Integrated and Non-Integrated Left Dislocation: A Comparative Study of LD in Avatime, Tundra Yukaghir & Whitesands

Dejan Matić, Saskia van Putten & Jeremy Hammond

Left-dislocated elements seem to be extremely frequent, if not universal, across languages. Despite their surface similarity, however, they tend to display strikingly different features, both structurally and functionally. Our purpose is to investigate the structural variability of left-dislocation constructions cross-linguistically. The study is based on a fine-grained analysis of naturally occurring instances of left dislocation in three geographically and genealogically distinct languages — Avatime (Kwa, Ghana), Tundra Yukaghir (isolate, north-eastern Siberia), and Whitesands (Oceanic, Vanuatu). The formal parameters of variation investigated include the frequency and type of resumption within the clause, integration within the clause via case/cross-referencing on the verb, recursivity, and connectivity effects such as island sensitivity. We show that the degree of integration of left-dislocation constructions in the sentence can vary from language to language. This variability is used as a basis for a discussion of the notions of *clause* and *sentence*, both within RRG and within general linguistic theory. We conclude that there is no simple formal paradigm (e. g. integrated vs. non-integrated) to satisfactorily explain the diversity we find in this language sample. Instead, we propose to differentiate different degrees and types of integration, which seem to provide a better account of cross-linguistic variation.

343

1 Introduction

The received wisdom about Left Dislocation (LD) is that it is a structure in which an element of the clause does not occur in its canonical position, but precedes its mother clause; the canonical position of the LD element is occupied by a coreferential pronominal (e. g. Givón 1976, Dik 1978, Haegeman 1991, Alexiadou 2006, Shaer 2009). This is illustrated in (1).

(1) (...) ***Bunyan,*** **you know,** *he***'s older than the rocks** (...) (Geluykens 1992: 90)

LD raises two major classification problems. First, it is not the only type of structure in which elements in a clause are displaced to the left. On the one hand, there is topicalisation (TOP), in which the left-displaced element is usually considered to reside within the confines of the clause. On the other hand, various types of online production phenomena, such as hesitations, false starts, self-corrections, etc., often do not formally differ from canonical LDs. Second, even if LD proper is somehow disentangled from the rest of detachment structures, there is a considerable variation within the category itself: some instances of LD are restricted as to the type of resumptive expression, while others are undifferentiated in this respect; the presence of a prosodic break between the dislocated element and the clause is frequent but optional; the same optionality characterises the presence of connectivity effects within LD structures.

These issues have bearing on both the practical question of classifying displaced elements in a corpus of any one language and the question of the correct syntactic analysis of LD. The necessity to differentiate LD from both speech disfluencies and clause-level phenomena (TOP) testifies to its perceived status as residing between a syntactic structure proper and a production-level phenomenon. LD displays features of a syntactic template (see Lambrecht 2001, Alexiadou 2006 and Shaer 2009 for recent overviews) and has conventionalised discourse functions (e. g. Prince 1997, 1998), but it does not seem to be fully integrated into the clause to which it relates.

One consequence of this in-between status is that, at the level of identifying and classifying individual instances of displacement in a corpus, there is no principled way of telling LD and speech disfluencies apart. This has to be done on an item-by-item basis with the help of various criteria such as presence of hesi-

tation and self-correction markers, prosodic signals, and the reconstruction of the general intention of the speaker.

   The problem of telling apart instances of LD from those of TOP is a more intricate and theory-laden one. The presence of a resumptive pronoun is considered to be the main property by which LD is to be distinguished from TOP, which does not permit resumptives (e. g. Gregory & Michaelis 2001: 1667). There are several other criteria, prosodic and syntactic, that are often adduced as characteristic for LD. It is claimed that LD forms a prosodic phrase marked by boundary tones and/or a prosodic break. However, this does not seem to be universally applicable: special prosodic phrasing is at best optional in a number of languages and can therefore not be used as a definitive criterion to identify LD (see e. g. Geluykens 1992: 92ff. for English, Avanzi et al. 2010 and Avanzi 2011 for French, Feldhausen 2008: 175ff. for Catalan). The syntactic criteria are based on the idea that elements outside of the clause proper cannot be sensitive to clause-internal constraints, while those that are within the clause do not react to discourse-level restrictions. It has been claimed that TOP is sensitive to island constraints, anaphor binding, and similar phenomena, while LD is not; and vice versa, LD is restricted with respect to syntactic contexts in which it can appear (ban on embedding) and with respect to word class (no LD adverbs, due to lack of reference), while TOP does not display restrictions of this kind (Shaer 2009). There are also problems with these types of criteria: first, they do not seem to be universally true (there are differences even between such closely related languages as English and German, Shaer 2009); second, connectivity tests, island tests etc. are applicable only after the fact, i. e. only after a construction has been already identified as an instance of LD or TOP, and are as such not useful in detecting LD structures in naturally occurring speech.

   This variability of the properties of LD — the second classificatory problem introduced above — has led linguists to postulate the existence of a number of different subtypes of LD, the major ones being Hanging Topic LD and 'Movement' LD[1], reanalysed as Integrated and Non-Integrated LD in recent research (Shaer 2009)[2]. The non-integrated type is prosodically separated, displays no connec-

---

[1] The latter includes *Contrastive LD* in Germanic and *Clitic LD* in Romance (Anagnostopoulou 1997, Cinque 1997, Alexiadou 2006).

[2] This distinction is often obfuscated by terminological confusion: some authors use the term *left dislocation* only for the non-integrated, hanging-topic type (e. g. Ross 1967[1986], Cinque 1990), while others (e. g. Altmann 1981, Frey 2004) restrict it to the integrated types. In this paper, the

tivity effects, etc., and is therefore clearly extraclausal. The integrated types are prosodically integrated, have a restricted repertoire of resumptives, show connectivity, etc., so that they count as somehow more integrated into the clause. The important point is that the common denominator for the assumed variable types is still the resumptive expression: since all other features are unevenly distributed across LD types, resumptives are the only reliable indicator of the LD structure.

However, this focus on resumptive pronouns is also problematic, as it leaves a number of phenomena outside the scope of LD studies. Firstly, it excludes the so-called Chinese-style topics (Chafe 1976), which share a number of features with typical LDs, but cannot be cross-referenced with resumptive pronouns. Secondly, the status of adjuncts and many types of adverbials is unclear, as some of these kinds of expressions seem to canonically appear in positions which resemble LD but are at best optionally followed by a resumptive pronoun (Lambrecht 2001: 1059). Thirdly, the focus on resumptive pronouns ignores languages which, due to the possibility of zero anaphora, do not need resumptive elements but may still have LD. This paper tackles these issues in some detail.

With respect to its syntactic structure, there is little agreement on how LD should be analysed. Much has been written about LD in the generative framework, and no consensus seems to have been reached. Some approaches seem to assume that LD elements have a specified position in the left periphery of the clause (see Section 3.1 for discussion of this approach). Since this kind of analysis incurs a number of problems, a different analysis has been proposed, according to which LD elements are so-called orphans, i.e. separate sentential fragments that are not part of the syntactic structure but are linked to the clause via discourse-level linking rules (Haegeman 1991). As we indicated above, not all LD constructions are the same, so that integrated LD elements are commonly treated as left-peripheral within the clausal domain, while non-integrated LD elements are analysed as orphans. Some recent proposals (De Cat 2007, Ott 2012) remove integrated LDs from the clausal domain, too, and consider them to be separate referential and/or clausal entities; their intra-clausal properties are explained via ellipsis or adjunction. In effect, these recent approaches bring the two postulated types of LD closer together, treating them both as extraclausal entities. For the sake of brevity, we shall use the term 'orphan' for all types of analysis that

---

term LD is used broadly to cover all types of structures that include pre-clausal elements, as has become common in the literature in the last decade.

assume an extra-clausal nature of LD elements, irrespective of the details of the analysis itself.

In Role and Reference Grammar (RRG), two syntactic domains are recognised: clause and sentence (Foley & Van Valin 1984, Van Valin & LaPolla 1997, Van Valin 2005). For the purposes of this paper, we assume quite simple definitions of these two domains. The clause is the domain that comprises everything that depends on the predicate plus clause-level adjuncts (see Van Valin & LaPolla 1997 on the notion of dependency, and Van Valin 2005 on types of adjuncts). The sentence comprises one or more coordinate or co-subordinate clauses, and/or extraclausal elements which have a determinate position in the syntactic architecture, such as sentence-level adjuncts (Van Valin 2005); in practice, sentences are often coextensive with clauses. The essential point is that the clause is not the only level at which syntactic rules operate, but that there is one more level, not directly bound by the predicate, but still subject to syntactic, as opposed to discourse, rules — the level of sentence. This allows for a dedicated position for LD elements, the left-detached position (LDP), which is, as it were, between integrated and non-integrated: within the sentence but outside of the clause (Van Valin & LaPolla 1997: 36).

The purpose of this paper is to assess the status of LD from a cross-linguistic perspective. We investigate whether LD exists as a grammatical phenomenon in three non-European languages and if so, whether this is the same grammatical phenomenon in each of the languages. The paper deals with LD in three genealogically unrelated and geographically distant languages — Avatime (Kwa, Niger-Congo; spoken in Ghana), Tundra Yukaghir (Paleo-Siberian, Isolate; spoken in north-eastern Siberia) and Whitesands (Oceanic, Austronesian; spoken in Vanuatu). We show that there is no single linguistic property that can reliably identify LD constructions crosslinguistically. Instead, we need a set of criteria which is different from language to language. Furthermore, the contradictory evidence of syntactic tests and morphological clues seem to imply that there might not be one or two right syntactic analyses of LD, but LDs have to be defined on a language-to-language basis. This means that different structures labelled Left Dislocation are in a relationship of family resemblance to each other rather than being mere instantiations of one or two universal syntactic structures.

To make comparison possible, we make use of a heuristic concept of LD, which does not have any ontological pretensions apart from enabling the researcher to

find structurally and/or functionally similar structures. This procedure resembles Haspelmath's (2010) idea of comparative concepts which are devised as a *tertium comparationis* for cross-linguistic research defined on the basis of minimal common denominators. Our heuristic concept of LD comprises all structures which include an element that precedes the clause and stands in some kind of relationship to it, excluding those instances which are readily identified as speech disfluencies. Importantly, we do not assume any definition of LD but rather define what LD is for each language and only subsequently try to determine the similarities and differences.

We have identified LD elements in corpora of spontaneous speech which were collected during fieldtrips to the communities where the languages are spoken. In addition, we have done some elicitation to collect grammaticality judgments in order to obtain negative evidence or check whether constructions not found in the corpora are possible. Our corpora comprise various genres, including conversations, public speeches and narratives. Each corpus contains roughly seven hours of text. The examples used in the paper come from these larger corpora. For statistical purposes, a smaller corpus of one to two hours from each language was chosen. The numerical data in the paper refer to the instances of LD found in these subsets of the larger corpora.

The paper is organised as follows. In the next section, we outline the syntactic features required to identify LD in the target languages. We then use the structures identified in this way as a basis for our comparison and discuss structural similarities and differences between the systems in Section 3. Section 4 concentrates on the nature of LD in the three languages. We also consider what additions or changes need to be made in the typology of LD, given the data presented in this chapter. We conclude in Section 5 with remarks on how RRG could best handle an updated typology, and highlight potential areas for future research.

2 Identifying left dislocation in three languages

In this section we describe the defining features of LD in Avatime, Tundra Yukaghir and Whitesands. Before we go into the details of each language, two general observations need to be made. Firstly, all three languages make use of zero anaphora to refer to definite specific referents, usually those that are active or activated in the given context. As the referent of the LD element is necessar-

ily activated at the moment when the clause that follows it is pronounced, these kinds of languages can practically always use NP gaps that are coreferential with the LD element. In other words, if the clause following the displaced element does not contain a resumptive, this does not necessarily mean that it is not LD. Resumptives, even though helpful in identifying LD structures, are thus not sufficient as diagnostics of LD in these languages. Secondly, prosody is not a reliable indicator of LD in any of the three languages: LD can be signalled by a prosodic break, but this is not necessarily the case and, more importantly, a prosodic break does not necessarily indicate LD.

2.1 Avatime

In Avatime, word order is rigidly SVO. Apart from LD, the only displacement construction is the focus construction, in which the focused element occurs clause-initially and is marked with the focus marker, which is a final extra-high tone. There is no topicalisation construction. Subjects are obligatorily marked on the verb (e. g. Avatime is head-marking with respect to the subject[3]) and there is no case marking. The sentence structure in Avatime looks as follows:

(2) LD&AdjunctClause_[Foc_[S_V_O]]

LD is most easily recognized in Avatime by the occurrence of a resumptive element. Of all identified cases of left-dislocated objects (N = 75), 91 % (N = 68) have a resumptive element, as in (3). The other 7 % are identified as LD simply by occurring before the verb (4), which is otherwise not possible for an object (unless marked as focused with the final extra-high tone). Of left-dislocated subjects (N = 58), 36 % (N = 21) are identified as LD by the occurrence of a resumptive element. An example can be seen in (5). This percentage is smaller than that for objects because independent subject pronouns, unlike object pronouns, are optional, due to the obligatory subject marking on the verb. The subject LD cases without independent resumptive element are identified as such because the LD element occurs before a focused element, which forms the left-edge of the clause (6), or because it

---

[3] An issue which we shall not be able to tackle here is the status of head marking. Many languages obligatorily mark arguments on the head of the phrase. The question is: Does head marking count as a resumptive element? If it does, either all arguments are double-instantiated, or all lexical arguments should be treated as LD (see Lambrecht 2001: 1056, and especially Van Valin 2013).

occurs before another LD element, an adjunct clause or a conjunction. Summing up, LD in Avatime is defined as follows:

(a)   any phrase followed by a co-referential resumptive pronoun or noun (3), (5);

(b)   any non-S argument preceding the subject and/or the verb which is not marked with the focus marker (4);

(c)   any argument preceding the focus-marked phrase, an adjunct clause, a conjunction or another LD phrase (5), (6).

(3) **lị̀-fịflị̀-nɛ**   *ɛ̀ɛ́-sɔlị̀*   **lɛ**
     C3s-porridge-DEF C1s.SBJ.PROG-catch C3s
     'The porridge, he was catching it.'

(4) **kpeve tsyɛ** *òmonò í-tanu*   *ze*   **Ø**
     Kpeve too today C2p.SBJ.NEG-be.able be.at
     'Kpeve too, today it (okra) was not there.'

(5) *kɔ* **li-bí**   **lɛ́-lɔ̀**   *gì* **lị̀-kpasí**   **wo**   **li-po-le**   **mè** *kɔ*
     then C3s-wound C3s-DIST REL C3s.SBJ-be.in 2SG.POS C3s-stomach-DEF inside then
     lɛ   *tsyɛ li-tse*
     C3s too C3s.SBJ-die
     'Then that wound in your stomach, then that too will heal (literally: die).'

(6) **wɔ** **tsyɛ** *á-dei-lá*   *wɛ̀ɛ́-ta*   *wìyàwìyà te*
     2SG too C3p-corn-DEF:FOC 2SG.SBJ.PROG-chew ID   like.that
     'You too, you are chewing corn like that.'

2.2   Tundra Yukaghir

Tundra Yukaghir is rigidly head-final, with a relatively free preverbal ordering of elements; the only exception is the focus-marked phrase, which is adjacent to the verb and forms a phrase with it (Matić & Nikolaeva 2014). The argument structure is indexed both on arguments, by means of a rich case system, and on the verb. Verbs are always fully head-marked for subjects; the object marking on the verb is less direct, based on two separate sets of suffixes for transitive and intransitive verbs. Tundra Yukaghir does not allow for centre-embedding of adjunct clauses, formed with converbal forms of the verb, which are clause-initial as a rule. The structure of the Tundra Yukaghir sentence can be summarised as follows:

(7) LD_[AdjunctClause_S&O_[Foc_V]]

In LD, resumptive elements are possible, but relatively rare: 36 % (N = 4) of LD subjects (N = 11) and 40 % (N = 13) of objects (N = 33) are re-instantiated with a resumptive (8). In cases without a resumptive, there are two criteria which can be adduced to identify LD. First, if an element precedes an adjunct clause, it must count as LD (9). Second, if an initial element does not carry the case which is governed by the verb, it can only be left-dislocated (10). In sum, three criteria for LD can be defined on the basis of these structural properties. LD is:

(a) any phrase followed by a co-referential resumptive pronoun or noun (8);
(b) any argument preceding the adjunct clause (9);
(c) any left-peripheral phrase which is not in the case governed by the predicate (10).

(8) ***pure:-ńə-j***   ***rukun, tudel*** *jawul əl=ičo:-čo:n*   *qajl'-ɣa*
    berry-COM-PTCP thing   he   road   NEG=look-PRIV stone-LOC
    jewgeč
    stumble.INTR.3SG
    'The one with the berries, he didn't look at the road and he stumbled on a stone.'

(9) ***mit aru:lək***   *l'e: [ńi=jo:-r]*   Ø   *mo:rqən anńa:-nun-d'əli*
    1PL tongue-INST PTL REC=see-SS.IMPF.CVB   only   speak-HAB-INTR.1PL
    'Our language, we speak it only when we meet.'

(10) ***mit aru:***   *könmə-l'ə*   Ø   *əl=kuril'i:,*   *met ten*
     1PL language other-NLZR   NEG=know.NEG.3SG 1SG DEM
     anńa:-nun-d'əŋ.
     speak-HAB-INTR.1SG
     'Our language, some don't know it, and I'm able to speak it.'
     (canonical case: accusative ***mit aru:-ɣanə***)

2.3   Whitesands

Whitesands is a verb-medial language, with a strict ordering of constituents (SVO). The argument structure is marked by word order, as there is no overt case marking for core arguments. There is no evidence to suggest that topicalisation or any other strategy has any influence on the ordering of constituents, so that LD is

the only construction that can alter the SVO order. Verbs are obligatorily head-marked for subjects through a series of agreement prefixes. This agreement prefix alone is often a sufficient argument (i. e. it does not require an additional nominal instantiation), and no referential information can intercede between the prefix and the predicate. All other verbal morphology is either clause operators, or manner/direction information. The Whitesands sentence structure can be summarised as:

(11) LD_[S_V_O_Obl]

LD is often resumed by a pronoun, but this is not obligatory, as there are cases of non-subject LD without resumption (14). For LD of subject arguments, resumption is the only clear criterion for identification, as there are no unique intonational or morphosyntactic cues that indicate that a constituent is no longer in the original pre-verbal position. For non-subject constituents identification of LD is clearer, since any pre-subject phrase must be considered a LD. These can occur with or without a resumptive element. If a non-subject, non-possessive constituent is resumed as a pronoun, this is usually in the special resumptive form *la-* (sg) or *e-* (NON-SG) (12). Thus, LD is:

(a) any phrase followed by a co-referential resumptive pronoun or noun (12);
(b) any non-subject argument that precedes the subject constituent or prefixed verb (13),(14).

(12) ***nepien mən u***   *kitah kostawafa*   *anah iien* ***la-n***
     bait   PL   PROX 1PL   1-PROG.NEG.PL-troll yet   NEG RES-3SG
     'These baits here, we haven't fished them yet.'

(13) ***tupunis*** *t-am-aro*   *m-ø-ətaŋit*   **Ø**
     tupunis 3SG-PST-why ER-SG-break.by.dropping
     'The tupunis, how did it break?'

(14) ***nanmetaw u***   *ne-k-ø-ek*   **Ø**
     hook   PROX 2-NPST-SG-touch
     'This (fishing) hook, you touch it.'

2.4   Summary

The positional and morphological criteria for all three languages are summarised in Table 1.

	Resumptive element	Non-governed case	Position		
			Pre-Subject	Pre-Focus	Pre-Adjunct-Clause
Avatime	+		+	+	+
Tundra Yukaghir	+	+			+
Whitesands	+		+		

Table 1: Criteria for LD in Avatime, Tundra Yukaghir and Whitesands

What is clear from the above is that there are no reliable criteria to determine LD cross-linguistically. Resumptives do universally indicate LD, but they are merely a sufficient, not a necessary condition for a structure to count as LD, since zero anaphora is frequent across languages. As adumbrated in the introduction, the criteria have to be established on a language-to-language basis, depending on the structural properties of the language, including features such as branching direction, head vs. dependent marking, structural positions for subjects, centre embedding of adjunct clauses, the phrasal nature of the focus, etc. In the next section, we will investigate structural similarities and dissimilarities between LD in the target languages. We demonstrate that all three languages do have a grammaticalised type of LD, i. e. a grammatically defined construction including some kind of extraclausal element occurring before the clause.

3 Similarities and dissimilarities

The investigation of the syntactic properties of the LD structures identified with the help of the criteria described in the previous section yields mixed results: while some of the features occur in all three languages and point to syntactic similarities, others are restricted to only one of the languages and imply underlyingly different structures. In this section, we enumerate the most important similarities and differences between LD structures in Avatime, Tundra Yukaghir and Whitesands and discuss them in an informal way; the results will be analysed in more detail in Section 4.

3.1 Shared features

Features that all three languages share shed some light on the nature of the syntactic positions of LD. First, LD can be iterated: multiple LDs are attested in all three languages, as illustrated in (15), (16) and (17).

(15) Avatime

mɔ   mɛ te   xunyɔɛ, mebi   ̣ṭiabà$_i$,   bɛ-k$_ḭ$ ba$_i$
1SG.CTR 1SG like.this CTR   1SG.POSS:children c_{1s}.two c_{1p}.SBJ-give c_{1p}

ku-plikpa.
c_{6p}-letter

'As for me, like this, my two children$_i$, they gave them$_i$ letters.'

(16) Tundra Yukaghir

taŋ aṅmil,   taŋ la:mə-p-lə, taŋ či:,   *taŋ, tittə   la:mə-pə-ɣanə*
that transport.dog that dog-PL-ACC that people that 3PL.POSS dog-PL-ACC

wa:j   tittəl ta:t   ńe:d'il ŋo-ri:-nun-ŋa:,   jaqta:-nun-ŋi.
again 3PL   thus story   be-APPL-HAB-3PL.TR sing-HAB-3PL.INTR

'Those sledge dogs, those dogs, those people, uhm, they tell stories about their dogs, they sing (about them).'

(17) Whitesands

swah u,   in u,   Nalaw *t-am-ol   trapol kam lah.*
man   PROX 3SG PROX Nalaw 3SG-PST-make trouble to   3PL

'This man here, he, Nalaw made the trouble for them.'

The question of multiple LDs has been a part of the discussion on the proper position of LD elements. A long-standing line of (mostly generative) research (e. g. Lasnik & Saito 1992: 78ff, Puskás 2000: 190) considers LD and TOP to occupy the same, left-peripheral position, and to differ only in the type of the movement operation involved. One indication for this is thought to be the alleged impossibility of having multiple LD elements in languages like English, Italian and Hungarian[4] — if there is only one position for LD elements, then there can be only one such element per clause. Our data show that this is not the case in our target languages, and, consequently, no special non-iterative left-peripheral position can be assumed.[5] Instead, the possibility of multiple LDs seems to indicate either a fully

---

[4] The empirical validity of this claim remains to be re-investigated.

[5] It is, of course, possible to construe a special clause-internal functional projection for each of the multiple LD elements in the spirit of linguistic cartography, as first conceived by Rizzi (1997). We reject this possibility on methodological grounds. The procedure of multiplying functional projections whenever a need occurs is, as pointed out by Bouchard (2009), among others, so generous to the analysing linguist that it can generate any surface word order, which makes it practically unfalsifiable.

separate sentential fragment (an 'orphan'), or a repeatable syntactic operation, such as adjunction (see Shaer 2009: 387 on the impossibility of differentiating orphans and adjuncts by way of multiple LDs). In any case, multiple LDs show that LD in Avatime, Whitesands and Tundra Yukaghir is not tied to one dedicated position in the left periphery of the clause.

Second, so-called Chinese-style topics occur in all three languages. We follow Chafe (1976) in using this term to denote preclausal elements which do not play an argument or an adjunct role in the clause, but are interpreted as delimiting the field of relevance of the proposition expressed by the clause, as in *The earth-quake, the government was slow to help.* Due to their clear connectedness to the clause and the relevance they have for its interpretation, these elements clearly fall under our heuristic concept of LD (Section 1); furthermore, as we show in Section 4.2, Chinese-style topics have the same syntactic distribution as LD with resumptive elements in at least one of the target languages, Avatime. Examples (18), (19) and (20) illustrate left dislocated Chinese-style topics in Avatime, Tundra Yukaghir and Whitesands.

(18) Avatime

ku-wò=e,   a-wlakpa-là lɛ-gbí
c_{5s}-fever=CM c_{3p}-leaf-DEF c_{3p}.SBJ-many
'Fever, there are many leaves (sc. that can be used to cure it).'

(19) Tundra Yukaghir

qajl', *tude*   *ugurčə puń-l'ʾəl-mələ*
stone REFL.POSS leg   hurt-EVID-OBJ.FOC.3SG
'Stone, he hurt his leg.'

(20) Whitesands

mani,   *parien t-ol*   *had, ah* **mani,**   *(0.5)*
money true   3SG.NPST-make hard HES money

metou sait raha nawanien, kapa
but   side POSS food   no
'Money, it is true that it is hard, money, but in terms of food, no.'

We see Chinese-style topics as a further indication that LD in our three languages is not tied to an argument-like position within the left periphery of the clause, as

LD elements in (18)–(20) are certainly not argument-like entities. The interesting fact is that, in their capacity as delimiters of relevance, Chinese-style topics take scope over the whole clause. This can once again be interpreted in two ways: Chinese-style topics in the three target languages are either separate clausal fragments whose scope-taking properties are a product of pragmatic inference at the level of discourse, or they are adjoined to the clause and are thus structurally high enough to be able to scope over the whole proposition.

Thirdly, in all three languages, LD structures display lack of sensitivity to island constraints, as shown by the dependencies between the LD and an argument of classical syntactic islands, adjunct (examples (21) and (22)) and relative clauses (23). Island violations with a resumptive pronoun in the island are attested in natural discourse in all three languages, and even without the resumptive pronoun, these constructions are still judged grammatical in Tundra Yukaghir and Whitesands and by some Avatime speakers.

(21) Tundra Yukaghir

taŋ kö:d'ədo:$_i$, awja: [*eńe:-gi* **(tude-ɣanə)**$_i$ *uba:nu-də-ɣa*]
that boy yesterday mother-3POSS 3SG-ACC kiss-3-DS.CVB
kewečəli.
leave.INTR.1PL
'That boy$_i$, we left while his mother was giving (him$_i$) a kiss.'

(22) Whitesands

John$_i$, *niamaha t-em-eru* [*mə in* **ko**$_i$ *k-am-eni nerek i*]
John anger 3SG-PST-see COMP 3SG that 3-PST-SAY poison TRNS
'John$_i$ was angry because they said that he$_i$ is poisonous.' [literally: John, the anger saw him because he was said to be poisonous.]

(23) Avatime

ɔ́-dzɛ́ **lɛ́-lɔ́=ɛ**$_i$ *mè-te ó-nyime* [*gì a-tá-gbanì* **yɛ**$_i$]
C$_{1s}$-woman C$_{1s}$-PROX=CM 1SG.SBJ-know C$_{1s}$-man REL C$_{1s}$.SBJ-INT-marry C$_{1s}$
'This woman$_i$, I know the man who will marry her$_i$.'

The way these island violations are interpreted depends on one's theory of syntactic islands (see Hoffmeister & Sag 2010 and Boeckx 2012 for good overviews). On syntactic accounts of islands, subjacency-based or otherwise, examples like

(21), (22) and (23) entail that the LD element does not belong to the same syntactic domain as the gap/pronoun with which it enters the filler-gap dependency. On a processing account, the LD element is processed separately from the clausal domain, so that its coreference with a gap does not overburden the human parser. On both accounts, the lack of island constraints in LD constructions in Avatime, Tundra Yukaghir and Whitesands shows again that LD elements cannot be considered elements of the clause-internal left periphery. As is the case with the previous two properties, island insensitivity is compatible with both adjunction and orphan analyses of the LD in the target languages. In RRG terms, this means that LD is either within the sentence, but outside of the clause, or that it is a separate sentence fragment.

Finally, all three languages allow noun phrases as resumptive elements. Examples are given in (24), (25) and (26).

(24) Avatime

> ***kipɛ***   *mó-te*   *ki-diyɛ*   *bɛ-kɔ*   *bi̠tɛ*
> t.o.medicine 1SG.SBJ.NEG-know c_{4s}-thing:PROX c_{1p}.SBJ-take make
>
> ***kipɛ***   *xunyɔ*
> t.o.medicine CTR
>
> 'Kipe (type of medicine), I don't know those things they use to make kipe.'

(25) Tundra Yukaghir

> ***la:mə-lək,*** *lewejmədəŋ qańd'əmədəŋ taŋ* ***la:mə-lək***
> dog-INST   in.summer in.winter   that dog-INST
>
> *ańmil-ńa:-nun-ŋi.*
> riding-PROP-HAB-3PL.INTR
>
> 'With dogs, in summer as in winter, they travel with dogs (sc. on dog sledges).'

(26) Whitesands

> ***penses raha-n,*** *ko* ***penses raha-n*** *t-askilim*
> pliers   POSS-3SG then pliers   POSS-3SG 3SG.NPST-hold
>
> 'His pliers, then his pliers have grabbed it.'

Note that full NPs that function as resumptive elements are not epithets, which according to some analyses can be bound clause-internally (see Alexiadou 2006).

These examples are therefore problematic for those accounts in which the LD element occurs in the left periphery of the clause in a dedicated position. The problem is that a clause-internal LD element would be co-indexed with an R-expression which it c-commands, in violation of Principle C of the Binding theory (Shaer 2009). If we treat LD as adjoined, the problem of the binding of resumptive NPs could be solved by assuming that adjuncts do not form an obligatory binding domain, as with certain types of English PPs (*They$_i$ hid the money behind them$_{i,j}$*). Of course, no stipulation is necessary if we assume that LD elements are orphans. For an RRG-based account, this does not necessarily follow. LD can be an orphan or an adjunct, but it can also be an element in the left-detached position (LDP) outside the clause but inside the sentence. In this way, it would be outside of the binding domain but still part of a larger syntactic unit.

The evidence of the features that all three target languages share suggests that LD in these languages does not have a dedicated clause-internal left-peripheral position. One possible way to save the clause-internal analysis of at least some instances of LD would be to try and differentiate different types of LD within each language, similar to the analyses that have been proposed for Germanic and Romance languages, so that one type of LD would be clause-internal and derived by some kind of movement, while the other would be of the clause-external hanging topic type (see Section 1). However, we were not able to find any clustering of features that would justify this division. For instance, it is not the case the tokens of LD carrying features that indicate extra-clausal LD, such as multiple LD, Chinese-style LD, or resumptive NPs, are characterised by one set of morphophonological properties (prosodic break, non-restricted resumptives, etc.), whereas those tokens that do not have these features have a different set of properties (no break, restricted type and position of the resumptive, etc.). We therefore assume that each of the target languages has only one type of LD, in absence of evidence to the contrary.

The shared features listed above indicate that the LD structure in Avatime, Tundra Yukaghir and Whitesands can be treated as (a) a sentence-internal element, outside of the clause proper but inside the sentence (in RRG terms) (b) adjoined to the clause at the sentential level (both in generative and in RRG terms), or (c) as an independent utterance ('orphan') connected to the clause inferentially. Options (a) and (c) seem more likely because of the occurrence of resumptive nominal phrases.

3.2 Structural differences

The three target languages differ in a number of respects, which seems to indicate that the position of LD elements in each of them might be similar, but is not identical. In what follows, we address the two most conspicuous differences.

   The first diverging feature is the possibility of embedded LD structures. As early as Ross (1967 [1987]), it was noted that LD structures cannot occur within embedded clauses. This has been considered a criterial property of the non-integrated type of LD, which immediately follows from its position outside of the clause proper: if it is the clause that is embedded, than elements that do not belong to it cannot undergo this syntactic operation (apart from some marginal exceptions identified by Ross; cf. Aissen 1992: 68). Among our target languages, Tundra Yukaghir and Whitesands fully conform to this restriction, as shown for Tundra Yukaghir in (27). In Avatime, however, LD elements can freely appear in all complement and some adjunct clauses, as in (28) and (29) (see Van Putten 2014 for more detail).

(27) Tundra Yukaghir

   * *awja:* **[Ivan,** *Po:d'ə tude-ɣanə paj-o:l-ɣanə]* *mə=jö:-ŋ*
   yesterday Ivan Pod'e 3SG-ACC hit-STAT.NLZR-ACC EX=see-TR.1SG
   Intended reading: * 'Yesterday, I saw that Ivan, Pod'e hit him'.

(28) Avatime

   sị ba sị **[ì-tsré lɔ gì ɛlɔm a-kɔ̀ maní ke-pa mɛ̀**
   tell c_{1p} COMP c_{2p}-okra DIST REL Elom c_{1s}.SBJ-take bring c_{6s}-house inside
   ba-da lɛ kị́ wɔ]
   c_{1p}.SBJ.SBJV-sell c_{3s} give 2SG
   'Tell them that the okra that Elom brought to the house, they should sell it to you.'

(29) Avatime

   blɔ petee kị́à-tsa **[tɔ kị-bɔ̀-ɛ** *bị́à-lɛ kɛ kị́*
   1PL all 1PL.SBJ.POT-meet PURP c_{4s}-money-DEF c_{1p}.SBJ.POT-share c_{4s} give
   ɔ̀-ma kákaa]
   c_{2s}-town every
   'We will all meet so that the money, we share it to every town.'

The obvious conclusion to be drawn out of this is that LD in Avatime is in some way more closely integrated with its host clause than is the case in Tundra Yukaghir and Whitesands. More precisely, it seems to be the case that in Avatime, but not in Tundra Yukaghir and Whitesands, LD and the clause form a unit which is a syntactic object that can be subject to embedding. This means that Avatime LD elements cannot be analysed as orphans, i. e. independent utterances that are linked to the clause via discourse linking.

The second feature in which the target languages differ is the nature of the resumptive expression. It has been noticed before that LD structures can differ in the type of the resumptive expression, both within and across languages. It is usually assumed that the integrated type tends to employ a specialised resumptive pronoun in well-defined positions, whereas non-integrated LD allows for all types of resumptive expressions in all positions (see Alexiadou 2006 and Shaer 2009). All our target languages allow for both nominal and pronominal resumptive elements, as well as zero anaphora (see Sections 2 and 3.1), but only Whitesands also has a specialised pronominal form which occurs only with LD. In Whitesands, this LD-specific pronominal expression is *la-* (SG) or *e-* (NON-SG), as illustrated in (30). Regular independent pronouns can also be used as resumptives, as in (31).

(30) Whitesands

    ra   *tiapen mən aha$_i$ n-am-at-ivi*   **la$_i$**

    POSS tuna   PL that 2-PST-PROG-SG-pull RES

    'Those tunas, you fish them.'

(31) Whitesands

    ka-Sauieh$_i$ **in$_i$** *t-apa*   *m-ø-uven m-ø-ol*   *tion e*

    DEIC-Sauieh 3SG 3SG.NPST-no ER-SG-go ER-SG-make join DAT

    'But this Sauieh, he didn't want to, so he went and joined in.'

Since syntactic tests yield identical results for both *la* and all other types of resumptives, we assume that Whitesands does not have two or more LD structures on a par with languages like German or Italian (see e. g. Cinque 1997, Grohmann & Boeckx 2004). In other words, Whitesands has developed a specialised marker of coreference with the LD element without concomitant syntactic specialisation. The existence of such a marker seems to imply a closer connection between the

LD and the clause than would be the case if LD were fully syntactically independent, i. e. it points in the same direction as the embedded LD structures in Avatime. Whitesands thus has one signal of closer integration of LD elements with their host clauses.

3.3 An interim summary

The similarities and dissimilarities of LD structures in Avatime, Tundra Yukaghir and Whitesands are summarised in Table 2.

	multiple LD	Chinese-style LD	island in-sensitivity	resumptive NPs	embedded LD	specialised resump-tive
Avatime	+	+	+	+	+	
Tundra Yukaghir	+	+	+	+		
Whitesands	+	+	+	+		+

Table 2: Features of LD in Avatime, Tundra Yukaghir and Whitesands

The evidence of the six criteria applied here is contradictory. On the one hand, the four shared features described in Section 3.1 point to an analysis of LD as non-integrated, be it as an orphan or adjunct (both in RRG and in generative terms) or as a sentential extraclausal element (RRG). On the other hand, embedded LDs in Avatime and the specialised resumptive *la* in Whitesands can be understood only if we assume a more integrated nature of LD in these two languages. In the following section, we shall try to tackle this dilemma from a more general perspective and offer a tentative conclusion.

4 The nature of LD

4.1 Status of LD in Tundra Yukaghir

Tundra Yukaghir provides the only relatively clear case of fully non-integrated LD. The features enumerated in Sections 2.2 and 3 point to syntactic independence of the LD element from the clause. Approaches that recognise only one syntactic level, the level of the clause, must treat non-integrated LD constructions as consisting of a separate sentential fragment (orphan) linked to the remainder of the utterance via discourse-based interpretive rules. If we analyse Tundra Yukaghir LD within the framework of RRG, its non-integrated nature can

be accounted for in two ways: (a) the left-dislocated element is an orphan, or (b) the left-dislocated element occurs in the left-detached position (LDP), outside of the clause but within the sentence, following the standard RRG analysis of LD. For the latter analysis to work, we would have to assume that consecutive (non-embedded) LDPs are allowed within one sentence, which is something of a moot point in the theory; we will elaborate on this further in Section 5.

4.2 Status of LD in Avatime

At first sight, most features of the Avatime LD construction characterise it as a non-integrated structure. However, its frequent occurrence in embedded contexts strongly points to an analysis as integrated LD. Importantly, there is no evidence that there are two distinct types of LD in Avatime, an integrated and a noninte-grated one. All the types of elements that can be left dislocated in main clauses can also be left dislocated in embedded clauses. Even Chinese-style LD, which could easily be conceived of as a separate type due to its loose connection to the remainder of the sentence, is subject to embedding, as shown in (32).

(32) Avatime

*nítemesì kui-do   sì   **ì-kɔɛ́**   **lɛ́-yá**   **tete**   ɔ-bìtɛ lɔ̀=ɛ   blɔ*
SO   1PL.SBJ-say COMP C3p-custom C3p-PROX like.this INF-do DIST=CM 1PL

kụ-lị̀   ní   committee líyὲ   abà ní   ɔvanɔ́ yà   blɔ petee blɔ
1PL.SBJ-be.at LOC committee C1s:PROX on   LOC Vane   here 1PL all   1PL

kíà-dzɛ   babiakpa
1PL.SBJ.POT-go Biakpa

'We said that the performance of those customs, we who are in this commit-tee here in Vane, all of us will go to Biakpa.'

This means that the integrated and non-integrated features of Avatime LD need to be reconciled within one analysis. To this end, we propose to add another type of LD to the typology: *loosely integrated LD*. The label 'loosely integrated' captures the fact that Avatime LD elements cannot be analysed as orphans. Their occurrence in embedded contexts implies that they form a unit together with the remainder of the sentence. At the same time, they do not behave like other integrated elements.

One possible solution would be to analyse Avatime LD as adjunction. We have seen in Section 3.1 that the occurrence of resumptive NPs is a possible problem for this analysis, but that it can be solved by redefining the binding domain in Avatime, i. e. by assuming that adjuncts are outside of the binding domain. If we take a RRG perspective, the syntactic account of loosely integrated LD becomes unproblematic. Avatime LD elements can occur in the left-detached position (LDP). In this way they form a unit with the remainder of the sentence, but are outside of the clause and thus not sensitive to clause-internal syntactic restrictions. The occurrence of resumptive nouns is not a problem on this account, as the LD element is outside of the clause, i. e. of the binding domain. Chinese-style LD is also unproblematic, as there are no restrictions on the semantic connectedness of the element in the LDP to the clause. The only assumptions we have to make are that sentences as a whole can be subject to embedding and again that the LDP can be iterated, both of which are not difficult to accommodate in the theory (see Section 5). In conclusion, the loosely integrated LD of the Avatime type can either be treated as an adjunct (if we assume a redefinition of the binding domain) or, with the assumptions mentioned above, as an element in the LDP within the sentence.

4.3   Status of LD in Whitesands

LD in Whitesands, like in the other two languages, looks non-integrated at first sight. However, like Avatime, it also shows one symptom of integration: the specialised resumptive pronoun. As was shown in Section 3.2, Whitesands has a specialised resumptive pronoun *la-/e-*, which occurs only in LD constructions. This pronoun is optional; regular independent pronouns, nouns and zero anaphora are also allowed as resumptive elements.

This specialised resumptive pronoun leads to a problem under the orphan analysis: how can a separate syntactic fragment trigger a pronoun that is only used with LD and not in other contexts? As mentioned in Section 3.2, there is no evidence that Whitesands has more than one type of LD, such that one has a special pronoun, while the other uses regular pronouns or NP gaps. This seems to leave us with the option to assume that all LD structures in Whitesands are integrated, in the way Italian Clitic LD or German Contrastive LD are integrated into the sentence and trigger specific pronouns (Cinque 1997; see Alexiadou 2006 and Shaer

2009 for an overview of the literature). However, we see no necessity to postulate that LD is of the integrated type in Whitesands because of *la*-resumptives. Given that all other types of evidence point to its lack of integration, we propose a fourth type of LD to account for the Whitesands data: *non-integrated bound LD*. This label is meant to refer to the type of LD in which LD elements can be partially syntactically bound, despite being outside of the sentence. The idea that independent syntactic units can be linked not only via discourse constraints, but also via syntactic rules, is supported by evidence that such clausal phenomena *par excellence* as VP-ellipsis (Van Valin 2005: 231ff) or case assignment (Shimojo 2008) can occur across conversational turns, where the preceding turn controls the syntax of the following turn. We thus presume that in Whitesands, under certain discourse conditions, a certain type of sentential fragment, LD, can trigger the use of the special pronoun *la-/e-* in the clause instead of the canonical personal pronoun. We still have no clear idea of the exact nature of these discourse conditions, but the principle according to which pronoun choice is triggered by previous clauses/sentences is well-attested cross-linguistically (see e. g. Kuno 1987)[6] and thus also a plausible explanation for the use of *la-/e-* in the Whitesands LD construction.

4.4 Expanding the LD typology

As the previous two sections have shown, a simple two-way typology of non-integrated LD versus integrated LD cannot capture all properties of the three languages discussed in this paper. In particular Avatime and Whitesands show properties of both integrated and non-integrated types of LD. There is no evidence in these languages that some LD structures are integrated and some are non-integrated. Therefore, we have proposed to add two more types of LD to the typology, which are between integrated and non-integrated.

The first type is loosely integrated LD. As the name suggests, this is a type of integrated LD that is more loosely connected to the remainder of the sentence than typical integrated LDs. This occurs in the situation in which an LD element is within the sentence proper (integrated), but fails to be sensitive to certain

---

[6] A well-known example is the choice between *er/sie* and *der/die* in German, which is (at least partly) determined by the grammatical role of the antecedent in the previous sentence (see e. g. Bosch & Umbach 2007, Bosch, Katz & Umbach 2007), such that *er/sie* will be preferably used with subject antecedents, *der/die* with objects.

phenomena within the sentence because of being in a position which cannot control them. That is, if an LD construction is not sensitive to some characteristic sentential processes, this still does not imply that it is outside of the sentence. As we have shown in Section 4.2, this type of LD can explain the phenomena we find in Avatime, where LD appears to be have non-integrated properties but is at the same time subject to embedding.

The second type we propose is non-integrated bound LD. This type can account for LD elements that are non-integrated but still control certain phenomena in the sentence. We assume that it is possible for non-integrated LD elements to entertain certain conventionalised, syntactic relationships with the following sentence/clause. This type of LD accounts for the Whitesands data, where LD shows all the signs of being non-integrated, even though there is a specialised resumptive pronoun *la* which occurs only in LD constructions.

Our revised LD typology thus contains four categories: (i) non-integrated, (ii) non-integrated bound, (iii) loosely integrated, and (iv) integrated. These are by no means all possible types that could be envisaged and it may well be that the typology needs to be further refined if data from more languages is taken into account.

5   Summary and conclusion

In this paper, we have investigated the LD constructions in three genealogically and geographically distant languages, Avatime, Whitesands and Tundra Yukaghir, in order to shed some light both on the possible variability of this seemingly simple construction and to contribute to a definition of the category of LD which can stand the test of data coming from languages not belonging to the Standard Average European type.

We limited our exploration only to those cases of LD-like phenomena which are reflected in the syntax, leaving out conversational phenomena such as false starts, hesitations, etc. We show that the major means of identifying LD, resumptive expressions, are indeed a reliable indicator of LD if they are present, but that, due to productive mechanisms of zero anaphora in many languages, their absence is not an indicator that no LD is involved. With this allegedly universal criterion for LD out of the way, the only method of identifying LD remains a definition of the construction on a language-to-language basis.

Once the properties of the structures identified as LD in the target languages are investigated, it turns out that the phenomena which look superficially similar correspond to distinct syntactic structures in each of the languages. Tundra Yukaghir LD seems to be an instance of fully non-integrated LD, most probably a separate syntactic unit, an orphan, lacking any clear syntactic connection to the clause to which it interpretatively relates. While the structure of LD in Whitesands seems to be similar, the connection to the clause is somewhat tighter, as indicated by a conventionalised use of certain pronominal types as resumptives within the clause. Thus, while Tundra Yukaghir LD is of the proper non-integrated kind, the corresponding structure in Whitesands belongs to a transitional category of non-integrated bound LD. The evidence of syntactic tests for Avatime shows that LD in this language belongs to the sentence and is thus of the integrated type. However, LD elements fail to control some syntactic phenomena which they should control if they had a dedicated position in the sentence hierarchy. Reasons for this can be different: LD elements might be sentence-level adjuncts, or have some other property which prevents them from participating in all sentence-level processes. However this phenomenon is explained, Avatime LD can be classified as special, loose integrated type of LD.

We have left our discussion so far at a rather theory-neutral level, assuming merely the existence of LDs separate from the sentence and those somehow integrated into it. The question is what this revised typology means for RRG. How can the distinction between integrated and non-integrated be captured in RRG terms and can the categories of non-integrated bound LD and loosely integrated LD fit in?

The standard analysis of LD elements in RRG is that they occur in the left-detached position (Van Valin 2005: 6), i. e. they are treated as a kind of entity between integrated and non-integrated in the classical generative division. Within RRG, no notion of orphan has been proposed, but the idea of a separate syntactic fragment could be easily accommodated in the theory. However, it is not immediately clear whether this is needed to account for LD within this framework. In the generative theory, the notion of orphan accounts for the extraclausal properties of non-integrated LD, since no intermediate level between clause and discourse exists. However, in RRG, the LDP is already extraclausal. In fact, the LDP has all properties that an orphan is supposed to have and at the same time it has the advantage of being within the syntactic structure. This means that the LDP can

also account for non-integrated LD, as well as for loosely integrated LD, as we have seen in Section 4.2. On the other hand, we think that the existence of independent sentential fragments could be a useful addition to the theory, because it would enable it to straightforwardly account for the difference between fully non-integrated LD structures, such as the one in Tundra Yukaghir, and various intermediate types, like the ones we found in Avatime and Whitesands.[7]

Our data can contribute to the theory in three further respects. First, it is unclear whether RRG can deal with multiple LD elements in one sentence. In the RRG sentence architecture, the assumption is usually made of only one LDP per sentence. This would be a problem in the analysis of LD in Avatime, which clearly includes the LDP and can nevertheless host more than one element. We do not see any inherent reason for the assumption that the LDP is not iterative, so that we would like to propose the possibility of multiple LDPs per sentence in RRG. Secondly, RRG does not freely permit sentential subordination, which would be necessary to account for Avatime subordinate LD. Sentential subordination is a part of the RRG architecture, but it is restricted to sentences occurring in the detached position of another sentence (Van Valin 2005: 192). We would propose to extend the possibility of sentential subordination to other types of subordinate clauses (for more details see Van Putten 2014). Finally, our Whitesands data demonstrate that the repertoire of syntactic dependencies across sentences, which has been limited to VP ellipsis and case assignment in the RRG literature (Section 4.3), can be expanded to include specialised resumptive pronouns triggered by LD elements across sentence boundaries, unless, of course, the Whitesands LD elements are also assumed to be placed in the LDP. On the methodological side, we aim to have demonstrated that the identification of LD across languages is no trivial matter, but requires a careful investigation of the potential LD structures on a language-to-language basis.

Our findings, preliminary as they might be, show that neither the simple dichotomy of integrated vs. non-integrated LD types nor the single LDP at the sentential level are sufficient to capture all the various disguises in which LD can occur in natural languages. For this reason we propose to view LD as a type of structure which can occur in three positions – as an orphan, in the LDP and, prob-

---

[7] The fully integrated type is not attested in our target languages, but it seems to be quite well documented in Romance, Germanic, and in Greek (Anagnastopoulou 1997, Cinque 1997, Alexiadou 2006, Shaer 2009), so that RRG would probably have to allow for a clause-internal type of LD, too.

ably, clause-initially (see footnote 7). These various positions are at least partly responsible for the differences we find across languages. We define two types of LD that are intermediate between fully integrated and non-integrated – non-integrated bound and loosely integrated – which help us classify our data. Future research, both on our target languages and on a wider sample of languages, will likely reveal further subtypes and result in a more fine-grained classification.

Abbreviations

A	actor	INTR	intransitive
ACC	accusative	LOC	locative
ASSOC	associative	NEG	negative
AUG	augmentative	NLZR	nominalizer
C-$_{S/P}$	nominal class-singular/plural	NPST	non-past
COM	comitative	PL	plural
COMP	complementizer	POSS	possessive
CONTR	contrast	POT	potential
CVB	converb	PRIV	privative
DAT	dative	PROG	progressive
DEF	definite	PROX	proximal
DEIC	deictic	PST	past
DIST	distal	PTCP	participle
DS	different subject	PURP	purposive
ER	echo referent	REC	reciprocal
EX	existential	REFL	reflexive
FOC	focus	REL	relativizer
HAB	habitual	RES	resumptive
ID	ideophone	SBJ	subject
IMPF	imperfective	SG	singular
INCL	inclusive	SS	same subject
INTJ	interjection	STAT	stative
INST	instrumental	TRNS	transitive
INTSF	intensifier	U	undergoer

References

Aissen, J. 1992. Topic and focus in Mayan. *Language* 68: 43–80.

Altmann, H. 1981. *Formen der 'Herausstellung' im Deutschen.* Tübingen: Niemeyer.

References

Alexiadou, A. 2006. Left dislocation (including CLLD). In M. Everaert & H. van Riemsdijk (eds.), *The Blackwell companion to syntax, Vol. II.* Malden: Blackwell.

Anagnostopoulou, E. (1997). Clitic left dislocation and Contrastive left dislocation. In E. Anagnostopoulou, H. van Riemsdijk & F. Zwarts (eds.), *Materials on left dislocation*, 668–699. Amsterdam: Benjamins.

Avanzi, M. 2011. La dislocation à gauche avec reprise anaphorique en français parlé. Etude prosodiqu. In H.-Y. Yoo & E. Delais-Roussarie (eds.), *Actes d'IDP 2009, Paris, Septembre 2009.* http://makino.linguist.jussieu.fr/idp09/actes_en .html/, accessed September 18th 2013.

Avanzi, M., C. Gendrot & A. Lacheret-Dujour. 2010. Is there a prosodic difference between left-dislocated and heavy subjects? Evidence from spontaneous French. *Proceedings of Speech Prosody 2010, Chicago* 100068. http:// speechprosody2010.illinois.edu/, accessed September 18th 2013.

Boeckx, C. 2012. *Syntactic islands.* New York: Cambridge University Press.

Bosch, P., G. Katz & C. Umbach. 2007. The non-subject bias of German demonstrative pronouns. In M. Schwarz-Friesel, M. Consten & M. Knees (eds.), *Anaphors in text: Cognitive, formal and applied approaches to anaphoric reference*, 145–165. Amsterdam: Benjamins.

Bosch, P. & C. Umbach. 2007. Reference determination for demonstrative pronouns. In D. Bittner (ed.), *Proceedings of the Conference on Intersentential Pronominal Reference in Child and Adult Language.* Special issue of *ZAS Papers in Linguistics*, Vol. 48, 39–51.

Bouchard, D. 2009. A solution to the conceptual problem of cartography. In J. van Craenenbroeck (ed.), *Alternatives to cartography*, 245–274. Berlin: Mouton de Gruyter.

Chafe, W. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (ed.), *Subject and topic*, 25–55. New York: AP.

Cinque, G. 1990. *Types of A' dependencies.* Cambridge, MA: MIT Press.

Cinque, G. 1997. 'Topic' constructions in some European languages. In E. Anagnostopoulou, H. van Riemsdijk & F. Zwarts (eds.), *Materials on left dislocation*, 93–118. Amsterdam: Benjamins.

De Cat, C. 2007. French dislocation without movement. *Natural Language & Linguistic Theory* 25: 485–534.

Dik, S. C. 1978. *Functional grammar.* Amsterdam: North Holland.

Feldhausen, I. 2008. *The prosody-syntax interface in Catalan*. Universität Potsdam dissertation. http://prosodia.upf.edu/aev/recursos/, accessed October 1ˢᵗ 2013.

Foley, W. A. & R. D. Van Valin, Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Frey, W. 2004. Notes on the syntax and the pragmatics of German left dislocation. In H. Lohnstein & S. Trissler (eds.), *The syntax and semantics of the left periphery*, 203–234. Berlin: Mouton de Gruyter.

Geluykens, R. 1992. *From discourse process to grammatical construction*. On leftdislocation in English. Amsterdam: Benjamins.

Givón, T. 1976. Topic, pronoun and grammatical agreement. In C. Li (ed.), *Subject and topic*, 149–188. New York: AP.

Gregory, M. L. & L. A. Michaelis. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33, 1665–1706.

Grohmann, K. K. & C. Boeckx. 2004. Left Dislocation in Germanic. In W. Abraham (ed.), *Focus on Germanic typology*, 139–152. Berlin: Akademie-Verlag.

Haegeman, L. 1991. Parenthetical adverbials: The radical orphan approach. In S. Chiba, A. Ogawa, Y. Fuiwara, N. Yamada, O. Koma & T. Yagi (eds.), *Aspects of Modern English Linguistics: Papers presented to Masatomo Ukaji on his 60th birthday*, 232–254. Tokyo: Kaitakushi.

Haspelmath, M. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86: 663–687.

Hoffmeister, P. & I. Sag. 2010. Cognitive constraints and island effects. *Language* 86: 366–415.

Kuno, S. 1987. *Functional syntax: Anaphora, discourse, and empathy*. Chicago: The University of Chicago Press.

Lambrecht, K. 2001. Dislocation. In M. Haspelmath, E. König, W. Oesterreicher & W. Raible (eds.), *Language typology and language universals: An international handbook*. Vol. 2, 1050–1078. Berlin: Walter de Gruyter.

Lasnik, H. & M. Saito. 1992. *Move Alpha: Conditions on Its application and output*. Cambridge, MA: MIT Press.

Matić, D. & I. Nikolaeva. 2014. Realis mood, focus, and existential closure in Tundra Yukaghir. *Lingua* 150: 202–231.

Ott, D. 2012. Movement and ellipsis in contrastive left-dislocation. In N. Arnett & R. Bennett (eds.), *Proceedings of WCCFL 30*, 281–291. Somerville, MA: Cascadilla.

References

Prince, E. 1997. On the functions of left-dislocation in English discourse. In A. Kamio (ed.), *Directions in functional linguistics*, 117–144. Amsterdam: Benjamins.

Prince, E. 1998. On the limits of syntax, with reference to topicalization and left-dislocation. In P. Cullicover & L. McNally (eds.), *Syntax and semantics 29: The limits of syntax*. New York: AP.

Puskás, G. 2000. *Word order in Hungarian*. Amsterdam: Benjamins.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.

Ross, H. 1967. *Constraints on variables in syntax*. MIT Dissertation (published 1987 as Infinite syntax! Norwood, NJ: ABLEX).

Shaer, B. 2009. German and English left-peripheral elements and the 'orphan' analysis of non-integration. In B. Shaer, P. Cook, W. Frey & C. Maienborn (eds.), *Dislocated elements in discourse,* 465–502. New York: Routledge.

Shimojo, M. 2008. How missing is the missing verb? The verb-less numeral quantifier construction in Japanese. In R. D. Van Valin, Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 285–304. Amsterdam: Benjamins.

van Putten, S. 2014. Left dislocation and subordination in Avatime (Kwa). In R. van Gijn, J. Hammond, D. Matić, S. van Putten & A. Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*, 71–98. Amsterdam: Benjamins.

Van Valin, R. D., Jr. & R. J. LaPolla. 1997. *Syntax : Structure, meaning, and function.* Cambridge: Cambridge University Press.

Van Valin, R. D., Jr. 2005. *Exploring the syntax-semantics interface.* Cambridge: Cambridge University Press.

Van Valin, R. D., Jr. 2013. Head-marking languages and linguistic theory. In B. Bickel, L. A. Grenoble, D. A. Peterson, & A. Timberlake (eds.), *What's where why? Language typology and historical contingency. A festschrift to honor Johanna Nichols*, 91–124. Amsterdam: Benjamins.

Authors

Dejan Matić, University of Graz
Saskia van Putten, Radboud University Nijmegen
Jeremy Hammond, University of Sydney