

Resumption and Extraction in an Implemented HPSG of Hausa

Berthold Crysmann

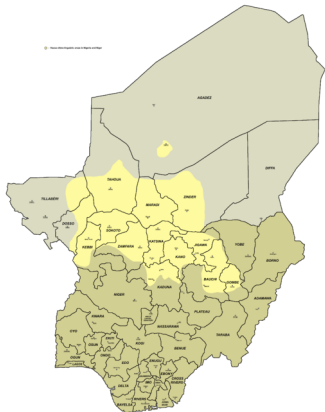
`crysmann@linguist.jussieu.fr`

CNRS, Laboratoire de linguistique formelle (UMR 7110), Paris-Diderot

DELPH-IN 2016

Introduction

- ▶ Hausa is a major Afroasiatic language (Chadic sub-branch) spoken by over 35 million speakers in Northern Nigeria and bordering Niger
- ▶ Unbounded dependency constructions (UDCs) in Hausa feature both
 - ▶ standard extraction (filler-gap dependencies)
 - ▶ resumptive pronoun strategy
- ▶ Resumptive elements include
 - ▶ free pronouns
 - ▶ bound pronominal affixes
 - ▶ zero anaphora (see below)



Resumption vs. gap strategy

- ▶ Choice of extraction strategy partially determined by the governing head
- ▶ Possessor complements of nouns only permit resumption

(1) $w\grave{a}_i$ ka àuri 'ya-r - $s\grave{a}_i$ / 'yā $*\emptyset_i$?
who 2.M.CMPL marry daughter.F-of.F -3S.M daughter
'Whose daughter did you marry?' (Jaggar, 2001)

- ▶ Complements of true prepositions equally do not permit gap strategy

(2) $s\grave{a}nd\bar{a}_i$ sukà d\`ok\`e shì dà $it\grave{a}_i$ / $*\emptyset_i$
stick 3P.CPL beat 3S.DO with 3S.F
'It was a stick they beat him with.' (Jaggar, 2001)

Human direct objects

- ▶ Direct objects of verbs, dynamic nouns, and verbal nouns can extract by way of a filler-gap dependency
- ▶ Overt resumptives are considered marginal

“Deletion is [...] the strongly preferred strategy for relativisation on direct objects.” (Jaggar, 2001, p. 534)

- (3) a. yāròn_i dà sukà dōkā \emptyset_i yanà asìbitì
boy REL 3P.CPL beat up 3.S.M.CONT hospital
‘The boy they beat up is in hospital’ (Jaggar, 2001, p. 534)
- b. gà yārinỳàr_i dà nakè sô \emptyset_i
there is girl REL 1.S.CONT want.VN
‘There’s the girl I love.’ (Jaggar, 2001, p. 534)
- c. ìnā littāfìn_i dà kakè màganà \emptyset_i
where book REL 2.S.M.CONT talking
‘Where is the book you’re talking about?’ (Jaggar, 2001, 534)

Human direct objects

- ▶ Marginality of resumption in highest clause familiar from subjects in Hebrew (Borer, 1984) and Irish (McCloskey, 1990)
- ▶ Resumption fine for more deeply embedded human objects (non-islands)

(4) mùtumìn_i dà dǎ̀libai sukà san [cêwā mālàma-r-sù
man REL students 3P.CPL know COMP teacher-L.F-3P.GEN
tanã sô-n-sà_i / sô Ø_i]
3.S.F.CONT like.VN-L-3.S.M.GEN / like.VN
'the man that the students know that their teacher likes' (Newman,
2000, 539)

Human direct objects

- ▶ Resumptives also found in across-the-board extraction from coordination
- ▶ ATB extraction in Hausa allows mixing of gap and resumptive strategy

- (5) [àbōkī-n-ā]_i dà [[na zìyartà \emptyset _i] àmmā [bàn sāmē shì_i
friend-L-1.S.GEN REL 1.S.CPL visit but 1.S.NEG.CPL find 3.S.M.DO
à gidā ba]]
at home NEG
'my friend that I visited but did not find at home' (Newman, 2000, p. 539)

Human direct objects

- ▶ Resumption required with long relativisation
 - ▶ from complements of non-bridge verbs
 - ▶ from relative (or wh) clauses

(6) gâ yârân_i dà Àli ya radâ minì [wai ya gan-sù_i
there are children REL Ali 3.S.CPL whisper 1.S.IO COMP 3.S.CPL see-3P.DO
/ *ganì Ø gida-n giyà]
/ see Ø house-L beer

‘Here are the children that Ali whispered to me that he saw in the bar.’
(Tuller, 1986, 169)

(7) gâ mùtumìn_j dà ka ga yārinyà_i [dà Ø_i ta
here.is man REL 2.S.M.CPL see girl REL 3.S.F.CPL
san shì_j / *sanì Ø_j]
know 3.S.M.DO / know Ø

‘Here’s the man that you saw the girl that knows him.’ (Tuller, 1986, 85)

Indirect objects

- ▶ Both resumption and gaps possible with indirect objects

(8) mutânên_i dà sukà ki sayar musù / wà \emptyset dà àbinci sukà fita
men REL 3P.CPL refuse sell 3P.IO / IOM with food 3P.CPL left
'the men they refused to sell food to left.' (Jaggar, 2001, 534)

- ▶ Resumption obligatory with long relativisation

(9) gâ tábōbîn_j dà Àli ya san mùtumìn_i [dà \emptyset _i zâi yī
here.is cigarettes REL Ali 3S.M.CPL know man REL 3S.M.FUT do
musù_j / *wà \emptyset _j kwālī]
3P.IO / IOM \emptyset box
'Here are the cigarettes that Ali knows the man that will make a box for.'
(Tuller, 1986, 84)

Null pronouns

- ▶ Hausa has null subjects and null **non-human** direct objects

(10) a. Kā ga littāfi-n Mūsa?
2S.M.CPL see book-of Musa
'Did you see Musa's book?'

b. Ī, nā gan shì. / Ī, nā ganī ∅
Yes 1.S.CPL see 3S.M Yes 1.S.CPL see
'Yes, I saw it.'

(Tuller, 1986, 61)

(11) a. Kā ga kanè-n Mūsa?
2S.M.CPL see brother-of Musa
'Did you see Musa's brother?'

b. Ī, nā gan shì. / *Ī, nā ganī ∅
Yes 1.S.CPL see 3S.M Yes 1.S.CPL see
'Yes, I saw him.'

(Tuller, 1986, 62)

- ▶ Interpretation of zero arguments is specific (Jaggar, 2001; Tuller, 1986)

Null resumptives

- ▶ Long relativisation out of relatives possible with pro-dropped arguments (subject and non-human direct object); cf. (Tuller, 1986)

(12) mùtumìn_i dà ka san littāfin_j [dà \emptyset _i ya rubùtā \emptyset _j]
man REL 2S.M.CPL know book REL 3S.M.CPL write
'the man that you know the book (he) wrote' (Tuller, 1986, 81)

(13) littāfin_i dà ka san mùtumìn_j [dà \emptyset _j ya rubùtā \emptyset _i]
book REL 2S.M.CPL know man REL 3S.M.CPL write
'the book that you know the man who wrote (it)' (Tuller, 1986, 81)

- ▶ Likewise, argument-drop permits relativisation out of wh-islands

(14) mùtumìn_i dà ka san [mè_j \emptyset _i ya rubùtā \emptyset _j]
man REL 2S.M.CPL know what 3S.M.CPL write
'the man that you know what (he) wrote' (Tuller, 1986, 80)

(15) littāfin_i dà ka san [wà_j \emptyset _j ya rubùtā \emptyset _i]
book REL 2S.M.CPL know who 3S.M.CPL write
'the book that you know who wrote (it)' (Tuller, 1986, 80)

Islands for wh-fronting

- ▶ By contrast, wh-extraction cannot escape islands, e.g. relatives

(16) * wà nè mùtùm_i ka bā nì littāfin_j dà Ø_i ya rubùtā
which man 2S.M.CPL give me book REL 3S.M.CPL write
Ø_j

‘Which man did you give me the book that wrote’ (Tuller, 1986, 81)

(17) * wà nè littāfi_j ka san wà_i Ø_i ya rubùtā Ø_j
which book 2S.M.CPL know who 3S.M.CPL write
‘which book do you know who wrote’ (Tuller, 1986, 80)

- ▶ Overt resumptives do not improve island sensitivity of wh-phrases

(18) wà_j ka yi màgànà dà shī_j
who 2S.M.CPL do talking with 3S.M
‘Who did you talk with?’ (Tuller, 1986, 158)

(19) * wà_j ka san mâtâr_i [dà Ø_i ta yi màgànà dà shī_j]
who 2S.M.CPL know woman REL 3S.F.CPL do talking with 3S.M

Triple relativisation

- ▶ Tuller (1986) cites a marginally acceptable example with triply nested relativisation

(20) ? gâ mâtâr_i dà ka bā nì littāfin_j dà mālāmai
here.is woman REL 2S.M.CPL give me book REL teachers
sukà san mùtumìn_k dà Ø_i ta rubùtā wà Ø_k Ø_j
3P.CPL know man REL 3S.F.CPL write for
'Here's the woman that you gave me the book the teachers
know the man she wrote it for.' (Tuller, 1986, 84)

SLASH dependencies in DELPH-IN grammars I

- ▶ Unbounded dependencies in HPSG are mediated via a non-local set-valued feature SLASH, relating properties of the filler to properties at the gap site
- ▶ Most HPSG practitioners (Alotaibi & Borsley, 2013; Taghvaipour, 2005; Crysmann, 2012) agree, based on ATB facts, that resumption should be regarded as a SLASH dependency
- ▶ Following Sag (1997); Ginzburg & Sag (2000), SLASH passing is
 - ▶ lexical:
gaps are introduced on the argument structure of the head
 - ▶ head-driven:
heads determine their SLASH value from those of their arguments

(21) SLASH amalgamation (Ginzburg & Sag, 2000)

$$\left[\begin{array}{l} \text{SYNSEM} \left[\text{NLOC} \left[\text{SL } \boxed{1} \cup \dots \cup \boxed{n} \right] \right] \\ \text{ARG-ST} \left\langle \left[\text{NLOC} \left[\text{SL } \boxed{1} \right] \right], \dots \left[\text{NLOC} \left[\text{SL } \boxed{n} \right] \right] \right\rangle \end{array} \right]$$

SLASH dependencies in DELPH-IN grammars II

- ▶ ERG (Copestake & Flickinger, 2000) and the Grammar Matrix Bender et al., 2002) closely follow the standard HPSG treatment
 - ▶ sets are approximated by difference lists
 - ▶ amalgamation is effected by four lexical types (e.g. *basic-two-arg*)

$$(22) \quad \left[\begin{array}{c} o\text{-diff-list} \\ \text{LIST} \quad [1] \\ \text{LAST} \quad [1] \end{array} \right] \left[\begin{array}{c} 1\text{-diff-list} \\ \text{LIST} \quad \left[\begin{array}{c} \text{FIRST} \quad [] \\ \text{REST} \quad [1] \end{array} \right] \\ \text{LAST} \quad [1] \end{array} \right] \left[\begin{array}{c} \textit{basic-two-arg} \\ \text{SYNSEM} \quad \left[\text{NLOC} \left[\text{SL} \left[\begin{array}{c} \text{LIST} \quad [1] \\ \text{LAST} \quad [3] \end{array} \right] \right] \right] \\ \text{ARG-ST} \quad \left\langle \left[\begin{array}{c} \text{LIST} \quad [1] \\ \text{LAST} \quad [2] \end{array} \right], \left[\begin{array}{c} \text{LIST} \quad [2] \\ \text{LAST} \quad [3] \end{array} \right] \right\rangle \end{array} \right]$$

- ▶ Restriction to at most one simultaneous SLASH dependency impedes treatment of long relativisation in Hausa
- ▶ Limitation even affects English:

$$(23) \quad [\text{A violin this well crafted}]_1 \text{ even } [\text{the most difficult sonata}]_2 \text{ will be easy } [\text{to play } _2 \text{ on } _1] ? \quad (\text{Pollard \& Sag, 1994, 169})$$

Multiple SLASH dependencies: Island sensitivity

- ▶ ATB facts suggest that resumptives and true gaps are compatible in principle (\Rightarrow both are SL elements)
- ▶ Island effects show that
 - ▶ only relatives footed by a resumptive can escape islands
 - ▶ both gap dependencies and phrasal fillers are island-sensitive
- ▶ Distinguish gaps/resumptives and wh/relatives by the amount being minimally shared

INDEX (resumptive, relative) vs. LOC (gap, wh-filler)

$$(24) \quad \left[\begin{array}{l} \text{gap-or-res} \\ \text{LOC} \quad \left[\text{CONT.HOOK.INDEX} \quad \boxed{1} \right] \\ \text{NLOC} \quad \left[\text{SL} \left\langle ! \left[\text{CONT.HOOK.INDEX} \quad \boxed{1} \right] ! \right\rangle \right] \end{array} \right] \quad (25)$$

$$\left[\begin{array}{l} \text{gap} \\ \text{LOC} \quad \boxed{1} \text{ full-local} \\ \text{NLOC} \quad \left[\text{SL} \left\langle ! \boxed{1} ! \right\rangle \right] \end{array} \right] \left[\begin{array}{l} \text{strict-res} \\ \text{NLOC} \quad \left[\text{SL} \left\langle ! \text{light-local} ! \right\rangle \right] \end{array} \right]$$

$$\left[\begin{array}{l} \text{local} \\ \text{CONT} \quad \text{mrs} \end{array} \right] \\ \text{full-local} \quad \text{light-local} \\ \left[\begin{array}{l} \text{CAT} \quad \text{cat} \end{array} \right]$$

Multiple SLASH dependencies: Launching

- ▶ Standard gap-type dependencies are introduced by CELR
- ▶ Full reentrancy of SL element with a dependent's LOC value

(26) Complement extraction

$$\left[\begin{array}{l} \text{SS} \left[\text{LOC} \left[\text{CAT} \left[\text{VAL} \left[\text{COMPS} \quad \boxed{I} \right] \right] \right] \right] \\ \text{DTR} \left[\text{SS} \left[\text{LOC} \left[\text{CAT} \left[\text{VAL} \left[\text{COMPS} \quad \langle \text{gap} \mid \boxed{I} \rangle \right] \right] \right] \right] \right] \end{array} \right]$$

- ▶ Lexical rules for resumption are crucially underspecified:
compatible with both full (wh) fillers and light index dependencies

(27) Resumption

$$\left[\begin{array}{l} \text{SS} \left[\text{LOC} \left[\text{CAT} \left[\text{VAL} \left[\text{COMPS} \quad \boxed{I} \right] \right] \right] \right] \\ \text{DTR} \left[\text{SS} \left[\text{LOC} \left[\text{CAT} \left[\text{VAL} \left[\text{COMPS} \quad \langle \text{gap-or-res} \mid \boxed{I} \rangle \right] \right] \right] \right] \right] \end{array} \right]$$

Multiple SLASH dependencies: Retrieval I

- ▶ Multiple simultaneous unbounded dependencies require recursive perusal of SLASH list
- ▶ HaG recognises exactly two constructions for retrieval:
 - ▶ classic filler-head structures (for wh- and focus fronting) identifies TO-BIND.FILL with filler's entire LOC value
 - ▶ relative complementiser (cf. Alotaibi & Borsley, 2013) identifies index in TO-BIND.FILL with index of antecedent noun (in REL)

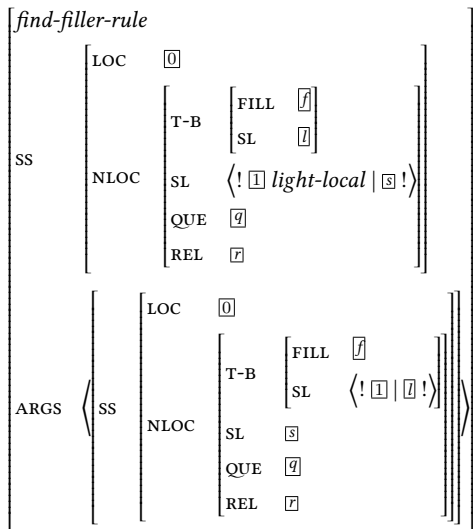
(28)

$$\left[\begin{array}{l} \text{filler-head-rule} \\ \text{SS.NLOC} \left[\begin{array}{l} \text{T-B} \left[\begin{array}{l} \text{FILL} \langle \boxed{I} \rangle \\ \text{SL} \boxed{S} \end{array} \right] \\ \text{SL} \langle \text{!!} \rangle \end{array} \right] \\ \text{FILLER-DTR} \left[\text{SS} \left[\text{LOC} \boxed{I} \right] \right] \\ \text{HD-DTR} \left[\text{SS} \left[\text{NLOC} \left[\text{SL} \boxed{S} \right] \right] \right] \end{array} \right] \left[\begin{array}{l} \text{rel-complementiser-lex} \\ \text{SS} \left[\begin{array}{l} \text{NLOC} \left[\begin{array}{l} \text{T-B} \left[\begin{array}{l} \text{FILL} \langle \text{CONT} \left[\text{HOOK} \left[\text{INDEX} \boxed{I} \right] \right] \rangle \right] \\ \text{SL} \boxed{S} \\ \text{REL} \langle \text{!} \boxed{I} \text{ ref-index !} \rangle \end{array} \right] \\ \text{SL} \langle \text{!!} \rangle \end{array} \right] \end{array} \right] \end{array} \right] \left[\begin{array}{l} \text{ARG-ST} \langle \text{S} \left[\text{NLOC} \left[\text{SL} \boxed{S} \right] \right] \rangle \end{array} \right]
 \end{array}$$

Multiple SLASH dependencies: Retrieval II

- ▶ *find-filler-rule*
recursively skips over
light-local T-B.SL
elements
- ▶ Skipped T-B.SL elements
are restored to main SL
- ▶ Rule effectively rules in
long relativisation with
resumption (*light-local*)
- ▶ Rule enforces extraction
island for gap
dependencies (*full-local*)
 - ▶ wh/focus fronting
 - ▶ relativisation
without resumptives

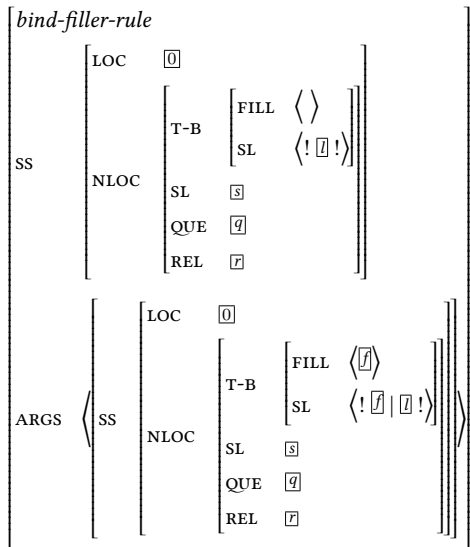
(29)



Multiple SLASH dependencies: Retrieval III

- ▶ *bind-filler-rule* performs actual identification of filler's properties with some first element on T-B.SL
- ▶ rule itself is agnostic as to the nature of the filler (i.e. *full-local* vs. *light-local*)
- ▶ bound SL element is not passed on

(30)



Adjunct extraction and termination

- ▶ As for argument extraction, number of simultaneous SLASH dependencies bounded by argument structure
- ▶ Number of adjunct gaps inherently unbounded: termination?
- ▶ Hausa permits at most one simultaneous gap dependency: long relativisation restricted to resumptive dependency
- ▶ Generally restrict remainder of SL list to *light-local* elements upon introduction of a *full-local* element

$$(31) \left[\begin{array}{l} \text{SS} \left[\text{NLOC.SL} \left[\begin{array}{l} \text{LIST} \langle \boxed{1} \text{ full-local} \mid \boxed{S} \rangle \\ \text{LAST} \boxed{I} \end{array} \right] \right] \\ \text{DTRS} \left\langle \left[\text{SS.NLOC.SL} \left[\begin{array}{l} \text{LIST} \boxed{S} \text{ list(index-local)} \\ \text{LAST} \boxed{I} \end{array} \right] \right] \right\rangle \end{array} \right]$$

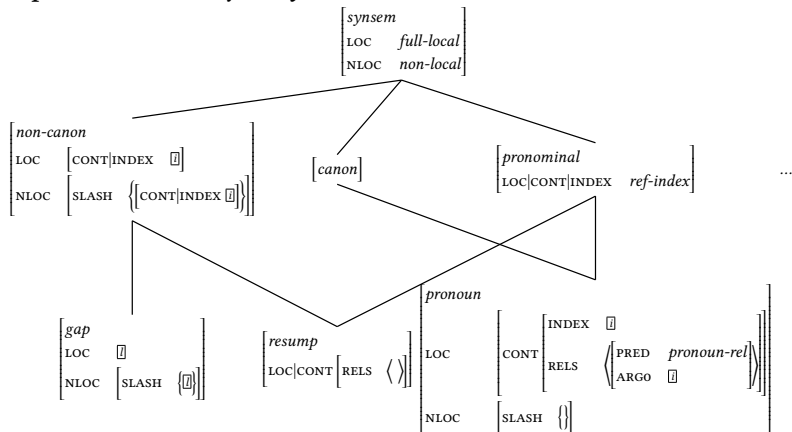
- ▶ Weaker constraint for English, cf. (23):
restrict SL to at most one adjunct gap (*list*([CAT.HD.MOD < >]))

Shortcomings of the ambiguity approach

- ▶ Duplication of
 - ▶ pronominal lexical items
 - ▶ pronominal affixation rules (accusative, dative, genitive)
 - ▶ zero pronoun rules
- ▶ Baseline implementation kept cost at bay by adopting an allomorphy approach to the 3 sets of pronominal affixes
- ▶ Lean formalisms force hidden costs out into the open
- ▶ Misses McCloskey's generalisation:
in languages with resumption, resumptives are always the ordinary pronouns of the language
- ▶ Alternatives: Borsley (2010); Alotaibi & Borsley (2013)
 - ▶ Locates choice between resumptive and pronoun function outside pronominal element :-)
 - ▶ Cannot account for difference in semantics :-)
 - ▶ Does not play nice with SLASH amalgamation

Resumption by underspecification

- ▶ Solution:
 - ▶ decision on function is property of governing head (cf. Borsley)
 - ▶ combine with underspecification of *pronominal-synsem* (generalises across bound and free pronouns)
 - ▶ Expanded hierarchy of *synsem*



Implementation

- ▶ Pronominals lexically underspecified (via synsem)
- ▶ Obligatory lexical rule layer performs choice between slashed and unslashed complements on governing head
 - ▶ disambiguation restricted to subject and direct object valencies
 - ▶ obliques turn out to be selected by markers or prepositions
- ▶ Generation
 - ▶ Problem: *pronoun_rel* not detectable by static predicate coverage check
 - ▶ Solution:
 - deterministically remove relation from internal MRS in pre-generation fixup
 - post-generation check performed on external MRS

Conclusion

- ▶ Linguistically complete treatment of resumption in Hausa necessitates low-level revisions to current treatment of extraction in HPSG grammar engineering:
 - ▶ island-insensitive long relativisation gives rise to multiple simultaneous SLASH dependencies
 - ▶ Hard-wired restriction to maximally singleton SLASH replaced by a restriction on number of simultaneously open gap dependencies
 - ▶ Constraint motivated by island effects provides a solution for termination and efficiency issues
- ▶ Underspecification significantly improves over ambiguity approach to resumptives (Crysmann, 2015)
 - ▶ Disjunctions between ordinary pronouns and resumptives fully eliminated
 - ▶ Provides an account of McCloskey's generalisation
 - ▶ Semantics identical for gap and resumptive strategies (in contrast to Asudeh, 2004)

Shi ke nan. Kurunkus.

References I

- Alotaibi, Mansour & Robert D. Borsley. 2013. Gaps and resumptive pronouns in Modern Standard Arabic. In Stefan Müller (ed.), *Proceedings of the 20th international conference on head-driven phrase structure grammar, freie universität berlin*, 6–26.
<http://cslipublications.stanford.edu/HPSG/2013/alotaibi-borsley.pdf>.
- Asudeh, Ash. 2004. *Resumption as resource management*: Stanford University dissertation.
- Bender, Emily M., Dan Flickinger & Stephan Oepen. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammar. In John Carroll, Nelleke Oostdijk & Richard Sutcliffe (eds.), *Proceedings of the workshop on grammar engineering and evaluation at the 19th international conference on computational linguistics*, 8–14.
- Borer, Hagit. 1984. Restrictive relatives in Modern Hebrew. *NLLT* 2. 219–260.
- Borsley, Robert D. 2010. An HPSG approach to Welsh unbounded dependencies. In Stefan Müller (ed.), *Proceedings of the 17th international conference on head-driven phrase structure grammar, université paris diderot, paris 7, france*, 80–100. Stanford, CA: CSLI Publications.
<http://cslipublications.stanford.edu/HPSG/2010/borsley.pdf>.
- Copetake, Ann & Dan Flickinger. 2000. An open-source grammar development environment and broad-coverage English grammar using HPSG. In *Proceedings of the second conference on language resources and evaluation (lrec-2000)*, Athens.
- Crysmann, Berthold. 2012. Resumption and island-hood in Hausa. In Philippe de Groote & Mark-Jan Nederhof (eds.), *Formal grammar. 15th and 16th international conference on formal grammar, fg 2010 copenhagen, denmark, august 2010, fg 2011 ljubljana, slovenia, august 2011*, vol. 7395 Lecture Notes in Computer Science, 50–65. Springer.

References II

- Crysmann, Berthold. 2015. Resumption and extraction in an implemented hpsg of Hausa. In *Proceedings of the acl-ijnlp workshop on grammar engineering across frameworks (geaf-2015), Beijing, China*, ACL.
- Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations. the form, meaning, and use of English interrogatives*. Stanford: CSLI Publications.
- Jaggar, Philip. 2001. *Hausa*. Amsterdam: John Benjamins.
- McCloskey, J. 1990. Resumptive pronouns, a'-binding and levels of representation in Irish. In R. Hendrick (ed.), *The syntax of the modern Celtic languages*, vol. 23 Syntax and Semantics, Academic Press.
- Newman, Paul. 2000. *The Hausa language. an encyclopedic reference grammar*. New Haven, CT: Yale University Press.
- Pollard, Carl & Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford: CSLI and University of Chicago Press.
- Sag, Ivan. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431-484.
ftp://ftp-csli.stanford.edu/linguistics/sag/rel-pap.ps.gz.
- Taghvaipour, Mehran. 2005. *Persian relative clauses in Head-driven Phrase Structure Grammar*: University of Essex dissertation.
- Tuller, Laurice A. 1986. *Bijjective relations in Universal Grammar and the syntax of Hausa*. Ann Arbor: UCLA dissertation.