

How to split friends, wives and boxes of books, in Vietnamese

Tue Trinh

University of Wisconsin–Milwaukee

Universität Frankfurt, 5/2/2016

Abstract

I derive the distribution of “doubling” in NP split constructions in Vietnamese from the way nouns, classifiers, modifiers and measure words compose syntactically and semantically, in conjunction with a syntax-phonology mapping rule.

1 Introduction

1.1 The copy theory of movement

The copy theory of movement accounts for our intuition about “double interpretation” but leaves a puzzle concerning pronunciation (cf. Chomsky 1993, 1995, Gärtner 1998, Sauerland 1998, 2004, Fox 1999, 2000, 2002, Corver and Nunes 2007, among others).

- (1) a. which pictures of himself did John see
b. *which pictures of himself did John see which pictures of himself

The puzzle is “solved” by postulating a rule, Copy Deletion, which deletes the lower copy of a chain.

- (2) which pictures of himself did John see ~~which pictures of himself~~

This leads to the following question.

- (3) What is the structural description of Copy Deletion?

1.2 Doubling phenomena

In many languages one finds constructions in which double interpretation is accompanied by double pronunciation (cf. Nunes 2003, 2004, Fanselow and Mahajan 1995, Fanselow 2001, Grohmann 2003, Grohmann and Nevins 2004, Grohmann and Panagiotidis 2004, Hiraiwa 2005, Landau 2006, Martins 2007, Cheng 2007, Vicente 2005, 2007, 2009, Kandybowicz 2006, 2007, 2008, 2009, Trinh 2009, 2010, among others).

- (4) Verb topicalization
a. mua thì nó mua hoa (Vietnamese, Trinh 2009, 2010)
buy TOP she bought flowers
b. liknot hi kanta et ha-praxim (Hebrew, Landau 2006)
buy she bought the flowers

Moreover, doubling seems to exhibit properties of movement.

- (5) a. what did she buy before she left the market
b. *what did she leave the market after she bought

- (6) a. mua thì nó mua hoa trước khi nó rời chợ
buy TOP she bought flower before she left market
b. *mua thì nó rời chợ sau khi nó mua hoa
buy TOP she left market after she bought flowers

This leads to the following hypothesis.

- (7) Doubling is non-application of Copy Deletion

1.3 The Edge Condition

It follows from the hypothesis in (7) that studying the distribution of doubling, across languages as well as across constructions within a single language, may lead to a partial answer of question (3). Trinh (2009, 2010) proposes (8).

- (8) The Edge Condition (EC)
For any chain (α, β) where α is the higher and β is the lower copy, phonological deletion of β requires that β end an XP
(9) Definition
A constituent α ends an XP iff the rightmost morpheme of α co-incides with the rightmost morpheme of a non-projecting category

As an example, EC predicts the following contrast between Vietnamese and German.

- (10) a. mua thì nó nên *(mua) hoa
buy TOP she should buy flowers
b. kaufen sollte sie Blumen (*kaufen)
buy should she flowers buy

1.4 The goal of the talk

In this talk I examine the distribution of doubling in NP-split constructions in Vietnamese and try to explain it in terms of EC. The explanation will require certain assumptions to be made about both the syntax and the semantics of nominals in Vietnamese, hence constitute arguments for these assumptions.

2 NP-Split as noun topicalization

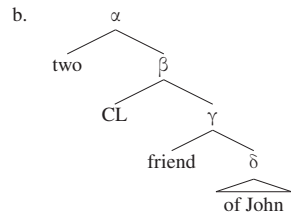
NP-Split constructions are those in which a subpart of a nominal complex is extracted from it, resulting in a “split” of the complex.

- (11) a. A book appeared about Chomsky
b. Bücher hat man damals interessante in den Osten keine mitnehmen dürfen
books has one then interesting in the East nó with-take may
‘As for books, one could not take any interesting ones to the East then’
(Fanselow and Ćavar 2002: 65, 67)

The nominal complexes in Vietnamese that are of interest to us consist of a numeral, a classifier, a noun, and possibly a complement or a modifier of the noun, in that order.

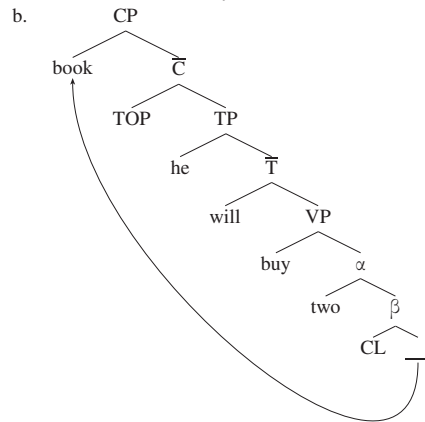
- (12) a. hai quyển sách
two CL book
b.
-

- (13) a. hai người bạn của John
two CL friend of John



The “split” involves topicalization of the noun from the nominal complex. I will assume that topicalization is A-movement to [Spec,C], and the topic marker **thì** is the C head (cf. Trinh 2007).

- (14) a. sách thì nó sẽ mua hai quyển
book TOP he will buy two CL



Noun topicalization obeys locality constraints observed for \bar{A} dependencies (Ross 1967, Chomsky 1977).

- (15) Clause unboundedness
a. tôi nghĩ [X rằng nó sẽ mua hai quyển sách]
‘I think that he will buy two CL book’
b. sách thì tôi nghĩ rằng [X nó sẽ mua hai quyển]
book TOP I think that he will buy two CL
- (16) Complex Noun Phrase Constraint
a. nó sẽ gặp [X một người có hai quyển sách]
he will meet one person have two CL book
‘He will meet a person who has two books’
b. *sách thì nó sẽ gặp [X một người có hai quyển]
book TOP he will meet one person have two CL
- (17) Adjunct Island Constraint
a. nó đi về [X sau khi nó mua hai quyển sách]
he go home after he buy two CL book
‘he went home after he bought two books’

- b. *sách thì nó đi về [X sau khi nó mua hai quyển]
book TOP he go home after he buy two CL

(18) Subject Island Constraint¹

- a. [X nó mua hai quyển sách] là tốt
he read two CL book be good
‘that he bought two books is good’
b. *sách thì [X nó mua hai quyển] là tốt
book TOP he buy two CL is good

(19) Non-bridge Verb Island Constraint

- a. nó thì thầm [X rằng nó sẽ mua hai quyển sách]
he whisper that he will buy two CL book
‘he whispered that he would buy two books’
b. *sách thì nó thì thầm [X rằng nó sẽ mua hai quyển]
book TOP he whisper that he will buy two CL

3 Relational vs. non-relational nouns

3.1 Preliminaries

I will assume that a non-relational noun such as **sách** ‘book’ or **mèo** ‘cat’ is of type $\langle e, t \rangle$, whereas a relational noun such as **bạn** ‘friend’ or **vợ** ‘wife’ is of type $\langle e, \langle e, t \rangle \rangle$.

- (20) Definition
a. $[[\text{sách}]] = [\lambda x.x \text{ is a book}]$
b. $[[\text{mèo}]] = [\lambda x.x \text{ is a cat}]$

- (21) Definition
a. $[[\text{bạn}]] = [\lambda y. [\lambda x.x \text{ is friends with } y]]$
b. $[[\text{vợ}]] = [\lambda y. [\lambda x.x \text{ is female and married to } y]]$

In addition, I will assume that the PP complement of a relational noun is of type e , which means the preposition has no semantic content.

- (22)
$$\left[\left[\begin{array}{c} \alpha \\ \text{friend} \\ \beta \\ \text{of John} \end{array} \right] \right] = \left[\left[\begin{array}{c} \alpha \\ \text{friend John} \end{array} \right] \right] = [\lambda x.x \text{ is friends with John}]$$

3.2 Obligatory cases

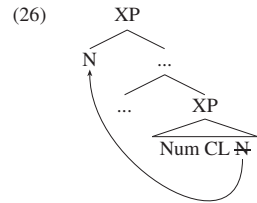
First observation: topicalization of a non-relational noun without modifiers results in obligatory non-doubling.

- (23) NP-Split Generalization 1
If N is non-relational: $N \dots \text{Num CL} (*N)$
- (24) a. sách thì nó sẽ mua hai quyển (*sách)
book TOP he will buy two CL book
b. mèo thì nó sẽ mua hai con (*mèo)
cat TOP he will buy two CL cat

This is predicted by EC, assuming the following economy principle.

¹ Note that subject sentences in Vietnamese are not introduced by an overt complementizer.

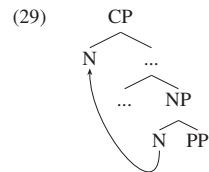
- (25) Pronunciation Economy (PE)
Copy Deletion must apply when it can



Second observation: topicalization of a relational noun with complement results in obligatory doubling.

- (27) NP-Split Generalization 2
If N is relational: N ... Num CL *(N) Compl
- (28) a. bạn thì nó sẽ gặp hai người *(bạn) của John
friend TOP he will meet two CL friend of John
b. vợ thì nó sẽ gặp hai người *(vợ) của John
wife TOP he will meet two CL wife of John

We predict this fact also.



Note that the contrasts in (28) can only be said to follow from EC if the topic noun and its double are related by movement. And there is evidence that they are.

- (30) *bạn thì nó bị ốm sau khi nó gặp một người bạn của John
friend TOP he got sick after he met one CL friend of John

Other examples can be given of other relational nouns and other island constraints, but for reason of space, I will not present them here.

3.3 Optional cases

3.3.1 Non-relational nouns

The generalization

Modifiers follow the head noun in Vietnamese, which means that even if N is non-relational and does not have a complement, it is still possible for N not to be the last constituent of the nominal complex.²

² The objection might be raised that the PP về vật lý 'about physics' which follows the noun sách 'book' in (32a) is a complement, not a modifier. An argument that the PP is actually a modifier is the fact that it can be placed after a copula verb and serve as a predicate of the noun, as (31a) shows. This is not possible for PPs which are clearly complements of nouns, as evidenced by the degradedness of (31b).

- (31) a. quyển sách này là về vật lý
CL book this is about physics
b. *người bạn này là của tôi
CL friend this is of me

- (32) a. nó sẽ mua hai quyển sách về vật lý
he will buy two CL book about physics
b. nó sẽ mua hai con mèo màu đen
he will buy two CL cat black

Third observation: topicalization of a non-relational noun which is followed by a modifier results in optional doubling.

- (33) a. sách thì nó sẽ đọc hai quyển (sách) về vật lý
book TOP he will read two CL book about physics
b. mèo thì nó sẽ mua hai con (mèo) màu đen
cat TOP he will buy two CL cat black
- (34) NP-Split Generalization 3
If N is non-relational: N ... Num CL (N) Mod

Excursus on classifiers and modifiers

In Vietnamese, a classifier must mediate between a numeral and an NP.

- (35) nó sẽ mua hai *(quyển) sách
he will buy two CL book

Following Chierchia (1998a,b, 2010), I assume that count nouns in languages such as Vietnamese denote "number neutral," i.e. cumulative, predicates.

- (36) [[sách]] = [$\lambda x.x$ is an atomic book or a plurality of books] = {a, b, c, a+b, a+c, b+c, a+b+c}

Suppose only atomic predicates can combine with numerals, we account for the obligatory presence of the classifier by assuming that the classifier maps a number-neutral predicate to a cumulative one, as follows.

- (37) [[CL]] = [$\lambda P \in D_{\langle e,t \rangle}. \lambda x \in D_e. x \in P \wedge x$ is atomic]

As for nominal modifiers, I assume that they are predicates of type $\langle e, t \rangle$ which are semantically integrated into the structure by way of Predicate Modification (Heim and Kratzer 1998).

- (38) Predicate Modification (PM)
If α and β are both of type $\langle e, t \rangle$, $[[\alpha \beta]] = [\lambda x. [[\alpha]](x) = [[\beta]](x) = 1]$

Deriving the generalization

From the definition of PM and CL it follows that both (39a) and (39b) are well-formed with respect to semantic type.

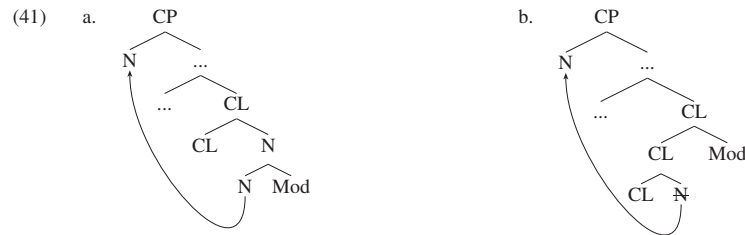
- (39) a. b.

I will assume that both structures in (39) are available in Vietnamese. The question now is what values should we give to α and β in (39). Suppose we make the following choices.

- (40) a. b.

Note that (31b) can be well-formed if we take it to mean something like 'this person, who is a friend of someone, belongs to me.' In this reading, the PP của tôi 'of me' is clearly not construed as the complement of bạn 'friend.'

We predict, then, that both (41a) and (41b) are possible.



EC and PE predicts Copy Deletion must apply in (41a) and cannot apply in (41b). Since the two structures are both available and semantically equivalent, NP-Split Generalization 3 is derived.

3.3.2 Relational nouns

The generalization

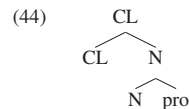
Fourth observation: when a relational noun without a complement is topicalized, doubling of the fronted noun is also optional.

- (42) a. vợ thì nó sẽ gặp hai người (vợ)
 wife TOP he will meet two CL wife
- b. bạn thì nó sẽ gặp hai người (bạn)
 friend TOP he will meet two CL friend

- (43) NP-Split Generalization 4
 If N is relational: N ... Num CL (N)

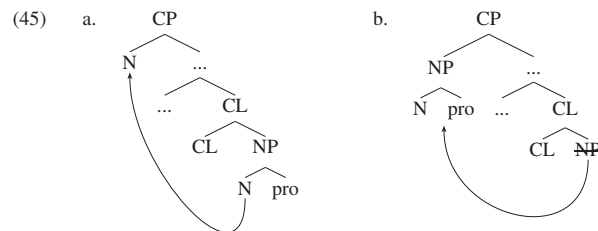
Deriving the generalization

Let us derive this fact in the following way. Suppose that the relational noun does have a complement, a silent pronoun, and consequently, that the classifier phrase has the structure in (44).



Note that **pro** is actually required: without it CL would not be able to compose with N due to type mismatch!

As **pro** is silent, topicalization of N will fill [Spec,C] with the same phonological material as topicalization of NP. Given EC, however, we expect that the phonological material at the base position will differ between these two cases: fronting N bleeds, while fronting NP feeds, Copy Deletion.

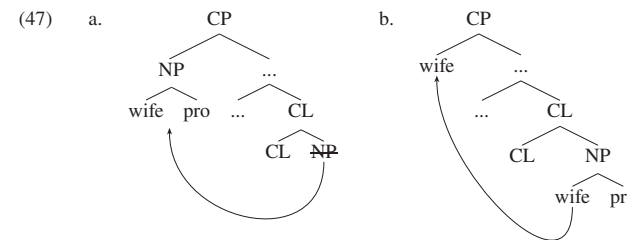


Independent evidence for **pro**

Note that although the topic constituents in (45a) and (45b) sound the same, they do not mean the same: the bare noun denotes a relation, the NP a predicate. Can an experiment be designed to show this difference?

- (46) a. Vợ thì nó gặp HAI người. Bạn thì nó gặp BA người.
 wife TOP he met two CL friend TOP he met three CL
- b. Vợ thì nó gặp HAI người vợ. Bạn thì nó gặp BA người bạn.
 wife TOP he met two CL wife friend TOP he met three CL friend
- c. #Vợ thì nó gặp HAI người. Bạn thì nó gặp BA người bạn.
 wife TOP he met two CL friend TOP he met three CL friend

EC predicts that the first and the second sentence of (46c) are (47a) and (47b), respectively.

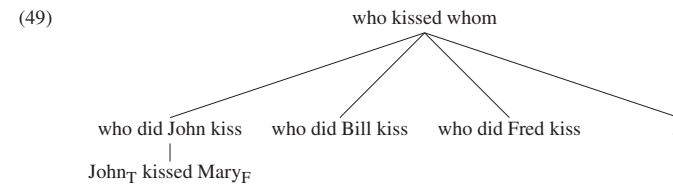


Question: why does the sequence consisting of (47a) followed by (47b) sound odd?

My answer will rely on the notion of “topic value,” as defined in Büring (1999, 2003), based on the notion of “focus value” of Root (1985, 1992, 1996).

- (48) a. $[[\text{John}_T \text{ kissed Mary}_F]]^o = \text{John kissed Mary}$
- b. $[[\text{John}_T \text{ kissed Mary}_F]]^f = \{\text{John kissed } y \mid y \in D_e\}$
- c. $[[\text{John}_T \text{ kissed Mary}_F]]^t = \{\{x \text{ kissed } y \mid y \in D_e\} \mid x \in D_e\}$

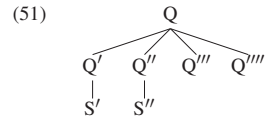
Given the proposition-set theory of questions (cf. Hamblin 1973), (48b) is the question ‘who did John kiss’ and (48c) the set containing such questions as ‘who did John kiss’, ‘who did Bill kiss’, ‘who did Fred kiss’, etc. These questions, in turn, can be seen as subquestions of a “superquestion,” namely ‘who kissed whom.’ This is represented in the following “discourse tree,” or “d-tree.”



I assume, following Büring (1999, 2003), that a sentence is felicitous only if it can be a node in an available d-tree, and that the following condition holds.

- (50) CT-Congruence
 A sentence S containing a topic and a focus can be a node in a d-tree D only if the question Q dominating S in D, and all of Q’s sisters, are elements of $[[S]]^t$

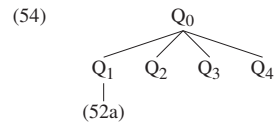
This means, for (51), that it has to be the case that $[[S']]^t = [[S'']]^t = \{Q', Q'', Q''', Q''''\}$.



Thus, (47a) instantiates (52) and (47b) instantiates (53).

- (52) a. he met two_F [wife of pro]_T
 b. $\{\{\text{he met } n \text{ P} \mid n \in \mathbb{N}\} \mid \text{P} \in D_{\langle e, t \rangle}\} = \{\text{how many wives of pro did he meet, how many friends of pro did he meet, how many linguistics students did he meet, how many female democrats did he meet}\}$
- (53) a. he met three_F [friend]_T of pro
 b. $\{\{\text{he met } n \text{ R of pro} \mid n \in \mathbb{N}\} \mid \text{R} \in D_{\langle e, e, t \rangle}\} = \{\text{how many friends of pro did he meet, how many wives of pro did he meet, how many children of pro did he meet, how many siblings of pro did he meet}\}$

Suppose that the utterance of (52a) reduces the set of “available d-trees” to those which are extensions of (54), i.e. which are derivable from (54) by plugging in the daughters of Q₂, Q₃ or Q₄.



It then follows that (53a) is infelicitous after (52a), as it cannot be a node in any of the available d-trees.

4 Measure words

In Vietnamese, container words such as **thùng** ‘box’ or **túi** ‘bag’ are systematically ambiguous between a “noun reading,” exemplified in (55a), and a “measure word reading,” exemplified in (55b).

- (55) a. John mua hai cái thùng
 John bought two CL box
 ‘John bought two boxes’
- b. John mua hai thùng sách
 John bought two MW_{box} book
 ‘John bought two boxes of books’

(56) $[[\text{MW}_{\text{box}}]] = [\lambda P \in D_{\langle e, t \rangle} . [\lambda x \in D_e . x \text{ is a box load of things that are P}]]$

Thus, measure words resemble classifiers in being functions of type $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$. However, they are, in a sense, more substantive than classifiers. This has interesting repercussions for modification.

- (57) a.
-
- b.
-

We predict, then, that (58) is ambiguous between ‘John will buy two boxes of large books’ and ‘John will buy two large boxes of books.’ This prediction is correct.

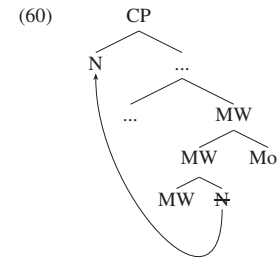
- (58) John mua hai thùng sách to
 John bought two MW_{box} book large
 ‘John bought two large boxes of books / John bought two boxes of large books’

Now consider the following sentence.

- (59) Sách thì John mua hai thùng to
 book TOP John bought two MW_{box} large
 ‘John bought two large boxes of books / *John bought two boxes of large books’

The observation is that (59) is unambiguous: it has the reading where the adjective modifies ‘boxes of books,’ but not the one where it modifies ‘books.’

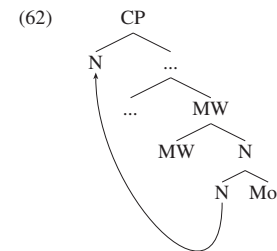
It turns out that this is exactly what we predict. Specifically, we predict, given EC, that the analysis of (59) must be (60).



Now let us consider the doubling variant of (59), i.e. the sentence which differs minimally from (59) in that the topicalized noun is doubled.

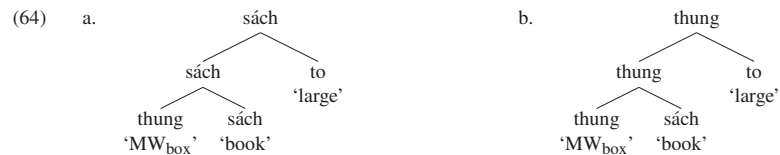
- (61) Sách thì John mua hai thùng sách to
 book TOP John bought two MW_{box} book large
 ‘John bought two large boxes of books / John bought two boxes of large books’

As the translation shows, (61) is ambiguous in exactly the same way as (58) is. This fact is not what we predict: we predict that (61) must have the structure in (62), and hence be unambiguous.



One solution is to assume that when a measure word and noun merge, projection is free choice: either the measure word or the noun can project. Thus, the string **thùng sách to** ‘MW_{box} book large’ would have four possible analyses.

- (63) a.
-
- b.
-



Among the four structures in (63) and (64), only one, namely (64b), is such that fronting **sách** 'book' from it will not result in doubling. The ambiguity of (61) follows from the possibility of analyzing it either as one of the structures in (63), or as (64a).

References

- Büring, Daniel. 1999. Topic. In *Focus — linguistic, cognitive, and computational perspectives*, ed. Peter Bosch and Rob van der Sandt, 142–165. Cambridge University Press.
- Büring, Daniel. 2003. On d-trees, beans, and b-accent. *Linguistics and Philosophy* 26:511–545.
- Cheng, Lisa Lai-Shen. 2007. Verb copying in Mandarin Chinese. In *The copy theory of movement*, ed. Norbert Corver and Jairo Nunes, 151–174. John Benjamins Publishing Company.
- Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of “semantic parameter”. In *Events and grammar*, ed. Susan Rothstein, 53–104. Kluwer Academic Publishers.
- Chierchia, Gennaro. 1998b. Reference to kinds across languages. *Natural Language Semantics* 6:339–405.
- Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174:99–149.
- Chomsky, Noam. 1977. On wh-movement. In *Formal syntax*, ed. P. Culicover, T. Wasow, and A. Akmajian. New York, New York: Academic Press.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. Ken Hale and Jay Keyser, 1–52. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1995. Categories and Transformations. In *The minimalist program*, chapter 4. MIT Press.
- Corver, Norbert, and Jairo Nunes, ed. 2007. *The Copy Theory of Movement*. John Benjamins Publishing Company.
- Fanselow, Gisbert. 2001. Partial movement. *University of Potsdam*.
- Fanselow, Gisbert, and Damir Čavar. 2002. Distributed deletion. In *Theoretical approaches to universals*, ed. Artemis Alexiadou, 65–107. John Benjamins Publishing Company.
- Fanselow, Gisbert, and Anoop Mahajan. 1995. Partial movement and successive cyclicity. In *Papers on Wh-Scope Marking*, ed. Uli Lutz and Gereon Müller, 131–161. University of Tübingen: Sprachtheoretische Grundlagen für die Computer Linguistik.
- Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.
- Gärtner, Hans-Martin. 1998. Review of “the copy theory of movement and linearization of chains in the Minimalist Program”. *GLOT International* 8:16–20.
- Grohmann, Kleanthes. 2003. *Prolific domains: On the anti-locality of movement dependencies*. John Benjamins.
- Grohmann, Kleanthes, and Andrew Nevins. 2004. Echo reduplication: When too-local movement requires pf-distinctness. *University of Maryland Working Papers in Linguistics* 13:84–108.
- Grohmann, Kleanthes, and Phoevos Panagiotidis. 2004. Demonstrative doubling in Greek. *University of Maryland Working Papers in Linguistics* 13:109–131.
- Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language* 10:41–53.
- Heim, Irene, and Angelica Kratzer. 1998. *Semantics in Generative Grammar*. Blackwell.
- Hiraiwa, Ken. 2005. Predicate clefts in Būli: Phase and category. *Linguistic Analysis* 32:544–583.

- Kandybowicz, J. 2007. Fusion and pf architecture. In *Proceedings of the 30th Annual Penn Linguistics Colloquium*, volume 13, 85–98.
- Kandybowicz, Jason. 2006. On fusion and multiple copy spell-out: The case of verb repetition. In *The copy theory of movement on the PF side*, ed. Jairo Nunes and Norbert Corver. Oxford University Press.
- Kandybowicz, Jason. 2008. *The grammar of repetition: Nupe grammar at the syntax-phonology interface*. John Benjamins Publishing Co.
- Kandybowicz, Jason. 2009. Embracing edges: syntactic and phono-syntactic edge sensitivity in Nupe. *Natural Language and Linguistic Theory* 27:305–344.
- Landau, Idan. 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9:32–66.
- Martins, Ana Maria. 2007. *Double realization of verbal copies in european portuguese emphatic affirmation*, 77–118. John Benjamins.
- Nunes, Jairo. 2003. *Head movement, remnant movement, and phonetic realization of chains*, 161–177. UCLA.
- Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.
- Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, UMASS-Amherst.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.
- Rooth, Mats. 1996. Focus. In *The handbook of contemporary semantic theory*, ed. Shalom Lappin, 271–297. Oxford: Blackwell Publishers.
- Ross, John. 1967. Constraints on Variables in Syntax. Doctoral Dissertation, MIT.
- Sauerland, Uli. 1998. The Meaning of Chains. Doctoral Dissertation, MIT.
- Sauerland, Uli. 2004. The interpretation of traces. *Natural Language Semantics* 12:63–127.
- Trinh, Tue. 2007. The case for no Case. MIT.
- Trinh, Tue. 2009. A constraint on copy deletion. *Theoretical Linguistics* 35:183–227.
- Trinh, Tue. 2010. Edges and linearization: A reply. *Theoretical Linguistics* 36:93–110.
- Vicente, Luis. 2005. Towards a unified theory of movement: an argument from spanish predicate clefts. *Leiden Papers in Linguistics* 2:43–67.
- Vicente, Luis. 2007. The syntax of heads and phrases: A study of verb (phrase) fronting. Doctoral Dissertation, Universiteit Leiden, Leiden, The Netherlands.
- Vicente, Luis. 2009. An alternative to remnant movement for partial predicate fronting. *Syntax* 12:158–191.