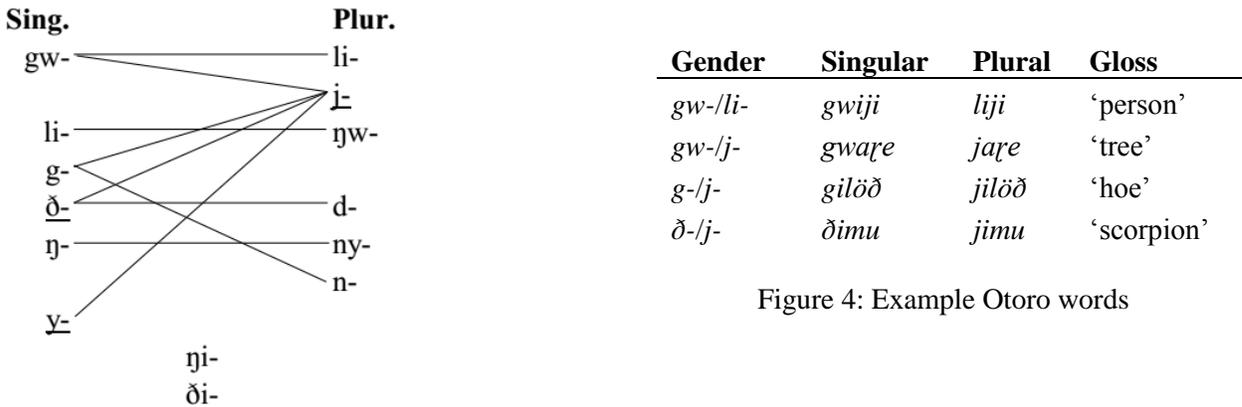


Figure 4: Example Otoro words

Figure 3: Otoro Noun Class System

The diagram in Figure 3 presents the noun class system of Otoro. The system is comprised of a set of noun classes, which are used to encode number, labeled here by the phonological shape of their prefixes. The lines indicate the pairing of classes into singular/plural genders. The classes on the left are singular classes, those on the right are plural classes, and those in the bottom center are used with nouns that have no number distinction (typically mass or abstract nouns). The classes which are underlined are those which appear with single class nouns in addition to appearing with singular/plural nouns. Note that there are singular classes which are paired with multiple plural classes, as well as plural classes paired with multiple singular classes. This is illustrated by the example words in Figure 4. In addition, the class marker *li-* functions as both a plural class marker in the *gw-/li-* gender and a singular class marker in the *li-/ŋw-* gender.

This irregularity of class pairing reaches an extreme, however, in languages such as Lumun (Smits 2011). Out of 11 phonologically distinct noun class markers, 6 have functions as both singular and plural markers in different genders, and 10 appear with single-class nouns. This irregularity is an example of a contradiction between criterion B (exceptionless application) and criterion C (semantic regularity), since most of the noun class markers are associated with different number semantics. To avoid this the analysis would have to appeal to widespread accidental

homophony. In addition, there is a second layer of semantic irregularity, in that the semantic effects of the class prefixes in Lumun go beyond the straightforward singular/plural opposition, often having semantic features such as singulative, collective, etc. This degree of irregularity has led Smits (2011) to follow Schadeberg's (2001) analysis of Swahili number, and suggests that number marking in Lumun is, surprisingly, a derivational, rather than inflectional, process. Since number marking is not optional in Lumun, this creates a contradiction between criterion B (exceptionless application) and criterion C (semantic regularity).

The second feature of Niger Congo noun class systems problematic to split morphology is the existence of *paradigm networks*. An illustration of one such network in the Atlantic language Baïnounk Gubêcher is shown below.

Table 1: Botanical paradigm network of Baïnounk Gubêcher (Adapted from Cobbinah 2013:319)

NC Paradigm	- <i>dóoma</i> 'kaba'	- <i>taat</i> 'annona'
<i>si-/mun-</i>	'kaba tree'	'annona tree'
<i>bu-/i-/di-</i>	'kaba fruit'	'annona fruit'
<i>ja-</i>	'leaves of kaba tree'	'leaves of annona tree'

In this botanical paradigm network roots referring to specific tree species enter into different noun class paradigms depending on what part of the plant is being referenced. One paradigm refers to the tree itself, a second refers to the fruit of the tree, and a third refers to the leaves of that tree. A second type of paradigm network is shown below.

Table 2: Cicipu ethnic group paradigm (McGill 2007:61)

Class	Acipu	Karishen	Kadonho	Hausa	Gloss
8	c-cípù	Ø-rísìnô	d-dípó	k-kógó	Person
2	à-cípù	ò-rísìnô	ò-dípó	ò-kógó	People
1		kò-rísìnô	kò-dípó		town/area
6	cì-cípù	tì-rísìnô	tì-dípó	tì-kógó	Language/dialect

This data from Cicipu, a Kainji language, is an example of an ethnic group paradigm network. Here, a root referring to a certain ethnicity can be associated with different paradigms to create different “words”, whether this word is referring to a person of this ethnic group, the language spoken by this group, or the area inhabited by this group. Paradigm networks such as this can be found throughout the Niger Congo family. Crucially, there is no *a priori* way to decide which of these paradigms, and the corresponding semantics, should be associated with a primary lexeme, and which others should be considered derived.² Nor are there any morphological or semantic criteria that would lead to a decision. A split-morphology analysis of these paradigm networks would then require the arbitrary selection of one paradigm as primary. Alternatively, one could imagine an analysis where none of the genders are chosen as primary, and instead all are derived.

This analysis, however, runs into its own set of problems. The roots for words which participate in these paradigm networks of course must be listed somehow in the lexicon, with their phonological information as well as the semantics relating to the particular plant species or ethnic

² Conceptually, it may at first seem appealing to assign the ‘tree’ as the basic lexeme of the botanical network, but there are numerous English compounds such as *coconut tree*, where *coconut* bears the semantic weight of species differentiation, and *tree* specifies the referent. In English, tree can be analyzed both the semantic and the morphological head. But an analogous analysis of the Niger-Congo data is not possible, as there would be a mismatch between the morphological head (i.e. the root) and the semantic head, which is determined by the noun class affix.

group. This, however, leads to a contradiction given the hierarchy of morphological construction types given above in Figure 2. On the one hand, these roots must be listed as lexemes, because derivational constructions take a sign which is a lexeme and produce a new lexeme. On the other hand, these roots *cannot* be lexemes, since these *never* enter into an inflectional construction to produce a word.

Koenig's (1999) Type-Underspecified Hierarchical Lexicon (TUHL) approach to morphology within HPSG could plausibly supply the formal architecture necessary to model the many irregularities of number marking in languages like Lumun. It is unclear, however, how it could be extended to cover the paradigm networks of languages like Baïnounk Gubëeher and Cicipu, in which a class marker seems to be simultaneously marking the 'derivation' of the noun and its associated semantics, and also marking the noun as 'inflected' for number. Moreover, TUHL is explicitly based on the assumption that, "What is stored [in the lexicon] is a set of underspecified, basic categories..." and that "fully type-specified entries... arise only while processing sentences" (53). There is however a growing body of historical evidence (Traugott and Trousdale 2013; Bybee 2006) that individuals *do* store more specific information about their linguistic experience. Bybee (2013) and Diessel (2015) have recently noted the natural compatibility of usage-based approaches with construction grammar. Moreover, Abbot-Smith and Tomasello (2006) provide evidence that exemplar representations exist alongside abstract schemas in syntax.

With this in mind, I propose implementing a theory of morphology within SBCG which is a hybrid of TUHL and an exemplar-based Word and Paradigm (Blevins 2006) model. In a W&P approach, no sub-word linguistic material is stored in the lexicon. Instead, exemplar paradigms are stored, and non-exemplar words are generated by analogy to the exemplar paradigms. Although W&P models generally distinguish between inflection and derivation, Bauer (1997) notes the paradigmatic nature of the relationships between 'derivationally' related forms. These 'derivational' paradigms can be analyzed in the exact same fashion in a W&P approach, because the same process of analogical formation underlies production of non-stored forms for both types of paradigm. This eliminates the problems associated with Niger-Congo noun class systems, since there is no need to decide how the noun roots are stored in the lexicon (they aren't!). It also allows for morphological functions such as number marking to display the same degree of irregularity as that seen in formerly derivational processes, such as nominalizations.

A language user's knowledge of the internal structure of paradigms, however, is modeled using the TUHL model, as Koenig demonstrates using the example of Latin verb paradigms (1999:142ff). The crucial difference is that, in the model proposed here, exemplar paradigms *are* stored fully specified in the lexicon. Non-exemplar word paradigms can be stored underspecified, only storing the cells necessary to determine the remainder of the paradigm (Blevins 2013, and citations therein).

References

- Abbot-Smith, Kirsten, and Micheal Tomasello. 2006. Exemplar-Learning and Schematization in a Usage-Based Account of Syntactic Acquisition. *The Linguistic Review* 23: 273–88.
- Bauer, Laurie. 1997. Derivational Paradigms. In *Yearbook of Morphology 1996*, edited by Jaap van Marle and Geert Booij, 243–56. Dordrecht: Kluwer.
- Blevins, James. 2006. Word-Based Morphology. *Journal of Linguistics* 42: 531–73.
- Blevins, James. 2013. The Information-Theoretic Turn. *Psihologija* 46 (3): 355–75.
- Bonami, Olivier, and Pollet Samvelian. 2015. The Diversity of Inflectional Periphrasis in Persian. *Journal of Linguistics* 51: 327–82.

- Bybee, Joan. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82 (4): 711–33.
- Bybee, Joan. 2013. Usage-Based Theory and Exemplar Representations of Constructions. In *Oxford Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale. Oxford: Oxford University Press.
- Cobbinah, A.Y. 2013. Noun Classification and Verbal Nouns in Bāinounk Gubēeher. Dissertation, London: University of London.
- Diessel, Holger. 2015. Usage-Based Construction Grammar. In *Handbook of Cognitive Linguistics*, 296–322. Berlin: Mouton de Gruyter.
- Koenig, Jean-Pierre. 1999. *Lexical Relations*. Stanford: CSLI Publications.
- McGill, Stewart. 2007. The Cicipu Noun Class System. *Journal of West African Languages* 34 (2): 51–90.
- Sag, Ivan. 2012. Sign-Based Construction Grammar: An Informal Synopsis. In *Sign-Based Construction Grammar*, edited by Ivan Sag and Hans Boas, 61–187. Stanford: CSLI Publications.
- Sag, Ivan, and Hans Boas, eds. 2012. *Sign-Based Construction Grammar*. Stanford: CSLI Publications.
- Schadeberg, Thilo. 2001. Number in Swahili Grammar. *Afrikanistische Arbeitspapiere* 68: 7–16.
- Smits, Heleen. 2011. Lumun Noun Classes and Number. In *Afrikanische Sprachen Im Fokus. Linguistische Beiträge Zum 19. Afrikanistentag, Mainz, 8.-10. April 2010*, 271–83. Cologne: Rüdiger Koppe Verlag.
- Stevenson, Roland C. 2009. A Grammar of Otoro. In *Tira and Otoro: Two Grammars of Kordofanian Languages by Roland C. Stevenson*, edited by Thilo Schadeberg. Cologne: Rüdiger Köppe Verlag.
- Stump, Gregory. 2001. *Inflectional Morphology*. Cambridge: Cambridge University Press.
- Stump, Gregory. 2005. Word Formation and Inflectional Morphology. In *Handbook of Word Formation*, 49–71. Dordrecht: Springer.
- Traugott, Elizabeth, and Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

Parallel affix blocks in Choctaw

George Aaron Broadwell, University of Florida

1. Introduction¹

Choctaw is a Muskogean language, spoken in Oklahoma and Mississippi. Verbal agreement works on an active/stative basis, while nominals show nominative/accusative case marking. Data is from Broadwell (2006) and the author's notes, unless otherwise indicated.

The simplest sentences in Choctaw consist of a verb plus a tense marker:

1. Písa-tok. 'S/he saw them.'
see:n-pt

When there is an overt NP argument, it is marked for case:

2. John-at niya-h. 'John is fat.'
John-nm fat-tns

Subject NPs are obligatorily marked with the nominative case /-at/. Object NPs are optionally marked with the accusative /-a/:

3. Gus-at John(-a) písa-tok. 'Gus saw John.'
Gus-nm John-ac see:n-pt

As these examples show, there is no subject or object agreement morphology for 3rd person.

Choctaw adds objects via applicative prefixes, as in the following example with an instrumental applicative.

4. Charles-at báshpo' nípi' isht-bashli-h.
Charles-nm knife meat instr-cut-tns
'Charles cut the meat with a knife.'

2. The agreement system

There are three main types of verbal agreement (called I, II, and III agreement). The overall agreement system is shown in the table below:²

¹ In the orthography used here, underlined vowels are nasal. Gloss abbreviations: ac = accusative, com = comitative, con = contrastive, instr = instrumental, n = nasal grade, nm = nominative, p = plural, pt = past, tns = default tense, s = singular.

² When two alternatives are shown, the first is before a vowel and the second is before a consonant. The 1st sg I affix -li is the only suffix in the system; all the other agreement is via prefix. I omit here a special set of agreement markers used in the negative.

	I	II	III
1sg	-li	sa-	(s)am-/ (s) <u>a</u> -
2sg	ish-	chi-	chim-/ch <u>i</u> -
1pl paucal	il-/ii-	pi-	pim-/p <u>i</u> -
1pl multiple	il-/ii-	hapi-	hapim-/hap <u>i</u> -
2pl	hash-	hachi-	hachim-/hach <u>i</u> -

There are many possible linkages between grammatical function and agreement type. Intransitive subjects, for example, can take any of the three kinds of agreement.

5. (An-akoosh) baliili-li-tok. 'I ran.' **(I subject)**
 (1sg-con:nm) run-1sI-pt
6. (An-akoosh) sa-niya-h. 'I am fat.' **(II subject)**
 (1sg-con:nm) 1sII-fat-tns
7. (An-akoosh) a-ponna-h. 'I am skilled.' **(III subject)**
 (1sg-con:nm) 1sIII-skilled-tns

Despite the different verb agreement, an overt subject for any of these clauses will be nominative. Transitive verbs also fall into several classes, depending on the sort of agreement with subjects and objects. The most frequent types of transitive show I agreement for the subject and either II or III agreement for object:

8. Chi-písa-li-h. 'I see you.' **(I subject/II object)**
 2sII-see:n-1sI-tns
9. Chi-paya-li-h. 'I call you.' **(I subject/III object)**
 2sIII-call-1sI-tns

A much smaller number of transitive verbs trigger II or III agreement for their subjects:

10. Chi-sa-banna-h. 'I want you' **(II subject, II object)**
 2sII-1sII-want-tns
11. Chi-sa-noklhakacha-h. 'I was startled by you.' **(II subject, III object)**
 2sIII-1sII-be:startled-tns

As an example like (10) shows, there may be multiple instances of a II prefix, and in such cases the object II agreement precedes the subject II agreement.

Multiple II prefixes may also arise when a verb has an applicative. In such cases the agreement for the direct object precedes the verb, while the agreement for the applied object precedes the applicative. Consider the example below where the verb contains the comitative applicative *ibaa*:

12. Chi-baa-sa-fama-h

'I was whipped with you.'

2sII-com-1sII-be:whipped-tns

Thus II prefixes may appear in three different places in the prefix template – as subject agreement, as (direct) object agreement, and as applied object agreement. Choctaw therefore displays what Stump (2001) calls the Parallel Block problem – identical exponence in different position classes.

3. Possessor raising

Choctaw also has a rule of possessor raising which makes a possessor an additional argument of a verb. There is both subject possessor raising and object possessor raising. Object possessor raising makes possessors of objects into applied objects of the verb. As applied objects they trigger III agreement:

13. A-shokha' nipi' apa-tok.

'He ate my bacon.'

1sIII-pig meat eat-pt

14. Shokha' nipi' am-apa-tok.

'He ate my bacon.'

pig meat 1sIII-eat-pt

The III prefixes associated with possessor raising come before any other prefixes in the stem. Note the following example, where the III agreement with 'doctor' precedes the instrumental applicative (Munro 1984:641).

15. Alikchi-ya ohooyo im-isht=anopoli-li-tok.

'I talked about the doctor's wife.'

doctor-ac woman III-instr=talk-1sI-pt

III prefixes, therefore, also exemplify Stump's Parallel Block problem.

4. The prefix ordering template

The final template for affix ordering in Choctaw is as follows:

III agr (raised possessors)	instrumental applicative	I agr (subj)	II (applic objects)	other applicatives	III (indirect objects)	II agr (objects)	II agr (subjects)	Verb
1	2	3	4	5	6	7	8	

As the template shows, II agreement shows up in three different position classes (4, 7, 8) and III agreement shows up in two different position classes (1, 6).

5. Information-Based Morphology

Using the conventions of Information-Based morphology (Crysmann & Bonami 2016), we can account for the parallel block problem in Choctaw by positing separate exponence and morphotactic feature sets.

The exponence of agreement with a second-person singular argument, for example, is one of the following:

$$\left[\begin{array}{c} \text{MUD} \left\{ \left[\begin{array}{l} \text{PER} \quad 2 \\ \text{NUM} \quad \text{sg} \\ \text{AGR} \quad \text{I} \end{array} \right] \right\} \\ \text{MPH} \left\{ \left[\text{PH} \quad \text{ish} \right] \right\} \end{array} \right] \quad \left[\begin{array}{c} \text{MUD} \left\{ \left[\begin{array}{l} \text{PER} \quad 2 \\ \text{NUM} \quad \text{sg} \\ \text{AGR} \quad \text{II} \end{array} \right] \right\} \\ \text{MPH} \left\{ \left[\text{PH} \quad \text{chi} \right] \right\} \end{array} \right] \quad \left[\begin{array}{c} \text{MUD} \left\{ \left[\begin{array}{l} \text{PER} \quad 2 \\ \text{NUM} \quad \text{sg} \\ \text{AGR} \quad \text{III} \end{array} \right] \right\} \\ \text{MPH} \left\{ \left[\text{PH} \quad \text{chim} \right] \right\} \end{array} \right]$$

The morphotactic feature sets will specify the position class of each affix, and here we can model the fact that II and III agreement appear in multiple positions. For instance, II agreement for subject, direct object, and applied object appear at three different position classes:

$$\left[\begin{array}{c} \text{MUD} \left\{ \left[\begin{array}{l} \text{applic-obj} \\ \text{AGR} \quad \text{II} \end{array} \right] \right\} \\ \text{MPH} \left\{ \left[\text{PC} \quad 4 \right] \right\} \end{array} \right] \quad \left[\begin{array}{c} \text{MUD} \left\{ \left[\begin{array}{l} \text{obj} \\ \text{AGR} \quad \text{II} \end{array} \right] \right\} \\ \text{MPH} \left\{ \left[\text{PC} \quad 7 \right] \right\} \end{array} \right] \quad \left[\begin{array}{c} \text{MUD} \left\{ \left[\begin{array}{l} \text{subj} \\ \text{AGR} \quad \text{II} \end{array} \right] \right\} \\ \text{MPH} \left\{ \left[\text{PC} \quad 8 \right] \right\} \end{array} \right]$$

This formulation, in which both the position class and the exponence of an affix can be specified and subject to restriction, leads to a natural solution to the Person Case Constraint, discussed in the next section.

6. The Person Case Constraint

Tyler (2017) explores in more complete detail a fact mentioned in Ulrich (1986). When a verb has II subject agreement (slot 8) and II (slot 7) or III object agreement (slot 6), there are severe restrictions on the person combinations that are allowed.

Consider the verbs in (16) and (17), which take II/II and II/III agreement

16. Chi-sa-banna-h. 'I want you'. vs. *Pi-chi-banna-h. 'You want us.'

2sII-1sII-want-tns

1pII-2sII-want-tns

17. I-sa-nokshoopa-h. 'I fear him.' vs. *A-chi-nokshoopa-h. 'You fear me'

III-1sII-fear-tns

1sIII-2sII-fear-tns

Tyler shows that the only grammatical combination are those in which the II subject agreement is 1sg /sa-/. All other combinations are ungrammatical.

However it is important to note that this constraint only applies to II subject agreement. If the II agreement is for direct (PC 7) or applied object (pc 4), then the constraint does not hold:

18. I-chi-tokcholi-tok.

'He tickled you for her.'

III-2sII-tickle-pt

The constraint also fails to hold if the III agreement is for a raised possessor (PC 1), rather than for an object:

19. Pallaska' a-chi-noktakali-h.

'You choked on my bread.'

bread

1sIII-2sII-choke-tns

Thus the Choctaw restriction must make reference not to the exponents of II and III agreement, but to a particular combination of morphotactic combination.

Within IbM, we may state the restriction on inflected words as follows:

$$20. \textit{word} [RR \{ [MPH \{ [PC6 \vee PC7] \}], [1] \{ MPH \{ [PC 8] \}, \dots \} \} \Rightarrow \\ [RR \{ [1] [MUD \{ [PER 1, NUM sg] \}], \dots \}]$$

That is, when a word has an affix in PC 6 or 7 plus an affix in PC 8, then the affix in PC 8 must be 1st person singular.

The type of constraint shown in the Choctaw pronominal prefixes is naturally modelled in a theory like that of IbM, which allows us to specify restrictions over combinations of morphotactic and morphosyntactic features.

7. References

- Broadwell, George Aaron. 2006. A Choctaw reference grammar. Lincoln: University of Nebraska Press.
- Crysmann, Berthold and Olivier Bonami. 2016. Variable morphotactics in Information-based morphology. *Journal of Linguistics* 52:311-374.
- Munro, Pamela. 1984. The syntactic status of object possessor raising in Western Muskogean. In *Proceedings of the Berkeley Linguistics Society* 10:634-64.
- Stump, Gregory. 2001. *Inflectional morphology. a theory of paradigm structure*. Cambridge: Cambridge University Press
- Stump, Gregory. 2012. Variable affix order and the amorphousness hypothesis. Paper presented at the First American International Morphology Meeting (AIMM), University of Massachusetts, Amherst, September 22, 2012
- Tyler, Matthew. 2017. Variation in Case-assignment by Appl heads: Evidence from Basque and Choctaw. Presented at Linguistic Society of America, Jan 2017.
- Ulrich, Charles. 1986. Choctaw morphophonology. PhD diss, UCLA

An incremental approach to gapping constructions

Petter Haugereid

Western Norway University of Applied Sciences

petter.haugereid@hvl.no

Abstract

In this paper I present an incremental approach to gapping where it is assumed that the first sentence in a gapping construction is fully parsed before the second sentence with the elided verb is parsed. The analysis is achieved by letting a construction type of the first conjunct be carried over to the second conjunct. This construction type imposes constraints on the arguments that the second conjunct can have. The analysis is implemented in an HPSG grammar of Norwegian.

1 Introduction

Gapping (see (1)) is one out of more phenomena referred to as non-constituent coordination (NCC) in the literature. It is a challenge to lexicalist approaches given the fact that the main verb of the second conjunct in these constructions is elided.

- (1) I eat fish, and you roast.

1.1 Regular gapping

Gapping is possible with a range of constructions in Norwegian. In this section, I will present a couple of the constructions that have been considered in the implementation of the HPSG grammar Norsyg.

In (2), two elements are gapped, the finite verb *tar* ('takes') and the particle *with* ('med'). The particle cannot appear in the gapping construction.

- (2) Jeg **tar med** mat, og du (*med) drikke.
I bring with food and you with drink
I will bring food, and you drinks.

In (5), the reflexive verb *ønske seg* 'wish for' is gapped. The reflexive cannot appear in the gapping construction.

- (3) Jeg **ønsker meg** fisk, og du (*deg) steik.
I wish REFL fish and you REFL roast
I want fish, and you roast.

In transitive idiomatic expressions, all the idiomatic words are elided in the second conjunct (see (4)). It is not possible to elide just parts of the idiom.

- (4) Jeg **brakte på bane** isen, og du
I brought on track ice-DEF and you
(*på) (*bane) sjokoladen.
on track chocolate-DEF
I brought up the ice cream, and you the chocolate.

Verbs with selected prepositions, however, behave slightly different. If a verb has a selected preposition, the gapping construction is very odd if it does not have the preposition, as shown in (5). However, when the gapping construction contains the selected preposition, as in (6), it is much better.

- (5) ??Jeg **hører på** Jon, og du Marit.
I listen to Jon and you Marit
I listen to Jon, and you Marit.
(6) Jeg **hører på** Jon, og du på Marit.
I listen to Jon and you to Marit
I listen to Jon, and you (listen) to Marit.

It is possible to have gapping with ditransitive verbs, as shown in (7). We then get three constituents in the second conjunct.

- (7) Per **serverte** meg fisk, og Kari deg steik.
Per served me fish and Kari you roast
Per served me fish, and Kari you roast.

We can also have gapping when a verb is passivized, as shown in (8). Then both the passive auxiliary and the verb are elided in the second conjunct.

- (8) Jeg **ble servert** fisk, og du steik.
I was served fish and you roast
I was served fish, and you roast.

The constituents in the gapping constructions may also be subordinate clauses or infinitival clauses.

1.2 Gapping with a shared topic

The examples we have looked at so far have been examples of gapping in sentence coordinations. (9) and (10) illustrate that it is possible to have gapping in cases where the topic is shared. In (9) the two conjuncts share the subject. I argue that (7) and (9) are examples of the same phenomenon, only that in (7), we have sentence coordination and in (10), we have coordination of sentences with a shared topic. As with other coordinations where the topic is shared, (10) shows that it is also possible to let an adjunct be shared in gapping constructions (*i går* ‘yesterday’). I will show in Section 3.3 that no extra machinery is needed in order to account for gapping in coordinations where the topic is shared.

- (9) Per **serverte** meg fisk og deg steik.
Per served me fish and you roast
Per served me fish, and you roast.
- (10) I dag **ble jeg servert** fisk og du steik.
Today was I served fish and you roast
Today I was served fish, and you roast.

2 Gapping in HPSG

In the HPSG theory, Immediate Dominance schemata allows a grammar writer to specify constraints on a phrase and its immediate daughters without specifying the order of the daughters (Pollard and Sag, 1994). This makes it possible to account for free word order phenomena, but is restricted to the immediate daughters of a phrase. In order to handle phenomena where the constituents involved are not immediate daughters of the same phrase, like discontinuous constituents and non-constituent coordination (NCC), the feature DOM(ain) has been introduced, where the linear order of the phonological items that a phrase consists

of, is represented (Reape, 1994). The elements on the DOM list may be arranged in an order that is not reflected in the derivation tree. This separation of the order of phonological items from the constituent structure is referred to as linearization. Most approaches to non-constituent coordination makes use of the linearization approach (Kathol, 1995; Beavers and Sag, 2004; Chaves, 2005; Crysmann, 2008). The use of DOM to handle linearization phenomena is powerful, and although relational constraints may be added to the grammar in order to impose restrictions on the order of the phonological items, it may put a heavy burden on the parser.

The distinction between phonological representation and constituent structure assumed in the linearization approach is not available in grammars written within the DELPH-IN network, like the ERG (Flickinger, 2000) and JACY (Siegel *et al.*, 2016). These grammars use regular phrase structure rules where the phonology is simply concatenated, and constituents are reflected in the derivation tree. This is efficient, but it poses a challenge to phenomena like non-constituent coordination since the valence information of the verb in the first conjunct is not accessible at the point where the coordination happens (the valence requirements have been cancelled off), and even if they were, there is no dummy verb in the second conjunct that can get these requirements.

3 An incremental approach to gapping

In this section, I will present an alternative, incremental approach to gapping, which makes use of regular phrase structure rules, but which has in common with the linearization approach that the derivation tree is separated from the constituent structure (although in a different way). The constituent structure, which is not a topic of this paper, is reflected by the entering and popping of structure onto a STACK (Haugereid and Morey, 2012).

3.1 Incremental parsing

The derivation tree in the incremental approach is assumed to consist of binary and unary phrase structure rules where the binary rules have a word as their second daughter. A simplified representation of a transitive clause is given in Figure 1.

There are two things worth mentioning in con-

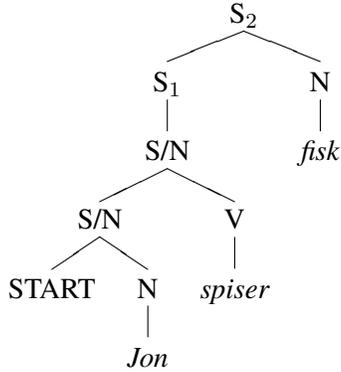


Figure 1: Simplified derivation tree of the transitive sentence *Jon spiser fisk* ('Jon eats fish')

nection to the tree in Figure 1. First, it is assumed that the topic of a main clause is extracted from its canonical position, also when the topic is a subject,¹ and that the SLASH feature percolates down the tree, rather than up, which is the common assumption in HPSG. In Figure 1, the rule S_1 is the subject extraction rule. It is a unary rule that enters an element onto the SLASH list of its daughter. The element on the SLASH list is filled in by the rule that combines START with the noun *Jon*.

Second, valence requirements are handled by means of types (Haugereid, 2009, 2015). The verb *spiser* ('eats') is listed in the lexicon with an underspecified valence type *spise_prd*. This type is a part of a hierarchy of valence types that constrain which constellations of arguments the verb is allowed to appear with. A small part of this type hierarchy is illustrated in Figure 2. It shows that a verb with the type *spise_prd* is compatible with a transitive frame (*_spise_12_rel*) and an intransitive frame (*_spise_1_rel*).

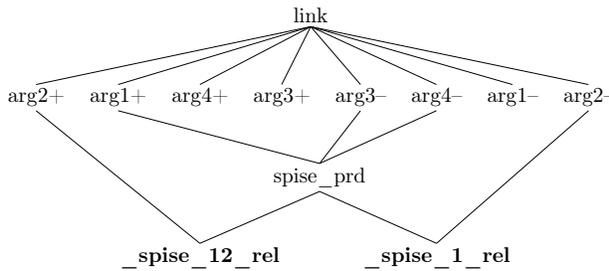


Figure 2: Type hierarchy of valence types

¹This is a common assumption in the literature on Scandinavian syntax.

The phrase at the top of the derivation of a clause is constrained to have negative valence types (*arg1-*, *arg2-*, *arg3-*, and *arg4-*), as shown in (11). As one goes down the tree, valence rules switch these types from negative in the mother to positive in the (first) daughter. In the tree in Figure 1, the types *arg1-* and *arg2-* are switched to *arg1+* and *arg2+* by the two valence rules (marked with subscript in the tree). At the bottom of the tree, the valence type of the verb, which is also the PRED value of its KEYREL (*spise_prd*) is unified with the four valence types *arg1+*, *arg2+*, *arg3-*, and *arg4-*, as shown in (12). This unification is allowed by the type hierarchy (the types *spise_prd*, *arg1+*, *arg2+*, *arg3-*, and *arg4-* have a common subtype), and yields the construction type *_spise_12_rel*, which also serves as the predicate of the relation introduced by the verb.

$$(11) \left[\begin{array}{l} \text{VAL} \\ \text{CMP1|LINK} \quad \text{arg1-} \\ \text{CMP2|LINK} \quad \text{arg2-} \\ \text{CMP3|LINK} \quad \text{arg3-} \\ \text{CMP4|LINK} \quad \text{arg4-} \end{array} \right]$$

$$(12) \left[\begin{array}{l} \text{START} \\ \text{VAL} \\ \text{KEYREL|PRED} \quad \boxed{\text{}} \\ \text{SLASH} \quad \langle \rangle \end{array} \right]$$

3.2 Incremental parsing and coordination

An obvious challenge for the incremental approach is coordination. In HPSG, coordination of full constituents is straightforward, at least as long as the constituents are of the same category. It is the regular coordination rule that holds: $XP \Rightarrow XP \text{ Conj } XP$. Whether the coordinated constituent is a sentence, a VP or an NP, coordination is assumed to be captured by the same rule. In an incremental approach, however, one is forced to start building the second constituent on top of the first, as shown in Figure 3. This means that the rules involved in coordination of full constituents no longer is the combination of two equal constituents. Rather, they mark the end of one constituent and the beginning of a new constituent. This is illustrated in Figure 3 where the rule that adds the coordinator, also marks the beginning of a new

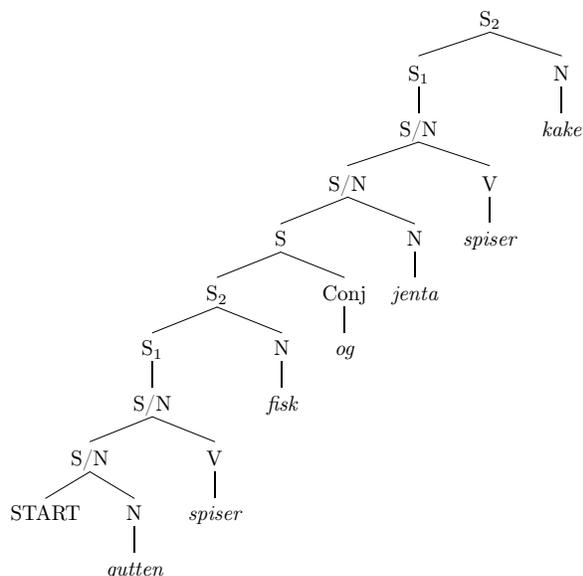


Figure 3: Derivation tree of two coordinated transitive sentences

clause. In practice, the mother of the coordinator has the same constraints as *START* (see (12)).

In order to account for coordination of main clauses and VPs, the hierarchy of phrase types in Figure 4 is created. Most of the constraints of the two types of coordination rules (*conj-s-struct* and *conj-top-struct*) are captured in a supertype *conj-struct*. *conj-struct* takes as its first daughter a structure that has realized all its arguments, that is, the valence types are all negative. The second daughter is a conjunction item, which can be either a conjunction or a comma, in case there are more than two conjuncts. The mother unifies the valence types of the sentence that is to be built next. The type links the two conjuncts with a conjunction relation that is entered onto the C(onstructional)-CONT RELS list. *conj-struct* is underspecified with regard to whether there is an element on the SLASH list or not.

The value of the SLASH list is specified on the two subtypes, *conj-s-struct* and *conj-top-struct*. The type *conj-s-struct* has an empty SLASH list. This means that it has the same status as *START* (see 12), and it initiates the building of a new sentence.

The second subtype, *conj-top-struct*, has an element on the SLASH list which is the topic. This gives it the status of a structure where the topic is realized, but where it is yet to be extracted. The topic of the first daughter is also the topic of the mother,

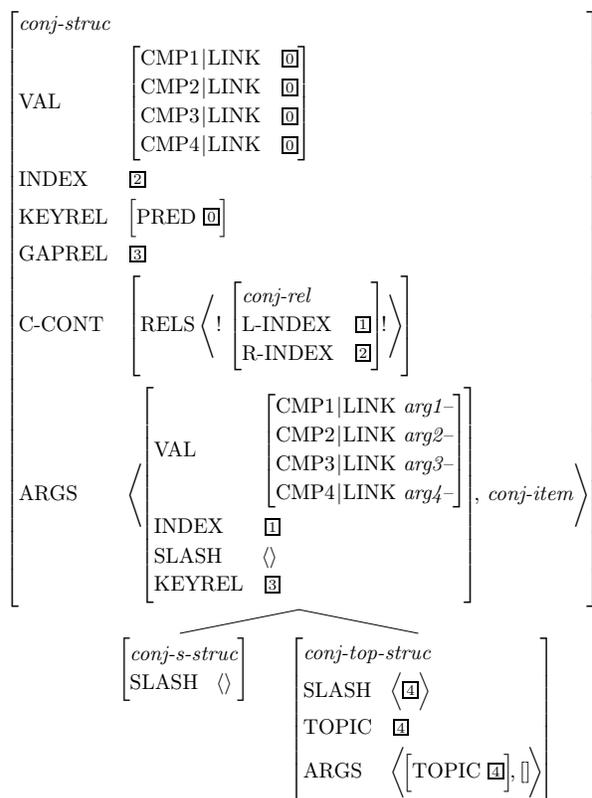


Figure 4: Hierarchy of coordination rules

which means that the two sentences will share topic. This accounts for VP coordination, where the shared topic is the subject, but also similar kinds of coordination where the shared topic is an object or an adjunct.

3.3 Analysis of gapping

In order to account for the gapping phenomena presented in Section 1, I introduce a set of unary rules corresponding to the rules that attach verbs, particles, reflexives and idiomatic words. The rule for eliding verbs is given in Figure 5. It takes as its only daughter a structure that requires a verb (the VBL value is *synsem*), and gives a new structure where there is no longer a verb requirement (the VBL value is *anti-synsem*). In addition, the value of GAPREL of the daughter is unified with the KEYREL. As shown in Figure 4, the GAPREL has as value the KEYREL relation of the first conjunct in a conjunction. This relation is the relation contributed by the main verb. This ensures that the gapping construction has the same relation (and syntactic construction) as the first

conjunct.

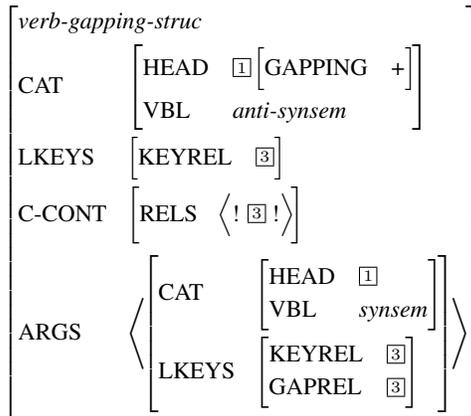


Figure 5: Type for ellided verbs

The incremental design where verbs are treated as a kind of obligatory adjuncts, makes an account of gapping constructions relatively straightforward. Since the contribution of a (main) verb in a regular main clause is to contribute a type which constrains what kinds of constructions it can appear in, the only addition needed is to make the construction type available in the gapping construction. As shown in Figure 5, this type comes from the GAPREL feature. In this way the gapping rule substitutes the verb. The construction type carried over from the first conjunct guarantees that the valence rules that apply in the first conjunct also apply in the second conjunct.

The examples of gapping in sentences with a shared topic (see (9) and (10)) are accounted for by the combination of the *conj-top-struct* rule and the *verb-gapping-struct* rule. The *conj-top-struct* rule takes a full clause as its first daughter and creates a structure with an element on the SLASH list that is unified with the TOPIC of the input clause.

References

Beavers, J. and Sag, I. A. (2004). Coordinate ellipsis and apparent non-constituent coordination. In S. Müller, editor, *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, pages 48–69. CSLI Publications, Stanford.

Chaves, R. P. (2005). A linearization-based approach to gapping. In G. Jger, P. Monachesi,

G. Penn, and S. Wintner, editors, *FG-MOL 2005: The 10th conference on Formal Grammar and The 9th Meeting on Mathematics of Language*, page 14. CSLI.

Crysmann, B. (2008). An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In G. Penn, editor, *Proceedings of FGVienna : The 8th Conference on Formal Grammar, Aug 16-17 2003, Vienna*, pages 45–64, Stanford. CSLI publications.

Flickinger, D. P. (2000). On building a more efficient grammar by exploiting types. *Natural Language Engineering*, 6(1), 15–28.

Haugereid, P. (2009). *Phrasal subconstructions: A constructionalist grammar design, exemplified with Norwegian and English*. Ph.D. thesis, Norwegian University of Science and Technology.

Haugereid, P. (2015). Unique lexical entries in a subconstructional grammar. In S. Müller, editor, *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar, Nanyang Technological University (NTU), Singapore*, pages 110–121, Stanford, CA. CSLI Publications.

Haugereid, P. and Morey, M. (2012). A left-branching grammar design for incremental parsing. In S. Müller, editor, *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon*, pages 181–194.

Kathol, A. (1995). *Linearization-Based German Syntax*. Ph.D. thesis, Ohio State University.

Pollard, C. J. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Reape, M. (1994). Domain union and word order variation in German. In J. Nerbonne, K. Netter, and C. J. Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, number 46 in CSLI Lecture Notes, pages 151–197. CSLI Publications, Stanford University.

Siegel, M., Bender, E. M., and Bond, F. (2016). *Jacy: An implemented grammar of Japanese*. CSLI Studies in Computational Linguistics. CSLI Publications, Stanford University.

The Syntax of the *not only...but also...* Construction: A Linearization Account

Sarah Hye-yeon Lee (sarahhl@usc.edu)
University of Southern California

In this study, the syntactic properties of the *not only...but also...* construction are examined and analyzed. *Not only* has free linearization possibilities in that it can appear in positions not directly adjacent to the focused constituent. This is an idiosyncratic construction which differs from both the *not...but...* construction and the adverbial *only*, because the position of *not only* is freer than that of *not*. Also, *not only*, but not *not* or *only*, triggers negative inversion in the clause-initial position. Therefore, the *not only...but also...* construction should be treated as a construction separate from *not...but...* or *only*.

I propose that the construction is a subtype of the *correlative-coord-ph*, in which each conjunct is headed by a conjunction. Within the first conjunct headed by *not only*, shuffling of *not only* is possible in the word order domain, as long as it precedes the focused element. And the second conjunct headed by *but also* can undergo deletion of shared materials. When *not only* in the first conjunct appears in the clause-initial position, negative inversion occurs. Finally, subject-auxiliary inversion triggered by the clause-initial *not only* is analyzed as a subtype of the *negative-inversion-ph*. Here, I propose a novel subtype: *negative-conjunction-ph*.

1. Linearization possibilities of *not only*

Not only can appear in positions not directly adjacent to the focused constituent. For example, it can appear at a position higher than its “original” position, as in (1). Sentences in (1) were found in corpus (Corpus of Contemporary American English).

- (1) *Not only* too high cases
 - a. Note that you can use this tool **not only** to upgrade Windows 7 or Windows 8.1 PCs, **but also** Windows 10 PCs. (COCA 2015 MAG PCWorld)
 - b. As a professional golfer, he became an international celebrity, known **not only** for his accomplishments on the golf course, **but** his extravagant lifestyle. (COCA 1997 MAG GolfMag)
 - c. Yet somehow I managed **not only** to rope myself into a ski trip, **but** a ski trip with my boss. (COCA 2015 FIC Bk:SlayedOnSlopes)

Not only can also appear at a position lower than its “original” position, as in (2).

- (2) *Not only* too low cases
 - a. they are here to poke and prod and asphyxiate me with the obligation to **not only** understand, to not only come to grips, **but also** to stand up there and explain... and in this case, explain what? (COCA 2015 FIC WarLitArts)
 - b. So definitely, my view of not only my parents, **but** of myself, has changed from, you know, having this son and my two daughters, too. (COCA 2014 SPOK NPR)

(3) shows possible positions of *not* in the *not...but...* construction, and (4) shows possible positions of *not only* in the *not only...but also...* construction.

- (3) Positions of *not* in the *not...but...* construction (Song 2012, p. 54)
Joe succeeded <not> by <? not> selling <*not> books <*not>, but by buying shoes.
- (4) Positions of *not only* in the *not only...but also...* construction
<not only (with subj-aux inversion)> Joe succeeded <not only> by <not only> selling <not only> books <*not only>, but (also) by selling shoes.

The contrast between (3) and (4) shows that *not* is more restricted in its positional possibilities than *not only*. The shaded positions in (4) are impossible in (3). Rather, the positions of *not only* are similar to *either* in disjunction constructions, which exhibits free linear possibilities (Hofmeister 2010), as in (5)-(8).

- (5) <Either> Thomas <either> will <either> write <either> a mystery or he'll write a romance.

- (6) <Either> You'll <either> need to <either> bring <either> a passport or a birth certificate.
- (7) <Either> Congress <either> will <either> pass the legislation or lose our confidence.
- (8) <Either> You <either> can <either> have <either> tea <either> from <either> China or from Tibet.

2. Negative Inversion

Particularly interesting is that unlike in the *not...but...* construction, *not only* (like *either*) can appear at the clause-initial position, in which case subject-auxiliary inversion is obligatorily triggered, as in (9).

- (9) **Not only** did Joe succeed by selling books, but by buying books.

Such a clause-initial position of the conjunction and subject-auxiliary inversion is clearly not possible in the *not...but...* construction, as can be seen from the ungrammaticality of the sentences in (10).

- (10) a. **Not* did Joe succeed by selling books, but by buying shoes.
- b. **Not* was Joe stupid, but lazy.

Note that (11) should be clearly distinguished from the “floating” cases of *not only* or *not* because in (11), the whole constituent (*not (only)* along with the focused constituent) is fronted.

- (11) *Not (only)* in its success does America's expectation lie but its failures.

One may intuitively think that the subject-auxiliary inversion triggered by clause-initial *not only* simply follows from properties of *only*-inversion (Huddleston & Pullum 2002). *Only* has been analyzed as a weak negator (Haegeman 1995; Maekawa 2012), and therefore *only*-inversion in (12) as a negative inversion phenomenon.

- (12) a. *Only his mother* will he obey.
- b. *Only on Sundays* do they eat with their children.

However, the ungrammaticality of the sentences in (13) shows that *only*-inversion and *not only*-inversion cannot be treated as the same phenomena because *only* by itself cannot float to the clause-initial position.

- (13) a. **Only* did Joe succeed by selling books.
- b. **Only* was Joe stupid.

Therefore, I propose that the “floating very high” possibility of *not only* is an idiosyncratic property of the *not only...but also...* construction, which is exhibited neither in the *not...but...* construction nor with *only*.

3. HPSG Analysis

3.1 A new subtype of the *coord-ph*: *correlative-coord-ph*

I analyze the *not only...but also...* construction as a coordination structure, in which each conjunct is marked by a conjunction, similar to Mouret's (2004) analysis of French conjunction doubling (*et* ‘and’...*et* ‘and’..., *soit* ‘either’... *soit* ‘or’). First, the constraints imposed on *coord-ph* are shown in (14).

$$(14) \text{ coord-ph} \rightarrow \left[\begin{array}{l} \text{CONJ} \\ \text{HEAD} \\ \text{VALENCE} \\ \text{SLASH} \\ \text{NON - HD - DTRS} \end{array} \begin{array}{l} \text{null} \\ [1] \\ [2] \\ [3] \\ < \left[\begin{array}{ll} \text{HEAD} & [1] \\ \text{VALENCE} & [2] \\ \text{SLASH} & [3] \end{array} \right], \dots, \left[\begin{array}{ll} \text{HEAD} & [1] \\ \text{VALENCE} & [2] \\ \text{SLASH} & [3] \end{array} \right] > \end{array} \right]$$

Mouret (2004) cross-classifies (French) coordinations into four types: *basic-coord-ph*, *iterative-coord-ph*, *asyndetic-coord* and *doubling-coord-ph*. I propose a new subtype of the *coord-ph* to account for correlative

coordination structures such as *not only...but also...*: **correlative-coord-ph** ((15)). Other types of “edge coordinations” such as *first...then...* can be explained via this phrase type as well.

(15) *correlative-coord-ph* → [NON-HD-DTRS / *nelist*([CONJ [1]]) ⊕ [CONJ [2]]]

I propose that *not only but also-ph* ((16)) is a subtype of the *correlative-coord-ph*.

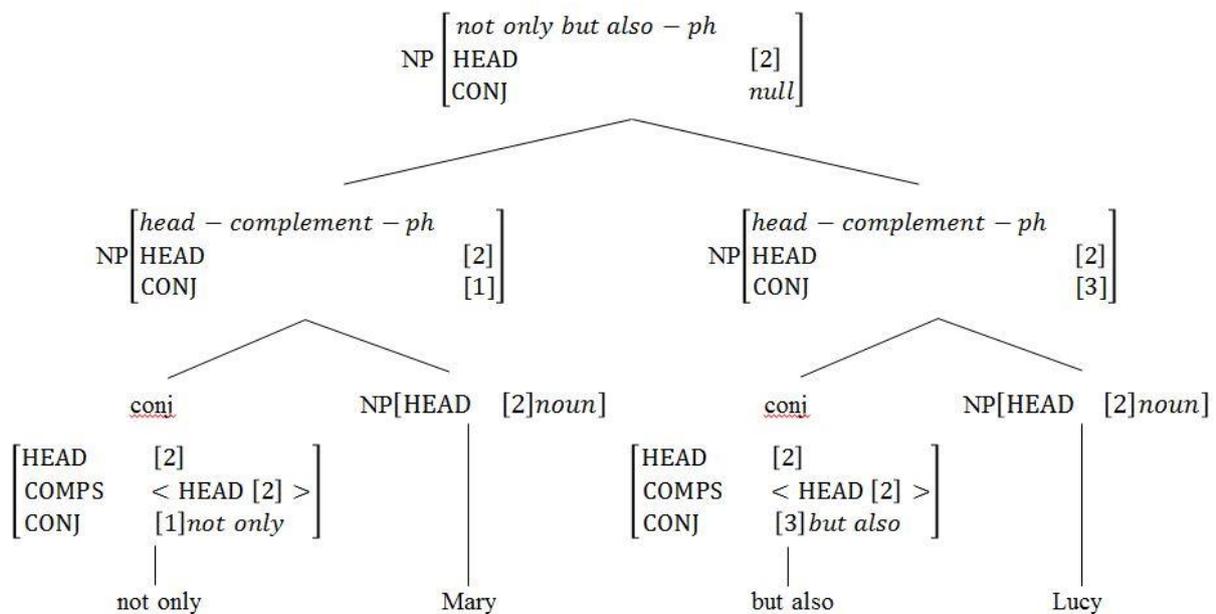
(16) *not only but also-ph* →
[NON-HD-DTRS *nelist*([CONJ *not only*]) ⊕ ([CONJ *but also*)]

The above constraint in (16) can adequately account for the fact that conjuncts headed by *not only* can appear multiple times, as in (17).

(17) they are here to poke and prod and asphyxiate me with the obligation to **not only** understand, to **not only** come to grips, but also to stand up there and explain... and in this case, explain what? (COCA 2015 FIC WarLitArts)

In (18), the structure of *not only Mary but also Lucy* is shown.

(18)



3.2 Linearization of *not only*

I explain the various “floating” positions via a licensing construction that enables shuffling of *not only* within the conjunct, a linear precedence (LP) rule that orders *not only* before the focused element, and coordinate ellipsis. This analysis is an extension of Hofmeister’s (2010) linearization analysis of *either* in disjunction structures.

3.2.1 *Not only* floating too low

In (19) are the constraints on the *not only-ph*.

$$(19) \text{ not only-ph} \rightarrow \left[\begin{array}{l} \text{MOTHER} \left[\begin{array}{l} \text{CONJ} \text{ not only} \\ \text{HEAD} [2] \\ \text{DOM} < [\text{DOM } \delta 1 \circ \delta 2 > \end{array} \right] \\ \text{DTRS} < \left[\begin{array}{l} \text{PHON} < \text{not only} > \\ \text{CONJ} \text{ not only} \\ \text{SYN|VAL} \text{ COMPS} < [1] > \\ \text{HEAD} [2] \\ \text{DOM} \delta 2 \end{array} \right], [1] \left[\begin{array}{l} \text{HEAD} [2] \\ \text{DOM} \delta 1 \end{array} \right] \end{array} \right]$$

The following linear precedence (LP) rule in (20) ensures that *not only* always precedes the focused element in the left conjunct, thereby blocking ungrammatical phrases like (21).

$$(20) \left[\begin{array}{l} \text{PHON} < \text{not only} > \\ \text{CONJ} \text{ not only} \end{array} \right] < [\text{INFO - STRUCT} [\text{FOC } \delta ne - list]]$$

(21) *by selling books not only

(22) shows how the licensing construction for combining *not only* with its complement and the linear precedence rule work.

$$(22) \left[\begin{array}{l} \text{DOM} < < [\text{not only}] > \circ < [\text{by}, [\text{selling}], [\text{books}] > > \\ \text{INFO - STRUCT|FOCUS} < [1] > \end{array} \right]$$

$$\left[\begin{array}{l} \text{PHON} < \text{not only} > \\ \text{DOM} < [\text{not only}] > \\ \text{SYN} \text{ conj} \end{array} \right] \quad \left[\begin{array}{l} \text{PHON} < \text{by, selling, books} > \\ \text{DOM} < [\text{by}, [\text{selling}], [\text{books}] > \\ \text{INFO - STRUCT|FOCUS} < [1][\text{books}] > \end{array} \right]$$

3.2.2 Not only floating too high: Ellipsis

Not only always attaches at the left edge of the second conjunct, without actual “floating”. Shared material at the non-initial conjunct(s) may undergo deletion, as in (23), resulting in what seems like upward floating of *not only* (which is, in fact, only an illusion).

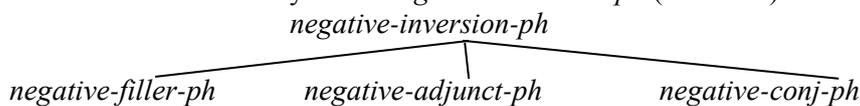
- (23) a. You can use this tool not only [to upgrade Windows 7 or Windows 8.1 PCs], but also [~~to~~
~~upgrade~~ Windows 10 PCs].
 b. I was able not only [to cross the street] but [~~to~~ make it down Fruit Street to Mass General].
 c. He is known not only [for his accomplishments on the golf course], but [~~for~~ his
 extravagant lifestyle].

3.3 Negative Inversion

I propose that the subject-auxiliary inversion that occurs when *not only* floats to the clause-initial position is a negative inversion phenomenon. Maekawa (2012) provides a detailed analysis of negative inversions within the HPSG framework. Maekawa assumes that there are at least two subtypes of the *negative-inversion-ph*: *negative-filler-ph* and *negative-adjunct-ph*.

However, neither of these phrase types can adequately account for the behavior of the *not only...but also...* construction because *not only* in the *not only...but also...* construction is neither a filler constituent nor an adjunct. Therefore, I propose that there is a third subtype of the *negative-inversion-ph*: *negative-conj-ph*. Now, there would be three subtypes of the *negative-inversion-ph*, as in (24).

(24) Constructional hierarchy of the *negative-inversion-ph* (modified)



I also modify constraints imposed on *negative-inversion-ph*, as in (25).

(25) *negative-inversion-ph* (modified) →

[SUBJ	< >
SLASH	{([3][NEG -])}
HD – DTR	[HEAD [INV +] [AUX +]
NON – HD – DTRS	/ [VAL [SUBJ < [1]NP > [COMPS < [2]VP[SLASH {([3]), ...}] >]]]]
	/< [NEG +], [1], [2] >

In (26) are constraints imposed on the new phrase type. Here, the head daughter is the negative conjunction (e.g. *not only*), which takes as its complement the entire following clause. The complement clause of the conjunct is headed by the inverted verb.

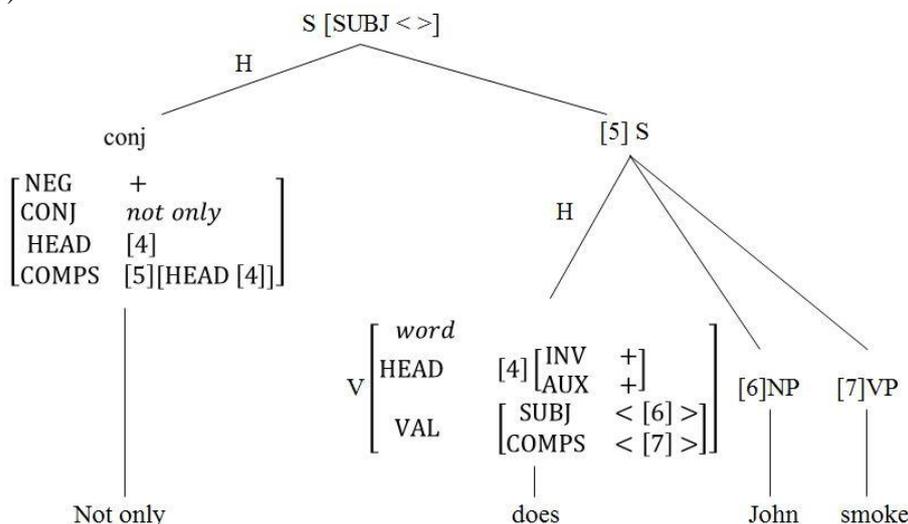
(26) *negative-conj-ph* →

HD – DTR	<i>conj</i>	[NEG + CONJ ¬ <i>null</i> HEAD [4] COMPS [5][HEAD [4]]]
NON – HD – DTRS	[5]S	[HD – DTR [<i>word</i> INV + AUX + SUBJ < [6]NP > COMPS < [7]VP[SLASH {([3]), ...}] >]]]
	NON – HD – DTRS	< [6], [7] >

The tree in (28) shows the structure for the italicized first conjunct in (27).

(27) *Not only does John smoke, but he also drinks.*

(28)



This new subtype—*negative-conj-ph*—can possibly account for other subject-auxiliary inversion phenomena with an initial negative conjunction such as *nor*. For example, the italicized second conjunct headed by *nor* in (29) can be analyzed in the same manner.

(29) John does not drink, *nor does he smoke.*

References

- Haegeman, L. 2000. Negative preposing, negative inversion, and the split CP. In *Negation and polarity*, ed. by Laurence R. Horn and Yasuhiko Kato, 21–61. Oxford: Oxford University Press.
- Hofmeister, P. 2010. A linearization account of either... or constructions. *Natural Language & Linguistic Theory*, 28(2), 275-314.
- Huddleston, R., & Pullum, G. K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Maekawa, T. 2012. An HPSG Approach to Negative Inversion Constructions. *北星学園大学短期大学部北星論集*, 10, 23-42.
- Mouret, F. 2004. The syntax of French conjunction doubling. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 40(2), 193-207.
- Song, Jina. 2012. On the *not...but* construction: A Constraint-based Approach. M.A. Thesis. Seoul National University.

EXPLANATIONS AND “ENGINEERING SOLUTIONS”? ASPECTS OF THE RELATION BETWEEN MINIMALISM AND HPSG

Robert D. Borsley
University of Essex

More than a quarter of a century after its emergence, Chomsky’s Minimalist framework remains the most influential approach to syntax. This is likely to continue if it is not challenged. It is important, therefore, to identify the central features of Minimalism and to show why they are dubious. The issues are clouded by rhetoric, but it is possible to make meaningful comparisons. There are a number of prominent features of Minimalism which are not essential in the sense that it would still be the same framework if they were abandoned. These include its lack of detailed formal analyses, its often cavalier attitude to data, and its procedural idiom. In contrast, the abstractness of Minimalist analyses seems an essential property, without which it would be a very different framework. Also essential is the idea that the properties of syntactic structures stem from their lexical ingredients. Since the grammar is just a few very general mechanisms, there is no real alternative. This contrasts sharply with HPSG with its complex hierarchies of phrase types/constructions. The Minimalist position is in fact not well developed, and there is no reason to think that it is preferable to the HPSG position. A further feature of Minimalism which seems essential is the Move/Internal Merge approach to unbounded dependencies. This works well with typical unbounded dependency examples, but compares poorly with HPSG’s SLASH mechanism where less typical examples are concerned. Thus, it is possible to compare Minimalism and HPSG, and the comparisons favour HPSG.