

2020 September 25, Harvard Linguistics Circle

Ambiguity avoidance vs. expectation sensitivity in ditransitive differential argument marking

MARTIN HASPELMATH

I. Functional-typological linguistics meets experimental psychology

Boas (diversity) + Jakobson (functionalism)
> Greenberg (functional-typological linguistics)

experimental psychology, e.g.

Fedzechkina et al. (2012)

Gibson et al. (2019)

Smith & Culbertson (2020)

discovering syntactic universals

Smith & Culbertson (2020): “efficiency/ambiguity avoidance” vs. “iconicity”

Haspelmath:

- Grammatical coding universals are often due to frequency-induced predictability and a pressure for efficient coding
- Speakers are sensitive to hearers’ expectations, and ambiguity avoidance is much less important than is often thought.

(or rather: **Producers** are sensitive to **comprehenders’** expectations, because the general principle applies to spoken languages as to signed languages)

2. Differential object marking: monotransitive and ditransitive

differential object flagging (“DOM”), more specifically: **differential patient flagging**

(I) Spanish

a. (inanimate P)

Vi la casa.

I.saw the house

‘I saw the house.’

b. (animate P)

*Vi **a** la mujer.*

I.saw **ACC** the woman

‘I saw the woman.’

(flagging = case-marking or adpositional marking)

differential recipient flagging

(2) Northeastern Neo-Aramaic of Telkepe (Coghill 2010)

a. (person-form R)

kəm-yāwəl-lə *hadiynə*
 PST-he.give-3SG.M.OBJ present
 'He gave him a present.' (=I4c)

b. (full nominal R)

wəl-lə pāre ta xa-məskənə
 gave-he money to a.certain-poor.person
 'He gave money to a certain poor person.' (=Coghill's I1b)

(3) Teop (Austronesian) (Mosel 2010: 490, 495)

a. (animate R)

Eam paa hee vaha nao a Adra bon-a moonii?
 2PL TAM give back DIR ART Adra OBJ-ART money
 'Did you give the money back to Adra?'

b. (inanimate R)

O-re paa no hee ni bona te-a address vai.
 3SG-then TAM go give APPL OBJ.it to-ART address DEM
 'He should give it to this address.'

differential theme flagging

(4) Akan (Atlantic-Congo; Osam 1996: 63-64)

a. (indefinite theme)

Kofi ma-a abofra no akokɔ
 Kofi give-COMPL child DEF chicken
 'Kofi gave the child a chicken.'

b. (definite theme)

**Kofi ma-a abofra no akokɔ no*
 Kofi give-COMPL child DEF chicken DEF
 ('Kofi gave the child the chicken.')

c. (definite theme)

Kofi de akokɔ no ma-a abofra no
 Kofi take chicken DEF give-COMPL child DEF
 'Kofi gave the chicken to the child.'

2. Generalizations: role-reference associations (cf. Haspelmath 2021b)

(5) Role-reference association universals

- If the P is more referentially prominent, it is more likely to be flagged.
- If the T is more referentially prominent, it is more likely to be flagged.
- If the R is more referentially prominent, it is **less** likely to be flagged.

(6) scales of referential prominence

a. inherent prominence

person scale: locuphoric (1st/2nd) > aliophoric (3rd person)
nominality scale: **person form (independent or index) > full nominal**
animacy scale: **human (> animal) > inanimate**

b. discourse prominence

definiteness scale: **definite (> specific indefinite) > indefinite nonspecific**
 givenness scale: discourse-given > discourse-new
 focus scale: background > focus

proposed explanation: a functional-adaptive pressure for efficient coding

- Languages tend to use zero or short coding for frequently occurring meanings and functions, and overt and long coding for rarely occurring functions.
- Frequent patterns are more expected, and rare patterns are less expected, so speakers (better: producers) tend to expend more effort on the rarer patterns.
- Through adaptability in language use, languages come to have or restore efficient patterns.
- agent (A) and recipient (R) roles are usually associated with referential prominence, while patient (P) and theme (T) roles are associated with non-prominence.

Note that this also explains the English ditransitive alternation:

(7)a. (N > N)

She gave Kim the money. (\approx *She gave the money **to** Kim.*)

c. (N > pers)

**She gave Kim it.*

d. (N > pers)

*She gave it **to** Kim.*

overall, the English recipient is differentially coded when it is less expected (referentially non-prominent)

3. The argument in a nutshell

I argue that **ambiguity avoidance** is not the main motivation behind differential flagging [these considerations were prompted by Smith & Culbertson 2020].

In general, differential coding is best explained by **expectation sensitivity**. This is crucially different from ambiguity avoidance, though in many cases, the differences are not immediately apparent.

Expectation sensitivity means that speakers are sensitive to the hearer's expectations and tend to give special marking to unexpected meanings.

This can account for a broader range of phenomena than ambiguity avoidance, and is thus independently needed. Apparently, ambiguity avoidance is **not independently needed** to explain general grammatical patterns and may thus be largely superfluous.

This is similar to my 2008 argument that frequency can account for everything that iconicity of complexity/cohesion can account for, and thus **iconicity is not independently needed**.

(I cannot rule out that ambiguity/iconicity are at play, but by Occam's razor, I assume that they are irrelevant until new evidence appears)

4. Differential P flagging in the earlier literature

There is a long tradition of saying that DOM (differential P flagging) has to do with distinguishing the patient from the agent, e.g.

- (8) a. “[. . .] the principle that it is more natural for rational beings to act than to be acted upon; and hence when they do happen to be acted upon – when the nouns by which they are denoted are to be taken objectively – it becomes necessary, in order to **avoid misapprehension**, to suffix to them the objective case-sign.” (Caldwell 1856: 271; Filimonova 2005: 78)
- b. “wenn die Sprache ein transitives Verb besitzt, in gewissen Fällen der Patiens als solcher durch sprachliche Mittel **zur Unterscheidung von Agens** gekennzeichnet werden muß, weil er sonst vom Hörer als Agens aufgefaßt werden würde. (Thomson 1912: 75; Filimonova 2005: 79)
- c. “Differential case-assignment to subjects and direct objects serves the function of **distinguishing subjects from direct objects**... [Some] languages have differential case-assignment only where **confusion** between subject and direct object is particularly likely...” (Comrie 1977: 16)
- d. “it is those direct objects which are most in need of being **distinguished from subjects** that get overtly case-marked” (Aissen 2003: 437)
- e. “Many analyses of asymmetric differential object marking ... argue that those objects which look too much like prototypical subjects are marked in order **to distinguish them from the subject**.” (Malchukov & de Swart 2009: 348)

At first glance, this appears to make sense because DOM tends to be used on those types of object that are most subject-like (animate, definite), and hence presumably most confusable with subjects.

5. Expectation sensitivity

Expectation sensitivity means that speakers are sensitive to the hearer's expectations and tend to give special marking to unexpected meanings.

Since frequently used meanings are **more expected** (= more probable, more predictable, less surprising), they tend to get **shorter coding** (= short markers, or no marking).

DOM: definite/animate nominals are **less expected in the P slot** than indefinite/inanimate objects, and hence they get longer coding, e.g.

(2) Spanish

- a. *Vi la casa.*
I.saw the house
'I saw the house.'
- b. *Vi a la mujer.*
I.saw ACCthe woman
'I saw the woman.'

In other words, the **usual association** is between referentially prominent arguments and highly ranked roles (A/R vs. P/T), and special coding is found **when the association is unusual** (as explored in detail in Haspelmath 2021b).

This explanation is thus a special case of the more general explanation of asymmetric coding:

If a pattern of asymmetric coding is systematic in the world's languages, then it is due to frequency-induced predictability: The less frequent construction type gets more coding. (Haspelmath 2021a)

e.g.

singular	plural	(<i>book – book-s</i>)
nominative (A/S)	accusative (P)	(<i>he – hi-m</i>)
allative	ablative	(<i>to – from</i>)
positive	comparative	(<i>small – small-er</i>)
present	future	(<i>go – will go</i>)
affirmative	negative	(<i>go – don't go</i>)
inanimate patient	animate patient	(Spanish <i>Ø la casa – a la mujer</i>)
3rd person	2nd person	(Spanish <i>canta</i> 3SG / <i>canta-s</i> 2SG 'sing(s)')
2nd person imperative	3rd person imperative	(<i>praise! – let her praise!</i>)

6. Ambiguity avoidance and expectation sensitivity

Ambiguity avoidance seems a worthy goal, but its importance seems to have been overrated (Wasow 2015).

Communication does not simply consist in the transmission of "sets of complete thoughts". Communicators constantly keep track of each other's mental states and the common ground. If we don't know what the others know, communication would be very difficult.

To a substantial extent, communication consists in filling in a few gaps in the interlocutor's knowledge state.

To a substantial extent, the structures of languages can be understood as resulting from an **efficient trade-off between speaker effort and hearer needs** – and what the hearer primarily needs is signals about unexpected parts of messages.

Speakers must be sensitive to the hearer's expectations.

But shouldn't speakers also worry about **ambiguity**?

(cf. Grice's maxim: "Avoid ambiguity!")

Well, ambiguity or vagueness is rampant in language anyway, e.g.

A: What did you do over the weekend?

B: I went to the zoo.

[one of thousands of things, but the least expected]

Subject-object ambiguity seems particularly hard to tolerate, because there's a huge difference between, e.g.

The dog bit the postman.

The dog bit the postman.
The postman bit the dog.

But is it so big?

cf. Kim **fuhr** nach Hause.

Kim **went** home [ambiguous, vague]

Kim **ging** nach Hause.

cf. *The four project members attended five conferences.*

(between them, vs. each of them)

7. Online disambiguation vs. anti-ambiguity as a general pressure

In most languages, DOM is conditioned by **general grammatical conditions** (e.g. definiteness, nominality, inflection class, or relative prominence in case of scenario splits).

e.g. an object is flagged in L_6 if it is definite & animate & inflection class 3

But occasionally, it seems to be conditioned by **online disambiguation** (Seržant 2019):

e.g. an object is flagged if it could be confused with the subject

(9) Yongren Lolo (Trans-Himalayan; Gerner 2008: 299-300)

ηο 6εμο -t^hie tʃo ʒi.

ISG	snake-OBJ	follow go
-----	-----------	-----------

'I will follow the snake'

Sika-t^hie *χek^hu ti* *na.*

tree-OB| house smash broken

'The house smashed the tree.'

But is this really **the rule that Lolo speakers learn?**

– **online** decisions taking the **full context** into account

vs. **grammatical** conditioning taking **a few factors** into account

(10) Malayalam (Dravidian: India; Asher & Kumari 1997: 204)

- a. *Tiiyyə kuṭil naṣippacu.*
 fire.NOM hut.NOM destroy.PST
 'Fire destroyed the hut.'
- b. *Kappal tiramaalaka[-e bheediccu.*
 ship.NOM wave.PL-ACC split.PST
 'The ship broke through the waves.'
- c. *Tiramaalakaḷ kappal-ine bheediccu.*
 wave.PL ship-ACC split.PST
 'The waves broke the ship.'

Anti-ambiguity can be hypothesized to be a **general pressure** that causes grammatically conditioned DOM – but general pressures are different from language-particular regularities.

Maybe there is no language where online disambiguation is really part of a convention – maybe conventions are of two types:

- (i) grammatically conditioned, or
- (ii) flexible (involving optionality)

Cf. (temporary) ambiguity in a different domain, subject clauses:

I don't believe (that) she knows Persian well.
 *(That) she knows Persian well I don't believe.

The conventions governing the presence or absence of *that* are not **conditioned by temporary ambiguity**. (Where the conventions are flexible, the actual occurrences may be influenced by ambiguity.)

8. Expectation sensitivity explains DOM better than ambiguity avoidance

8.1. DOM may occur when the subject is ergative (and there is no ambiguity)

e.g. Hindi-Urdu

laṛke=ne gāṛī calāyī hai
 boy:OBL=ERG car drive:PRF:FEM:SG be:PRES:3.SG

'The boy has driven the car.'

laṛke=ne gāṛī=ko calāyā hai
 boy:OBL=ERG car=ACC drive:PRF:MASC:SG be:PRES:3.SG

'The boy has driven the car.'

e.g. Dyirbal

[p̪ana-na] [a̰uma-ŋgu] bura-n
we-ACC father-ERG see-NONFUT
'Father saw us.' (Dixon 1994: 130)

8.2. DOM is occasionally manifested in shorter vs. longer accusative marking

e.g. Evenki	definite accusative	-va	
	indefinite accusative	-(j)a	(I. Nedjalkov 1997)

8.3. DOM may be innovated for cases where there is an existing accusative distinction

e.g. Portuguese preposition *a*+ only with personal pronouns, e.g. *a mim* ‘me’

e.g. Ge'ez preposition *la+* for animate objects,
even though the old Semitic accusative *-a* was preserved

9. Expectation sensitivity also explains differential R and T marking (Haspelmath 2021)

9.1. Special R marking when the R is not topical

English *She gave me the money.* vs. *She gave the money **to me.***

(no ambiguity because English R-T order is rigid)

9.2. Special R marking when the T is not a full nominal

English *She gave me the money.* vs. *She gave it **to** me.*
 (**She gave me it; would not be ambiguous*)

9.3. Special R marking when the T is not 3rd person

Modern Greek *su ton éðose* vs. **tu se éðose*
 you.DAT him.ACC gave him.DAT you.ACC gave
 'she gave him to you' 'she gave you to him'
 (OK: *tu éðose eséna*)

(This is a “PCC effect”; cf. Haspelmath 2004, and it involves scenario-based predictability; Haspelmath 2021b)

Sometimes, differential R marking does seem to help avoid ambiguity, e.g.

French *il me te présentera 'he will present me to you'
OK: il me présentera **à** toi

10. Interlude: Some remarks on generative approaches to differential marking

Generative approaches typically **assume** that the explanation must come from the formal machinery of generative grammar, e.g. Aissen (2003).

They do not even consider alternative approaches in terms of functional pressures.

I believe that this is

- because they **assume** that the formal machinery is innate (because otherwise it could be different for different languages, as Boas urged his students)
- and because they **assume** that the formal machinery must be restrictive not only in enabling language acquisition (despite the poverty of the stimulus), but also in limiting the range of possible languages

I do not see sufficient grounds for making these assumptions. I find it very plausible that **the range of possible languages may be primarily limited by functional factors**.

blogposts on differential object marking:

(2018-07) <https://dlc.hypotheses.org/1119> (on Levin 2018)

(2018-10) <https://dlc.hypotheses.org/1496> (on Kalin 2018)

on differential place marking:

(2020-06) <https://dlc.hypotheses.org/2385> (on Matushansky 2019)

11. Expectation sensitivity also explains differential place marking (Haspelmath 2019)

11.1. Special place marking when the place is not inanimate

Italian	<i>vado a-lla chiesa</i>	vs.	<i>vado da-l poliziotto</i>
	'I go to the church'		'I walk up to the policeman'

11.2. Differential zero-marking when the place is a place-name

e.g. in Maltese (Stolz et al. 2017: 463)

<i>Jgħallem</i>	<i>Għawdex.</i>
3SG.M.IMPFV.teach	Gozo
'He teaches on Gozo (an island).'	
<i>Jgħallem</i>	<i>f-l-iskejjel ta-l-Gvern.</i>
3SG.M.IMPFV.teach	in-DEF-schools of-the government
'He teaches in the schools of the government.'	

Place-names do not seem to be *less* confusable with subjects than common place nouns.

12. Expectation sensitivity also explains other cases of differential coding

e.g. special possessive coding with alienable nouns (Haspelmath 2017)

Maltese	<i>id-i</i>	'my hand'	
	<i>*ktieb-i</i>	'my book'	(OK: <i>il-ktieb tiegħ-i</i>)

e.g. special independent coding with possessor pronouns (Michaelis 2019)

English	<i>my house</i>	<i>is here</i>	
	<i>*my</i>	<i>is here</i>	(OK: <i>mine is here</i>)

e.g. special past marking with stative verbs

Haitian	<i>krazé</i>	'destroyed'	
	<i>*malad</i>	'was ill'	(OK: <i>te malad</i>)

13. What is “communicative efficiency”?

Smith & Culbertson (2020: 6):

“What leads to these patterns of differential marking? One possibility, advanced by e.g. Comrie (1989) ..., is that Differential Case Marking represents a trade-off between **communicative function** and **efficiency**. Using explicit argument marking reduces the possibility of **miscommunication**, specifically reducing the likelihood of the listener **confusing** the roles of the arguments in the event being described....”

“A related explanation is that differential marking is not motivated by ambiguity avoidance per se, but represents an example of a more general **iconicity preference**, a “grand isomorphism” (Givón 1991), where **unusual events**/concepts/structures/constituents tend to be **associated with special (standardly, more weighty) linguistic material**, which Haspelmath (2008) dubs iconicity of markedness matching, where the term **marked** does double duty to refer both to atypicality at the conceptual level and weightiness in the surface signal.”

But Comrie (1978) had already said the right thing:

“There seems to be a general supposition in human discourse that certain entities are inherently more agentive than others, and as such **inherently more likely to appear as A of a transitive verb** and less likely to appear as P of a transitive verb. The mainstay of this supposition is the animacy (agentivity) hierarchy, which claims basically that more animate entities will tend to act upon less animate entities rather than vice versa.”

Bossong 1991:

The most general basic principle underlying DOM can be termed the principle of *interaction between slot and filler*:⁴² the case role of direct object, which can be considered as a syntactic and semantic slot, is subdivided into two subcategories according to the specific semantic nature of the NP which fulfills this function. It is the semantics of slot and filler which interact.

(This is similar to the idea that **roles and referential prominence are associated**; Aissen 2003; Haspelmath 2021b)

Efficiency is best defined as an optimal trade-off between **speaker effort** (articulatory energy) and **hearer needs** (explicitness) (see also Levshina 2018).

It seems that the hearer needs are primarily determined by **expectations**, and not so much by **avoiding ambiguity**.

References

- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483.
- Asher, R. E & Kumari, T. C. 1997. *Malayalam*. London: Routledge.
- Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Kibbee, Douglas & Wanner, Dieter (eds.), *New analyses in Romance linguistics*, 143–170. Amsterdam: Benjamins.
- Caldwell, Robert. 1856. *A comparative grammar of the Dravidian or South-Indian family of languages*. London: Harrison.
- Coghill, Eleanor. 2010. Ditransitive constructions in the Neo-Aramaic dialect of Telkepe. In Malchukov, Andrej L. & Haspelmath, Martin & Comrie, Bernard (eds.), *Studies in ditransitive constructions: A comparative handbook*, 221–242. Berlin: De Gruyter Mouton.
- Comrie, Bernard. 1977. Subjects and direct objects in Uralic languages: A functional explanation of case marking systems. *Études Finno-Ougriennes* 12. 5–17.
- Comrie, Bernard. 1978. Ergativity. In Lehmann, Winfred P. (ed.), *Syntactic typology: Studies in the phenomenology of language*, 329–394. Austin: University of Texas Press.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Fedzechkina, Maryia & Jaeger, T. Florian & Newport, Elissa L. 2012. Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences* 109(44). 17897–17902. (doi:10.1073/pnas.1215776109)
- Filimonova, Elena. 2005. The noun phrase hierarchy and relational marking: Problems and counterevidence. *Linguistic Typology* 9(1). 77–113.
- Gerner, Matthias. 2008. Ambiguity-driven differential object marking in Yongren Lolo. *Lingua* 118(3). 296–331. (doi:10.1016/j.lingua.2007.06.002)
- Gibson, Edward & Futrell, Richard & Piandadosi, Steven T. & Dautriche, Isabelle & Mahowald, Kyle & Bergen, Leon & Levy, Roger. 2019. How efficiency shapes human language. *Trends in Cognitive Sciences* 23(5). 389–407. (doi:10.1016/j.tics.2019.02.003)
- Givón, T. 1991. Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 15(2). 335–370. (doi:10.1075/sl.15.2.05giv)
- Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive linguistics* 19(1). 1–33.
- Haspelmath, Martin. 2019. Differential place marking and differential object marking. *STUF – Language Typology and Universals*.

- Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. to appear. (<https://ling.auf.net/lingbuzz/004531>)
- Haspelmath, Martin. 2021b. Role-reference associations and the explanation of argument coding splits. *Linguistics*, to appear (<https://ling.auf.net/lingbuzz/004047>)
- Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. New York: Oxford University Press.
- Levshina, Natalia. 2018. *Towards a theory of communicative efficiency in human languages*. Leipzig: Leipzig University.
- Malchukov, Andrej & de Swart, Peter. 2009. Differential case marking and actancy variations. In Malchukov, Andrej & Spencer, Andrew (eds.), *The Oxford handbook of case*, 339–355. Oxford: Oxford University Press.
- Michaelis, Susanne Maria. 2019. Support from creole languages for functional adaptation in grammar: Dependent and independent possessive person-forms. In Schmidtke-Bode, Karsten & Levshina, Natalia & Michaelis, Susanne Maria & Seržant, Ilja A. (eds.), *Explanation in typology*, 179–201. Berlin: Language Science Press. (<http://langsci-press.org/catalog/book/220>)
- Mosel, Ulrike. 2010. Ditransitive constructions and their alternatives in Teop. In Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard (eds.), *Studies in ditransitive constructions*, 486–509. Berlin: Mouton de Gruyter.
- Nedjalkov, Igor. 1997. *Evenki*. London: Routledge.
- Osam, E. Kweku. 1996. The object relation in Akan. *Afrika und Übersee* 79.
- Smith, Kenny & Culbertson, Jennifer. 2020. *Communicative pressures shape language during communication (not learning): Evidence from casemarking in artificial languages*. PsyArXiv. (<https://osf.io/5nwhq>)
- Stolz, Thomas & Levkovich, Nataliya & Urdze, Aina. 2017. When zero is just enough ... In support of a Special Toponymic Grammar in Maltese. *Folia Linguistica* 51(2). 453–482. (doi:[10.1515/flin-2017-0016](https://doi.org/10.1515/flin-2017-0016))