Resumption and Extraction in an Implemented HPSG of Hausa

Berthold Crysmann
CNRS, Laboratoire de linguistique formelle (UMR 7110)
Université Paris Diderot (Paris 7),
Case 7031, 5 rue Thomas, F-75205 Paris cedex 13
crysmann@linguist.univ-paris-diderot.fr

Abstract

In this paper, we describe the treatment of extraction in HaG, an emerging computational grammar of Hausa, concentrating on the intricate patterns of interaction between resumptive and gap strategies. We shall argue with Tuller (1986) that Hausa resumption (both overt and covert) opens up the possibility for relativisation only to escape well-attested extraction islands in the language. As suggested by the mutual compatibility of gaps and resumptives in ATB extraction, however, we shall conclude that both strategies must be regarded as unbounded dependencies (UDCs) to be modelled via the SLASH feature in HPSG. We shall discuss how the treatment of UDCs has been generalised, in HaG, to permit more than one simultaneous SLASH dependency, and focus in particular on how the distinction between true gaps and resumptive SLASH elements can be exploited to address efficiency issues.

1 Resumptives and gaps in Hausa extraction

Like many other languages, Hausa makes use of extraction in a variety of constructions, including relative clause formation, matrix and embedded wh-questions, and focus fronting. Alongside gap strategies, familiar from English, Hausa also employs resumption, marking the extraction site with a pronominal.

(1) wǎ ka àuri 'ya *(-r -sà) ?
who 2.M.CMPL marry daughter.F -of.F-3.S.M
'Whose daughter did you marry?' (Jaggar, 2001)

(2) sàndā sukà dǒkē shì dà *(ita)
stick 3P.CPL beat 3S.DO with 3S.F
'It was a stick they beat him with.' (Jaggar, 2001)

The distribution of gaps and resumptives partly overlap: while in some contexts only a resumptive strategy is possible, e.g. with possessors of nouns (1), or with complements of non-locative prepositions (2), extraction of core arguments (direct/indirect objects) in general permits both strategies, with a clear preference for zero expression in the case of direct objects (Newman, 2000; Jaggar, 2001) for short extraction.

(3) mutǎnên dà sukà ƙi sayar musù / wà ∅ dà
men REL 3.P.CPL refuse sell to.them / to with
àbinci sukà fìta
food 3.P.CPL left
'the men they refused to sell food to left.' (Jaggar, 2001, p. 534)

As stated by Crysmann (2012), the preference for direct object gaps, however, is much reduced in slightly more complex cases, involving, e.g. Across-The-Board (ATB) extraction or long-distance relativisation (see the discussion below), making resumption a natural, if not the only option. As shown in (4), an overt resumptive is retained in the second conjunct.

(4) [àbōkī-n-ā]$_i$ dà [[na zìyartǎ ∅$_i$] àmmā
friend-L-1.S.GEN REL 1.S.CPL visit but
[bàn sǎmē shì$_i$ à gidā ba]]
1.S.NEG.CPL find 3.S.M.DO at home NEG
'my friend that I visited but did not find at home' (Newman, 2000, p. 539)

Example (4) further illustrates that extraction from coordinate structures in Hausa appears to treat resumptives on a par with gaps, as far as the ATB constraint is concerned. Another important observation relates to the possibility of ATB extraction to target different grammatical functions in both conjunct, as illustrated in (5).

(5) mùtumìn$_i$ dà na bā shì$_i$ aro-n
man REL 1.S.CPL give 3.S.M.DO lending-L
bàrgō-nā àmmā duk dà hakà ∅$_i$ yakè
blanket-L.1.S.G but in spite of that ∅ 3.S.M.CONT
jîn sanyī
feel-L cold
'the man whom I lent my blanket but who still felt cold' (Newman, 2000, p. 540)

Proceedings of the Grammar Engineering Across Frameworks (GEAF) Workshop, 53rd Annual Meeting of the ACL and 7th IJCNLP, pages 65–72,
Beijing, China, July 26-31, 2015.

A central property of Hausa resumption is that it permits long distance relativisation out of extraction islands: these include relative clauses, embedded wh-clauses, subject clauses, and complement clauses of non-bridge verbs (see Tuller (1986) for the full set of data). We illustrate here on the basis of embedded relative clauses: as shown below, relativisation of an indirect (6) or human direct object (7) out of relatives is fine, provided there is a resumptive in situ.

(6) Gà tābōbîn$_j$ dà Àli ya san mùtumìn$_i$
here.is cigarettes REL Ali 3.S.M.CPL know man
dà $\emptyset_i$ zâi yī musù$_j$ / *wà $\emptyset_j$ kwālī
REL 3.S.M.FUT do to.them / to $\emptyset$ box
'Here are the cigarettes that Ali knows the man that will make a box for.' (Tuller, 1986, p. 84; tone added)

(7) Gà mùtumìn$_j$ dà ka ga yārinyàr$_i$ dà $\emptyset_i$
here.is man REL 2.S.M.CPL see girl REL
ta san shì$_j$ / *sanī $\emptyset_j$
3.S.F.CPL know him / know
'Here's the man that you saw the girl that knows him.' (Tuller, 1986, p. 85; tone added)

The grammar of extraction in Hausa heavily interacts with argument drop: as discussed by Tuller (1986), Hausa allows pro drop not only with subjects, but also with non-human direct objects, which receive a specific, i.e. non-generic interpretation (Jaggar, 2001). Subject properties are identified by agreement marking on the discrete TAM markers.

(8) a. Kā ga littāfì-n Mūsa?
2.S.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 3.S.M Yes 1.S.CPL see
'Yes, I saw it.' (Tuller, 1986, p. 61; tone added)

(9) a. Kā ga ƙanè-n Mūsa?
2.S.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 3.S.M Yes 1.S.CPL see
'Yes, I saw him.' (Tuller, 1986, p. 62; tone added)

As discovered by Tuller (1986), the possibility for relativisation to escape what are otherwise extraction islands in the language extends from overt resumptives to zero pronominals. I.e., she observes that non-human direct objects, which can be freely pro-dropped, do permit long relativisation out of islands even without an overt resumptive, whereas direct objects with human reference do so only if realised overtly by a direct object pronominal affix.

(10) mùtumìn$_i$ dà ka san littāfìn$_j$ [dà $\emptyset_i$
man REL 2.S.M.CPL know book REL
ya rubùtā $\emptyset_j$]
3.S.M.CPL write
'the man that you know the book (he) wrote' (Tuller, 1986, p. 81)

(11) littāfìn$_i$ dà ka san mùtumìn$_j$ [dà $\emptyset_j$
book REL 2.S.M.CPL know man REL
ya rubùtā $\emptyset_i$]
3.S.M.CPL write
'the book that you know the man who wrote (it)' (Tuller, 1986, p. 81)

The very same can be shown to hold for wh islands: again, relativisation out of wh clauses is possible for subjects, and for non-human direct objects, even without an overt resumptive.

(12) mùtumìn$_i$ dà ka san [mẽ$_j$ $\emptyset_i$ ya
man REL 2.S.M.CPL know what 3.S.M.CPL
rubùtā $\emptyset_j$]
write
'the man that you know what (he) wrote' (Tuller, 1986, p. 80)

(13) littāfìn$_i$ dà ka san [wà̀$_j$ $\emptyset_j$ ya rubùtā
book REL 2.S.M.CPL know who 3.S.M.CPL write
$\emptyset_i$]
'the book that you know who wrote (it)' (Tuller, 1986, p. 80)

The converse, however, is not true: wh phrases never extract out of either relative or embedded wh clauses, regardless of the presence of overt or covert resumptives. Examples (14) and (15) illustrates this for subjects and non-human direct objects, whereas (17) provides evidence that resumption (here with an oblique) does not improve acceptability.

(14) * wànè mùtûm$_i$ ka bā nì littāfìn$_j$ dà $\emptyset_i$
which man 2.S.M.CPL give me book REL
ya rubùtā $\emptyset_j$
3.S.M.CPL write
'Which man did you give me the book that wrote' (Tuller, 1986, p. 81; tone added)

(15) * wànè littāfì$_j$ ka san wà̀$_i$ $\emptyset_i$ ya
which book 2.S.M.CPL know who 3.S.M.CPL
rubùtā $\emptyset_j$
write
'which book do you know who wrote' (Tuller, 1986, p. 80; tone added)

(16) wà̀$_j$ ka yi màganà̀ dà shī$_j$
who 2.S.M.CPL do talking with 3.S.M
'Who did you talk with?' (Tuller, 1986, p. 158)

(17) * wà̀$_j$ ka san mà̀târ$_i$ [dà $\emptyset_i$ ta yi
who 2.S.M.CPL know woman REL 3.S.F.CPL do
màganà̀ dà shī$_j$]
talking with 3.S.M
'Who do you know the woman that talked to him' (Tuller, 1986, p. 159)

The most complex case of long-distance relativisation cited in the literature involves triply embedded relatives, with all three extraction sites contained within the inner-most sub-clause.

(18) ? gȁ mȁtâr$_i$ dà ka bā nì littāfìn$_j$ dà
here.is woman REL 2S.M.CPL give me book REL
mȁlàmai sukà san mùtumìn$_k$ dà $\emptyset_i$ ta
teachers 3P.CPL know man REL 3S.F.CPL
rubȕtā wà $\emptyset_k$ $\emptyset_j$
write for

'Here's the woman that you gave me the book the teachers know the man she wrote it for.' (Tuller, 1986, p. 84; tone added)

To summarise the empirical data, relativisation in Hausa is insensitive to extraction islands, provided the presence of a resumptive pronoun at the extraction site. Other types of extraction, like wh or focus fronting, do not exhibit this property, regardless of the presence of resumptives. As suggested by extraction from coordinate structures, however, resumptives are fully compatible with gaps, as far as the ATB constraint is concerned. We therefore conclude that both processes should be considered unbounded dependency constructions (UDCs), yet the specific constraints on locality and on the use of gaps vs. resumptives should be associated with properties of the elements at the top or the bottom of the dependency: i.e. the difference between a single relative marker merely mediating coreference with the antecedent noun vs. a full displaced constituent, as well as the nature of the governing head at the extraction site, i.e. verbs vs. prepositions.

1.1 Previous approaches

The first extensive formal study of Hausa extraction and resumption certainly is Tuller's (1986) doctoral dissertation on the language. Using a GB framework she suggests to account for the difference between island-insensitive resumptive relativisation and wh extraction by means of a distinction between base-generation and binding of a pronominal for relativisation vs. Ā movement for wh-extraction. Resumptives found in wh extraction as complements of obliques or possessors of nouns, by contrast, are treated as instances of phonetic trace (Koopman, 1984). The multitude of analytic devices (both base generation and movement with phonetic trace) for what appears to be a single phenomenon (resumption) has been criticised in Crysmann (2012).

Within HPSG, one of the first studies are the works of Nathan Vaillette on resumption in Hebrew (2001a) and Irish (2001b), proposing two separate features for gap and resumptive extractions. This separation has been criticised repeatedly in the HPSG literature, including Taghvaipour (2004; Taghvaipour (2005b; Taghvaipour (2005a), (Alotaibi and Borsley, 2013), and (Crysmann, 2012), mainly based on the known compatibility of gaps and resumptives in ATB extraction.

More specifically, Crysmann (2012) argues, on the basis of the Hausa data, for the compatibility between the two types of extractions. He shows further that no HPSG treatment available at the time was capable to capture the differences with respect to extraction islands. He suggests that both types of unbounded dependencies should be regarded as SLASH dependencies, distinguishing gap and resumptive dependencies in terms of the properties of the SLASH elements. More precisely, he argues that gaps require sharing of entire LOCAL values, whereas sharing of INDEX values is sufficient for resumptives (see Borsley (2010; Alotaibi and Borsley (2013) for a similar proposal). Since the description of resumptives subsumes that of gaps, the ATB facts are readily explained. The differences in locality, however, are due to constraints imposed at the retrieval site: while wh and focus fronting require full sharing of their LOCAL values, relatives merely require index sharing. If retrieval sites are transparent to indices, but not to full local values, the empirical pattern can be explained with a single mechanism.

A previous implementation of resumption in HaG has treated these elements essentially like gaps, including the restriction of SLASH to contain at most one element at any time. In this paper, we shall explore how the empirically and theoretically more desirable approach advanced in Crysmann (2012) can be put to use in a computational grammar of the language.

2 Implementation in LKB & friends

The implementation in HaG follows quite closely the theoretical proposal made in Crysmann (2012). Thus, both gap and resumptive dependencies are represented on SLASH, HPSG's feature for extraction, distinguishing them for the purposes of island effects in terms of the elements rather than by virtue of a distinct unbounded dependency.

2.1 The Grammar Matrix

The LinGO Grammar Matrix (Bender et al., 2002) is a starter kit for the development of HPSG grammars running on the LKB (Copestake, 2002), Pet (Callmeier, 2000) and Ace (by Woodley Packard (Crysmann and Packard, 2012)) platforms. Grammars running on these platforms use a conjunctive subset of TDL (Krieger, 1996) as their description language and Minimal Recursion Semantics (Copestake et al., 2005) for meaning representation. The Grammar Matrix not only makes for fast bootstrapping of new grammars, it also ensures a high degree of parallelism, owing to a carefully worked out constraint set on meaning construction combined

with a type hierarchy of rule types, suitable for a wider range of syntactic constructions.

The Matrix has been distilled to a great extent from the LinGO ERG (Copestake and Flickinger, 2000). As for extraction, both the ERG and the Matrix are highly faithful to the theory of unbounded dependencies advanced by Sag (1997) and Ginzburg and Sag (2001): thus, passing of non-local features (most notably SLASH) proceeds in a head-driven fashion, with heads amalgamating the NON-LOCAL values of their arguments.

(19) SLASH amalgamation

$$\begin{bmatrix} \textsc{synsem}\left[\textsc{nloc}\left[\textsc{sl}\ \boxed{1} \cup \ldots \cup \boxed{n}\right]\right] \\ \textsc{arg-st}\left\langle\left[\textsc{nloc}\left[\textsc{sl}\ \boxed{1}\right]\right], \ldots \left[\textsc{nloc}\left[\textsc{sl}\ \boxed{n}\right]\right]\right\rangle \end{bmatrix}$$

In the ERG and the Matrix, amalgamation is broken down into four constraints depending on the arity of the argument structure list, one of which any lexical head will inherit from. Owing to the absence of sets (and set union) in the underlying formalism, set-valued features are represented by means of difference lists (Clocksin and Mellish, 1981) instead,[1] instead, as shown in the example for two-element argument structure lists in (20).

(20)

$$\begin{bmatrix} \textit{0-diff-list} \\ \textsc{list}\ \boxed{1} \\ \textsc{last}\ \boxed{1} \end{bmatrix} \begin{bmatrix} \textit{1-diff-list} \\ \textsc{list}\ \begin{bmatrix} \textsc{first}\ [\,] \\ \textsc{rest}\ \boxed{1} \end{bmatrix} \\ \textsc{last}\ \boxed{1} \end{bmatrix}$$

$$\begin{bmatrix} \textit{basic-two-arg} \\ \textsc{synsem}\ \left[\textsc{nloc}\left[\textsc{sl}\ \begin{bmatrix} \textsc{list}\ \boxed{1} \\ \textsc{last}\ \boxed{3} \end{bmatrix}\right]\right] \\ \textsc{arg-st}\ \left\langle\begin{bmatrix} \textsc{list}\ \boxed{1} \\ \textsc{last}\ \boxed{2} \end{bmatrix}, \begin{bmatrix} \textsc{list}\ \boxed{2} \\ \textsc{last}\ \boxed{3} \end{bmatrix}\right\rangle \end{bmatrix}$$

Among the lexical amalgamation types, there is already one definition in the Matrix specifically aimed at resumptive pronouns, i.e. the possibility to launch a non-local dependency that does not correspond to an argument. However, this constraint will only ever be suitable for free pronouns, not bound ones, as we find in Hausa, since the type constraint is defined on the level of the lexical entry. Furthermore, resumption is still treated as entirely identical to gap-type extraction.

Besides these more technical issues, there is, however, a more fundamental difference between the treatment of SLASH dependencies in the theoretical HPSG literature and its implementation in the ERG and the Matrix: while Pollard and Sag (1994) clearly argue that more than one element can be in SLASH at the same time, the ERG and the Matrix both limit the length of the SLASH list to at most one, thereby ruling out the combination of strong and weak UDCs witnessed in (21). The reason behind this restriction is most certainly related to processing efficiency.

(21) [A violin this well crafted]$_1$ even [the most difficult sonata]$_2$ will be easy to play $\_2$ on $\_e_1$? (Pollard and Sag, 1994, 169)

While for English, cases of multiple simultaneous SLASH dependencies can possibly be marginalised without jeopardising overall coverage on natural language data, this is certainly not the case in a variety of other languages, including multiple wh-fronting in Slavic, or long-distance relativisation in Hausa. Thus, a more systematic solution is called for that we shall develop in the following section.

2.2 Multiple SLASH dependencies

Since HaG is based on the Grammar Matrix, the current general approach to extraction is already head-driven, in accordance with the current consensus amongst HPSG scholars. Since restrictions on the size of SLASH are imposed on introduction and retrieval, we can concentrate on these two critical points in our discussion of extraction in HaG.

2.2.1 Launching

With the exception of adjunct extraction,[2] gaps in HaG are introduced by means of unary *lexical* rules suppressing a valency corresponding to an argument introduced into SLASH, essentially following Pollard and Sag (1994), Sag (1997), Ginzburg and Sag (2001), as well as common practice in the ERG and the Matrix. In addition, a unary *phrase structure* rule permits launching of adjunct extractions.

As for resumptives, the current implementation maintains two sets of lexical rules, one for bound pronominals and one for bound resumptives, as well as two sets of lexical entries for pronominals and resumptives. Making systematic use of the type hierarchy of rules and lexical types, shared properties of resumptive and non-resumptive uses, including autosegmental morphophonological properties, are abstracted out into common supertypes. Most crucially, in the true pronominal case, a semantic relation is inserted into the MRS RELS list and the SLASH

[1] Difference lists permit list concatenation by means of unification: essentially, such lists maintain a pointer (LAST) to the open end of the list. We shall use exclamation marks to distinguish these from ordinary lists, as is the convention in DELPHIN grammars (Copestake, 2002, cf.).

[2] See Levine (2003) for arguments to distinguish adjunct and complement extraction in English along the syntax/lexicon divide.

value is restricted to be empty, whereas in the resumptive case, the RELS list is empty, but SLASH contains an element the referential index of which is shared with that of the resumptive. Owing to the absence of internal disjunction from the underlying feature formalism (rules and lexical entries are in disjunctive normal form), specification of separate rules and entries for both uses turned out to be unavoidable.[3] In order to keep the number of disjunctive specifications to an absolute minimum, we have therefore generalised the existing 3 sets of morphological rules for pronominal affixation (objective, genitive, dative), capturing the difference in shape by reference to the segmental make-up of the base, rather than in terms of the syntactic category of the base, enabling us to collapse all three sets into one. This move was greatly facilitated by the fact that nouns and verbs, as well as the applicative marker *wà* independently undergo characteristic inflection for the type of argument realisation of their first complement (Crysmann, 2005), distinguishing inter alia realisation by a pronominal affix: e.g., pronominal affixes from the genitive set are always preceded by the gender differentiated linker *-n/-r* (cf. example (1)), a segment that is crucially absent in final positions of all verbs taking pronominal affixes from the direct object set. To account for differences in tonal specification (genitive set is low, whereas objective set alternates), we generalised our previous treatment of "polar" tone with objective pronouns, representing the tonal specification of the pronominal affix as a floating tone of the base. In fact,

Following Crysmann (2012), the main difference between gap and resumptive SLASH values is that the former require reentrancy with a full local value, whereas the latter are underspecified in this respect: minimally, they only require identity of INDEX. Elaborating on the hierarchy of *synsem* proposed in Sag (1997), we have complemented the *gap* subtype of *synsem* with a type for resumptives and abstracted out shared minimal requirements into a common super type.

(22)
$$\begin{bmatrix} \textit{gap-or-res} \\ \text{LOC} \quad \left[\text{CONT.HOOK.INDEX} \quad \boxed{1}\right] \\ \text{NLOC} \quad \left[\text{SL}\left\langle ! \left[\text{CONT.HOOK.INDEX} \quad \boxed{1}\right] ! \right\rangle\right] \end{bmatrix}$$

$$\begin{bmatrix} \textit{gap} \\ \text{LOC} \quad \boxed{1}\ \textit{full-local} \\ \text{NLOC} \quad \left[\text{SL} \quad \left\langle ! \ \boxed{1} \ ! \right\rangle\right] \end{bmatrix} \quad \begin{bmatrix} \textit{resump} \\ \text{NLOC} \quad \left[\text{SL} \quad \left\langle ! \ \textit{light-local} \ ! \right\rangle\right] \end{bmatrix}$$

[3] Underspecification techniques using list types do not provide a solution either, since we need to use difference lists for which this abstraction is only available to a limited extent.

(23)
$$\begin{bmatrix} \textit{local} \\ \text{CONT} \quad \textit{mrs} \end{bmatrix}$$

$$\begin{bmatrix} \textit{full-local} \\ \text{CAT} \quad \textit{cat} \end{bmatrix} \qquad \textit{light-local}$$

Using the types just introduced, extraction and resumption rules are the defined as follows:

(24) Complement extraction

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

(25) Resumption

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

Both rules perform a valence reduction, but do so imposing constraints of different strength on the locally suppressed complement. Since elements of the COMPS valence list are reentrant with ARG-ST, the restriction towards SLASH will be picked up correctly by SLASH amalgamation. For resumption, the grammar distinguishes variants for null realisation (subject and non-human direct object) and pronominal affixation.

On the basis of the distinction between *full-local* and *light-local* values, we have furthermore defined typed list constraints that permit to restrict what kind of dependencies can be active simultaneously. As we have seen above, only relativisation can escape wh-islands in Hausa, provided the dependency involves a resumptive. Furthermore, relative clause formation, in contrast to wh-extraction and focus fronting, does not permit pied-piping. Thus, multiple simultaneous SLASH dependencies can involve at most one single gap type dependency at any node. Most importantly, this state of affairs enables us to ensure termination in the light of adjunct extraction which involves true gaps: whenever a true gap is inserted into SLASH, the remainder of the SLASH list is constrained to consist entirely of elements of type *light-local*. Complement extraction lexical rules are constrained in a similar fashion.

2.2.2 Retrieval

Given that SLASH values may contain multiple elements, retrieval at the top of the dependency marks a more clear departure from common practice in the Grammar Matrix: in essence, we need to search the SLASH list for a suitable element to be bound off, and

pass on any *light-local* elements to be retrieved further up the tree.

The grammar has exactly two constructions where retrieval can take place, the first one being a classical filler-head construction to be used for binding wh and focus-fronted fillers, both of which allow pied-piping. As for relatives, we follow Borsley and assume that Hausa relative "pronouns" are actually (inflected) relative complementisers that take the clause containing a gap or resumptive as its complement. This assumption not only takes care of the impossibility of pied-piping in relative clauses, but it also captures nicely the similarity of the uninflected relative complementiser *dà* to its homonymous non-relative counterpart. What is common to both constructions is that they define a non-empty TO-BIND value (cf. Pollard and Sag (1994)).[4] In filler-head structures, the SLASH dependency to be retrieved is constrained to be of type *full-local* by virtue of structure sharing with the filler's LOCAL value, whereas no such constraint is imposed by the relative complementiser which only requires a referential index.

(26)
$$\left[\begin{array}{l}\textit{filler-head-rule}\\ \text{SS.NLOC}\ \left[\begin{array}{ll}\text{T-B} & \left[\begin{array}{ll}\text{FILL} & \langle \boxed{l} \rangle\\ \text{SL} & \boxed{s}\end{array}\right]\\ \text{SL} & \langle\, !\ !\,\rangle\end{array}\right]\\ \text{FILLER-DTR}\ \left[\text{SS}\left[\text{LOC}\ \boxed{l}\right]\right]\\ \text{HD-DTR}\ \left[\text{SS}\left[\text{NLOC}\left[\text{SL}\ \boxed{s}\right]\right]\right]\end{array}\right]$$

(27)
$$\left[\begin{array}{l}\textit{rel-complementiser-lex}\\ \text{SS}\ \left[\text{NLOC}\left[\begin{array}{ll}\text{T-B} & \left[\begin{array}{ll}\text{FILL} & \left\langle\left[\text{CONT}\left[\text{HOOK}\left[\text{INDEX}\ \boxed{i}\right]\right]\right]\right\rangle\\ \text{SL} & \boxed{s}\\ \text{REL} & \langle\, !\ \boxed{i}\ \textit{ref-index}\ !\,\rangle\end{array}\right]\\ \text{SL} & \langle\, !\ !\,\rangle\end{array}\right]\right]\\ \text{ARG-ST}\ \left\langle \text{S}\left[\text{NLOC}\left[\text{SL}\ \boxed{s}\right]\right]\right\rangle\end{array}\right]$$

Two unary retrieval rules then take care of binding the filler to an appropriate percolated SLASH element and to pass on any elements of type *light-local*.

(28)
$$\left[\begin{array}{l}\textit{bind-filler-rule}\\ \text{SS}\ \left[\begin{array}{ll}\text{LOC} & \boxed{0}\\ \text{NLOC} & \left[\begin{array}{ll}\text{T-B} & \left[\begin{array}{ll}\text{FILL} & \langle\,\rangle\\ \text{SL} & \langle\, !\ \boxed{l}\ !\,\rangle\end{array}\right]\\ \text{SL} & \boxed{s}\\ \text{QUE} & \boxed{q}\\ \text{REL} & \boxed{r}\end{array}\right]\end{array}\right]\\ \text{ARGS}\ \left\langle\left[\text{SS}\left[\begin{array}{ll}\text{LOC} & \boxed{0}\\ \text{NLOC} & \left[\begin{array}{ll}\text{T-B} & \left[\begin{array}{ll}\text{FILL} & \langle \boxed{f} \rangle\\ \text{SL} & \langle\, !\ \boxed{f} \mid \boxed{l}\ !\,\rangle\end{array}\right]\\ \text{SL} & \boxed{s}\\ \text{QUE} & \boxed{q}\\ \text{REL} & \boxed{r}\end{array}\right]\end{array}\right]\right\rangle\end{array}\right]$$

(29)
$$\left[\begin{array}{l}\textit{find-filler-rule}\\ \text{SS}\ \left[\begin{array}{ll}\text{LOC} & \boxed{0}\\ \text{NLOC} & \left[\begin{array}{ll}\text{T-B} & \left[\begin{array}{ll}\text{FILL} & \boxed{f}\\ \text{SL} & \boxed{l}\end{array}\right]\\ \text{SL} & \langle\, !\ \boxed{1}\ \textit{light-local} \mid \boxed{s}\ !\,\rangle\\ \text{QUE} & \boxed{q}\\ \text{REL} & \boxed{r}\end{array}\right]\end{array}\right]\\ \text{ARGS}\ \left\langle\left[\text{SS}\left[\begin{array}{ll}\text{LOC} & \boxed{0}\\ \text{NLOC} & \left[\begin{array}{ll}\text{T-B} & \left[\begin{array}{ll}\text{FILL} & \boxed{f}\\ \text{SL} & \langle\, !\ \boxed{1} \mid \boxed{l}\ !\,\rangle\end{array}\right]\\ \text{SL} & \boxed{s}\\ \text{QUE} & \boxed{q}\\ \text{REL} & \boxed{r}\end{array}\right]\end{array}\right]\right\rangle\end{array}\right]$$

While the *bind-filler-rule* performs the actual instantiation and retrieval of the unbounded dependency, the *find-filler-rule* merely iterates over the original SLASH difference list and puts back one by one the original elements, constrained to *light-local*. In essence, these rules jointly ensure the island restriction towards non-resumptive SLASH dependencies: since extraction out of relative and wh islands is restricted to relativisation footed by a resumptive, the constraints on *light-local* for further percolation after the first retrieval of a SLASH element accounts for the ungrammaticality of, e.g. (14) and (15), while still permitting relativisation out of relatives, as witnessed in (18).[5]

[4] All lexical entries other than the relative complementiser require their entire TO-BIND value to be empty (i.e. both TO-BIND.FILL and TO-BIND.SLASH). Furthermore, we constrain the type *head-nexus-phrase* (Sag, 1997) as well as standard unary phrase structure rules to effect structure sharing of TO-BIND between the mother and the (head) daughter. Similarly, elements on ARG-ST are equally restricted to have empty TO-BIND features. As a net effect, no other syntactic rule can interfere in the middle of retrieval.

[5] One might wonder why we insist on full perusal of SL(ASH), even after a filler has been found, instead of merely constraining the remainder of the list using the aforementioned list types. First, this recursion does not add any complexity factor beyond the possibility of that introduced by considering alternative instantiations for the filler. Second, recreating the SLASH list recursively step by step from the unretrieved elements enables us to get rid of any latent constraints on the open end of the list regarding *local* type: once we have retrieved a *full-local* dependency, we want to be able to add new gap dependencies further

2.2.3 ATB extraction

The underspecification approach to resumptives, i.e. their compatibility with both *light-local* and *full-local*, already ensures the compatibility between true gaps and resumption in ATB extraction from coordinated structures. Furthermore, given the ambiguity of pronominals between resumptive and true pronoun uses, identity requirements will only have to hold for those pronominals that actually enter in a non-local dependency.

To this end, we restrict coordinating constructions to enforce identity of entire SLASH lists, as shown in (30).

(30)
$$\left[\begin{array}{ll} \textit{s-binary-coord} & \\ \textsc{synsem}\left[\textsc{nloc}\left[\textsc{sl}\ \boxed{s}\right]\right] & \\ \textsc{lconj-dtr} & \left[\textsc{synsem}\left[\textsc{nloc}\left[\textsc{sl}\ \boxed{s}\right]\right]\right] \\ \textsc{rconj-dtr} & \left[\textsc{synsem}\left[\textsc{nloc}\left[\textsc{sl}\ \boxed{s}\right]\right]\right] \end{array}\right]$$

A complicating factor, however, comes in owing to the use of lists, instead of sets, imposed by the underlying formalism. As detailed in Newman (2000), ATB extraction may target different grammatical functions in both conjuncts. Since Hausa also permits multiple relativisation from the same clause (Tuller, 1986), we expect multiple ATB relativisation to be possible also with reversal of grammatical functions, as illustrated by the (constructed) example in (31).

(31) gā̀ mùtûm dà kukà san mā̀târ
there.is man.M REL you.PL.CPL know woman.DEF.F
dà yakȅ sôn -tà àmmā takȅ
REL 3.S.M.CONT like.VN.of -3.S.F but 3.S.F.CONT
ƙîn -sà
hate.VN.of -3.S.M
'Here's the man that you know the woman who he likes but (who) hates him.'

In order to allow for this possibility, we complement the standard coordination schema sketched above with an alternative one that has the first elements of the right conjunct reversed.

For efficiency reasons, I am currently limiting myself to permutation of the first two SLASH elements. This decision, however, is supported by the observation that triple relativisation in itself is already considered marked to some extent: see Tuller's question mark on the relevant example in (18). While these data clearly contrast with the unacceptability of island violations, I seriously doubt that their marked acceptability will improve when combined with ATB extraction from non-parallel conjuncts, thereby further increasing complexity. Thus, until we have evidence to the contrary, I shall refrain, for the time being, from full permutation of SLASH lists greater than 2, assuming parallelism of dependencies except for the first two elements.

up the tree, independently of whether this new SLASH element will be prepended or appended to our current SLASH list.

3 Conclusion

We have argued in this paper that Hausa extraction militates for an extension of current practice in HPSG grammar implementation to permit multiple simultaneous SLASH dependencies. Based on the theoretical proposal by Crysmann (2012), we have provided an implementation of the Hausa extraction facts. In essence, we have generalised the constraints on SLASH to permit multiple members at any time, but have systematically exploited the distinction between light and full local values to constrain multiple extraction to involve at most on gap simultaneously. This not only correctly captures the island constraints in Hausa, but it also provides a straightforward means to ensure efficiency, including termination of adjunct extraction. Furthermore, to keep disjunctive specifications of pronominal and resumptive uses to a minimum, we have developed a more generalised treatment of pronominal affixation in the language, collapsing morphological rules for genitive accusative and dative pronominal affixes. In future work, we shall explore how the systematic ambiguity between resumptive and non-resumptive uses of pronominals may be captured without disjunctive specification at all, in order to provide a complete answer to McCloskey's (2002) generalisation.

Acknowledgements

Work on HaG is partially supported by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" programme (reference: ANR-10-LABX-0083), project ResHau [re:'zo:] of the Laboratoire d'excéllence "Empirical Foundations of Linguistics".

References

Mansour Alotaibi and Robert D. Borsley. 2013. Gaps and resumptive pronouns in Modern Standard Arabic. In Stefan Müller, editor, *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, pages 6–26.

Emily M. Bender, Dan Flickinger, and 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, and Richard Sutcliffe,

editors, *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, pages 8–14.

Robert D Borsley. 2010. An HPSG approach to Welsh unbounded dependencies. In Stefan Müller, editor, *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot, Paris 7, France*, pages 80–100, Stanford, CA. CSLI Publications.

Ulrich Callmeier. 2000. PET — a platform for experimentation with efficient HPSG processing techniques. *Journal of Natural Language Engineering*, 6(1):99–108.

William F. Clocksin and Christopher S. Mellish. 1981. *Programming in Prolog*. Springer, Heidelberg.

Ann Copestake and 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.

Ann Copestake, Dan Flickinger, Carl Pollard, and Ivan Sag. 2005. Minimal recursion semantics: an introduction. *Research on Language and Computation*, 3(4):281–332.

Ann Copestake. 2002. *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford.

Berthold Crysmann and Woodley Packard. 2012. Towards efficient HPSG generation for German, a non-configurational language. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012)*, pages 695–710, Mumbai, India.

Berthold Crysmann. 2005. An inflectional approach to Hausa final vowel shortening. In Geert Booij and Jaap van Marle, editors, *Yearbook of Morphology 2004*, pages 73–112. Kluwer.

Berthold Crysmann. 2012. Resumption and island-hood in Hausa. In Philippe de Groote and Mark-Jan Nederhof, editors, *Formal Grammar. 15th and 16th International Conference on Formal Grammar, FG 2010 Copenhagen, Denmark, August 2010, FG 2011 Ljubljana, Slovenia, August 2011*, volume 7395 of *Lecture Notes in Computer Science*, pages 50–65. Springer.

Jonathan Ginzburg and Ivan Sag. 2001. *Interrogative Investigations: the Form, Meaning and Use of English Interrogatives*. CSLI publications, Stanford.

Philip Jaggar. 2001. *Hausa*. John Benjamins, Amsterdam.

Hilda Koopman. 1984. *The Syntax of Verbs. From Verb Movement Rules in the Kru Languages to Universal Grammar*. Foris, Dordrecht.

Hans-Ulrich Krieger. 1996. *TDL — A Type Description Language for Constraint-Based Grammars*, volume 2 of *Saarbrücken Dissertations in Computational Linguistics and Language Technology*. DFKI GmbH, Saarbrücken.

Robert D. Levine. 2003. Adjunct valents: cumulative scoping adverbial constructions and impossible descriptions. In Jongbok Kim and Stephen Wechsler, editors, *The Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, pages 209–232, Stanford. CSLI Publications.

James McCloskey. 2002. Resumptives, successive cyclicity, and the locality of operations. In Samuel David Epstein and T. Daniel Seely, editors, *Derivation and Explanation in the Minimalist Program*, pages 184–226. Blackwell, Oxford.

Paul Newman. 2000. *The Hausa Language. An Encyclopedic Reference Grammar*. Yale University Press, New Haven, CT.

Carl Pollard and Ivan Sag. 1994. *Head–Driven Phrase Structure Grammar*. CSLI and University of Chicago Press, Stanford.

Ivan Sag. 1997. English relative clause constructions. *Journal of Linguistics*, 33(2):431–484.

Mehran Taghvaipour. 2004. An HPSG analysis of Persian relative clauses. In Stefan Müller, editor, *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, pages 274–293. CSLI Publications, Stanford.

Mehran Taghvaipour. 2005a. *Persian Relative Clauses in Head-driven Phrase Structure Grammar*. Ph.D. thesis, University of Essex.

Mehran A Taghvaipour. 2005b. Persian free relatives. In Stefan Müller, editor, *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, pages 364–374, Stanford. CSLI Publications.

Laurice A. Tuller. 1986. *Bijective Relations in Universal Grammar and the Syntax of Hausa*. Ph.D. thesis, UCLA, Ann Arbor.

Nathan Vaillette. 2001a. Hebrew relative clauses in HPSG. In Dan Flickinger and Andreas Kathol, editors, *The Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, pages 305–324, Stanford. CSLI Publications.

Nathan Vaillette. 2001b. Irish gaps and resumptive pronouns in HPSG. In Frank Van Eynde, Dorothee Beermann, and Lars Hellan, editors, *The Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, pages 284–299, Stanford. CSLI Publications.